



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

MODELOS CHAIN-LADDER CLÁSICOS Y BAYESIANOS
PARA EL CÁLCULO DE RESERVAS AUXILIADO CON R

T E S I S

QUE PARA OBTENER EL TÍTULO DE
ACTUARIO

PRESENTA
DANIEL DE JESÚS ALFARO FUENTES.

DIRECTOR DE TESIS:
ACT. EDUARDO SELIM MARTÍNEZ MAYORGA.

Ciudad Universitaria, Cd. Mx., Diciembre 2016





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*A mis papás,
mis hermanos,
mis abuelos,
Mariana,
mi familia,
mi tutor,
mis amigos.*

Agradecimientos

Con el mayor de mis agradecimientos a mis padres por su apoyo incondicional...

Daniel de Jesús Alfaro Fuentes

Índice general

| | |
|---|-----------|
| Introducción | 1 |
| 1. El proceso de reservas | 5 |
| 1.1. Proceso de registro de siniestros | 5 |
| 1.1.1. Tipos de siniestros | 6 |
| 1.1.1.1. Siniestros ocurridos pero no reportados | 6 |
| 1.1.1.2. Siniestros reportados pero no liquidados | 6 |
| 1.1.2. Pérdidas pendientes de cumplir | 7 |
| 1.1.2.1. Algunos componentes del monto total de pérdidas | 7 |
| 1.2. Reservas | 8 |
| 1.2.1. Procedimiento de estimación de las reservas | 9 |
| 1.2.2. Estimación de las reservas de acuerdo a la Regulación Mexicana | 10 |
| 1.2.2.1. Solvencia II | 11 |
| 1.2.3. Tipos de Reservas | 12 |
| 1.2.3.1. Reservas Técnicas | 12 |
| 1.2.3.2. Reserva de Riesgos en Curso | 14 |
| 1.2.3.3. Reserva de Obligaciones Pendientes de Cumplir | 15 |
| 1.2.3.4. Reserva Catastrófica | 18 |
| 1.2.4. Método estatutario para reservas | 18 |
| 2. Modelos Chain-Ladder clásicos | 23 |
| 2.1. Tipos de datos para el estudio de reservas | 23 |
| 2.2. Métodos básicos de estimación/pronóstico | 30 |
| 2.2.1. Método de la razón | 31 |
| 2.2.2. Método Chain-Ladder (algorítmico) | 34 |
| 2.2.3. Método Bornhuetter-Ferguson | 41 |
| 2.2.3.1. Comparación entre los estimadores BF y CL | 43 |
| 2.3. Medición del error en modelos Chain-Ladder | 48 |
| 2.3.1. Error cuadrático medio de predicción | 48 |
| 2.3.1.1. Error cuadrático medio de predicción no-condicional | 49 |

| | | |
|-----------|--|-----------|
| 2.3.1.2. | Error cuadrático medio de predicción condicional | 50 |
| 2.3.2. | Modelo Chain-Ladder con componente de varianza | 51 |
| 2.3.2.1. | Estimación del proceso de varianza condicional | 57 |
| 2.3.2.2. | Estimación del error para años accidente individuales | 58 |
| 2.4. | Algunos modelos con hipótesis distribucionales | 70 |
| 2.4.1. | Modelo Poisson para número de reclamaciones | 70 |
| 2.4.1.1. | Obtención del algoritmo CL a partir del modelo Poisson | 74 |
| 2.4.2. | Modelo log-normal para reclamaciones acumuladas | 79 |
| 2.4.2.1. | Varianza σ_j^2 conocida | 82 |
| 2.4.2.2. | Varianza desconocida | 88 |
| 2.4.2.3. | Intervalo de confianza modelo log-normal para reclamaciones acumuladas | 89 |
| 2.4.3. | Reclamaciones incrementales | 91 |
| 2.4.3.1. | Modelo Poisson de sobredispersión | 91 |
| 2.4.3.2. | Intervalo de confianza del modelo Poisson para las reclamaciones incrementales, $X_{i,j}$ | 92 |
| 2.4.3.3. | Modelo binomial-negativo | 92 |
| 2.4.3.4. | Modelo log-normal | 94 |
| 2.4.3.5. | Intervalo de confianza del modelo log-normal para las reclamaciones incrementales, $X_{i,j}$ | 95 |
| 2.4.3.6. | Modelo Gamma | 96 |
| 3. | Modelos Chain-Ladder Bayesianos | 99 |
| 3.1. | Modelos Bayesianos | 99 |
| 3.1.1. | El método de Benktander-Hovinen y el modelo Cape-Cod | 99 |
| 3.1.1.1. | Método de Benktander-Hovinen | 99 |
| 3.1.1.2. | Modelo Cape-Cod | 102 |
| 3.1.2. | Métodos basados en credibilidad | 105 |
| 3.1.2.1. | Minimización de funciones de pérdida cuadráticas | 106 |
| 3.1.2.2. | Ejemplos distribucionales para la reserva de reclamaciones basados en credibilidad | 111 |
| 3.1.2.3. | Modelo log-normal/log-normal | 114 |
| 3.2. | Métodos <i>Bootstrap</i> | 116 |
| 3.2.1. | <i>Bootstrap</i> no-paramétrico de Efron | 116 |
| 3.2.1.1. | Intervalos de confianza | 119 |
| 3.2.2. | <i>Bootstrap</i> paramétrico | 119 |
| 3.2.3. | Modelo log-normal para reclamaciones acumuladas | 120 |
| 3.2.4. | Modelos lineales generalizados | 122 |
| 3.2.5. | Método <i>bootstrap</i> Chain-Ladder | 123 |
| 3.2.5.1. | Estimación de error no-condicional | 125 |
| 3.2.5.2. | Estimación de error condicional | 126 |

| | | |
|-----------|--|------------|
| 3.2.6. | Suposiciones matemáticas sobre métodos <i>bootstrap</i> | 126 |
| 3.3. | Métodos Bayesianos exactos | 129 |
| 3.3.1. | Modelo Poisson de sobredispersión con distribución Gamma a <i>priori</i> | 130 |
| 3.4. | Credibilidad Bayesiana para el modelo Chain-Ladder | 136 |
| 3.4.1. | Chain-Ladder Bayesiano | 138 |
| 3.4.2. | Credibilidad para el modelo Chain-Ladder | 144 |
| 3.4.3. | Credibilidad exacta para el modelo Chain-Ladder | 151 |
| 3.4.3.1. | Relación con el modelo Chain-Ladder clásico | 157 |
| 3.5. | Metodología Bayesiana de cálculo aproximado (ABC) | 160 |
| 3.5.1. | Modelo Chain-Ladder Bayesiano de distribución libre | 160 |
| 3.5.2. | Distribución predictiva completa y Valor en Riesgo | 164 |
| 3.5.3. | <i>Bootstrap</i> y el error cuadrático medio de predicción | 165 |
| 3.5.3.1. | <i>Bootstrap</i> no-paramétrico condicional | 165 |
| 3.5.3.2. | Estimación <i>bootstrap</i> frecuentista | 166 |
| 3.5.3.3. | Estimación Bayesiana del error cuadrático medio de predicción | 167 |
| 3.5.3.4. | Estimador de credibilidad para el error cuadrático medio de predicción | 168 |
| 3.6. | Metodología ABC (<i>Approximate Bayesian Computation</i>) para verosimilitudes complejas | 168 |
| 3.6.1. | Justificación del algoritmo MCMC-ABC | 170 |
| 4. | Generalizaciones del modelo Chain-Ladder de Mack | 173 |
| 4.1. | Introducción | 173 |
| 4.2. | Modelo Munich Chain Ladder | 173 |
| 4.2.1. | Introducción al método Munich Chain Ladder | 173 |
| 4.2.1.1. | El problema (P/I) en cálculos del Chain-Ladder separado | 174 |
| 4.2.1.2. | Correlación entre datos pagados e incurridos | 176 |
| 4.2.2. | El método Chain-Ladder de Mack | 177 |
| 4.2.3. | El modelo Munich Chain-Ladder | 179 |
| 4.2.3.1. | Análisis de las suposiciones del modelo Munich Chain-Ladder | 182 |
| 4.2.4. | Estimación de parámetros | 183 |
| 4.2.4.1. | Parámetros del Chain-Ladder por separado | 184 |
| 4.2.4.2. | Parámetros del Munich Chain-Ladder | 184 |
| 4.3. | Modelo Double Chain Ladder | 190 |
| 4.3.1. | Suposiciones del modelo Double Chain-Ladder | 190 |
| 4.3.1.1. | La extensión del modelo de Verrall <i>et. al.</i> (2010) | 191 |
| 4.3.2. | Algoritmo Double Chain-Ladder | 194 |
| 4.3.2.1. | Estimación del retraso de reporte | 195 |
| 4.3.2.2. | Estimación de los parámetros de la distribución de los pagos individuales | 196 |
| 4.3.3. | Estimadores DCL para las reservas RBNS e IBNR | 197 |

| | | |
|--|--|------------|
| 4.3.4. | Estimación de las reclamaciones pendientes ignorando colas | 198 |
| 4.3.5. | Método Double Chain-Ladder-Bornhutter&Ferguson (DCL-BF) | 199 |
| 4.3.5.1. | Pronóstico de las reservas IBNR y RBNS en el modelo DCL-BF | 201 |
| 4.4. | Modelo Chain-Ladder continuo | 208 |
| 4.4.1. | El problema de estimación de la densidad para datos “agregados” | 209 |
| 4.4.2. | Histograma con suavizamiento de <i>kernel</i> | 212 |
| 4.4.3. | El estimador lineal local no estructurado en las observaciones del triángulo | 214 |
| 4.4.4. | El estimador de la densidad lineal local | 214 |
| 4.4.5. | Corrección multiplicativa sesgada | 216 |
| 4.4.6. | Elección del grado de suavizamiento | 216 |
| 4.4.7. | Predicción/pronóstico de las reclamaciones pendientes | 217 |
| Conclusión | | 219 |
| Apéndices | | 221 |
| A. Aspectos generales del algoritmo Metropolis-Hastings | | 223 |
| A.1. | Cadenas de Markov | 223 |
| A.2. | Algoritmo Metropolis-Hastings | 224 |
| A.2.1. | El algoritmo Monte Carlo para cadenas de Markov | 224 |
| A.2.2. | El algoritmo Metropolis-Hastings independiente | 225 |
| A.3. | Muestreo de Gibbs | 225 |
| A.3.1. | Muestreo Gibbs de dos etapas | 226 |
| A.3.2. | Muestreo Gibbs multi-etapas | 226 |
| B. Familias conjugadas | | 229 |
| C. Librería Chain-Ladder de R | | 235 |
| C.1. | Triángulos de datos | 236 |
| C.1.1. | Importación de triángulos de otras fuentes | 238 |
| Referencias | | 241 |

Introducción

Desde el surgimiento de los seguros, las primeras compañías aseguradoras tuvieron la necesidad de constituir una provisión, que se conoce como reserva, que hiciera frente a las futuras reclamaciones del asegurado, por tal motivo el actuario, además de calcular la prima que la aseguradora cobra en la póliza de seguro, desarrolla métodos estadísticos que le permita (a partir de datos históricos proporcionados por la aseguradora, como primas, conteos de siniestros, siniestros incurridos, siniestros pagados, flujos de efectivo, etc.), hacer un estimado de los siniestros futuros a los que la aseguradora está obligada a hacerse cargo.

Todavía en la década de 1970, la estimación de la reserva se realizaba sólo para reclamaciones conocidas, la técnica consistía en que un actuario evaluaba individualmente cada siniestro en la todas las etapas de su existencia. Esa estimación individual sería agregada al total de la reserva de obligaciones pendientes de cumplir. Con el paso del tiempo, y la creciente capacidad de desarrollar nuevas herramientas que pudiesen someter los resultados de este proceso a un análisis estadístico, se crearon otros métodos de aproximación de la reserva que fuesen más apropiados. Sin embargo, todavía existen casos en algunas líneas de negocios, donde el número de siniestros son pequeños, o donde sea particularmente complejo la conjunción con otros métodos.

Existen por lo menos 30 diferentes métodos analíticos en los que los actuarios se pueden basar en el momento de estimar la reserva de reclamaciones, y éstos son los responsables de elegir los métodos con los que se trabajará así como de interpretar y evaluar los resultados según su experiencia.

Uno de los primeros métodos que comenzó a utilizarse para la estimación de la reserva fue el Poisson para conteo de siniestros pero, por su estructura, en ocasiones no resultaba apropiado si se presentaban flujos negativos en las reclamaciones, a partir de aquí surge el método Chain-Ladder (CL) que actualmente se considera como uno de los más populares métodos de estimación de reclamaciones futuras.

Evidentemente, una de las primeras preguntas que surge es, por qué este método es el más común.

La respuesta rápida es por la facilidad de uso. Pero otra razón (en principio), es simplemente por costumbre.

Muchas personas creen que el método Chain-Ladder al ser un algoritmo determinista no genera una buena estimación de la reserva pero en el transcurso del Capítulo 2 de este trabajo se estudiará qué, bajo ciertas suposiciones, el método Chain-Ladder sí tiene una génesis estocástica y de aquí profundizar su desarrollo, i.e. por ejemplo, estimar el error de estimación, la varianza o bajo la misma metodología pero con diferentes suposiciones dar lugar a nuevos métodos, como el Bornhuetter-Ferguson, o bien con suposiciones distribucionales y a partir de estas suposiciones estimar sus parámetros por métodos estadísticos clásicos, así como dar intervalos de confianza para dichos parámetros.

Uno de los objetivos será (de acuerdo a los resultados) decidir cuál método es la mejor estimación de la reserva ó si efectivamente el método CL clásico es la mejor opción.

En el capítulo 3, se hace uso de metodologías Bayesianas para estudiar métodos como el Cape Code, Benktander-Hovinen y se analizan métodos de varianza mínima como los de credibilidad, además se describirá el conocido algoritmo recursivo *bootstrap*, que es de gran utilidad para el muestreo de los datos y la estimación de su distribución. Finalmente, se analizará el método *approximate Bayesian computation* el cual consiste en realizar simulaciones para estimar la distribución predictiva.

El capítulo 4, extiende ideas de los capítulos 2 y 3 ya que se estudiarán métodos, como el Munich Chain-Ladder, Double Chain-Ladder y Chain-Ladder continuo, basados en la estructura Chain-Ladder pero que requieren de información adicional, ya que se trabajará con dos tipos de datos que se describirán más adelante.

Muchos de estos métodos tienen dos versiones, para siniestros incurridos y para pagados, y puede pasar que cada estimación tenga una diferencia en el resultado, será ahí cuando el actuario decidirá el monto de siniestralidad última que la aseguradora tendrá que registrar en sus estados financieros. Por lo tanto, la pregunta sería como obtener este monto si hasta el momento sólo se habla de métodos teóricos. Una de la soluciones que se encontró para este problema es dar una representación práctica, auxiliándose del software estadístico R y una base de datos incluida dentro de las librerías del software, donde al ejecutarlo, realizar una comparación numérica sobre cual método sería el más conveniente de utilizar.

Tanto la reserva para riesgos en curso como la de obligaciones pendientes de cumplir son las provisiones más importantes en una aseguradora y su cálculo debiese ser “exacto”, pero esto dependerá de los datos obtenidos y de los criterios que por experiencia el actuario decida ejecutar.

Por su parte, la última reforma a la regulación mexicana se dio precisamente para incorporar un sistema que le permitiese a la institución reguladora tener más control sobre el límite en la capacidad de solvencia de una compañía aseguradora, así como de comprobar que los métodos, parámetros y suposiciones utilizados en la estimación de la reserva sean los correctos.

En el capítulo 1, se da una descripción del tipo de datos necesarios para la estimación de la reserva y como se clasifican, también se explica las diferentes reservas que, según la actual regulación mexicana, son obligatorias de constituir, por lo tanto también se revisa el método estatutario para el cálculo de las reservas proporcionado por la Comisión Nacional de Seguros y Fianzas.

Capítulo 1

El proceso de reservas

1.1. Proceso de registro de siniestros

Una póliza de seguro es un contrato entre dos partes: el **asegurado**, que mediante el pago de una remuneración, la **prima**, recibe a cambio, por parte del **asegurador**, la promesa de cubrir los daños económicos que pueda causar la ocurrencia de un evento fortuito.

Al conjunto de pagos de siniestros, pagos por posibles pérdidas y siniestros pagados sobre los que el asegurador está obligado a hacerse cargo parcial o totalmente se le conoce como **monto de pérdidas**.

Para que un siniestro se considere parte del monto de pérdidas la compañía debió ser informada por el asegurado y ésta tuvo que saldar el monto reclamado.

La Figura 1.1, representa una línea del tiempo del proceso que se sigue al reportar la ocurrencia de un siniestro. Donde el tiempo t_1 , es la fecha exacta en la que ocurrió el siniestro y t_2 es el momento en el que se le notificó a la aseguradora.

En general, los tiempos t_3 y t_4 muestran como es común que el monto del siniestro no se pague inmediatamente, esto debido a los retrasos administrativos por parte de la aseguradora, diferentes cotizaciones al siniestro por parte del ajustador o los trámites lentos por los que pasa el asegurado. Por ejemplo, cuando el pago es por la reposición de ingresos y es necesario un comprobante, o también cuando el daño es a terceros y existe una demanda al respecto. Así, nuevos retrasos podrían surgir de la investigación, disputa, litigio o cualquier otro proceso que tenga relación con la determinación del monto del siniestro.

La Figura 1.1 indica que al tiempo t_5 , después de varios pagos del siniestro, la aseguradora considera que el proceso de reclamación está completo y da por concluido su expediente.

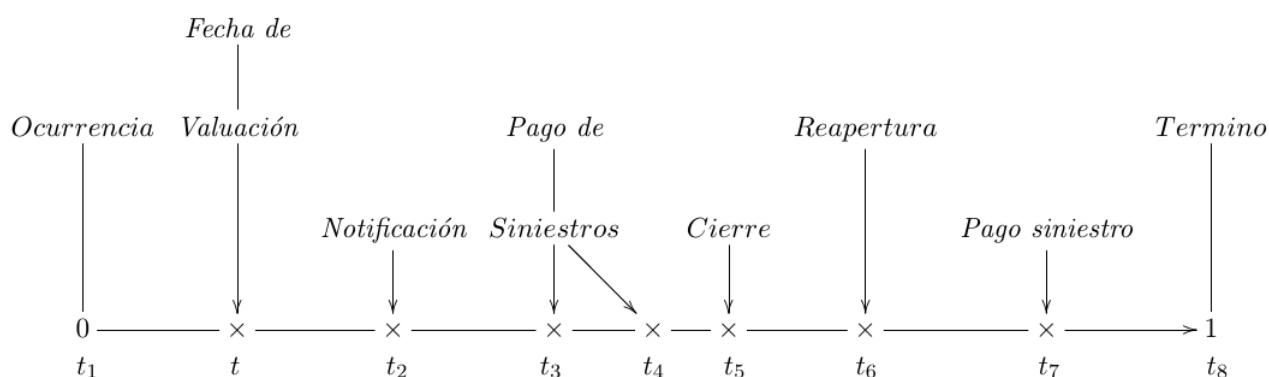


Figura 1.1: Línea del tiempo de un siniestro

Pero hay casos (t_6) en los que se considera que el monto del siniestro no fue correcto, por lo que se continúa con la investigación y el caso es reabierto; esto podría generar más pagos (t_7), finalmente si no existe ninguna otra acción que afecte al monto definido se da por terminado el evento (t_8).

1.1.1. Tipos de siniestros

1.1.1.1. Siniestros ocurridos pero no reportados

En el intervalo de tiempo (t_1, t_2) , la aseguradora ya es responsable del monto del siniestro ocurrido -ya que es la promesa que le hizo al asegurado- pero ésta aún no es consciente de la existencia del evento. Durante este intervalo, se dice que el siniestro fue ocurrido pero todavía no se reporta a la aseguradora, por tal motivo la aseguradora debe estar preparada para este suceso acumulando una provisión o reserva que más adelante se describirá.

Por ejemplo, supóngase que se realiza una valuación del siniestro al tiempo $t \geq t_1$ (siguiendo la secuencia de la Figura 1.1). Si $t \geq t_2$, entonces el siniestro será conocido por la aseguradora, y su aportación consistirá en la estimación de los pagos restantes. Para valores mayores que t , esta estimación podría ser cero. Por ejemplo, podría ser el caso que $t = t_8$ o cuando $t_5 \leq t \leq t_6$. Si se considera el caso $t_1 < t < t_2$, se tiene un siniestro ocurrido pero no reportado.

1.1.1.2. Siniestros reportados pero no liquidados

De acuerdo a la Figura 1.1, obsérvese el intervalo de tiempo (t_2, t_7) donde la aseguradora ya fue informada del siniestro, a partir de esto comienza la estimación del monto a pagar (t_3, t_4), en t_5 se realizó un primer cierre del monto final estimado, pero para t_6 el caso fue reabierto y nuevas estimaciones surgieron (t_7), mientras este lapso de tiempo transcurre se

dice que el siniestro fue reportado pero aún no es liquidado al asegurado. En general, ningún siniestro es liquidado inmediatamente esto se debe a las diferentes estimaciones sobre el monto final.

Los siniestros reportados pero no liquidados se pagan de la prima devengada y difieren de los incurridos pero no reportados en la forma en que han sido “reportados” a la aseguradora, pero coinciden en que ninguno de los dos han sido completamente saldados durante el cierre del periodo contable.

1.1.2. Pérdidas pendientes de cumplir

Las pérdidas pendientes de cumplir son el monto de siniestros por los que el asegurador se hará cargo en un tiempo futuro a partir de la fecha de corte. Se refiere a los siniestros futuros estimados para calcular el monto total de pérdidas.

1.1.2.1. Algunos componentes del monto total de pérdidas

Los flujos de efectivo son uno de los componentes que influyen en la estimación de siniestros futuros con periodo de origen menor al último año contable, ya que de éstos depende el pago de dichos siniestros. Por tanto, existen procedimientos de proyecciones financieras utilizados para generar estimaciones de estos flujos de efectivo en valores monetarios indexados a una fecha de evaluación particular.

Ajustes a los flujos de efectivo se pueden requerir debido a:

- **Gastos de ajuste** considerados en dos categorías:
 - **ALAE** (*Allocated Loss Adjustment Expenses*): Gastos médicos asociados, gastos de peritos, abogados externos, identificados en cada siniestro. Están incluidos en el monto de siniestros.
 - **ULAE** (*Unallocated Loss Adjustment Expenses*): Gastos internos de ajuste, sistema de manejo de siniestros, salarios, etc. No están incluidos en el monto del siniestro y se tienen que estimar por separado.
 - **Rendimientos de inversiones:** Las pérdidas pendientes de cumplir pueden disminuir si la aseguradora obtiene un rendimiento positivo de inversiones futuras. Es decir, si al invertir parte de su capital genera ganancias, sumado al total de ingresos obtenidos en una fecha contable, será suficiente para cumplir con los flujos de efectivo de los pasivos.
-

- **Importes recuperables:** El pago de siniestros de la aseguradora es común compensarlo con “recuperaciones”. Por ejemplo, de los reaseguradores¹, o de terceras personas que ejercieron culpabilidad para la ocurrencia del evento, o en el caso de bienes materiales, el valor de salvamento. Entonces, el monto total de pérdidas de la aseguradora comprenderá el pasivo neto fuera de estas recuperaciones.

Algunos tipos de recuperaciones se consideran como contribuciones negativas para los siniestros acumulados, y por lo tanto no es necesario separarlos al momento de la valuación. Sin embargo, hay casos en los que no pueden ser incluidos al siniestro y se deben considerar por separado.

1.2. Reservas

El marco teórico de seguros y los componentes que son necesarios para llevar a cabo el proceso de reserva se describieron anteriormente, ahora, a partir del procedimiento de reclamaciones se dará una descripción de cómo surge el problema de reservar.

Cómo ya se mencionó antes, existe un retraso entre el evento o siniestro y la fecha de liquidación lo que significa que el asegurador deberá constituir una reserva respecto a esas reclamaciones que aún no han sido liquidadas. El tiempo que tarda un siniestro (Figura 1.1) en registrarse contablemente, hace que la estimación de la reserva no sea calculada con precisión ya que el flujo de efectivo está variando continuamente lo que provoca alteraciones en la estimación. En consecuencia, se necesita información adicional sobre la suficiencia de la reserva. Por lo tanto, se debe realizar un cuidadoso análisis sobre los métodos y supuestos utilizados para las estimaciones.

La reserva requerida en ese momento es el recurso necesario para cubrir los costos de los daños ocasionados por el asegurado a medida en que se van liquidando. La aseguradora deberá poder cuantificar estas responsabilidades si evalúa correctamente su posición financiera, para ambos propósitos, normativo e interno.

La necesidad de calcular el monto de las reservas está motivada de diferentes circunstancias dentro de la operación del seguro. Por ejemplo:

- Evaluar la situación financiera de una compañía de seguros, ya que los movimientos en las reservas en más de un período es esencial para conocer su desarrollo.

¹El *reasegurador* es una compañía a la cual el asegurador por medio de un contrato transfiere una parte o la totalidad de sus riesgos, así como una parte de la prima, a cambio de devolver el monto por los siniestros ocurridos de los riesgos en posesión.

- Tarificación en distintas líneas de negocios, en el sentido de estimar el costo de siniestros futuros en riesgos que aún no han considerado en su cartera.
- Evaluar la solvencia de la aseguradora, en términos de la probabilidad de cumplimiento de la obligaciones (i.e., se requiere una evaluación de las pérdidas máximas probables).
- Poner un precio al valor neto de la compañía de seguros, en particular para el caso de ventas o adquisiciones.
- Contratos de reaseguro por cerrar, es decir, poner un valor financiero al *run-off* de la cartera de negocios de seguros.

A partir de estas circunstancias, el análisis de reservas se convirtió en un área de mucha importancia ya que las cantidades de dinero que se constituyen representan el patrimonio de las compañías aseguradoras, y en su conjunto forman parte significativa del PIB de un país, esto fue uno de los motivos para desarrollar un plan en el que cada componente involucrada en esta transacción financiera juegue un papel importante.

Las compañías aseguradoras tienen la obligación de realizar una “buena” estimación de la reserva ya que es una de las cantidades más importantes en sus estados financieros y deberá maximizar su viabilidad y rentabilidad, por lo que, además de la experiencia en el momento de su estimación es importante un análisis científico riguroso.

El agente regulador, interesado en la solvencia de las aseguradoras, fija una serie de normas estrictas. Dichas normas se tienen que cumplir y apegarse al momento de calcular la reserva, con la finalidad de reducir la probabilidad de incumplimiento.

El agente fiscal, a cargo de obligar a mantener actualizada la declaración de los ingresos obtenidos de las compañías aseguradora, entonces ésta deberá estimar la reserva de tal modo que refleje los pagos reales o lo más parecido a ellos.

El asegurado es el más interesado en conocer la solvencia de una aseguradora en el transcurso de la vigencia de la póliza, por lo tanto, puede apoyarse de la información transparente que los agentes, regulador y fiscal, obligan a las compañías aseguradoras a publicar o basarse de estudios realizados por empresas privadas dedicadas a otorgar calificaciones crediticias para empresas financieras y así convencerse de contratar el seguro.

1.2.1. Procedimiento de estimación de las reservas

La evaluación/estimación de las reservas, en general, involucra los siguientes pasos:

1. Construir un modelo de estimación, donde se establezcan hipótesis y suposiciones que se quiera que cumplan.

2. Ajustar el modelo antes propuesto, utilizando observaciones pasadas.
3. Evaluar el modelo ajustado con las suposiciones, aceptarlo, rechazarlo y/o ajustarlo un número considerable de veces.
4. Utilizar el modelo para hacer predicciones sobre futuras estadísticas de interés.
5. Aplicar la experiencia profesional para la interpretación y selección de los resultados.

El diseño de un “modelo” apropiado implica la selección de uno entre todos los modelos posibles. Naturalmente, lo ideal sería elegir un modelo tan simple como sea posible. Desafortunadamente, en la operación de la industria aseguradora, el acceso a los datos es muy restringido, lo que reduce el número de modelos que se pueden aplicar, pues no sería apropiado el uso de un modelo que no sea respaldado por los datos disponibles.

1.2.2. Estimación de las reservas de acuerdo a la Regulación Mexicana

Actualmente en México, una reserva se refiere a la mejor estimación de las pérdidas pendientes esperadas (*BEL, Best Estimate Liability*). Generalmente se estima como el promedio de reclamaciones futuras, pero supóngase que se considera un conjunto de reclamaciones con distribución normal, entonces la media de este conjunto de variables correspondería a decir que la aseguradora tendrá una solvencia máxima del “50 %” del monto total de los siniestros esperados, lo ideal sería estimar una reserva más robusta que les permita la solvencia económica a las compañías aseguradoras, por tal motivo se optó por agregar un margen de riesgo (*MR*). Este margen de riesgo se refiere a la posible desviación en eventos desfavorables, por lo tanto, el margen de riesgo será el monto que garantice que la reserva sea suficiente. Entonces

$$\text{Reserva} = BEL + MR.$$

Cuando un actuario determina el monto de las reservas, está calculando el monto de los pasivos de la aseguradora y a su vez la utilidad. Además, al determinar las reservas de obligaciones sumadas las indemnizaciones ya pagadas, se genera información que el actuario necesita para calcular las primas de los siguientes años.

En el año 2013, en México se implementa una nueva regulación, Solvencia II, con la finalidad de supervisar que las instituciones dedicadas a las operaciones de seguros y fianzas cumplan con las obligaciones que acordaron con los asegurados. El cumplimiento de tales obligaciones consiste fundamentalmente en hacer frente a las reclamaciones futuras que hagan los asegurados, por lo tanto, la aseguradora debe contar con los recursos financieros

suficientes para solventar los siniestros futuros. El principal recurso con que cuentan las aseguradoras para tales efectos son las reservas técnicas, por lo que es fundamental establecer criterios generales para la constitución de reservas en las compañías de seguros.

1.2.2.1. Solvencia II

Solvencia II fue una iniciativa que surgió en la Unión Europea para establecer un esquema en la administración de riesgos de las compañías de seguros y reaseguros, a través de la definición del requerimiento de capital de solvencia, así como la instalación de procesos y procedimientos para identificar, medir y gestionar los niveles de riesgo asumidos. Su objetivo es el desarrollo y establecimiento de un sistema que permita medir los recursos necesarios, para garantizar la solvencia de una aseguradora en función de los riesgos asumidos por ésta.

Solvencia II se diseñó sobre tres pilares: el primero consiste en un proceso de análisis de las reservas, activos y pasivos necesarios para cubrir las obligaciones aceptadas en las pólizas, así como cuantificar los requerimientos de capital para enfrentar los riesgos asumidos; el segundo se ocupa de definir las reglas de supervisión, control interno y gobierno corporativo; y el tercero busca establecer las obligaciones de información que las aseguradoras deberán presentar al mercado.

De forma más general:

El Pilar I es de naturaleza cuantitativa y se puede decir que se ocupa de tres principales elementos: Determinación de Fondos Propios, valuación de Reservas y requerimientos de Capital.

El esquema de Solvencia II está basado en la valuación económica del riesgo y el capital de las aseguradoras, lo que llevará a las aseguradoras a aplicar principios económicos cuando calculen el capital obligatorio y sus fondos propios. La diferencia entre el Valor de Mercado de los Activos (VMA) y el Valor de Mercado de los Pasivos (VMP) dará como resultado los Fondos Propios, los cuales deben ser suficientes para cubrir el Requerimiento de Capital de Solvencia.

El Requerimiento de Capital de Solvencia busca garantizar que habrá recursos patrimoniales suficientes para hacer frente a los riesgos y responsabilidades asumidas, en función de las operaciones y los riesgos a los que esté expuesta la institución.

El Pilar II pretende la promoción de estándares mejorados y consistentes de gestión de riesgos. Estos requerimientos tendrán un gran efecto en las diversas actividades o aspectos de gestión de las instituciones. Los aspectos clave afectados drásticamente son los siguientes:

- Sistema de gestión y seguimiento del riesgo.
-

- Estrategia y apetito al riesgo.
- Autoevaluación del Riesgo y la Solvencia (ORSA por sus siglas en inglés).
- Función de control interno.
- Papel relevante de la función actuarial.
- El uso de la externalización (outsourcing).

La mayoría de los expertos consideran que cumplir con los requerimientos del Pilar II, será un reto mucho mayor para las aseguradoras que el contar con un modelo para la determinación de los requerimientos de capital.

El Pilar III busca la transparencia a través de la publicación de información periódica sobre la situación financiera y de solvencia de la compañía. Supone establecer una disciplina de mercado para todas las instituciones de seguros y reaseguros, cuyo objetivo último es el apoyo para obtener metas regulatorias. Éste es un elemento muy importante dentro de Solvencia II y las compañías deben prepararse para publicar la información en base a la periodicidad y detalle establecido por el regulador.

La gestión de la transparencia y el reporte de riesgos al mercado configuran un aspecto que las entidades deben analizar con un enfoque amplio, ya que su éxito se basa en gran medida en el grado en que se integren internamente los diferentes elementos que conforman la gestión de riesgos.²

1.2.3. Tipos de Reservas

1.2.3.1. Reservas Técnicas

Las reservas técnicas al ser el pasivo más importante de una aseguradora, deberán evaluar sus activos que las respaldan a precio de mercado. En general, no existe un mercado financiero cuyos instrumentos permitan replicar los flujos de estos pasivos.

Las reservas técnicas, en Solvencia II, hace referencia al monto que las aseguradoras pagarían por transferir de inmediato las obligaciones a otra empresa, entendido como el valor presente esperado de los flujos futuros, tomando en cuenta el valor del dinero en el tiempo, más un margen de riesgo (costo de capital). Deben de ser consistentes con la valuación de sus activos y pasivos en el mercado y deben establecerse con respecto a todas las obligaciones de los asegurados y beneficiarios de contratos de seguros.

²Comité de Solvencia II de Amis. (2014). *Solvencia II*. Asociación Mexicana de Instituciones de Seguros 11-29.

La Ley de Instituciones de Seguros y Fianzas (LISF) establece que el valor de las reservas técnicas será igual a la suma del mejor estimador de obligaciones futuras (BEL) y un margen de riesgo (MR), evaluados por separado.

El BEL se define como el valor esperado de los flujos futuros del portafolio de riesgos, entendido como la media ponderada por la probabilidad de dichos flujos, considerando el valor temporal del dinero con base en las curvas de tasas de interés libres de riesgo de mercado. Por su parte, el Margen de Riesgo representa el costo de asegurar que el capital requerido estará disponible para mantener las obligaciones de seguros para los años subsecuentes, es decir, será el monto que, sumado al BEL, garantice que el monto total de las reservas técnicas sea el equivalente al que las aseguradoras requerirán para asumir y hacer frente a sus obligaciones.

Por lo tanto, la suma de ambos elementos indicará el valor del portafolio de riesgos asegurados, dado que incluye los flujos de posibles ingresos y egresos asociados al negocio suscrito y el costo del capital necesario para continuar la operación del mismo hasta su extinción.

Formalmente la LISF establece

“ARTÍCULO 216.- Las Instituciones de Seguros deberán constituir las siguientes reservas técnicas:

- I. *Reservas de riesgos en curso;*
- II. *Reservas para obligaciones pendientes de cumplir;*
- III. *Reserva matemática especial;*
- IV. *Reserva para fluctuación de inversiones;*
- V. *Reserva de contingencia;*
- VI. *Reserva de riesgos catastróficos”*³

Cuando los flujos futuros asociados a las obligaciones de seguro y de reaseguro, puedan replicarse utilizando instrumentos financieros con un valor de mercado directamente observable, el valor de las reservas técnicas respectivas se determinará a partir del valor de mercado de dichos instrumentos financieros. En tales casos, no será necesario calcular por separado la mejor estimación y el margen de riesgo.

³Ley de Instituciones de Seguros y Fianzas

1.2.3.2. Reserva de Riesgos en Curso

La reserva de riesgos en curso se puede definir técnicamente como la parte de la prima que se debe utilizar para el cumplimiento de las obligaciones futuras por concepto de reclamaciones, a lo que también se le llama “Prima no Devengada”.

Tiene por objeto que en el estado de resultados se reflejen los ingresos, es decir, el efecto de las primas devengadas al final de un ejercicio contable. Se deducen las primas cedidas en reaseguro, para representar solamente la porción retenida por la aseguradora. Además, se excluyen los reaseguros de riesgos catastróficos por que éstos se destinan a cubrir la retención del asegurador. Cuando se constituye la reserva se hace un cargo a los gastos y un abono a la cuenta de reservas. El abono, en realidad no es un gasto, porque no representa una pérdida de valor de ningún activo. El saldo acreedor que se genera con la constitución de la reserva no es un pasivo, porque la empresa no espera desprenderse de esos recursos.

Esta definición se traduce como una serie de cálculos actuariales que pueden resultar de mayor o menor complejidad dependiendo del tipo de seguro, por lo que para hacer una mejor explicación es importante hacer la distinción de los diferentes tipos de seguros y referirnos a los métodos de constitución de reservas correspondientes a cada uno de ellos. Es importante también aclarar que cuando se trata de seguros de vida o pensiones es más común llamar a esta reserva, “Reserva Matemática”, en tanto que, para los seguros de daños, así como de salud, se le conoce como “Reserva de Riesgos en Curso”.

Según la LISF

“ARTÍCULO 217.- Las reservas técnicas a que se refiere el artículo 216 de esta Ley, tendrán como propósito:

- I. *En el caso de las reservas de riesgos en curso, cubrir el valor esperado de las obligaciones futuras derivadas del pago de siniestros, beneficios, valores garantizados, dividendos, gastos de adquisición y administración, así como cualquier otra obligación futura derivada de los contratos de seguro.*
 - a) *En la operación de vida:*
 - 1) *Para los seguros de vida con temporalidad mayor a un año;*
 - 2) *Para los seguros de vida con temporalidad menor o igual a un año;*
 - 3) *Para los seguros de pensiones o de supervivencia relacionados con la edad, jubilación o retiro de personas bajo esquemas privados complementarios a la seguridad social, y*
 - 4) *Para los seguros de pensiones derivados de las leyes de seguridad social;*
 - b) *Para los seguros comprendidos en la operación de accidentes y enfermedades, y*
-

c) *Para los seguros comprendidos en la operación de daños;*³

Seguro de Daños

Los seguros de daños son en su mayor parte de vigencia anual o inferior, por lo que la constitución de la reserva de riesgos en curso es similar y puede llevarse a cabo de manera exacta, póliza por póliza. También se pueden establecer fórmulas de cálculo de reserva que en lugar de días sean por meses, trimestres, quincenas o semanas, para el devengamiento de la reserva.

Es importante aclarar que, el cobro de una prima única en la cual están incluidos los costos de administración adquisición y margen de utilidad futuros. En este caso el asegurado no tiene obligación futura de pago de primas por lo cual la reserva se constituye sólo con el valor esperado de las obligaciones futuras de la aseguradora, que corresponde a las primas no devengadas que deberán ser guardadas para el pago de siniestros de años futuros o devueltas al asegurado en caso de que el bien asegurado desaparezca.

Otro aspecto que es importante aclarar es que el pago fraccionado de la prima no modifica el procedimiento de cálculo de la reserva, por lo que el cálculo debe realizarse como si la prima se hubiese pagado en su totalidad al inicio de vigencia del plan, sin embargo, para evitar una distorsión en la contabilidad se debe permitir a la aseguradora que reconozca en estados financieros el saldo deudor de las fracciones de primas que faltan por cobrar.

1.2.3.3. Reserva de Obligaciones Pendientes de Cumplir

Según la LISF,

II. . . . *“En el caso de las reservas para obligaciones pendientes de cumplir, cubrir el valor esperado de siniestros, beneficios, valores garantizados o dividendos, una vez ocurrida la eventualidad prevista en el contrato de seguro. Las reservas para obligaciones pendientes de cumplir serán:*

- a) *Por pólizas vencidas y siniestros ocurridos pendientes de pago;*
- b) *Por dividendos y repartos periódicos de utilidades;*
- c) *Por siniestros ocurridos y no reportados, así como por los gastos de ajuste asignados a los siniestros, y*
- d) *Por las operaciones de que trata la fracción XXI del artículo 118 de la presente Ley;”*

Cuando un asegurado sufre un riesgo y este riesgo está amparado por la póliza que contrató, la compañía de seguros comienza con el servicio correspondiente generando procesos operativos, administrativos y contables, los cuales toman una gran relevancia para

la compañía ya que esto le permitirá cumplir con sus pasivos denominados “Obligaciones pendientes por cumplir”.

La reserva de obligaciones pendientes de cumplir es una provisión contable en los estados financieros de una compañía de seguros que cuantifica los pasivos de las pérdidas aún no cubiertas o liquidadas, de todas aquellas reclamaciones de seguros que hayan sido comunicadas o que se hayan generado pero aún no estén declaradas, i.e., corresponde a la suma de dos reservas distintas, por una parte, la reserva de siniestros pendientes de pago, y por otra, la reserva por siniestros ocurridos y no reportados; ambas reservas constituyen obligaciones pendientes de cumplir. Sin embargo, su metodología de constitución es distinta. Se trata de una reserva que se ha establecido para prever la futura responsabilidad de los siniestros que se han producido, pero que aún no han sido liquidados.

Las reservas para afrontar los siniestros futuros o pendientes de pago se generan de las primas ya devengadas para la aseguradora, es decir, aquellos periodos de aseguramiento que ya fueron pagados por el asegurado y que no incurrió en ningún siniestro, por lo que con esta, la aseguradora es capaz de resolver y liquidar todas las reclamaciones que son causadas por estas pólizas de seguro.

En la primera parte, se distinguieron dos categorías de siniestros ocurridos para una fecha en particular (fecha de valuación), que son los siniestros ocurridos pero no reportados y los siniestros ocurridos pero no liquidados, y corresponden a la estimación de la reserva de siniestros incurridos pero no reportados (IBNR, *Incurred But Not Reported*) y la reserva de siniestros reportados pero no liquidados (RBNS, *Reported But Not Settled*), respectivamente.

Reserva de Siniestros Ocurridos Pendientes de Pago

La reserva de siniestros ocurridos pendientes de pago es, junto con la reserva de riesgos en curso, esencial para asegurar la solvencia de la compañía de seguros.

Esta reserva se constituye con los saldos de las obligaciones derivadas de reclamaciones recibidas por la compañía y que por alguna razón no se han liquidado o pagado. En los casos más comunes, se constituye cuando la institución tiene conocimiento de la ocurrencia de un siniestro, pero su valor no ha quedado totalmente definido por el ajustador o no se ha logrado determinar si la reclamación es procedente. Cuando se origina el siniestro, la compañía, por práctica de mercado o por norma regulatoria, debe constituir la reserva provisional correspondiente a la reclamación, hasta en tanto se realice el ajuste respectivo.

La reserva de siniestros ocurridos pendientes de pago, no tiene un procedimiento de cálculo específico, ya que ésta se constituye con los saldos estimados por los ajustadores

de siniestros que hayan ocurrido y se tengan que liquidar en el futuro. Asimismo, dentro de esta reserva deben incluirse otros gastos ligados al siniestro, como son gastos de ajuste, intereses por mora, devoluciones de primas, penalizaciones, entre otros.

Reserva de Siniestros Ocurridos y No Reportados

Los siniestros ocurridos pero no reportados son eventos que ocurren en un intervalo de tiempo, durante la vigencia de la póliza, pero que se conocen con posterioridad a la fecha de cierre o de valuación de un periodo contable.

Así, la reserva por siniestros ocurridos y no reportados (IBNR por sus siglas en inglés “Incurred But Not Reported”) es una reserva técnica que corresponde al pasivo que se produce cuando los siniestros que ocurren en un determinado año, y que por diversas causas no son reclamados en el mismo, sino en años posteriores, con la consecuente obligación que debe reconocerse en estados financieros.

El problema es cómo construir una reserva de reclamaciones de las que aún no se tiene conocimiento, de lo único que se tiene certeza es que tales reclamaciones llegaran y se debe contar con una provisión para estas dado que en el futuro pueden manifestarse. En esta etapa, se puede disponer de datos como pagos, número de siniestros y montos incurridos reportados que describan un patrón de desarrollo.

El proceso de “reportar” no es instantáneo. Las reclamaciones generalmente se dan a través de sucursales, intermediarios o agentes de seguros. Entonces, habrá un intervalo de tiempo entre el primer reporte y el proceso en ser registrado en la base de datos de la oficina central.

Para el caso de los seguros de Responsabilidad Civil, para la determinación el monto total que se pagará se requieren períodos posteriores a la terminación de la vigencia, incluyendo largos procesos legales (como se especifique en cada caso). Por lo que estos siniestros se llaman de “cola larga”. Es importante notar que la legislación sobre responsabilidad civil de cada país influye en que tan pesada es la “cola” de cada siniestro.

Pasada la etapa de reporte, los datos obtenidos son utilizados para construir estimadores de pagos futuros y este consiste de un triángulo de reclamaciones incrementales agrupadas por año de origen y tiempo de desarrollo, el objetivo es completar el triángulo hasta formar un cuadro.

Existen diversos métodos clásicos como razón, Chain-Ladder, Bornhuetter-Fergusson, Poisson para conteo de reclamaciones, bootstrap (que se estudiarán más adelante) para la constitución de esta reserva, sin embargo, el problema esencial que hay que resolver es el

riesgo asociado con el futuro.

Reserva de gastos de ajuste al siniestro

Se considera como reserva de gastos de ajuste al siniestro de una institución de seguros, al monto por la atención de los siniestros ocurridos y no reportados, tales como honorarios de abogados y de ajustadores externos contratados para tal propósito, obtenido conforme a la valuación trimestral que realice para cada una de las operaciones y ramos indicados en la LISF.

1.2.3.4. Reserva Catastrófica

Solvencia II, contempla la existencia de un requerimiento de capital de solvencia por riesgos catastróficos. Por su parte, la LISF incluye en los requerimientos de suscripción de No Vida, un requerimiento de capital para eventos extremos.

“VI. En el caso de la reserva de riesgos catastróficos, cubrir el valor de la pérdida máxima probable derivada de la ocurrencia de siniestros de naturaleza catastrófica.”

La LISF define también las Reservas Catastróficas y su operación se manda a normatividad secundaria, debido a que éstas se determinan sobre la base de que corresponden a un nivel de pérdida en un percentil superior al valor esperado (Pérdida Máxima Probable). En este mismo sentido también se contempla, que pueden existir riesgos que por su naturaleza requieran niveles de confianza distintos al 99.5 % que fija el regulador.

1.2.4. Método estatutario para reservas

Dentro de la regulación mexicana si el método o modelo utilizado para el cálculo de la reserva de riesgos en curso y obligaciones pendientes por cumplir no es el correcto, las compañías tienen un tiempo de prórroga para corregirlo y ajustarlo de no ser así la Comisión Nacional de Seguros y Fianzas (CNSF) asigna un método estatutario desarrollado en la Circular Única de Seguros y Fianzas (CUSF) al cual se tienen que apegar y considerar métodos, parámetros y estimaciones sin rechazos.

A continuación se dará una descripción del método estatutario para el cálculo de la reserva de riesgos en curso:

- I) Se determinarán, por cada ramo o tipo de seguro , con la información estadística del mercado, los índices de reclamaciones futuras, mediante la metodología siguiente:
-

- a) Se calculará el monto de primas emitidas en cada año calendario i (PE_i), entendiendo como prima emitida, al monto de las primas emitidas que provienen de las pólizas que iniciaron su vigencia en el año calendario i y que hayan estado en vigor en dicho año calendario;
- b) Se determinarán y clasificarán los montos brutos de las reclamaciones, así como los montos de dividendos, ajustes, salvamentos y recuperaciones, registrados por año de origen i y año de desarrollo j ($S_{i,j}$),
- c) Se calcularán los índices de siniestralidad registradas por año de origen i y año de desarrollo j ($IS_{i,j}$), como el cociente que resulte de dividir el monto de las reclamaciones $S_{i,j}$, entre el monto de prima emitida en el año calendario i . Esto es:

$$IS_{i,j} = \frac{S_{i,j}}{PE_i}.$$

- d) Se obtendrá una estadística de índices de siniestralidad, identificados por ramo o tipo de seguro, en su caso, años de origen y de desarrollo, con los cuales se llevará a cabo la simulación de reclamaciones, dividendos, ajustes, salvamentos y recuperaciones futuras para efectos de la valuación de la reserva de riesgos en curso y/o para la reserva para obligaciones pendientes de cumplir por siniestros ocurridos no reportados y de gastos de ajuste asignados al siniestro.
- II) Se realizará la simulación y estimación de reclamaciones futuras para efectos de la valuación de la reserva de riesgos en curso y/o de la reserva para obligaciones pendientes de cumplir por siniestros ocurridos no reportados y de gastos de ajuste asignados al siniestro, de una Institución de Seguros mediante el siguiente procedimiento:

- a) Mediante los índices de siniestralidad, calculados anteriormente, se simularán los siniestros futuros de la aseguradora, provenientes de cada uno de los años de origen i y año de desarrollo j ($\hat{S}_{i,j}$), calculados como el producto del índice de siniestralidad, elegido aleatoriamente del conjunto de índices ($IS_{i,j}^{sim}$), por el monto de primas emitidas en el año. Esto es

$$\hat{S}_{i,j} = IS_{i,j}^{sim} * PE_i.$$

- b) Con el monto de las reclamaciones simuladas por cada año de origen, se estimará el valor de las reclamaciones totales para cada año de origen i , como la suma de las reclamaciones tanto conocidas como simuladas:

$$\hat{S}_i = \sum_j S_{i,j} + \hat{S}_{i,j}.$$

- c) Con las reclamaciones totales simuladas se obtendrá una estadística de índices de siniestralidad última ($IS_{i,j}$), dividiendo, en cada simulación, el valor de las

reclamaciones totales provenientes de un determinado año de origen , entre la prima emitida en dicho año. Es decir:

$$IS_{i,J} = \frac{\widehat{S}_i}{PE_i}.$$

- d) Con base en la estadística de índices de siniestralidad última indicada en el inciso anterior, se determinará la mejor estimación del índice de siniestralidad última ($\widehat{IS}_{i,J}^{RRC}$), como:

$$\widehat{IS}_{i,J}^{RRC} = \frac{1}{N * n} \left(\sum_{k=1}^N IS_{1,J}^k + \sum_{k=1}^N IS_{2,J}^k + \cdots + \sum_{k=1}^N IS_{I,J}^k \right),$$

donde:

- N es el número de escenarios simulados;
- k se refiere a cada uno de los escenarios simulados;
- I es el número de años de origen considerados en la simulación, y
- $IS_{i,J}^k$ es el valor del índice de siniestralidad última obtenido en cada escenario k , para el año de origen i .

Nota: la simulación del índice de siniestralidad última de la aseguradora, determinado conforme al procedimiento señalado en el inciso c) anterior, deberá considerar el número necesario de iteraciones para asegurar que la mejor estimación de dicho índice no difiera en más del 1.0 % de su verdadero valor.

- III) Se determinará para cada ramo, el porcentaje de gastos de administración, α , como el promedio ponderado de los porcentajes de gastos de administración que resulten de dividir, para cada año de origen, los gastos anuales de administración observados en cada aseguradora, entre los montos correspondientes de prima emitida.
- IV) La reserva de riesgos en curso (RRC) que deberá constituir la Institución de Seguros, será la que se obtenga de multiplicar la prima de tarifa no devengada ($PTND$) de cada póliza en vigor por el índice de siniestralidad última total ($\widehat{IS}_{i,J}^{RRC}$), más el porcentaje de gastos de administración, y al resultado se le sumará el margen de riesgo (MR) calculado conforme a lo establecido en el Capítulo 5.4 de las presentes Disposiciones⁴. Es decir,

$$RRC = PTND(\widehat{IS}_{i,J}^{RRC} + \alpha) + MR.$$

Observación 1.2.1.

⁴Circular Única de Seguros y Fianzas

Se entenderá como prima de tarifa no devengada, a la prima de tarifa que corresponda a la póliza que se trate, multiplicada por la proporción de tiempo de vigencia no transcurrido. En los casos en que el riesgo no disminuya proporcionalmente en el tiempo o no mantenga proporcionalidad respecto de la prima, la CNSF asignará, caso por caso, un criterio específico para definir la forma en que deberá realizarse el devengamiento de obligaciones, dependiendo de las características de cada riesgo. ∇

El procedimiento para la reserva de obligaciones pendientes por cumplir es similar ya que, con las reclamaciones simuladas conforme a lo señalado en el inciso a) anterior, se obtendrá una estadística de índices de siniestralidad última que capture el efecto de los flujos de siniestros ocurridos no reportados o que no hayan sido completamente reportados, así como sus gastos de ajuste, salvamentos y recuperaciones ($IS_{i,J}^{IBNR}$), dividiendo, en cada simulación, el monto de aquellos flujos provenientes de un determinado año de origen que se pagan en años de desarrollo posteriores al año de su ocurrencia, entre la prima emitida en dicho año de origen.

A partir de esta estadística, se determinará la mejor estimación del índice de siniestralidad última ($\widehat{IS}_{i,J}^{IBNR}$), como el valor medio de dichos índices.

Se determinará la reserva para obligaciones pendientes de cumplir por siniestros ocurridos no reportados y de gastos de ajuste asignados al siniestro ($RIBNR$), como el monto que se obtenga de multiplicar la prima emitida devengada (PTD) en cada uno de los últimos cinco años de operación de la Institución de Seguros, por el índice de siniestros ocurridos pero no reportados o que no hayan sido completamente reportados, así como sus gastos de ajuste ($\widehat{IS}_{i,J}^{IBNR}$), y por el factor de devengamiento correspondiente a cada año, más el margen de riesgo calculado conforme a lo establecido en estas Disposiciones. Es decir:

$$RIBNR = \sum_{t=1}^5 \left(PTD_t * \widehat{IS}_{i,J}^{IBNR} \right) * FD_t^{IBNR} + MR.$$

Capítulo 2

Modelos Chain-Ladder clásicos

En este capítulo se describirán algunos métodos algorítmicos básicos para la estimación de la reserva, como lo son el método de la razón, el Chain-Ladder clásico y a partir de éste, bajo ciertas suposiciones, el Bornhuetter-Ferguson.

Además de presentar la forma mecánica del modelo Chain-Ladder, se harán suposiciones, condicionales y no condicionales que reflejen la forma estocástica que subyace de este modelo, y así poder medir el error de estimación y la varianza en las predicciones obtenidas con el método.

Finalmente, bajo hipótesis distribucionales en las observaciones, se obtendrán estimadores e intervalos de confianza de los parámetros de cada distribución así como los estimadores del monto total de siniestralidad última, para el cálculo de la reserva.

2.1. Tipos de datos para el estudio de reservas

En esta sección se introducirá la notación clásica para el cálculo/estimación de la reserva de reclamaciones pendientes de cumplir. Donde, generalmente se estudia a partir del desarrollo de triángulos que separan en dos ejes de tiempo los siniestros pagados en los que la aseguradora ha participado.

Entonces, los datos del desarrollo de las reclamaciones se clasifican por periodo de exposición (año de origen) y periodo de desarrollo. Un año de origen se refiere al año en el que la póliza fue suscrita, emitida o el periodo de ocurrencia de un siniestro. Es claro que el periodo de origen no necesariamente es anual, es decir, puede ser trimestral, mensual, etc. El periodo de desarrollo de un año de origen se le conoce como edad o retraso y es el tiempo que tarda una reclamación en ser saldada. Los datos en las diagonales del triángulo son los pagos en el año calendario que se tome como punto inicial.

Observación 2.1.1.

La información de pólizas individuales se homogeneiza en una misma línea de negocio, ramo, riesgo, etc. Es decir, en estos triángulos generalmente hay información siniestral de un ramo, un tipo particular de asegurados o cualquier criterio que haga que intuitivamente se tenga el mismo tipo de datos. ∇

En este tipo de triángulos puede haber información con respecto al número de siniestros ocurridos, monto de los siniestros o montos de los siniestros pagados. Los métodos que se desarrollarán en este trabajo se harán para cualquiera de estos tipos de datos y en caso de que un método se desarrolle para un tipo en particular se hará explícita dicha especificación. Se utilizará simplemente el término “reclamación” para referirse a cualquiera de estos tres tipos de datos.

A continuación se presenta un esquema con dicha estructura triangular.

| Año de accidente i | Año de desarrollo j | | | | | | | | | |
|----------------------|-----------------------|---|---|---|---|-----|-----|-----|-----|--|
| | 0 | 1 | 2 | 3 | 4 | ... | j | ... | J | |
| 0 | | | | | | | | | | |
| 1 | | | | | | | | | | |
| ⋮ | | | | | | | | | | |
| $I + 1 - J$ | | | | | | | | | | |
| ⋮ | | | | | | | | | | |
| ⋮ | | | | | | | | | | |
| i | | | | | | | | | | |
| ⋮ | | | | | | | | | | |
| ⋮ | | | | | | | | | | |
| $I - 2$ | | | | | | | | | | |
| $I - 1$ | | | | | | | | | | |
| I | | | | | | | | | | |

Observaciones $C_{i,j}, X_{i,j}$
($i + j \leq I$)

Predicciones $\tilde{C}_{i,j}, \tilde{X}_{i,j}$
($i + j > I$)

Notación 2.1.1. (Notación Clásica)

Para las siguientes secciones y capítulos se utilizará la siguiente notación:

- $X_{i,j}$: son los siniestros incurridos en el periodo de desarrollo j con año de accidente i , i.e., $X_{i,j}$ corresponde a las reclamaciones en el año de accidente i realizados j periodos posteriores a la ocurrencia del evento o en el año contable $i + j$.

- $C_{i,j}$: son las reclamaciones acumuladas en el año de accidente i después de j años de desarrollo

$$C_{i,j} = \sum_{k=0}^j X_{i,k},$$

i.e., son todos los siniestros incurridos en el año i con un retraso de liquidación de j periodos.

- i : año de accidente, año de ocurrencia (eje vertical), $i \in \{0, 1, \dots, I\}$, donde I es el año de accidente más reciente.
- j : año de desarrollo, periodo de desarrollo (eje horizontal), $j \in \{0, 1, \dots, J\}$, donde J es el último año de desarrollo.

▽

En el triángulo de desarrollo de siniestros, los años de accidente, i , se especifican en el eje vertical y los periodos de desarrollo, j , en el eje horizontal. Al tiempo I , el cuadro de desarrollo de siniestros se dividen en dos partes: el triángulo/trapezoide superior que contiene todas las observaciones $X_{i,j}$ ($i + j \leq I$) y el triángulo inferior con estimados o valores pronosticados de las obligaciones pendientes de cumplir $\tilde{X}_{i,j}$ ($i + j > I$).

Definición 2.1.1. (*Triángulo de Desarrollo de Siniestros*)

Se define al conjunto de reclamaciones acumuladas

$$\mathcal{D}_I = \{X_{i,j} : i + j \leq I, j \in \{0, 1, \dots, J\}\}, \quad (2.1.1)$$

como el triángulo de datos observados (triángulo superior) en los años de desarrollo j .

$$\mathcal{D}_I^c = \{\tilde{X}_{i,j} : i + j > I, i \leq I, j \leq J\}, \quad (2.1.2)$$

es el triángulo de siniestros futuros a estimar/predecir (triángulo inferior).

Las reclamaciones incrementales en el año contable $k \geq 0$ se denotan por $X_k := \sum_{i+j=k} X_{i,j}$. Nótese por ejemplo que $X_1 = X_{1,0} + X_{0,1}$ y $X_2 = X_{2,0} + X_{0,2} + X_{1,1}$, es decir, las reclamaciones incrementales en el año contable k se muestran en la $(k + 1)$ -ésima diagonal del triángulo de desarrollo de siniestro.

Las **reclamaciones incrementales** $X_{i,j}$, representan los pagos incrementales en la celda (i, j) , el número de siniestros reportados con un retraso de j periodos de reporte y

año de accidente i , ó el ajuste en la cantidad de siniestro reportada/pagada en la celda (i, j) .

Las **reclamaciones acumuladas** $C_{i,j}$, representan los siniestros pagados acumulados, el total de reclamaciones reportadas ó los siniestros incurridos (para reclamaciones acumuladas reportadas).

$C_{i,J}$ se conoce como la **el monto de siniestralidad última** del año de accidente i ó cantidad total de siniestros en el año i .

| Reclamaciones Incrementales | | Reclamaciones Acumuladas |
|---|--------|---|
| $X_{i,j}$ pagos incrementales | \iff | $C_{i,j}$ pagos acumulados |
| $X_{i,j}$ número de siniestros reportados con retraso j | \iff | $C_{i,j}$ número total de siniestros reportados |
| $X_{i,j}$ ajuste en la cantidad del siniestro reportado | \iff | $C_{i,j}$ siniestros incurridos |

Como $X_{i,j}$ denota cantidades incrementales, entonces las pérdidas pendientes (*outstanding loss liabilities*) para el año de accidente i al tiempo j están dadas por

$$R_{i,j} := \sum_{k=j+1}^J X_{i,k} = C_{i,J} - C_{i,j},$$

estas $R_{i,j}$ se necesitan pronosticar y se les conoce como **reserva de obligaciones pendientes de cumplir**. Constituye una cantidad que se tiene que estimar a partir de las reclamaciones pasadas $C_{i,j}$ y se obtiene un estimador/predictor para la siniestralidad última, $C_{i,J}$, del año de accidente i .

El paquete `ChainLadder` de R es una serie de rutinas estadísticas que proporciona varios métodos que generalmente se utilizan para la estimación de la reserva de obligaciones pendientes por cumplir, así como pronosticar el desarrollo de siniestros futuros.

Implementar en R algunos de los métodos para el cálculo de reservas tiene varias ventajas, por ejemplo:

- Un amplio lenguaje para modelos estadísticos y manejo de datos permitiendo una rápida aplicación del modelo.
- Distintas interfaces para bases de datos y otras aplicaciones, como Excel.
- Facilidad de construcción de funciones para crear reportes de investigación en combinación con otras herramientas como L^AT_EX.

A continuación se ejemplifica cómo hacer uso de la paquetería `ChainLadder` tomando como ejemplo la base de datos llamada “GenIns” que es un triángulo (matriz de 10×10) con datos de los siniestros acumulados de una aseguradora para la operación de daños.

Código de R

```
R>data(GenIns) #base de datos
R> GenIns
  dev
origin  1    2    3    4    5    6    7    8    9    10
  1  357848 1124788 1735330 2218270 2745596 3319994 3466336 3606286 3833515 3901463
  2  352118 1236139 2170033 3353322 3799067 4120063 4647867 4914039 5339085    NA
  3  290507 1292306 2218525 3235179 3985995 4132918 4628910 4909315    NA    NA
  4  310608 1418858 2195047 3757447 4029929 4381982 4588268    NA    NA    NA
  5  443160 1136350 2128333 2897821 3402672 3873311    NA    NA    NA    NA
  6  396132 1333217 2180715 2985752 3691712    NA    NA    NA    NA    NA
  7  440832 1288463 2419861 3483130    NA    NA    NA    NA    NA    NA
  8  359480 1421128 2864498    NA    NA    NA    NA    NA    NA    NA
  9  376686 1363294    NA    NA    NA    NA    NA    NA    NA    NA
 10 344014    NA    NA    NA    NA    NA    NA    NA    NA    NA
```

Código de R

El triángulo muestra los valores de las pérdidas conocidas a partir de cada año de origen y de las evaluaciones anuales posteriores. Por ejemplo, las pérdidas expuestas en el año de origen 3 son 290507, 1292306, 2218525 para los años de desarrollo 1,2,3 respectivamente. La última diagonal, i.e., el vector 3901463, 5339085, 4909315,...,344014, muestra la evaluación más reciente (de arriba hacia abajo).

El objetivo de este trabajo es pronosticar el desarrollo de futuras reclamaciones al completar la esquina inferior derecha del triángulo y posibles desarrollos. En cierto punto todas las reclamaciones se tendrán que saldar, pero no siempre se conoce el tiempo que en que esto se lleve a cabo (sobretudo en seguros de cola larga).

La librería `ChainLadder` contiene dos funciones útiles para la construcción de los triángulos de desarrollo, `cum2incr` y `incr2cum` para transformar el triángulo de reclamaciones acumuladas al triángulo de reclamaciones incrementales y viceversa, respectivamente.

Código de R

```
R>#Transformación de un triángulo de reclamaciones acumuladas
R<#a un triángulo de siniestros incrementales
R> GenIns.inc <- cum2incr(GenIns)
R>#Transformación de un triángulo de con siniestros incrementales
R>#a un triángulo de siniestros acumulados
R> GenIns.cum <- incr2cum(GenIns.inc)
R>#Reclamaciones Incrementales
R> GenIns.inc

  dev
```

```

origin      1      2      3      4      5      6      7      8      9     10
  1  357848  766940  610542  482940  527326  574398  146342  139950  227229  67948
  2  352118  884021  933894  1183289  445745  320996  527804  266172  425046    NA
  3  290507  1001799  926219  1016654  750816  146923  495992  280405    NA    NA
  4  310608  1108250  776189  1562400  272482  352053  206286    NA    NA    NA
  5  443160  693190  991983  769488  504851  470639    NA    NA    NA    NA
  6  396132  937085  847498  805037  705960    NA    NA    NA    NA    NA
  7  440832  847631  1131398  1063269    NA    NA    NA    NA    NA    NA
  8  359480  1061648  1443370    NA    NA    NA    NA    NA    NA    NA
  9  376686  986608    NA    NA    NA    NA    NA    NA    NA    NA
 10  344014    NA    NA    NA    NA    NA    NA    NA    NA    NA

R> GenIns.inc[1,] #Muestra el desarrollo del primer año del triángulo incremental

      1      2      3      4      5      6      7      8      9     10
357848 766940 610542 482940 527326 574398 146342 139950 227229 67948

R> GenIns.cum[1,] #Muestra el desarrollo del primer año del triángulo acumulado

      1      2      3      4      5      6      7      8      9     10
357848 1124788 1735330 2218270 2745596 3319994 3466336 3606286 3833515 3901463

```

Código de R

Además, se puede generar un resumen de cada periodo de desarrollo, como:

```

Código de R
R> summary(GenIns)
      1      2      3      4      5
Min.   :290507   Min.   :1124788   Min.   :1735330   Min.   :2218270   Min.   :2745596
1st Qu.:346040   1st Qu.:1236139   1st Qu.:2159608   1st Qu.:2941786   1st Qu.:3474932
Median :358664   Median :1292306   Median :2187881   Median :3235179   Median :3745390
Mean   :367139   Mean   :1290505   Mean   :2239043   Mean   :3132989   Mean   :3609162
3rd Qu.:391271   3rd Qu.:1363294   3rd Qu.:2268859   3rd Qu.:3418226   3rd Qu.:3939263
Max.   :443160   Max.   :1421128   Max.   :2864498   Max.   :3757447   Max.   :4029929
      NA's :1      NA's :2      NA's :3      NA's :4
      6      7      8      9     10
Min.   :3319994   Min.   :3466336   Min.   :3606286   Min.   :3833515   Min.   :3901463
1st Qu.:3873311   1st Qu.:4307785   1st Qu.:4257800   1st Qu.:4209908   1st Qu.:3901463
Median :4120063   Median :4608589   Median :4909315   Median :4586300   Median :3901463
Mean   :3965654   Mean   :4332845   Mean   :4476547   Mean   :4586300   Mean   :3901463
3rd Qu.:4132918   3rd Qu.:4633649   3rd Qu.:4911677   3rd Qu.:4962692   3rd Qu.:3901463
Max.   :4381982   Max.   :4647867   Max.   :4914039   Max.   :5339085   Max.   :3901463
NA's   :5      NA's :6      NA's :7      NA's :8      NA's :9

```

Código de R

El primer acercamiento con el triángulo de datos será gráficamente. La siguiente gráfica muestra el desarrollo de las reclamaciones que se tuvo en cada año de origen, identificado por cada curva con su respectivo número, el eje de las abscisas corresponde a los periodos de desarrollo y el eje de las ordenadas a los montos acumulados de las reclamaciones.

```

Código de R
R> plot(GenIns)

```

Código de R

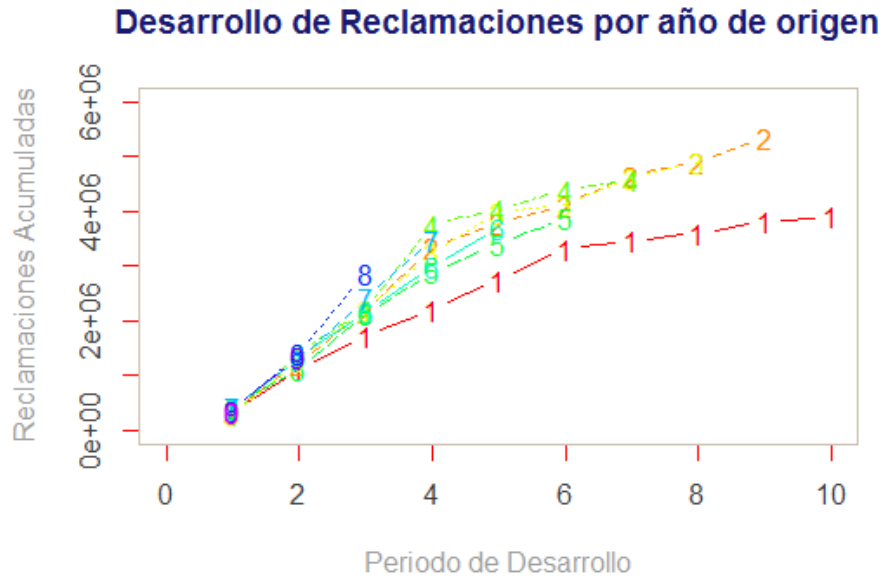


Figura 2.1: *Desarrollo de Reclamaciones*

Lo ideal para una aseguradora, sería que conforme aumente el periodo de desarrollo las curvas tiendan a mantenerse constantes eso indicaría que no se esperan más reclamaciones.

Para un análisis individual, i.e., para cada año de origen, basta con incluir el argumento `lattice=TRUE` para obtener una gráfica dividida con el desarrollo de cada año de origen.

| | |
|---|-------------|
| | Código de R |
| <pre>R> plot(GenIns, lattice=TRUE)</pre> | |
| | Código de R |

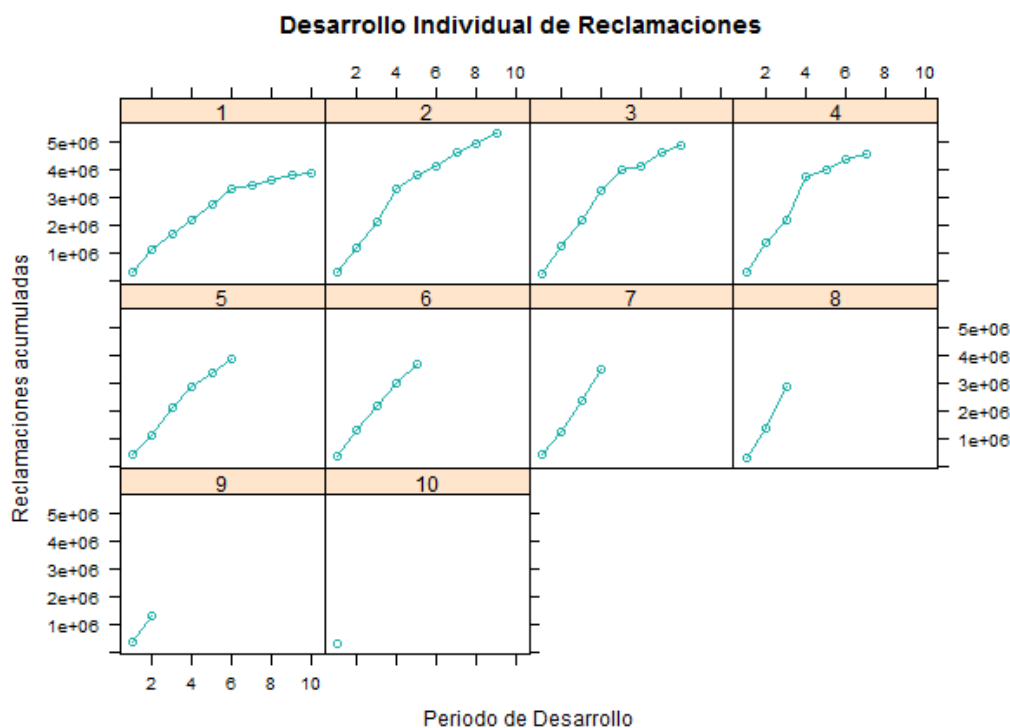


Figura 2.2: *Desarrollo individual de reclamaciones*

2.2. Métodos básicos de estimación/pronóstico

Los métodos clásicos que más se utilizan para estimar la reserva de obligaciones pendientes de cumplir son:

- I. Método de la razón.
- II. Chain-Ladder (CL).
- III. Bornhuetter-Ferguson (BF).

Observación 2.2.1.

Se supondrá que $I = J$ y si $j > J$ entonces $X_{i,j} = 0$.

∇

A continuación se describirá cada uno de los métodos clásicos, adicionalmente se ejemplificará su procedimiento haciendo uso de la paquetería `ChainLadder` de R, partiendo de una base de datos contenida en el paquete de nombre `GenIns`, que ya se introdujo en las secciones anteriores.

2.2.1. Método de la razón

En este método se estima el monto total de los siniestros para cada año de ocurrencia, por medio del producto de la última siniestralidad conocida por un respectivo factor de desarrollo.

A partir del triángulo de siniestros acumulados se calcula el crecimiento que se presentó entre cada periodo de desarrollo para posteriormente determinar los factores de desarrollo promedio (*age-to-age*) que permitan estimar el monto total de los siniestros futuros para cada año de origen.

| Año de accidente i | Año de desarrollo j | | | | | | | |
|----------------------|-----------------------|-------------|-----------|-----|-----------|-----|-------------|-----------|
| | 0 | 1 | 2 | ... | j | ... | $J-1$ | J |
| 0 | $C_{0,0}$ | $C_{0,1}$ | $C_{0,2}$ | ... | $C_{0,j}$ | ... | | $C_{0,J}$ |
| 1 | $C_{1,0}$ | $C_{1,1}$ | $C_{1,2}$ | ... | $C_{1,j}$ | ... | $C_{1,J-1}$ | |
| ⋮ | | | | | | ... | | |
| ⋮ | | | | | | | | |
| i | | | | ... | | | | |
| ⋮ | | | | | | | | |
| $I-1$ | | $C_{I-1,1}$ | | | | | | |
| I | $C_{I,0}$ | | | | | | | |

Una vez que se calculan los factores promedio, se estima $\hat{\alpha}$, de los siniestros que aún faltan por pagar. Cabe mencionar que este factor es muy importante para la estimación final, sin embargo, no existe una metodología objetiva para determinarlo, una forma podría ser a través de un ajuste lineal de la siniestralidad o con base en la experiencia que se tenga.

| Año de accidente i | Periodos de desarrollo | | | | | |
|-------------------------------|--|--|-----|--|-----|-----------------------------|
| | 0 al 1 | 1 al 2 | ... | $(j-1)$ al j | ... | $(J-1)$ al J |
| 0 | $C_{0,1}/C_{0,0}$ | $C_{0,2}/C_{0,1}$ | ... | $C_{0,j}/C_{0,j-1}$ | ... | $C_{0,J}/C_{0,J-1}$ |
| 1 | $C_{1,1}/C_{1,0}$ | $C_{1,2}/C_{1,1}$ | ... | $C_{1,j}/C_{1,j-1}$ | ... | |
| ⋮ | ... | ... | ... | ... | ... | |
| i | $C_{i,1}/C_{i,0}$ | $C_{i,2}/C_{i,1}$ | ... | $C_{i,j}/C_{i,j-1}$ | ... | |
| ⋮ | | | | | | |
| $I-1$ | $C_{I-1,1}/C_{I-1,0}$ | | | | | |
| Factores promedio (M_j) | $\frac{1}{I-1} \sum_{i=0}^{I-1} \frac{C_{i,1}}{C_{i,0}}$ | $\frac{1}{I-2} \sum_{i=0}^{I-2} \frac{C_{i,2}}{C_{i,1}}$ | ... | $\frac{1}{I-j} \sum_{i=0}^{I-j} \frac{C_{i,j}}{C_{i,j-1}}$ | ... | $\frac{C_{0,J}}{C_{0,J-1}}$ |
| Factores ponderados (W_j) | $(1 + \hat{\alpha}) \prod_{i=1}^J M_i$ | $(1 + \hat{\alpha}) \prod_{i=2}^J M_i$ | ... | $(1 + \hat{\alpha}) \prod_{i=j}^J M_i$ | ... | $(1 + \hat{\alpha}) M_J$ |

La estimación de los siniestros futuros a través de los factores promedio es

$$\widehat{C}_{i,J}^{raz} = C_{i,I-i} M_{I-i} \cdots M_J,$$

o si se opta por la estimación con los factores ponderados

$$\widehat{C}_{i,J}^{raz} = C_{i,I-i} W_{I-i} \cdots W_J.$$

Por lo tanto, la estimación de la reserva con el método de la razón para ambos factores se obtiene como

$$\widehat{R}^{raz} = \widehat{C}_{i,J}^{raz} - C_{i,I-i}.$$

Ejemplo 2.2.1.

A partir de la función `ata()` de la librería `ChainLadder` de R se genera una matriz con el cálculo de los incrementos de siniestros acumulados entre periodos de desarrollo para cada año de origen. Además, se calculan los factores promedio y los factores ponderados para cada periodo de desarrollo.

A continuación se presenta para el conjunto de datos `GenIns`.

Código de R

```
R> #Matriz de razones y factores de desarrollo
R> razon<-ata(GenIns)
R> razon
      dev
origin 1-2  2-3  3-4  4-5  5-6  6-7  7-8  8-9  9-10
  1    3.143 1.543 1.278 1.238 1.209 1.044 1.040 1.063 1.018
  2    3.511 1.755 1.545 1.133 1.084 1.128 1.057 1.086   NA
  3    4.448 1.717 1.458 1.232 1.037 1.120 1.061   NA   NA
  4    4.568 1.547 1.712 1.073 1.087 1.047   NA   NA   NA
  5    2.564 1.873 1.362 1.174 1.138   NA   NA   NA   NA
  6    3.366 1.636 1.369 1.236   NA   NA   NA   NA   NA
  7    2.923 1.878 1.439   NA   NA   NA   NA   NA   NA
  8    3.953 2.016   NA   NA   NA   NA   NA   NA   NA
  9    3.619   NA   NA   NA   NA   NA   NA   NA   NA
smpl 3.566 1.746 1.452 1.181 1.111 1.085 1.053 1.075 1.018
vwtd 3.491 1.747 1.457 1.174 1.104 1.086 1.054 1.077 1.018
```

Código de R

Los atributos “`smpl`” y “`vwtd`”, son los factores de desarrollo promedio y los factores de desarrollo ponderados, respectivamente. Se pueden extraer los factores utilizando la función `attr()`.

Código de R

```
R> #Factores de desarrollo promedio
R> f.promedio<-attr(razon,"smpl")
```

```
R> f.promedio
  1-2      2-3      3-4      4-5      5-6      6-7      7-8      8-9      9-10
3.566143 1.745557 1.451961 1.180984 1.111247 1.084818 1.052739 1.074753 1.017725
R> #Factores de desarrollo ponderados
R> f.ponderados<-attr(razon,"vwtd")
R> f.ponderados
  1-2      2-3      3-4      4-5      5-6      6-7      7-8      8-9      9-10
3.490607 1.747333 1.457413 1.173852 1.103824 1.086269 1.053874 1.076555 1.017725
```

Código de R

Observación 2.2.2.

R asigna el peso de los factores ponderados automáticamente pero un buen ejercicio sería estimarlo. ▽

A partir de los factores promedio, aplicados al triángulo de siniestros acumulados, se estiman los siniestros futuros, i.e., se completa el triángulo siguiendo el siguiente procedimiento

```
Código de R
```

```
R> trianguloCompleto<-GenIns
R> #Número de factores o cambios
R> n=length(GenIns[,1])-1
R> #Construcción del triángulo completo
R> for(k in 1:n){
+   trianguloCompleto[(n-k+1):n+1, k+1] <-trianguloCompleto[(n-k+1):n+1,k]*f.promedio[k]}
R> #Triángulo completo
R> round(trianguloCompleto)
      dev
origin  1      2      3      4      5      6      7      8      9     10
  1  357848 1124788 1735330 2218270 2745596 3319994 3466336 3606286 3833515 3901463
  2  352118 1236139 2170033 3353322 3799067 4120063 4647867 4914039 5339085 5433719
  3  290507 1292306 2218525 3235179 3985995 4132918 4628910 4909315 5276300 5369821
  4  310608 1418858 2195047 3757447 4029929 4381982 4588268 4830251 5191325 5283340
  5  443160 1136350 2128333 2897821 3402672 3873311 4201836 4423439 4754103 4838368
  6  396132 1333217 2180715 2985752 3691712 4102403 4450360 4685070 5035291 5124540
  7  440832 1288463 2419861 3483130 4113520 4571136 4958850 5220377 5610614 5710061
  8  359480 1421128 2864498 4159139 4911875 5458306 5921267 6233552 6699527 6818274
  9  376686 1363294 2379707 3455241 4080584 4534536 4919145 5178578 5565691 5664341
 10 344014 1226803 2141454 3109308 3672042 4080545 4426648 4660107 5008462 5097236
```

Código de R

Por lo tanto, con el uso de la función `getLatestCumulative()`, que obtiene la siniestralidad última del triángulo de datos observados y estimados, la reserva IBNR estimada a partir de los factores de desarrollo promedio se obtiene como:

```
Código de R
```

```
R> reserva<-sum(getLatestCumulative(trianguloCompleto)-getLatestCumulative(GenIns))
R> round(reserva)
[1] 18883073
```

Código de R

Análogamente, siguiendo el mismo procedimiento pero ahora con los factores de desarrollo ponderados, se obtiene la siguiente estimación:

```

Código de R
R> #Con factores ponderados
R> triangulocompleto<-GenIns
R> for(k in 1:n){
+ triangulocompleto[(n-k+1):n+1, k+1]<-triangulocompleto[(n-k+1):n+1,k]*f.ponderados[k]}
R> #Muestra triángulo completo
R> round(triangulocompleto)
      dev
origin  1    2    3    4    5    6    7    8    9   10
  1  357848 1124788 1735330 2218270 2745596 3319994 3466336 3606286 3833515 3901463
  2  352118 1236139 2170033 3353322 3799067 4120063 4647867 4914039 5339085 5433719
  3  290507 1292306 2218525 3235179 3985995 4132918 4628910 4909315 5285148 5378826
  4  310608 1418858 2195047 3757447 4029929 4381982 4588268 4835458 5205637 5297906
  5  443160 1136350 2128333 2897821 3402672 3873311 4207459 4434133 4773589 4858200
  6  396132 1333217 2180715 2985752 3691712 4074999 4426546 4665023 5022155 5111171
  7  440832 1288463 2419861 3483130 4088678 4513179 4902528 5166649 5562182 5660771
  8  359480 1421128 2864498 4174756 4900545 5409337 5875997 6192562 6666635 6784799
  9  376686 1363294 2382128 3471744 4075313 4498426 4886502 5149760 5544000 5642266
 10 344014 1200818 2098228 3057984 3589620 3962307 4304132 4536015 4883270 4969825

R> reserva<-sum(getLatestCumulative(triangulocompleto)-getLatestCumulative(GenIns))
R> reserva
[1] 18680856
Código de R

```

La diferencia entre las reservas de perdidas pendientes estimada para cada tipo de factores no es significativa pero ésta depende del factor $\hat{\alpha}$ estimado.

2.2.2. Método Chain-Ladder (algorítmico)

En principio, el modelo clásico Chain-Ladder (CL) no se fundamenta en un modelo estocástico. Más bien, es un algoritmo recursivo que se utiliza para estimar la reserva de obligaciones por cumplir. Sin embargo, existen varios modelos estocásticos que justifican el método CL. Ahora se estudiará el de “distribución libre” (Mack (1993)).

Bajo el método CL de distribución libre se relacionan las reclamaciones acumuladas con *link-ratios* adecuados basados en el siguiente modelo.

Suposición 2.2.1. (*Modelo Chain-Ladder de distribución libre*)

- (i) Las reclamaciones acumuladas, $C_{i,j}$, para diferentes años accidente i son independientes, i.e. $C_{0,j}, C_{1,j}, \dots, C_{I,j}$ son independientes.
- (ii) Existen factores de desarrollo $f_0, f_1, \dots, f_{J-1} > 0$ tales que para todo $i \in \{0, 1, \dots, I\}, j \in \{1, \dots, J\}$

$$\mathbb{E}[C_{i,j}|C_{i,0}, \dots, C_{i,j-1}] = \mathbb{E}[C_{i,j}|C_{i,j-1}] = C_{i,j-1} \cdot f_{j-1}. \quad (2.2.1)$$

La Suposición 2.2.1 consiste de las primeras dos suposiciones del modelo CL propuesto por Mack (1993). Sólo hay suposición con respecto al primer momento en (2.2.1). Esto ya es suficiente para estimar condicionalmente las reclamaciones y por lo tanto, describir el algoritmo CL. Los momentos de orden superior sólo serán relevantes cuando se quiera cuantificar incertidumbres de estos estimadores y pronósticos, por ejemplo a partir de la varianza.

La suposición de independenciam entre reclamaciones en diferente año de accidente es común en casi todos los métodos de reservas. Una implicación de esta suposición es que se necesita eliminar los efectos de los años contables en los datos. Se pueden hacer suposiciones más fuertes sobre la sucesión $C_{i,0}, C_{i,1} \dots$, por ejemplo, que sea una cadena de Markov.

Además, $\left\{ C_{i,j} \prod_{k=0}^{j-1} f_k^{-1} \right\}_{j=1}^{\infty}$ es una martingala.

Los factores f_j se conocen como *link-ratios*, factores de desarrollo, factores Chain-Ladder o factores *age-to-age*, éstos son el objeto central de interés en el método CL y describen cómo se relacionan las reclamaciones acumuladas sucesivas. El principal reto es dar estimadores apropiados para estos factores CL.

Proposición 2.2.1. (*Reclamación esperada bajo el método CL*)

Bajo la Suposición 2.2.1 para todo $i \in \{1, \dots, I\}$

$$\mathbb{E}[C_{i,J}|\mathcal{D}_I] = \mathbb{E}[C_{i,J}|C_{i,I-i}] = C_{i,I-i} f_{I-i} \cdots f_{J-1}. \quad (2.2.2)$$

Demostración:

$$\begin{aligned}
\mathbb{E}[C_{i,J}|\mathcal{D}_I] &= \mathbb{E}[C_{i,J}|C_{i,0}, \dots, C_{i,I-i}] \\
&= \mathbb{E}[\mathbb{E}(C_{i,J}|C_{i,J-1})|C_{i,0}, \dots, C_{i,I-i}], \text{ por el Teorema de Esperanza Iterada} \\
&= \mathbb{E}[f_{J-1}C_{i,J-1}|C_{i,0}, \dots, C_{i,I-i}] \\
&= f_{J-1}\mathbb{E}[C_{i,J-1}|\mathcal{D}_I].
\end{aligned}$$

Si se repite este procedimiento hasta que se alcance la diagonal $i + j = I$, se obtiene la siniestralidad última, es decir,

$$\begin{aligned}
\mathbb{E}[C_{i,J}|\mathcal{D}_I] &= f_{J-1} \cdot \mathbb{E}[C_{i,J-1}|\mathcal{D}_I] = f_{J-1} [f_{J-2} \mathbb{E}[C_{i,J-2}|\mathcal{D}_I]] \\
&\quad \vdots \\
&= f_{J-1} \cdot f_{J-2} \cdots f_{I-i} \mathbb{E}[C_{i,I-i}|\mathcal{D}_I] \\
&= f_{J-1} \cdot f_{J-2} \cdots f_{I-i} C_{i,I-i}.
\end{aligned}$$

□

La Proposición 2.2.1 proporciona un algoritmo recursivo para estimar/predecir la siniestralidad última $C_{i,J}$, dadas las observaciones \mathcal{D}_I . Para factores CL conocidos, f_j , las obligaciones pendientes (*outstanding claims liabilities*) del año de accidente i , dado \mathcal{D}_I se estiman como

$$\mathbb{E}[C_{i,J}|\mathcal{D}_I] - C_{i,I-i} = C_{i,I-i}(f_{I-i} \cdots f_{J-1} - 1). \quad (2.2.3)$$

Esto corresponde a un “mejor estimador” de la reserva para el año de accidente I (basados en la información \mathcal{D}_I y los factores CL conocidos f_j). Es decir, se utiliza la esperanza condicional (2.2.3) para pronosticar el resultado de la variable aleatoria $C_{i,J} - C_{i,I-i}$ dado \mathcal{D}_I .

Sin embargo, en la mayoría de las aplicaciones prácticas, los factores CL, f_j , no se conocen y se tienen que estimar. Así, para $j = 0, \dots, J - 1$ los factores Chain-Ladder, f_j , se estiman como

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}. \quad (2.2.4)$$

Observación 2.2.3.

Proponer a f_j como se hizo en (2.2.4) podría parecer muy artificial, sin embargo, se probará que con este estimador se obtendrán pronósticos insesgados y de varianza mínima. ▽

Nótese que

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{i,j+1}}{C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} F_{i,j+1},$$

con $F_{i,j+1} = \frac{C_{i,j+1}}{C_{i,j}}$, i.e., \hat{f}_j es un “promedio ponderado” de los factores $F_{i,j+1}$ con los pesos

$w_i := \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}}$. Una vez que se dio una manera de estimar los factores f_j , se puede definir

el estimador Chain-Ladder .

Estimador 2.2.1. (*Estimador Chain-Ladder*)

Para $i + j > I$, el estimador CL para $\mathbb{E}[C_{i,j}|\mathcal{D}_I]$ está dado por

$$\hat{C}_{i,j}^{CL} = \hat{\mathbb{E}}[C_{i,j}|\mathcal{D}_I] = C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{j-1}, \quad i + j > I. \quad (2.2.5)$$

Si se considera al método CL sólo de forma mecánica, la expresión (2.2.5) proporciona un algoritmo para calcular la reserva CL.

Observación 2.2.4.

En general, este algoritmo se utiliza sin considerar un modelo estocástico subyacente, por eso se conoce como de “distribución libre”. ∇

Notación 2.2.1.

Defínase el conjunto de observaciones hasta el tiempo k como

$$\mathcal{B}_k := \{C_{i,j} : i + j \leq I, 0 \leq j \leq k\} \subseteq \mathcal{D}_I. \quad (2.2.6)$$

En particular $\mathcal{B}_J = \mathcal{D}_I$. Además, \mathcal{B}_0 corresponde a la primera columna del triángulo de reclamaciones. ∇

Proposición 2.2.2.

Bajo la Suposición 2.2.1

- (a) $\mathbb{E}[\hat{f}_j | \mathcal{B}_j] = f_j$. Es decir, dado \mathcal{B}_j , \hat{f}_j es un estimador insesgado de f_j .
- (b) $\mathbb{E}[\hat{f}_j] = f_j$, i.e. \hat{f}_j es un estimador insesgado de f_j .
- (c) $\mathbb{E}[\hat{f}_0 \cdots \hat{f}_j] = \mathbb{E}[\hat{f}_0] \cdots \mathbb{E}[\hat{f}_j]$. En particular, $\hat{f}_0, \dots, \hat{f}_j$ no están correlacionados.
- (d) $\mathbb{E}[\hat{C}_{i,J}^{CL} | C_{i,I-i}] = \mathbb{E}[C_{i,J} | \mathcal{D}_I]$, dado $C_{i,I-i}$, el estimador $\hat{C}_{i,J}^{CL}$, es insesgado para

$$\mathbb{E}[C_{i,J} | \mathcal{D}_I] = \mathbb{E}[C_{i,J} | C_{i,I-i}].$$

- (e) $\mathbb{E}[\hat{C}_{i,J}^{CL}] = \mathbb{E}[C_{i,J}]$, i.e. $\hat{C}_{i,J}^{CL}$ es un estimador insesgado de $\mathbb{E}[C_{i,J}]$.

Demostración:

- (a) Como $C_{0,j}, C_{1,j}, \dots, C_{I-j-1,j}$ son medibles con respecto a \mathcal{B}_j entonces $\sum_{i=0}^{I-j-1} C_{i,j}$ es medible con respecto a \mathcal{B}_j . Entonces,

$$\begin{aligned} \mathbb{E}[\hat{f}_j | \mathcal{B}_j] &= \mathbb{E}\left[\frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} \middle| \mathcal{B}_j\right] = \frac{1}{\sum_{i=0}^{I-j-1} C_{i,j}} \mathbb{E}\left[\sum_{i=0}^{I-j-1} C_{i,j+1} \middle| \mathcal{B}_j\right] \\ &= \frac{\sum_{i=0}^{I-j-1} \mathbb{E}[C_{i,j+1} | \mathcal{B}_j]}{\sum_{i=0}^{I-j-1} C_{i,j}} = \frac{\sum_{i=0}^{I-j-1} f_j C_{i,j}}{\sum_{i=0}^{I-j-1} C_{i,j}} = f_j. \end{aligned}$$

- (b)

$$\mathbb{E}[\hat{f}_j] = \mathbb{E}\left[\mathbb{E}(\hat{f}_j | \mathcal{B}_j)\right] = \mathbb{E}[f_j] = f_j.$$

(c)

$$\begin{aligned}\mathbb{E}[\hat{f}_0 \cdots \hat{f}_j] &= \mathbb{E}[\mathbb{E}(\hat{f}_0 \cdots \hat{f}_j | \mathcal{B}_j)] = \mathbb{E}[\hat{f}_0 \cdots \hat{f}_{j-1} \mathbb{E}(\hat{f}_j | \mathcal{B}_j)] \\ &= \mathbb{E}[\hat{f}_0 \cdots \hat{f}_{j-1}] f_j = \mathbb{E}[\hat{f}_0 \cdots \hat{f}_{j-1}] \mathbb{E}[\hat{f}_j].\end{aligned}$$

Iterativamente,

$$\mathbb{E}[\hat{f}_0 \cdots \hat{f}_j] = \mathbb{E}[\hat{f}_0] \cdots \mathbb{E}[\hat{f}_j].$$

(d)

$$\begin{aligned}\mathbb{E}[\hat{C}_{i,J}^{CL} | C_{i,I-i}] &= \mathbb{E}[C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{J-1} | C_{i,I-i}] \\ &= \mathbb{E}[C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{J-2} \mathbb{E}(\hat{f}_{J-1} | \mathcal{B}_{J-1}) | C_{i,I-i}] \\ &= f_{J-1} \mathbb{E}[\hat{C}_{i,J-1}^{CL} | C_{i,I-i}].\end{aligned}$$

Iterativamente,

$$\mathbb{E}[\hat{C}_{i,J}^{CL} | C_{i,I-i}] = C_{i,I-i} f_{I-i} \cdots f_{J-1} = \mathbb{E}[C_{i,J} | \mathcal{D}_I].$$

Por lo tanto,

$$\mathbb{E}[\hat{C}_{i,J}^{CL} | C_{i,I-i}] = \mathbb{E}[C_{i,J} | \mathcal{D}_I].$$

(e)

$$\mathbb{E}[\hat{C}_{i,J}^{CL}] = \mathbb{E}[\mathbb{E}(C_{i,J} | \mathcal{D}_I)] = \mathbb{E}[C_{i,J}].$$

□

Observación 2.2.5.

Los estimadores CL, \hat{f}_j , no están correlacionados, aún cuando se estiman a partir de factores *age-to-age* dependen parcialmente de los mismos datos (en el numerador y denominador). Los factores estimados *age-to-age* \hat{f}_j están no correlacionados pero esto no implica que sean independientes. De hecho, no lo son. Se puede demostrar que \hat{f}_j^2 y \hat{f}_{j+1}^2 están correlacionados negativamente. Esta correlación negativa se estudiará más adelante cuando se quiera medir el error de las estimaciones. ▽

Los incisos (a) y (b) de la Proposición 2.2.2 establecen que los factores Chain-Ladder, f_j , se estiman de manera insesgada a partir de \hat{f}_j aún sin hipótesis distribucional. Esto motiva la elección de \hat{f}_j de esa manera. Por supuesto hay otros estimadores insesgados pero más adelante se verá que satisface un criterio de optimalidad bajo ciertas restricciones de

la varianza.

El inciso (d) de la Proposición 2.2.2 muestra que $\widehat{C}_{i,J}^{CL}$ es un estimador insesgado de $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$ y justifica lo “artificial” de los estimadores \hat{f}_j . Esto también justifica al algoritmo CL desde la perspectiva de “distribución libre”.

Este algoritmo no permite cuantificar observaciones inciertas para predecir la reserva. Para analizar la aleatoriedad asociada, existen modelos estocásticos que calculan la reserva con el método Chain-Ladder, por ejemplo, el modelo de Mack de distribución libre o el modelo Chain-Ladder Bayesiano. Para este último, suele usarse métodos como *bootstrap* con suposiciones Bayesianas.

Ejemplo 2.2.2.

A continuación, considerando el conjunto de datos del ejemplo anterior (GenIns), se muestra una aplicación mecánica del método Chain-Ladder para obtener los factores de desarrollo (ó *age-to-age ratios*).

Código de R

```
R> #Método CL
R> n <- 10 #Número de periodos de desarrollo
R> factores <- sapply(1:(n-1),
+                   function(i){
+                     sum(GenIns[c(1:(n-i)),i+1])/sum(GenIns[c(1:(n-i)),i])})
R> factores
[1] 3.490607 1.747333 1.457413 1.173852 1.103824 1.086269 1.053874 1.076555 1.017725
```

Código de R

La paquetería `ChainLadder` contiene la función `chainladder()`, ésta calcula los factores de desarrollo de una forma más eficiente, ya que recibe como argumento el triángulo de reclamaciones acumuladas.

Código de R

```
R> factores<-chainladder(GenIns)
R> coef(factores)
      1      2      3      4      5      6      7      8      9
3.490607 1.747333 1.457413 1.173852 1.103824 1.086269 1.053874 1.076555 1.017725
```

Código de R

Lo anterior es la estructura no-estocástica del método Chain-Ladder para el cálculo de la reserva de reclamaciones.

Entonces, siguiendo el algoritmo para obtener el total de pérdidas pendientes por cumplir, por lo tanto, se obtiene el total de la última columna que son las pérdidas esperadas (resultado del algoritmo Chain-Ladder), a dicho monto se le restará la suma de la última diagonal de las observaciones, i.e.

```

Código de R
R> m=9
R> completo.GenIns <- GenIns
R> for(k in 1:m){
+   completo.GenIns[(m-k+1):n+1, k+1] <- completo.GenIns[(m-k+1):n+1, k]*factoresCL[k]
+ } #llena el triángulo
R> completo.GenIns
      dev
origin 1      2      3      4      5      6      7      8      9     10
  1 357848 1124788 1735330 2218270 2745596 3319994 3466336 3606286 3833515 3901463
  2 352118 1236139 2170033 3353322 3799067 4120063 4647867 4914039 5339085 5433719
  3 290507 1292306 2218525 3235179 3985995 4132918 4628910 4909315 5285148 5378826
  4 310608 1418858 2195047 3757447 4029929 4381982 4588268 4835458 5205637 5297906
  5 443160 1136350 2128333 2897821 3402672 3873311 4207459 4434133 4773589 4858200
  6 396132 1333217 2180715 2985752 3691712 4074999 4426546 4665023 5022155 5111171
  7 440832 1288463 2419861 3483130 4088678 4513179 4902528 5166649 5562182 5660771
  8 359480 1421128 2864498 4174756 4900545 5409337 5875997 6192562 6666635 6784799
  9 376686 1363294 2382128 3471744 4075313 4498426 4886502 5149760 5544000 5642266
 10 344014 1200818 2098228 3057984 3589620 3962307 4304132 4536015 4883270 4969825
R> reservaCL<-sum(getLatestCumulative(completo.GenIns)-getLatestCumulative(GenIns))
R> reservaCL
[1] 18680856
Código de R

```

Nótese que la reserva de pérdidas pendientes del modelo CL es menor que la obtenida a partir del método de la razón por factores promedio. Es decir, la compañía aseguradora en este caso optaría por considerar el algoritmo CL ya que el monto que resultó de la diferencia entre reservas es capital que se puede usar libremente y esto genera mayores utilidades para los inversionistas, a diferencia de la reserva estimada con los factores ponderados el factor $\hat{\alpha}$ resultó ser equivalente con los factores de desarrollo CL.

2.2.3. Método Bornhuetter-Ferguson

El método Bornhuetter-Ferguson (BF) en general es más robusto ya que no considera *outliers* en las observaciones. El método fue desarrollado por Bornhuetter & Ferguson (1972) en el artículo "The Actuary and IBNR".

Al igual que el Chain-Ladder, el método BF generalmente se entiende de manera mecánica. Pero también existen varias formas de definir un modelo estocástico subyacente que motive su definición.

Suposición 2.2.2.

- (i) Las reclamaciones acumuladas $C_{i,j}$ son independientes para diferentes años accidente i .
- (ii) Existen parámetros $\mu_0, \dots, \mu_I > 0$ y un patrón $\beta_0, \dots, \beta_J > 0$ con $\beta_J = 1$ tales que para todo $i \in \{0, \dots, I\}, j \in \{0, \dots, J-1\}, k \in \{1, 2, \dots, J-j\}$
 - (a) $\mathbb{E}[C_{i,0}] = \beta_0 \cdot \mu_i$.
 - (b) $\mathbb{E}[C_{i,j+k} | C_{i,0}, \dots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i$.

Bajo la Suposición 2.2.2

$$\mathbb{E}[C_{i,j}] = \beta_j \cdot \mu_i \quad \text{y} \quad \mathbb{E}[C_{i,J}] = \mu_i.$$

La sucesión $\{\beta_j\}_{j=0}^J$, se le conoce como patrón de desarrollo de las reclamaciones.

Suposición 2.2.3.

- Las reclamaciones acumuladas $C_{i,j}$ de diferentes años accidente i son independientes.
- Existen parámetros $\mu_0, \dots, \mu_I > 0$ y un patrón $\beta_0, \dots, \beta_J > 0$ con $\beta_J = 1$ tales que para todo $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$.

$$\mathbb{E}[C_{i,j}] = \beta_j \cdot \mu_i. \tag{2.2.7}$$

Observación 2.2.6.

La Suposición 2.2.2 es más fuerte que la Suposición 2.2.3 ya que agrega la condición de que $\mathbb{E}[C_{i,j+k} | C_{i,0}, \dots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j)\mu_i$. ∇

La Suposición 2.2.3 dificulta un poco al algoritmo BF ya que

$$\begin{aligned} \mathbb{E}[C_{i,J} | \mathcal{D}_I] &= \mathbb{E}[C_{i,J} | C_{i,0}, \dots, C_{i,I-i}] \\ &= C_{i,I-i} + \mathbb{E}[C_{i,J} - C_{i,I-i} | C_{i,0}, \dots, C_{i,I-i}]. \end{aligned} \tag{2.2.8}$$

Si no se hacen suposiciones adicionales con respecto a la estructura de dependencia entre la reclamaciones incrementales, entonces no se sabe exactamente qué hacer con el

término $\mathbb{E}[C_{i,J} - C_{i,I-i} | C_{i,0}, \dots, C_{i,I-i}]$ de la ecuación (2.2.8). Si se sabe que $C_{i,J} - C_{i,I-i}$ es independiente de $C_{i,0}, \dots, C_{i,I-i}$, entonces

$$\begin{aligned} \mathbb{E}[C_{i,J} | \mathcal{D}_I] &= \mathbb{E}[C_{i,J} | C_{i,0}, \dots, C_{i,I-i}] \\ &= C_{i,I-i} + \mathbb{E}[C_{i,J} - C_{i,I-i}] \\ &= C_{i,I-i} + (1 - \beta_{I-i})\mu_i \quad (\text{de la Suposición 2.2.2}). \end{aligned} \quad (2.2.9)$$

Bajo ambas suposiciones, falta estimar el último término del lado derecho de las ecuaciones (2.2.8) y (2.2.9).

Estimador 2.2.2. (*Estimador Bornhuetter-Ferguson*)

Para $i \in \{1, \dots, I\}$, el estimador BF de $\mathbb{E}[C_{i,J} | \mathcal{D}_I]$ está dado por

$$\widehat{C}_{i,J}^{BF} = \widehat{\mathbb{E}}[C_{i,J} | \mathcal{D}_I] = C_{i,I-i} + (1 - \widehat{\beta}_{I-i})\widehat{\mu}_i, \quad (2.2.10)$$

donde $\widehat{\beta}_{I-i}$ es un estimador apropiado de β_{I-i} y $\widehat{\mu}_i$ es un estimador *a priori* de la siniestralidad última, $\mathbb{E}[C_{i,J}]$.

La expresión (2.2.10) constituye el algoritmo BF, simplemente de manera mecánica, sin considerar un modelo estocástico subyacente. La pregunta que surge naturalmente es cómo se determinan los estimadores $\widehat{\beta}_j$ y $\widehat{\mu}_i$.

2.2.3.1. Comparación entre los estimadores BF y CL

A partir de la Suposición 2.2.1 del modelo CL

$$\begin{aligned} \mathbb{E}[C_{i,j}] &= \mathbb{E}[\mathbb{E}(C_{i,j} | C_{i,j-1})], \text{ por el Teorema de Esperanza iterada} \\ &= \mathbb{E}[f_{j-1} \cdot C_{i,j-1}] = f_{j-1}\mathbb{E}[C_{i,j-1}] \end{aligned}$$

De manera iterativa,

$$= \mathbb{E}[C_{i,0}] \prod_{k=0}^{j-1} f_k.$$

Entonces,

$$\mathbb{E}[C_{i,0}] = \mathbb{E}[C_{i,j}] \prod_{k=0}^{j-1} f_k^{-1}. \quad (2.2.11)$$

Ahora, para la siniestralidad última $\mathbb{E}[C_{i,J}] = \mathbb{E}[C_{i,0}] \prod_{k=0}^{J-1} f_k$, por lo tanto

$$\mathbb{E}[C_{i,0}] = \mathbb{E}[C_{i,J}] \prod_{k=0}^{J-1} f_k^{-1}. \quad (2.2.12)$$

Igualando las expresiones (2.2.11) y (2.2.12),

$$\mathbb{E}[C_{i,J}] \prod_{k=0}^{J-1} f_k^{-1} = \mathbb{E}[C_{i,j}] \prod_{k=0}^{j-1} f_k^{-1},$$

entonces,

$$\mathbb{E}[C_{i,j}] = \mathbb{E}[C_{i,J}] \prod_{k=0}^{J-1} f_k^{-1} \prod_{k=0}^{j-1} f_k = \mathbb{E}[C_{i,J}] \prod_{k=j}^{J-1} f_k^{-1}.$$

Si se compara con el modelo BF de la Suposición 2.2.2 donde $\mathbb{E}[C_{i,j}] = \beta_j \cdot \mu_i$, entonces

$$\prod_{k=j}^{J-1} f_k^{-1} \text{ se puede considerar como } \beta_j. \quad (2.2.13)$$

Entonces, bajo algunas hipótesis ocurre que $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$, que es la proporción de $\mu_i = \mathbb{E}[C_{i,J}]$ ya pagado después de j periodos de desarrollo, suponiendo el modelo CL. Esto se puede hacer bajo la Suposición 2.2.3 pero no bajo la Suposición 2.2.2. Es decir, si se conoce los factores CL, f_j , entonces se puede construir el patrón de desarrollo, a partir de

$$\beta_j = \prod_{k=j}^{J-1} f_k^{-1}.$$

Entonces, el estimador BF se puede interpretar como

$$\widehat{C}_{i,J}^{BF} = C_{i,I-i} + \widehat{\mu}_i \left(1 - \left(\frac{1}{\widehat{\prod_{j=I-i}^{J-1} f_j}} \right) \right). \quad (2.2.14)$$

Por otro lado, para el estimador CL

$$\begin{aligned}
\widehat{C}_{i,J}^{CL} &= C_{i,I-i} \prod_{j=I-i}^{J-1} \widehat{f}_j = C_{i,I-i} + C_{i,I-i} \left(\prod_{j=I-i}^{J-1} \widehat{f}_j - 1 \right) \\
&= C_{i,I-i} + \frac{\widehat{C}_{i,J}^{CL}}{\prod_{j=I-i}^{J-1} \widehat{f}_j} \left(\prod_{j=I-i}^{J-1} \widehat{f}_j - 1 \right) \\
&= C_{i,I-i} + \widehat{C}_{i,J}^{CL} - \frac{\widehat{C}_{i,J}^{CL}}{\prod_{j=I-i}^{J-1} \widehat{f}_j} = C_{i,I-i} + \widehat{C}_{i,J}^{CL} \left(1 - \left(\frac{1}{\prod_{j=I-i}^{J-1} \widehat{f}_j} \right) \right). \quad (2.2.15)
\end{aligned}$$

Entonces, si se comparan las expresiones (2.2.14) y (2.2.15) la diferencia entre los patrones de desarrollo de los métodos CL y BF, es que para el método BF se confía completamente en el estimador *a priori* $\widehat{\mu}_i$, mientras que en el método CL el estimador *a priori* se remplaza por el estimado $\widehat{C}_{i,J}^{CL}$ (que se basa sólo en las observaciones).

Desde esta perspectiva, los métodos CL y BF son dos posiciones extremas en el problema de *reserving* y por su puesto si se pueden combinar éstos.

Para μ_i se necesita un estimador *a priori* $\widehat{\mu}_i$. Generalmente es un valor del plan de negocios o un valor que usa para el cálculo de primas. Este valor se debe estimar antes de tener cualquier observación, i.e.; este debe ser un estimado *a priori* “puro” basado en una opinión experta.

Para el factor $(1 - \beta_{I-i})$, también se debe utilizar un estimador *a priori* si se desea aplicar mecánicamente el método BF. Esto se debe hacer independientemente de las observaciones. Sin embargo, en la mayoría de las aplicaciones prácticas, se desvía de la trayectoria del BF puro y se estima el factor $(1 - \beta_{I-i})$ de los datos con los factores estimados CL .

Si \widehat{f}_k es el factor estimado de (2.2.4), entonces

$$\widehat{\beta}_j^{(CL)} = \widehat{\beta}_j = \left(\frac{1}{\prod_{k=j}^{J-1} \widehat{f}_k} \right) = \prod_{k=j}^{J-1} \frac{1}{\widehat{f}_k}. \quad (2.2.16)$$

En este caso, el método BF y el método CL difieren sólo en la elección del estimador

de la reclamación última $C_{i,J}$, i.e., el estimador a *priori* μ_i v.s. el estimador CL $\hat{C}_{i,J}^{CL}$

$$\begin{aligned}\hat{C}_{i,J}^{BF} &= C_{i,I-i} + \left(1 - \hat{\beta}_{I-i}^{(CL)}\right) \hat{\mu}_i, \\ \hat{C}_{i,J}^{CL} &= C_{i,I-i} + \left(1 - \hat{\beta}_{I-i}^{(CL)}\right) \hat{C}_{i,J}^{CL}.\end{aligned}\quad (2.2.17)$$

Más adelante se estudiará un método para la estimación del error de predicción condicional del método BF si se usa $\{\hat{\beta}_j^{(CL)}\}$ para la estimación del patrón de desarrollo $\{\beta_j\}$.

Ejemplo 2.2.3.

Se comenzará por hacer el estimado de los parámetros $\hat{\beta}_j$ y $\hat{\mu}_i$, para el conjunto de datos GenIns.

Considérese a *priori* $\hat{\mu}_i$ como el promedio de las reclamaciones observadas por año accidente multiplicado por un factor de experiencia (en este caso, 1.05), entonces

```

Código de R
R> estimador_a_priori <- cbind(GenIns, mu_i = sapply(1:m,function(i){
+   1.05*mean(GenIns[i,m-i+1])}))
R> mu_i<-estimador_a_priori[,11]

R> mu_i
   1         2         3         4         5         6         7         8         9        10
3691889 4173886 4536619 4853016 5083933 5393264 5691518 5876399 5522832 5218316
Código de R

```

De acuerdo al Ejemplo 2.2.2, se obtuvieron los estimadores CL para el mismo conjunto de datos, i.e.,

```

Código de R
R> factores.CL
   1         2         3         4         5         6         7         8         9
3.490607 1.747333 1.457413 1.173852 1.103824 1.086269 1.053874 1.076555 1.017725
Código de R

```

Entonces, supóngase que $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$, por tanto, los estimadores $\hat{\beta}_j$, son

```

Código de R
R> beta_j<-c(rev(cumprod(rev(factores.CL^(-1))))),1)
R> beta_j
   1         2         3         4         5         6         7         8
0.06922055 0.24162171 0.42219349 0.61531022 0.72228295 0.79727292 0.86605315 0.91271120
   9        10
0.98258397 1.00000000
Código de R

```

De acuerdo al estimador BF (2.2.10), la estimación de la siniestralidad última es

```

Código de R
R> BF.J<-sapply(1:m,function(i){
+   getLatestCumulative(GenIns)[i]+(1-beta_j[m-i+1])*mu_i[i])
R> BF.J
      1      2      3      4      5      6      7      8      9     10
3901463 5411778 5305311 5238314 4903962 5189513 5672599 6259919 5551690 5201115
Código de R

```

Entonces, la reserva estimada es

```

Código de R
R> reservaBF<-sum(BF.J-getLatestCumulative(GenIns))
R> reservaBF
[1] 18277574
Código de R

```

La reserva obtenida por este método es menor que la del Ejemplo 2.2.2 ya que si se observa el triángulo de reclamaciones acumuladas (base de datos GenIns) durante los primeros periodos de desarrollo se tienen *outliers*, es decir, grandes incrementos en las reclamaciones acumuladas que de acuerdo a la consideración de $\hat{\mu}_i$, en este caso el promedio, la reserva estimada con el método BF resulta benéfica para la compañía.

A continuación se muestra una comparación de los estimados de la siniestralidad última y el cálculo de la reserva utilizando cada método descrito anteriormente

| <i>i</i> | Estimador | | | Reserva | | |
|----------|----------------------|-----------------------|----------------------|----------|----------|---------|
| | $\hat{C}_{i,J}^{CL}$ | $\hat{C}_{i,J}^{raz}$ | $\hat{C}_{i,J}^{BF}$ | CL | Raz | BF |
| 1 | 3901463 | 3901463 | 3901463 | 0 | 0 | 0 |
| 2 | 5433719 | 5433719 | 5432071 | 94634 | 94634 | 92986 |
| 3 | 5378826 | 5369821 | 5337843 | 469511 | 460506 | 428528 |
| 4 | 5297906 | 5283340 | 5202852 | 709638 | 695072 | 614584 |
| 5 | 4858200 | 4838368 | 4658536 | 984889 | 965057 | 785225 |
| 6 | 5111171 | 5124540 | 4823055 | 1419459 | 1432828 | 1025251 |
| 7 | 5660771 | 5710061 | 4519624 | 2177641 | 2226931 | 1339925 |
| 8 | 6784799 | 6818274 | 4519624 | 3920301 | 3953776 | 1655126 |
| 9 | 5642266 | 5664341 | 2397187 | 4278972 | 4301047 | 1033893 |
| 10 | 4969825 | 5097236 | 664215 | 4625811 | 4753222 | 320201 |
| Total | 53038946 | 53241163 | 41653809 | 18680856 | 18883073 | 7295718 |

Se observa que el estimador *a priori* para el método BF eliminó los posibles *outliers* de los datos y redujo (comparado a los otros métodos) la reserva estimada.

2.3. Medición del error en modelos Chain-Ladder

Anteriormente sólo se dio un estimado para el valor esperado de la siniestralidad última. Naturalmente, se desea saber qué tan bien predice los valores de las variables aleatorias, $(X_{i,j})$. Para poder responder esta pregunta surge la idea de un modelo estocástico subyacente, i.e. ya no sólo se verán como algoritmos.

Para medir la calidad de las reservas de reclamaciones pendientes estimadas, se considerarán los segundos momentos. Entones, se calculará el “Error Cuadrático Medio de Predicción” (ECMP).

2.3.1. Error cuadrático medio de predicción

Supóngase que se tiene una variable aleatoria X y un conjunto de observaciones \mathcal{D} . Se supondrá que \hat{X} es un estimador de $\mathbb{E}[X|\mathcal{D}]$, \mathcal{D} -medible, i.e. \hat{X} es un predictor de X que también es \mathcal{D} -medible.

Definición 2.3.1. (*Error Cuadrático Medio de Predicción Condicional*)

Dada la información \mathcal{D} , se define el error cuadrático medio de predicción condicional (ECMP) del predictor \hat{X} de X como

$$ECMP_{X|\mathcal{D}}(\hat{X}) := \mathbb{E} \left[(\hat{X} - X)^2 \middle| \mathcal{D} \right].$$

Si \hat{X} es un estimador/predictor \mathcal{D} -medible entonces

$$\begin{aligned} ECMP_{X|\mathcal{D}}(\hat{X}) &= \mathbb{E} \left[(\hat{X} - X)^2 \middle| \mathcal{D} \right] = \mathbb{E} \left[\hat{X}^2 - 2X\hat{X} + X^2 \middle| \mathcal{D} \right] \\ &= \hat{X}^2 - 2\hat{X}^2 \mathbb{E}[X|\mathcal{D}] + \mathbb{E}[X^2|\mathcal{D}] \\ &= \mathbb{E}[X^2|\mathcal{D}] - \mathbb{E}^2[X|\mathcal{D}] + \hat{X}^2 - 2\hat{X}^2 \mathbb{E}[X|\mathcal{D}] + \mathbb{E}^2[X|\mathcal{D}] \\ &= Var(X|\mathcal{D}) + \left(\hat{X} - \mathbb{E}[X|\mathcal{D}] \right)^2. \end{aligned}$$

Al término $Var(X|\mathcal{D})$ se le conoce como proceso de varianza condicional (error estocástico). Éste describe la variación en el modelo estocástico, es decir, la aleatoriedad pura que no puede ser eliminada. El término $\left(\hat{X} - \mathbb{E}[X|\mathcal{D}] \right)^2$ es el error de estimación del parámetro que refleja la incertidumbre en la estimación de los parámetros y la esperanza condicional, respectivamente. En general, este error de estimación se vuelve más pequeño si se tienen más observaciones pero éste no desaparece completamente porque está prediciendo el com-

portamiento esperado futuro con información pasada.

Con el fin de estimar el error de estimación se necesita calcular explícitamente $(\hat{X} - \mathbb{E}[X|\mathcal{D}])^2$. Esto se podría hacer fácilmente si $\mathbb{E}[X|\mathcal{D}]$ fuese conocido, pero primero es necesario estimar a \hat{X} .

Una posible manera de garantizar la calidad de \hat{X} es estudiar las posibles fluctuaciones alrededor de $\mathbb{E}[X|\mathcal{D}]$.

2.3.1.1. Error cuadrático medio de predicción no-condicional

Supóngase que X es independiente de \mathcal{D} (por ejemplo, si se tienen experimentos independientes e idénticamente distribuidos). En este caso

$$\mathbb{E}[X|\mathcal{D}] = \mathbb{E}[X] \quad \text{y} \quad \text{Var}(X|\mathcal{D}) = \text{Var}(X).$$

Si se considera el error cuadrático del predictor no-condicional de \hat{X}

$$ECMP_X(\hat{X}) = \mathbb{E} \left[ECMP_{X|\mathcal{D}}(\hat{X}) \right] = \text{Var}(X) + \mathbb{E} \left[(\hat{X} - \mathbb{E}[X])^2 \right].$$

Además, si \hat{X} es un estimador insesgado de $\mathbb{E}[X]$, ($\mathbb{E}[\hat{X}] = \mathbb{E}[X]$), entonces

$$ECMP_X(\hat{X}) = \text{Var}(X) + \text{Var}(\hat{X}).$$

Por tanto, el error de estimación de parámetro se estima a partir de la varianza de \hat{X} .

Ejemplo 2.3.1.

Sean X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas con media μ y varianza $\sigma^2 < \infty$. Para el estimador $\hat{X} = \frac{1}{n} \sum_{i=1}^n X_i$ de μ se tiene que

$$ECMP_{X|\mathcal{D}}(\hat{X}) = \sigma^2 + \left(\frac{1}{n} \sum_{i=1}^n X_i - \mu \right)^2.$$

Nótese que $\lim_{n \rightarrow \infty} ECMP_{X|\mathcal{D}}(\hat{X}) = \sigma^2$. Si se quisiera obtener el número de observaciones, n , para que este error se minimizara, se necesita calcular la “distancia” entre $\frac{1}{n} \sum_{i=1}^n X_i$ y μ . Sin embargo, en general μ es desconocido; por tanto, sólo se puede dar un estimado de dicha distancia. El ECMP no-condicional está dado por

$$ECMP_X(\hat{X}) = \sigma^2 + \frac{\sigma^2}{n}.$$

Por lo tanto, se puede decir que la desviación de $\frac{1}{n} \sum_{i=1}^n X_i$, alrededor de μ , es de orden $\frac{\sigma^2}{n}$. Sin embargo, esto no dice nada acerca del error de estimación del parámetro para una realización específica de $\frac{1}{n} \sum_{i=1}^n X_i$.

2.3.1.2. Error cuadrático medio de predicción condicional

Supóngase que X no es independiente de las observaciones \mathcal{D} .

En este caso,

$$\begin{aligned} ECMP_X(\hat{X}) &= \mathbb{E} \left[ECMP_{X|\mathcal{D}}(\hat{X}) \right] \\ &= \mathbb{E} [Var(X|\mathcal{D})] + \mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}] \right)^2 \right] \\ &= Var(X) - Var(\mathbb{E}[X|\mathcal{D}]) + \mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X|\mathcal{D}] \right)^2 \right] \\ &= Var(X) - 2\mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X] \right) \left(\mathbb{E}[X|\mathcal{D}] - \mathbb{E}[X] \right) \right] + \mathbb{E} \left[\left(\hat{X} - \mathbb{E}[X] \right)^2 \right]. \end{aligned}$$

Si además, \hat{X} es un estimador insesgado de $\mathbb{E}[X]$ entonces

$$ECMP_X(\hat{X}) = Var(X) - 2Cov(\hat{X}, \mathbb{E}[X|\mathcal{D}]) + Var(\hat{X}).$$

De nuevo, esto dice información con respecto al error de estimación promedio, pero no dice nada acerca de la calidad del estimador \hat{X} para una realización específica.

En adelante, se medirá la calidad de los estimadores y predictores para la siniestralidad última a partir de los segundos momentos tales como el ECMP y el coeficiente de variación (condicionales). Sin embargo, lo más deseable de los métodos estocásticos de reservas es obtener la distribución predictiva completa de las reservas (England & Verrall (2007)). Sin embargo, en la mayoría de los casos no es posible calcular la distribución predictiva analíticamente y se utilizarán métodos numéricos tales como el *bootstrap* y cadenas de Markov Monte Carlo (MCMC) para simular una distribución predictiva para las reservas. Una vez que se tenga la distribución predictiva simulada se podrán aproximar los primeros dos momentos de las reservas para reclamaciones, intervalos de predicción, cuantiles, medidas de riesgo, etc.

Sin embargo, algunos cálculos prácticos como el cálculo del requerimiento de capital, es suficiente la estimación de los segundos momentos tales como el ECMP condicional, los componentes del proceso de varianza/estimación de error y el coeficiente de variación. Es decir, no se requiere conocer explícitamente la distribución predictiva, sólo algunas de sus características numéricas, dada su fácil interpretación.

2.3.2. Modelo Chain-Ladder con componente de varianza

Un aspecto importante en el desarrollo de modelos estocásticos subyacentes al método CL fue la aportación de Kremer (1982) ya que descubrió que la estructura del modelo CL parametrizada es idéntica a la de un modelo lineal que involucra una variable de log-respuesta con dos covariables que no interactúan (*two-way analysis of variance*).

Hachemeister & Stanard (1975), Kremer (1985) y Mack (1991) demostraron que el estimador CL de siniestralidad última se puede obtener maximizando un verosimilitud Poisson. Renshaw (1989) y Renshaw & Verrall (1998) relacionaron directamente el método CL con modelos lineales generalizados; Verrall (1989) dio una representación espacial del método CL y utilizó un filtro de Kalman para predecir la siniestralidad última.

Todos estos modelos son diferentes desde el punto de vista estocástico pero tienen ciertas propiedades y con éstos se obtiene la misma reserva que el algoritmo CL ya sea mediante estimadores máximo verosímiles, estimadores por momentos o estimadores Bayesianos. A continuación, se estudiará el modelo de distribución libre para el algoritmo CL como el que se hizo en la Sección 2.2.2 pero ahora incorporando una hipótesis sobre la varianza de las reclamaciones acumuladas o incrementales.

Entonces, el método CL tiene algunas deficiencias; entre las cuales se encuentran que:

- Se debe satisfacer la propiedad de homogeneidad, es decir, no debe haber tendencia en los factores de desarrollo, o bien los datos se tienen que subdividir en sub-portafolios de tal forma que para cada sub-portafolio satisfaga la propiedad de homogeneidad (al menos empíricamente).
 - Para los factores de desarrollo “viejos” (f_j con j grande), hay muy poca información disponible, y esta puede no ser representativa para años de accidente más recientes (para i grande).
 - Las primeras observaciones de los años de accidente a veces no son representativas de la evolución de las reclamaciones y generan problemas en años de accidente más recientes. En caso de que haya un *outlier* en la última diagonal contable y este *outlier* se proyecte a la siniestralidad última, llevará a un resultado no muy adecuado. Por lo tanto, para años de accidente más recientes o en caso de alguna observación sospechosa en la última diagonal, es preferible usar el método BF.
 - Para ramos de cola larga, la diferencia entre los pagos acumulados y la reclamaciones ocurridas es muy grande, debido a que no cumplen con la propiedad de homogeneidad. Además, si ocurre un fenómeno en los datos, los métodos basados en siniestros ocurridos generalmente sobre-estiman su efecto, mientras que las estimaciones de pagos realizados lo subestiman. Generalmente se debe a los ajustes de siniestros o porque
-

los nuevos fenómenos en la siniestralidad que se paga sólo se observan con conforme van ocurriendo con respecto al tiempo.

Además, de la Suposición 2.2.1, se definirá una vez más el modelo CL; pero esta vez se extenderá la definición incluyendo los segundos momentos para estimar el ECMP condicional del estimador CL.

Suposición 2.3.1. (*Modelo CL de distribución libre*)

- 1) Las reclamaciones acumuladas, $C_{i,j}$, de diferentes años de accidente i son independientes.
- 11) $\{C_{i,j}\}_{j=0}^{\infty}$ es una cadena de Markov entonces, existen factores f_0, \dots, f_{J-1} y parámetros (de varianza) $\sigma_0^2, \dots, \sigma_{J-1}^2$ tales que para cualquiera $i \in \{0, \dots, I\}, j \in \{1, \dots, J\}$

$$\begin{aligned}\mathbb{E}[C_{i,j}|C_{i,j-1}] &= f_{j-1}C_{i,j-1}, \\ \text{Var}(C_{i,j}|C_{i,j-1}) &= \sigma_{j-1}^2 C_{i,j-1}.\end{aligned}$$

Observación 2.3.1.

En el modelo de Mack (1993), la hipótesis de cadena de Markov se reemplaza con supuestos con respecto a los dos primeros momentos de $\{C_{i,j}\}_{j=0}^{\infty}$. La propiedad de Markov no es fundamental, pero simplificará las deducciones y la notación. Además, no se hace ninguna suposición sobre la distribución de $C_{i,j}$ dado $C_{i,j-1}$, se hacen suposiciones con respecto a su primer y segundo momentos condicionales. ∇

Recordatorio:

1. Los parámetros f_j y σ_j^2 se estiman como

$$\begin{aligned}\hat{f}_j &= \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \frac{C_{i,j+1}}{C_{i,j}} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} F_{i,j+1}, \\ \hat{\sigma}_j^2 &= \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2,\end{aligned}$$

donde, $F_{i,j+1} := \frac{C_{i,j+1}}{C_{i,j}}$ son los factores de desarrollo individuales.

2. \hat{f}_j es insesgado para f_j .

3. $\widehat{f}_j | \mathcal{B}_j$ es condicionalmente insesgado para f_j .
4. $\widehat{f}_0, \dots, \widehat{f}_{J-1}$ no están correlacionados.
5. Los factores *age-to-age*, \widehat{f}_j , son promedios ponderados de los factores $F_{i,j+1}$, i.e.

$$\widehat{f}_j = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} F_{i,j+1}. \quad (2.3.1)$$

6. $F_{i,j+1}$ es condicionalmente insesgado para f_j , dado $C_{i,j}$.

Lema 2.3.1.

Sean ρ_1, \dots, ρ_H estimadores independientes e insesgados de μ con varianzas $\sigma_1^2, \dots, \sigma_H^2 > 0$, respectivamente. Entonces, el estimador que es una combinación lineal de ρ_1, \dots, ρ_H , insesgado y de mínima varianza es

$$\widehat{\rho} = \frac{\sum_{h=1}^H \frac{\rho_h}{\sigma_h^2}}{\sum_{h=1}^H \frac{1}{\sigma_h^2}}, \quad \text{con} \quad \text{Var}(\widehat{\rho}) = \frac{1}{\sum_{h=1}^H \frac{1}{\sigma_h^2}}.$$

Demostración:

Defínase a los vectores $\rho = (\rho_1, \dots, \rho_H)^\top \in \mathbb{R}^H$ y $\alpha = (\alpha_1, \dots, \alpha_H)^\top \in \mathbb{R}^H$.

Si V es la matriz de covarianzas de ρ , entonces para $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^H$ se satisface que

$$\mathbb{E}[\alpha^\top \rho] = \mu \alpha^\top \mathbf{1} \quad \text{y} \quad \text{Var}(\alpha^\top \rho) = \alpha^\top V \alpha = \sum_{h=1}^H \alpha_h^2 \sigma_h^2.$$

A partir de multiplicadores de Langrange se obtendrá el estimador insesgado de varianza mínima, i.e.

$$\mathcal{L}(\alpha, \lambda) = \frac{1}{2} \alpha^\top V \alpha - \lambda (\alpha^\top \mathbf{1} - 1).$$

La optimización requiere la solución al siguiente sistema de ecuaciones

$$\begin{cases} \frac{\partial}{\partial \alpha} \mathcal{L}(\alpha, \lambda) = V \alpha - \alpha \mathbf{1} = 0, \\ \frac{\partial}{\partial \lambda} \mathcal{L}(\alpha, \lambda) = \alpha^\top \mathbf{1} - 1 = 0. \end{cases}$$

Si $\alpha^\top \mathbf{1} = \mathbf{1}\alpha^\top$ y $\alpha = \lambda V^{-1}\mathbf{1}$, entonces $\lambda = (\mathbf{1}^\top V^{-1}\mathbf{1})^{-1}$.

Por lo tanto, si se resuelve para α se obtiene que

$$\alpha = (\mathbf{1}^\top V^{-1}\mathbf{1})^{-1} V^{-1}\mathbf{1}.$$

De esta forma, $Var(\alpha^\top \rho) = (\mathbf{1}^\top V^{-1}\mathbf{1})^{-1}$ por el hecho que V es simétrica y positivo-definida. Entonces,

$$\alpha = \left(\sum_{h=1}^H \frac{1}{\sigma_h^2} \right)^{-1} \left(\frac{1}{\sigma_1^2}, \dots, \frac{1}{\sigma_H^2} \right)^\top.$$

De aquí que $\hat{\rho} = \alpha^\top \rho$. □

Con el lema 2.3.1 se puede establecer la siguiente propiedad de optimalidad de los factores CL estimados.

Proposición 2.3.1.

Bajo la Suposición 2.3.1. El estimador \hat{f}_j es un estimador insesgado de f_j , que tiene varianza condicional mínima entre las combinaciones lineales insesgadas de los estimadores insesgados $\{F_{i,j+1}\}_{i=0}^{I-j-1}$ para f_j , condicionados sobre \mathcal{B}_j . i.e.

$$Var(\hat{f}_j | \mathcal{B}_j) = \min_{\sum_i \alpha_i = 1} \left\{ Var \left(\sum_{i=0}^{I-j-1} \alpha_i F_{i,j+1} \mid \mathcal{B}_j \right) \right\}.$$

Además,

$$Var(\hat{f}_j | \mathcal{B}_j) = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}.$$

Demostración:

Considérese los factores de desarrollo individuales $F_{i,j+1} = \frac{C_{i,j+1}}{C_{i,j}}$. Si se condiciona $\{F_{i,j+1}\}_{i=0}^{I-j-1}$ con \mathcal{B}_j se obtienen los estimadores insesgados e independientes para f_j tales que

$$Var(F_{i,j+1} | \mathcal{B}_j) = Var(F_{i,j+1} | C_{i,j}) = \frac{\sigma_j^2}{C_{i,j}},$$

con la hipótesis inicial y de acuerdo al Lema 2.3.1, se puede deducir que

$$Var(\hat{f}_j | \mathcal{B}_j) = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}.$$

□

Proposición 2.3.2.

Bajo la Suposición 2.3.1, se satisface que $\hat{\sigma}_j^2$ es un estimador insesgado de σ_j^2 , i.e.

(a) $\mathbb{E}(\hat{\sigma}_j^2 | \mathcal{B}_j) = \sigma_j^2.$

(b) $\mathbb{E}(\hat{\sigma}_j^2) = \sigma_j^2.$

Demostración:

(a) Nótese que

$$\begin{aligned} \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j\right)^2 \middle| \mathcal{B}_j\right] &= \mathbb{E}\left[\left(\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right) - (\hat{f}_j - f_j)\right)^2 \middle| \mathcal{B}_j\right] \\ &= \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 \middle| \mathcal{B}_j\right] - 2\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j\right)(\hat{f}_j - f_j) \middle| \mathcal{B}_j\right] \\ &\quad + \mathbb{E}\left[(\hat{f}_j - f_j)^2 \middle| \mathcal{B}_j\right]. \end{aligned}$$

Sin embargo, el primer término está dado por

$$\mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - f_j\right)^2 \middle| \mathcal{B}_j\right] = Var\left(\frac{C_{i,j+1}}{C_{i,j}} \middle| \mathcal{B}_j\right) = \frac{\sigma_j^2}{C_{i,j}}.$$

A partir de la suposición de independencia entre los años accidente, para el segundo término se tiene que

$$\begin{aligned} \mathbb{E}\left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j\right)(\hat{f}_j - f_j) \middle| \mathcal{B}_j\right] &= Cov\left(\frac{C_{i,j+1}}{C_{i,j}}, \hat{f}_j \middle| \mathcal{B}_j\right) \\ &= \frac{C_{i,j}}{\sum_{i=0}^{I-j-1} C_{i,j}} Var\left(\frac{C_{i,j+1}}{C_{i,j}} \middle| \mathcal{B}_j\right) = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}. \end{aligned}$$

Para el último término

$$\mathbb{E} \left[\left(\hat{f}_j - f_j \right)^2 \middle| \mathcal{B}_j \right] = \text{Var} \left(\hat{f}_j \middle| \mathcal{B}_j \right) = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}}.$$

Si se juntan los resultados obtenidos, entonces

$$\mathbb{E} \left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2 \middle| \mathcal{B}_j \right] = \sigma_j^2 \left(\frac{1}{C_{i,j}} - \frac{1}{\sum_{i=0}^{I-j-1} C_{i,j}} \right).$$

Por lo tanto,

$$\begin{aligned} \mathbb{E} [\hat{\sigma}_j^2 | \mathcal{B}_j] &= \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \mathbb{E} \left[\left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j \right)^2 \middle| \mathcal{B}_j \right] \\ &= \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \sigma_j^2 \left(\frac{1}{C_{i,j}} - \frac{1}{\sum_{i=0}^{I-j-1} C_{i,j}} \right) \\ &= \frac{\sigma_j^2}{I-j-1} \sum_{i=0}^{I-j-1} \left(1 - \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \right) \\ &= \frac{\sigma_j^2}{I-j-1} \left[I-j - \frac{\sum_{i=0}^{I-j-1} C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \right] \\ &= \frac{\sigma_j^2}{I-j-1} (I-j-1) = \sigma_j^2. \end{aligned}$$

$$(b) \mathbb{E} \left(\hat{\sigma}_j^2 \right) = \mathbb{E} \left[\mathbb{E} \left(\hat{\sigma}_j^2 | \mathcal{B}_j \right) \right] = \sigma_j^2.$$

□

Observación 2.3.2.

Las proposiciones 2.3.1 y 2.3.2 hacen evidente la forma en la que se eligió a los estimadores $\hat{\sigma}_j^2$ y \hat{f}_j ya que estos son estimadores insesgados de σ_j^2 y f_j , respectivamente. ∇

La siguiente igualdad tiene un papel importante en la obtención de un estimador para el error de estimación condicional.

$$\mathbb{E} \left[\hat{f}_j^2 \middle| \mathcal{B}_j \right] = \text{Var} \left(\hat{f}_j \middle| \mathcal{B}_j \right) + f_j^2 = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2. \quad (2.3.2)$$

Ya se dijo cómo predecir la siniestralidad última, $C_{i,J}$, en el método CL dada la información \mathcal{D}_I a partir de la expresión

$$\widehat{C}_{i,J}^{CL} = \widehat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I] = C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{J-1}.$$

El objetivo ahora es obtener, en un contexto del modelo CL de distribución libre, un estimador del ECMP condicional de $\widehat{C}_{i,J}^{CL}$ para los años de accidente $i \in \{1, \dots, I\}$

$$\begin{aligned} ECMP_{C_{i,J}|\mathcal{D}_I}(\widehat{C}_{i,J}^{CL}) &= \mathbb{E} \left[\left(\widehat{C}_{i,J}^{CL} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\ &= \text{Var}(C_{i,J}|\mathcal{D}_I) + \left(\widehat{C}_{i,J}^{CL} - \mathbb{E}[C_{i,J}|\mathcal{D}_I] \right)^2. \end{aligned} \quad (2.3.3)$$

También para los años de accidente agregados

$$ECMP_{\sum_i C_{i,J}|\mathcal{D}_I} \left(\sum_{i=1}^I \widehat{C}_{i,J}^{CL} \right) = \mathbb{E} \left[\left(\sum_{i=1}^I \widehat{C}_{i,J}^{CL} - \sum_{i=1}^I C_{i,J} \right)^2 \middle| \mathcal{D}_I \right].$$

A partir de la ecuación (2.3.3), se ve clara la necesidad de dar un estimado del proceso de varianza condicional y del error de estimación condicional, suponiendo que f_j se está estimando a partir de \hat{f}_j .

2.3.2.1. Estimación del proceso de varianza condicional

Primero se considerará el término $\text{Var}(C_{i,J}|\mathcal{D}_I)$ de la ecuación (2.3.3), que se conoce como proceso de varianza condicional. Para $i > 0$,

$$\begin{aligned} \text{Var}(C_{i,J}|\mathcal{D}_I) &= \text{Var}(C_{i,J}|C_{i,I-i}) \\ &= \mathbb{E}[\text{Var}(C_{i,J}|C_{i,J-1}|C_{i,I-i})] + \text{Var}(\mathbb{E}[C_{i,J}|C_{i,J-1}|C_{i,I-i}]) \\ &= \sigma_{J-1}^2 \mathbb{E}[C_{i,J-1}|C_{i,I-i}] + f_{J-1}^2 \text{Var}(C_{i,J-1}|C_{i,I-i}) \\ &= \sigma_{J-1}^2 C_{i,I-i} \prod_{m=I-i}^{J-2} f_m + f_{J-1}^2 \text{Var}(C_{i,J-1}|C_{i,I-i}). \end{aligned}$$

Por lo tanto, se obtendrá una fórmula recursiva del proceso de varianza condicional para un año de accidente i ,

$$\begin{aligned} \text{Var}(C_{i,J}|C_{i,I-i}) &= C_{i,I-i} \sum_{j=I-i}^{J-1} \prod_{n=j+1}^{J-1} f_n^2 \sigma_j^2 \prod_{m=I-i}^{j-1} f_m \\ &= \sum_{j=I-i}^{J-1} \prod_{n=j+1}^{J-1} f_n^2 \sigma_j^2 \mathbb{E}[C_{i,j}|C_{i,I-i}] \\ &= (\mathbb{E}[C_{i,j}|C_{i,I-i}])^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2 / f_j^2}{\mathbb{E}[C_{i,j}|C_{i,I-i}]}. \end{aligned}$$

Proposición 2.3.3. (*Proceso de Varianza para un año de accidente i*)

Bajo la Suposición 2.3.1, el proceso de varianza condicional para la reclamación última de un solo año de accidente i , $i \in \{1, \dots, I\}$, está dado por

$$\text{Var}(C_{i,J}|\mathcal{D}_I) = (\mathbb{E}[C_{i,j}|C_{i,I-i}])^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2/f_j^2}{\mathbb{E}[C_{i,j}|C_{i,I-i}]}.$$

De aquí que el estimador condicional para el proceso de varianza condicional para un año de accidente i es

$$\begin{aligned} \widehat{\text{Var}}(C_{i,J}|\mathcal{D}_I) &= \widehat{\mathbb{E}} \left[(C_{i,J} - \mathbb{E}[C_{i,J}|\mathcal{D}_I])^2 \middle| \mathcal{D}_I \right] \\ &= \left(\widehat{C}_{i,J}^{CL} \right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2/\widehat{f}_j^2}{\widehat{C}_{i,j}^{CL}}. \end{aligned} \quad (2.3.4)$$

El estimador en la ecuación (2.3.4) se puede reescribir de forma recursiva. Para $j \in \{I-i+1, \dots, J\}$, como

$$\widehat{\text{Var}}(C_{i,j}|\mathcal{D}_I) = \widehat{\text{Var}}(C_{i,j-1}|\mathcal{D}_I)\widehat{f}_{j-1}^2 + \widehat{\sigma}_{j-1}^2\widehat{C}_{i,j-1}^{CL},$$

con, $\widehat{\text{Var}}(C_{i,I-i}|\mathcal{D}_I) = 0$ y $\widehat{C}_{i,I-i}^{CL} = C_{i,I-i}$.

Bajo el supuesto de independencia para cada año accidente i , se cumple además que

$$\text{Var} \left(\sum_{i=1}^I C_{i,J} \middle| \mathcal{D}_I \right) = \sum_{i=1}^I \text{Var}(C_{i,J}|\mathcal{D}_I).$$

Por lo tanto, el estimador del proceso de varianza condicional para años accidente agregados está dado por

$$\widehat{\text{Var}} \left(\sum_{i=1}^I C_{i,J} \middle| \mathcal{D}_I \right) = \sum_{i=1}^I \widehat{\text{Var}}(C_{i,J}|\mathcal{D}_I).$$

2.3.2.2. Estimación del error para años accidente individuales

Ahora se obtendrá un estimador para el error condicional de estimación del parámetro, i.e. se requiere obtener un estimador para la exactitud de los factores CL estimados: \widehat{f}_j .

El parámetro de error para un año accidente individual en el estimado CL es

$$\begin{aligned} \left(\widehat{C}_{i,J} - \mathbb{E}[C_{i,J}|\mathcal{D}_I]\right)^2 &= C_{i,I-i}^2 \left(\widehat{f}_{I-i} \cdots \widehat{f}_{J-1} - f_{I-i} \cdots f_{J-1}\right)^2 \\ &= C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \widehat{f}_j^2 - 2 \prod_{I-i}^{J-1} \widehat{f}_j f_j + \prod_{j=I-i}^{J-1} f_j^2 \right]. \end{aligned} \quad (2.3.5)$$

Es decir, el objetivo es calcular los productos en la ecuación (2.3.5).

Las realizaciones de los estimadores $\widehat{f}_{I-i} \cdots \widehat{f}_{J-1}$ son conocidos al tiempo I ; sin embargo, los “verdaderos” factores CL, $f_{I-i} \cdots f_{J-1}$, son desconocidos (de hecho son los que se intenta estimar). Por lo tanto (2.3.5) no se puede calcular directamente.

Con esta idea intuitiva de medir el grado en el que los posibles factores CL estimados, \widehat{f}_j , varían alrededor de los verdaderos valores f_j se obtendrá el error condicional de estimación.

Con este razonamiento resulta muy natural que las técnicas que se estudiarán son Bayesianas o de re-muestreo para intentar saber qué otros valores de \widehat{f}_j se pueden tomar.

Hay varias metodologías para re-muestrear estos valores \widehat{f}_j . Para explicar estas metodologías se fijará un año accidente $i \in \{1, \dots, I\}$.

A partir del lado derecho de la expresión (2.3.5) se puede notar que la principal dificultad cuando se estudia la volatilidad de los estimadores proviene del cálculo del cuadrado de los estimadores CL.

El término $\prod_{j=I-i}^{J-1} f_j \widehat{f}_j$ de la expresión (2.3.5) se puede calcular como un promedio gracias al insesgamiento y no-correlación de los estimadores de los factores CL.

Por lo tanto, se centrará la atención en remuestrear el producto

$$\widehat{f}_{I-i}^2 \cdots \widehat{f}_{J-1}^2.$$

Notación 2.3.1.

Se define a la esquina superior de las observaciones \mathcal{D}_I con respecto al año de desarrollo $j = I - i + 1$ como

$$\mathcal{D}_{I,i}^\circ = \{C_{k,j} \in \mathcal{D}_I : j > I - i\} \subseteq \mathcal{D}_I.$$

| Año de accidente i | Año de desarrollo j | | | | |
|----------------------|-----------------------|-----|---------|-----|---------------------------|
| | 0 | ... | $I - i$ | ... | J |
| 0 | | | | | |
| ⋮ | | | | | |
| i | | | | | $\mathcal{D}_{I,i}^\circ$ |
| ⋮ | | | | | |
| I | | | | | |

▽

Observación 2.3.3.

Ya se estableció que para $j \in \{1, \dots, J - 1\}$, \hat{f}_j es \mathcal{B}_{j+1} -medible.

▽

A continuación, se evaluarán tres propuestas para el re-muestreo y estimación de la exactitud de los factores CL. Asimismo, se verán algunas fortalezas y debilidades/inconvenientes de dichas propuestas,

- *Propuesta 1:* Re-muestreo no condicional en $\mathcal{D}_{I,i}^\circ$

Este método calcula la esperanza

$$\mathbb{E} \left[\hat{f}_{I-i}^2 \cdots \hat{f}_{J-1}^2 \mid \mathcal{B}_{I-i} \right], \quad (2.3.6)$$

que es un “promedio completo” sobre la distribución multi-dimensional después del tiempo $I - i$.

Como $\mathcal{D}_{I,i}^\circ \cap \mathcal{B}_{I-i} = \emptyset$, entonces el valor de (2.3.6) no depende de las observaciones en $\mathcal{D}_{I,i}^\circ$. Esto significa que las observaciones en la esquina superior $\mathcal{D}_{I,i}^\circ$ no tiene influencia en la estimación del parámetro de error. Es por esta razón que se le llama al método de esta manera, pues proporciona el error de estimación promedio independientemente de las observaciones en $\mathcal{D}_{I,i}^\circ$.

- *Propuesta 2:* Re-muestreo condicional parcial en $\mathcal{D}_{I,i}^\circ$

En este método se calcula el valor

$$\hat{f}_{I-i}^2 \cdots \hat{f}_{J-2}^2 \cdot \mathbb{E} \left[\hat{f}_{J-1}^2 \mid \mathcal{B}_{J-1} \right]. \quad (2.3.7)$$

Es decir, el “promedio” se esta obteniendo parcialmente.

En este caso como $\mathcal{D}_{I,i}^\circ \cap \mathcal{B}_{J-1} \neq \emptyset$ entonces la expresión (2.3.7) depende de las observaciones en $\mathcal{D}_{I,i}^\circ$.

- *Propuesta 3*: Re-muestreo condicional en $\mathcal{D}_{I,i}^\circ$

En este método se calculará

$$\mathbb{E} \left[\hat{f}_{I-i}^2 | \mathcal{B}_{I-i} \right] \mathbb{E} \left[\hat{f}_{I-i+1}^2 | \mathcal{B}_{I-i+1} \right] \cdots \mathbb{E} \left[\hat{f}_{J-1}^2 | \mathcal{B}_{J-1} \right]. \quad (2.3.8)$$

A diferencia de la ecuación (2.3.7), aquí se está promediando en cada posición $j \in \{I-i, \dots, J-1\}$ sobre la estructura condicional.

Como $\mathcal{D}_{I,i}^\circ \cap \mathcal{B}_j \neq \emptyset$, si $j > I-i$ entonces las realizaciones observadas en $\mathcal{D}_{I,i}^\circ$ tienen una influencia directa en el estimado y (2.3.8) depende de las observaciones en $\mathcal{D}_{I,i}^\circ$. A diferencia de (2.3.6), el promedio sólo se hace sobre las distribuciones condicionales y no sobre la distribución multi-dimensional después de $I-i$, de aquí que éste se conozca como re-muestreo condicional. Desde un punto de vista numérico, la propuesta 3 permite una estructura multiplicativa de la medida de volatilidad.

Una vez que se especificó que tipo de cálculo realizan cada una de las 3 propuestas, se darán estimaciones del error bajo las tres propuestas.

Propuesta 1: Re-muestreo no-condicional

Por la no-correlación e insesgamiento de los factores CL, entonces el error se obtiene mediante

$$\begin{aligned} \mathbb{E} \left[\left(\hat{C}_{i,J}^{CL} - \mathbb{E}(C_{i,J} | \mathcal{D}_I) \right)^2 \middle| \mathcal{B}_{I-i} \right] &= C_{i,I-i}^2 \mathbb{E} \left[\prod_{j=I-i}^{J-1} \hat{f}_j^2 - 2 \prod_{I-i}^{J-1} \hat{f}_j f_j + \prod_{j=I-i}^{J-1} f_j^2 \middle| \mathcal{B}_{I-i} \right] \\ &= C_{i,I-i}^2 \left[\mathbb{E} \left(\prod_{j=I-i}^{J-1} \hat{f}_j^2 \middle| \mathcal{B}_{I-i} \right) - 2 \mathbb{E} \left(\prod_{I-i}^{J-1} \hat{f}_j f_j \middle| \mathcal{B}_{I-i} \right) + \prod_{j=I-i}^{J-1} f_j^2 \right] \\ &= C_{i,I-i}^2 \left[\mathbb{E} \left(\prod_{j=I-i}^{J-1} \hat{f}_j^2 \middle| \mathcal{B}_{I-i} \right) - 2 \prod_{I-i}^{J-1} f_j f_j + \prod_{j=I-i}^{J-1} f_j^2 \right] \\ &= C_{i,I-i}^2 \left[\mathbb{E} \left(\prod_{j=I-i}^{J-1} \hat{f}_j^2 \middle| \mathcal{B}_{I-i} \right) - \prod_{j=I-i}^{J-1} f_j^2 \right]. \end{aligned} \quad (2.3.9)$$

Entonces para dar un estimado del error de estimación bajo esta idea no-condicional, se tiene que calcular la esperanza $\mathbb{E} \left[\prod_{j=I-i}^{J-1} \hat{f}_j^2 | \mathcal{B}_{I-i} \right]$ de la expresión (2.3.9). Evidentemente, si los factores CL, \hat{f}_j , fuesen independientes este cálculo sería trivial.

Aunque los factores no son independientes, sí son no-correlacionados. Eso se establece en el siguiente resultado.

Proposición 2.3.4.

Bajo la Suposición 2.3.1, para $j \in \{1, \dots, J - 1\}$ se cumple que $Cov(\hat{f}_{j-1}^2, \hat{f}_j^2 | \mathcal{B}_{j-1}) < 0$.

Demostración:

Nótese que para $j \in \{1, \dots, J - 1\}$, \hat{f}_{j-1} es \mathcal{B}_j -medible. Defínase

$$S_j^{[k]} := \sum_{i=0}^k C_{i,j}, \quad k \geq 0. \quad (2.3.10)$$

Por la ecuación (2.3.2),

$$\mathbb{E}[\hat{f}_j^2 | \mathcal{B}_j] = \frac{\sigma_j^2}{\sum_{i=0}^{I-j-1} C_{i,j}} + f_j^2.$$

Entonces,

$$\begin{aligned} Cov(\hat{f}_{j-1}^2, \hat{f}_j^2 | \mathcal{B}_{j-1}) &= \mathbb{E}\left[Cov(\hat{f}_{j-1}^2, \hat{f}_j^2 | \mathcal{B}_j) | \mathcal{B}_{j-1}\right] + Cov\left(\mathbb{E}[\hat{f}_{j-1}^2 | \mathcal{B}_j], \mathbb{E}[\hat{f}_j^2 | \mathcal{B}_j] | \mathcal{B}_{j-1}\right) \\ &= Cov(\hat{f}_{j-1}^2, \mathbb{E}[\hat{f}_j^2 | \mathcal{B}_j] | \mathcal{B}_{j-1}) + 0 \\ &= Cov\left(\hat{f}_{j-1}^2, \frac{\sigma_j^2}{S_j^{[I-j-1]}} + f_j^2 | \mathcal{B}_{j-1}\right) \\ &= \frac{\sigma_j^2}{(S_{j-1}^{[I-j]})^2} Cov\left(\left(\sum_{i=0}^{I-j} C_{i,j}\right)^2, \frac{1}{S_j^{[I-j-1]}} | \mathcal{B}_{j-1}\right). \end{aligned}$$

Sin embargo,

$$\left(\sum_{i=0}^{I-j} C_{i,j}\right)^2 = (S_j^{[I-j-1]})^2 + 2S_j^{[I-j-1]}C_{I-j,j} + C_{I-j,j}^2,$$

y también $\mathbb{E}[C_{I-j,j} | \mathcal{B}_{j-1}] = f_{j-1} \cdot C_{I-j,j-1}$. Entonces por la independencia de los diferentes

años accidente se tiene que

$$\begin{aligned} Cov\left(\hat{f}_{j-1}^2, \hat{f}_j^2 \middle| \mathcal{B}_{j-1}\right) &= \frac{\sigma_j^2}{\left(S_{j-1}^{[I-j]}\right)^2} \left[Cov\left(\left(S_j^{[I-j-1]}\right)^2, \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1}\right) \right. \\ &\quad \left. + 2 f_{j-1} C_{I-j, j-1} \cdot Cov\left(S_j^{[I-j-1]}, \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1}\right) + 0 \right]. \end{aligned}$$

Sin embargo, para $\alpha = 1, 2$

$$\begin{aligned} Cov\left(\left(S_j^{[I-j-1]}\right)^\alpha, \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1}\right) &= \mathbb{E}\left[\left(S_j^{[I-j-1]}\right)^{\alpha-1} \middle| \mathcal{B}_{j-1}\right] \\ &\quad - \mathbb{E}\left[\left(S_j^{[I-j-1]}\right)^\alpha \middle| \mathcal{B}_{j-1}\right] \mathbb{E}\left[\left(S_j^{[I-j-1]}\right)^{-1} \middle| \mathcal{B}_{j-1}\right] \\ &< \mathbb{E}\left[\left(S_j^{[I-j-1]}\right)^{\alpha-1} \middle| \mathcal{B}_{j-1}\right] \\ &\quad - \mathbb{E}^\alpha\left[S_j^{[I-j-1]} \middle| \mathcal{B}_{j-1}\right] \mathbb{E}^{-1}\left[S_j^{[I-j-1]} \middle| \mathcal{B}_{j-1}\right] = 0. \end{aligned}$$

(por la desigualdad de Jensen)

Por lo tanto,

$$Cov\left(\hat{f}_{j-1}^2, \hat{f}_j^2 \middle| \mathcal{B}_{j-1}\right) < 0.$$

□

Aunque ya se pudo afirmar algo con respecto a la estructura de dependencia de $\{\hat{f}_j^2\}_{j=1}^{J-1}$ aún no se puede calcular fácilmente $\mathbb{E}\left[\prod_{j=I-i}^{J-1} f_j^2 \middle| \mathcal{B}_{I-i}\right]$.

Es por esto que con la Propuesta 1 no se pueda encontrar una expresión cerrada para el error de estimación.

Teorema 2.3.1.

Bajo la Suposición 2.3.2,

(i)

$$C_{i,I-i}^2 \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} \hat{f}_j - \prod_{j=I-i}^{J-1} f_j \right)^2 \middle| \mathcal{B}_{I-i} \right] \geq C_{i,I-i}^2 \prod_{j=I-i}^{J-1} f_j^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2 / f_j^2}{\mathbb{E} \left[S_j^{[I-j-1]} \middle| \mathcal{B}_{I-i} \right]}.$$

(ii)

$$C_{i,I-i}^2 \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} \hat{f}_j - \prod_{j=I-i}^{J-1} f_j \right)^2 \middle| \mathcal{B}_{I-i} \right] \leq C_{i,I-i}^2 \left(\mathbb{E} \left[\prod_{j=I-i}^{J-1} \left(\frac{\sigma_j^2}{S_j^{[I-j-1]}} + f_j^2 \right) \middle| \mathcal{B}_{I-i} \right] - \prod_{j=I-i}^{J-1} f_j^2 \right).$$

Observación 2.3.4.

El Teorema 2.3.1 proporciona cotas para el error de estimación, sin embargo la cota superior (ii) no se puede calcular directamente dado que aparece el término $\mathbb{E} \left[\prod_{j=I-i}^{J-1} \left(\frac{\sigma_j^2}{S_j^{[I-j-1]}} + f_j^2 \right) \right]$ pero se puede estimar a partir de simulación y *bootstrapping*. ∇

Para demostrar el Teorema 2.3.1, se necesitan algunos lemas auxiliares.

Lema 2.3.2.

Bajo la Suposición 2.3.2

$$\mathbb{E} \left[\hat{f}_{j-1}^2 \hat{f}_j^2 \middle| \mathcal{B}_{j-1} \right] \geq \left(\frac{\sigma_{j-1}^2}{S_{j-1}^{[I-j]}} + f_{j-1}^2 \right) f_j^2 + \frac{\sigma_j^2 f_{j-1}}{S_{j-1}^{[I-j-1]}}.$$

Demostración:

$$\begin{aligned}
\mathbb{E} \left[\widehat{f}_{j-1}^2 \widehat{f}_j^2 \middle| \mathcal{B}_{j-1} \right] &= \mathbb{E} \left[\widehat{f}_{j-1}^2 \mathbb{E} \left(\widehat{f}_j^2 \middle| \mathcal{B}_j \right) \middle| \mathcal{B}_{j-1} \right] \\
&= \mathbb{E} \left[\widehat{f}_{j-1}^2 \left(\frac{\sigma_j^2}{S_j^{[I-j-1]}} + f_j^2 \right) \middle| \mathcal{B}_{j-1} \right] \\
&= \left(\frac{\sigma_{j-1}^2}{S_{j-1}^{[I-j]}} + f_{j-1}^2 \right) f_j^2 + \frac{\sigma_j^2}{\left(S_{j-1}^{[I-j]} \right)^2} \mathbb{E} \left[\left(S_j^{[I-j]} \right)^2 \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right].
\end{aligned} \tag{2.3.11}$$

Sin embargo,

$$\begin{aligned}
\mathbb{E} \left[\left(S_j^{[I-j]} \right)^2 \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right] &= \mathbb{E} \left[\left(S_j^{[I-j-1]} + C_{I-j,j} \right)^2 \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right] \\
&= \mathbb{E} \left[S_j^{[I-j-1]} \middle| \mathcal{B}_{j-1} \right] + 2 \mathbb{E} \left[C_{I-j,j} \middle| \mathcal{B}_{j-1} \right] + \mathbb{E} \left[\frac{C_{I-j,j}^2}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right] \\
&= f_{j-1} \left(S_{j-1}^{[I-j]} + C_{I-j,j-1} \right) + \mathbb{E} \left[\frac{C_{I-j,j}^2}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right].
\end{aligned}$$

Obsérvese que los términos del cociente de la esperanza condicional tienen diferente año accidente y por tanto son independientes. Entonces, de la desigualdad de Jensen se tiene que

$$\begin{aligned}
\mathbb{E} \left[\left(S_j^{[I-j]} \right)^2 \frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right] &= f_{j-1} \left(S_{j-1}^{[I-j]} + C_{I-j,j-1} \right) + \mathbb{E} \left[C_{I-j,j}^2 \middle| \mathcal{B}_{j-1} \right] \cdot \mathbb{E} \left[\frac{1}{S_j^{[I-j-1]}} \middle| \mathcal{B}_{j-1} \right] \\
&\geq f_{j-1} \left(S_{j-1}^{[I-j]} + C_{I-j,j-1} \right) + \frac{\mathbb{E}^2 \left[C_{I-j,j} \middle| \mathcal{B}_{j-1} \right]}{\mathbb{E} \left[S_j^{[I-j-1]} \middle| \mathcal{B}_{j-1} \right]} \\
&= f_{j-1} \left(S_{j-1}^{[I-j]} + C_{I-j,j-1} + \frac{C_{I-j,j-1}^2}{S_{j-1}^{[I-j-1]}} \right) = f_{j-1} \frac{\left(S_{j-1}^{[I-j]} \right)^2}{S_{j-1}^{[I-j-1]}}.
\end{aligned}$$

Sustituyendo este resultado en la ecuación (2.3.11), se obtiene

$$\mathbb{E} \left[\widehat{f}_{j-1}^2 \widehat{f}_j^2 \middle| \mathcal{B}_{j-1} \right] \geq \left(\frac{\sigma_{j-1}^2}{S_{j-1}^{[I-j]}} + f_{j-1}^2 \right) f_j^2 + \frac{\sigma_j^2 f_{j-1}^2}{S_{j-1}^{[I-j-1]}}.$$

□

Método 3: *Re-muestreo condicional*

En el método 3 se re-muestran los factores CL observados \hat{f}_j sobre la estructura condicional. Para describir a detalle esta metodología se harán suposiciones más fuertes a partir de un modelo de serie de tiempo.

Suposición 2.3.2. (*Modelo de serie de tiempo*)

1. Las reclamaciones acumuladas $C_{i,j}$ de diferentes años de accidente i son independientes.
2. Existen constantes $f_j > 0$, $\sigma_j > 0$ y variables aleatorias $\varepsilon_{i,j+1}$ tales que para todo $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J-1\}$

$$C_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \varepsilon_{i,j+1},$$

donde, dado \mathcal{B}_0 , $\varepsilon_{i,j+1}$ son independientes y $\mathbb{E}[\varepsilon_{i,j+1} | \mathcal{B}_0] = 0$, $\mathbb{E}[\varepsilon_{i,j+1}^2 | \mathcal{B}_0] = 1$ y $\mathbb{P}(C_{i,j+1} > 0 | \mathcal{B}_0) = 1$ para cualesquiera $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J-1\}$.

Dado \mathcal{B}_0 , las variables aleatorias $\varepsilon_{i,j+1}$ se definen para garantizar que las reclamaciones acumuladas $C_{i,j+1}$ sean positivas $\mathbb{P}(\cdot | \mathcal{B}_0)$ -c.s., i.e., todos los resultados se harán bajo la medida $\mathbb{P}(\cdot | \mathcal{B}_0)$. Esta condición de positividad también se hizo en el modelo CL de libre distribución (Suposición 2.3.1) para que se tuviera una condición de varianza válida.

Observación 2.3.5.

Se puede demostrar que la Suposición 2.3.2 implica la Suposición 2.3.1 ∇

Entonces, bajo el método 3, se remuestreará en el modelo de series de tiempo, i.e., se remuestrearán observaciones de $\hat{f}_{I-i}, \dots, \hat{f}_{J-1}$, dado el triángulo superior \mathcal{D}_I .

De esta forma se está considerando la posibilidad de que, dado \mathcal{D}_I , las observaciones de \hat{f}_j pueden haber sido diferentes para los valores observados. Para considerar esta fuente de incertidumbre se puede hacer lo siguiente: Dado \mathcal{D}_I , se generan “nuevas” observaciones $\tilde{C}_{i,j+1}$ para $i \in \{0, \dots, I\}$, $j \in \{0, \dots, J-1\}$ a partir de la expresión

$$\tilde{C}_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \tilde{\varepsilon}_{i,j+1}, \quad (2.3.12)$$

donde $\sigma_j > 0$ y, dado \mathcal{B}_0 , $\varepsilon_{i,j+1}$ y $\tilde{\varepsilon}_{i,j+1}$ son independientes y tienen la misma distribución. En este sentido, $C_{i,j}$ es una “medida de volumen” fija y se simula $\tilde{C}_{i,j+1} \stackrel{(d)}{=} C_{i,j+1}$ dado \mathcal{B}_j .

Observación 2.3.6.

Se utiliza la notación “ \sim ” para enfatizar que $\tilde{C}_{i,j+1}$ son variables aleatorias y $C_{i,j}$ son deterministas (dado \mathcal{D}_I). ∇

Entonces se remuestran las observaciones \hat{f}_j sólo remuestrando las observaciones del año de desarrollo $j + 1$, y con la suposición de remuestreo (2.3.12) se obtienen la siguiente representación remuestreada para los estimadores de los factores de desarrollo

$$\hat{f}_j = \frac{\sum_{i=0}^{I-j-1} \tilde{C}_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}} = f_j + \frac{\sigma_j}{S_j^{[I-j-1]}} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}} \tilde{\varepsilon}_{i,j+1}, \quad (2.3.13)$$

donde $S_j^{[I-j-1]} = \sum_{i=0}^{I-j-1} C_{i,j}$.

Notación 2.3.2.

- En la expresión (2.3.13) y el siguiente análisis se utilizará sólo la notación \hat{f}_j para los estimados remuestreados de los factores de desarrollo f_j para no complicar la notación.
- Se denotará por $\mathbb{P}_{\mathcal{D}_I}^*$ a la medida de probabilidad inducida de estos estimados CL remuestreados. ∇

Dado \mathcal{B}_I , los estimados remuestreados de los factores de desarrollo \hat{f}_j tienen la misma distribución que los factores estimados originales.

A diferencia de las observaciones $\{C_{i,j} : i + j \leq I\}$, los valores $\{\tilde{C}_{i,j} : i + j \leq I\}$ y los estimadores remuestreados \hat{f}_j son variables aleatorias, dado \mathcal{D}_I . Además, dado $\mathcal{B}_0 \subseteq \mathcal{D}_I$ las variables aleatorias $\tilde{\varepsilon}_{i,j}$ son independientes. Entonces,

1. Los estimados remuestreados $\hat{f}_0, \dots, \hat{f}_{J-1}$ son independientes con respecto a $\mathbb{P}_{\mathcal{D}_I}^*$.
2. Para $j \in \{0, \dots, J-1\}$, $\mathbb{E}_{\mathcal{D}_I}^* [\hat{f}_j] = f_j$.
3. Para $j \in \{0, \dots, J-1\}$, $\mathbb{E}_{\mathcal{D}_I}^* \left[(\hat{f}_j)^2 \right] = f_j^2 + \frac{\sigma_j^2}{S_j^{[I-j-1]}}$.

Entonces, en el método 3, el error de estimación condicional se estima mediante la siguiente expresión

$$\begin{aligned}
C_{i,I-i}^2 \cdot \mathbb{E}_{\mathcal{D}_I}^* \left[\left(\widehat{f}_{I-i} \cdots \widehat{f}_{J-1} - f_{I-i} \cdots f_{J-1} \right)^2 \right] &= C_{i,I-i}^2 \cdot \text{Var}_{\mathbb{P}_{\mathcal{D}_I}^*} \left(\widehat{f}_{I-i} \cdots \widehat{f}_{J-1} \right) \\
&= C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \mathbb{E}_{\mathcal{D}_I}^* \left[\left(\widehat{f}_j \right)^2 \right] - \prod_{j=I-i}^{J-1} f_j^2 \right] \\
&= C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(f_j^2 + \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} f_j^2 \right].
\end{aligned} \tag{2.3.14}$$

Observación 2.3.7.

El cálculo en la expresión (2.3.14) es exacto. Dicha “facilidad” de cálculo se debe a que bajo el método 3, el estimador se obtiene bajo la medida condicional $\mathbb{P}_{\mathcal{D}_I}^*$. ∇

Remplazando los parámetros $\sigma_{I-i}^2, \dots, \sigma_{J-1}^2$ y f_{I-i}, \dots, f_{J-1} con sus estimados, se obtiene el siguiente estimador para el error de estimación condicional para el año de accidente $i \in \{1, \dots, I\}$

$$\begin{aligned}
\widehat{\text{Var}} \left(\widehat{C}_{i,J}^{CL} \mid \mathcal{D}_I \right) &= \widehat{\mathbb{E}}_{\mathcal{D}_I}^* \left[\left(\widehat{C}_{i,J}^{CL} - \mathbb{E}[C_{i,J} \mid \mathcal{D}_I] \right)^2 \right] \\
&= C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \right],
\end{aligned}$$

que para $j \in \{I-i+1, \dots, J\}$ se puede escribir recursivamente como

$$\begin{aligned}
\widehat{\text{Var}} \left(\widehat{C}_{i,j}^{CL} \mid \mathcal{D}_I \right) &= \widehat{\text{Var}} \left(\widehat{C}_{i,j-1}^{CL} \mid \mathcal{D}_I \right) \widehat{f}_{j-1}^2 + C_{i,I-i}^2 \frac{\widehat{\sigma}_{j-1}^2}{S_{j-1}^{[I-j-1]}} \prod_{l=I-i}^{j-2} \left(\widehat{f}_l^2 + \frac{\widehat{\sigma}_l^2}{S_l^{[I-l-1]}} \right) \\
&= \widehat{\text{Var}} \left(\widehat{C}_{i,j-1}^{CL} \mid \mathcal{D}_I \right) \left(\widehat{f}_{j-1}^2 + \frac{\widehat{\sigma}_{j-1}^2}{S_{j-1}^{[I-j-1]}} \right) + C_{i,I-i}^2 \frac{\widehat{\sigma}_{j-1}^2}{S_{j-1}^{[I-j-1]}} \prod_{l=I-i}^{j-2} \widehat{f}_l^2,
\end{aligned}$$

donde $\prod_{j=m}^m C_j := 1$ y $\widehat{\text{Var}} \left(\widehat{C}_{i,I-i}^{CL} \mid \mathcal{D}_I \right) = 0$.

Estimador 2.3.1. *Error cuadrático medio de predicción condicional para años accidente individuales*

Bajo la Suposición 2.3.2, el estimador del error cuadrático medio de predicción condicional de la siniestralidad última para un año de accidente individual $i \in \{1, \dots, I\}$ está dado por

$$\begin{aligned} \widehat{ECMP}_{C_{i,J}|\mathcal{D}_I}(\widehat{C}_{i,J}^{CL}) &= \widehat{\mathbb{E}} \left[\left(\widehat{C}_{i,J}^{CL} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\ &= \left(\widehat{C}_{i,J}^{CL} \right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2 / \widehat{f}_j^2}{\widehat{C}_{i,J}^{CL}} \\ &\quad + C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \right]. \end{aligned} \quad (2.3.15)$$

Se puede reescribir (2.3.15) como

$$\widehat{ECMP}_{C_{i,J}|\mathcal{D}_I}(\widehat{C}_{i,J}^{CL}) = \left(\widehat{C}_{i,J}^{CL} \right)^2 \left[\sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2 / \widehat{f}_j^2}{\widehat{C}_{i,J}^{CL}} + \prod_{j=I-i}^{J-1} \left(\frac{\widehat{\sigma}_j^2 / \widehat{f}_j^2}{S_j^{[I-j-1]}} + 1 \right) - 1 \right].$$

Y se puede hacer una aproximación lineal del error de estimación mediante

$$C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \right] \approx C_{i,I-i}^2 \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2 / \widehat{f}_j^2}{S_j^{[I-j-1]}}. \quad (2.3.16)$$

En particular, $C_{i,I-i}^2 e \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2 / \widehat{f}_j^2}{S_j^{[I-j-1]}}$ es una cota inferior de

$$C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(\widehat{f}_j^2 + \frac{\widehat{\sigma}_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^2 \right].$$

Esto motiva al siguiente estimador

Estimador 2.3.2. *Error cuadrático medio de predicción para años accidente individuales*

Bajo la Suposición 2.3.2, el estimador del error cuadrático medio de predicción condicional de la siniestralidad última para un año de accidente individual $i \in \{1, \dots, I\}$ está dado por

$$\widehat{ECMP}_{C_{i,J}|D_I}(\widehat{C}_{i,J}^{CL}) = \left(\widehat{C}_{i,J}^{CL}\right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^2}{\widehat{f}_j^2} \left(\frac{1}{\widehat{C}_{i,J}^{CL}} + \frac{1}{S_j^{[I-j-1]}} \right). \quad (2.3.17)$$

2.4. Algunos modelos con hipótesis distribucionales

Hasta el momento se ha trabajado con modelos de “distribución libre” ya que no se han dado suposiciones explícitas con respecto a la distribución de las reclamaciones (ya sean acumuladas o incrementales), aunque sí se hicieron suposiciones con respecto a la estructura de dependencia de dichas reclamaciones. En esta sección se trabajarán “de nuevo” los modelos que ya se estudiaron pero haciendo algunas suposiciones distribucionales explícitas.

2.4.1. Modelo Poisson para número de reclamaciones

El modelo Poisson se utiliza para el conteo de siniestros, pues se considera como el primer modelo para el cálculo de la reserva de reclamaciones, se basa en una suposición distribucional, que permite calcular los estimadores máximo verosímiles (ML) de los parámetros necesarios para la estimación de la reserva. El modelo Poisson es un modelo estocástico alternativo que se usa como motivación para el cálculo de la reserva CL.

Suposición 2.4.1. (*Modelo Poisson*)

Existen parámetros $\mu_0, \dots, \mu_I > 0$ y $\gamma_0, \dots, \gamma_J > 0$ tales que las reclamaciones incrementales $X_{i,j}$ son independientes y con distribución Poisson donde para cualesquiera $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$,

$$\mathbb{E}[X_{i,j}] = \mu_i \gamma_j \quad \text{con} \quad \sum_{j=0}^J \gamma_j = 1.$$

Observación 2.4.1.

La Suposición 2.4.1 considera que las reclamaciones incrementales $X_{i,j}$ son no-negativas. Sin embargo, en varias aplicaciones se observan incrementos negativos. En dichos casos, el modelo Poisson no es apropiado. ∇

$$\text{Además, si } C_{i,J} = \sum_{j=0}^J X_{i,j}, \text{ entonces } \mathbb{E}[C_{i,J}] = \sum_{j=0}^J \mathbb{E}[X_{i,j}] = \sum_{j=0}^J \mu_i \gamma_j = \mu_i \sum_{j=0}^J \gamma_j = \mu_i.$$

Por lo tanto, $C_{i,J} \sim \text{Poisson}(\mu_i)$.

Es decir, la reclamación acumulada para el año de accidente i , $C_{i,J}$, tiene distribución Poisson con media $\mathbb{E}[C_{i,J}] = \mu_i$. Así, el parámetro μ_i representa la siniestralidad total esperada en el año de accidente i y el parámetro γ_j es un “patrón de flujos” de efectivo/reportes durante los diferentes periodos de desarrollo j .

Nótese que, $\frac{\mathbb{E}[X_{i,j}]}{\mathbb{E}[X_{i,0}]} = \frac{\gamma_j}{\gamma_0}$, es decir, $\frac{\mathbb{E}[X_{i,j}]}{\mathbb{E}[X_{i,0}]}$ no depende de i .

Proposición 2.4.1.

El modelo Poisson satisface la Suposición 2.2.2. Es decir,

- Las reclamaciones acumuladas, $C_{i,j}$, son independientes para diferentes año accidente i .
- Existen parámetros $\mu_0, \dots, \mu_I > 0$ y un patrón $\beta_0, \dots, \beta_J > 0$ con $\beta_J = 1$ tales que para todo $i \in \{0, \dots, I\}, j \in \{0, \dots, J-1\}, k \in \{1, 2, \dots, j\}$

(a) $\mathbb{E}[C_{i,0}] = \beta_0 \cdot \mu_i.$

(b) $\mathbb{E}[C_{i,j+k} | C_{i,0}, \dots, C_{i,j}] = C_{i,j} + (\beta_{j+k} - \beta_j) \mu_i.$

Demostración:

Primero obsérvese que $\mathbb{E}[C_{i,0}] = \mathbb{E}[X_{i,0}] = \mu_i \gamma_0 := \mu_i \beta_0$ con $\gamma_0 = \beta_0$.

Además,

$$\begin{aligned}
 \mathbb{E}[C_{i,j+k}|C_{i,0}, \dots, C_{i,j}] &= C_{i,j} + \sum_{l=1}^k \mathbb{E}[X_{i,j+l}|C_{i,0}, \dots, C_{i,j}] \\
 &= C_{i,j} + \sum_{l=1}^k \mathbb{E}[X_{i,j+l}] \\
 &= C_{i,j} + \sum_{l=1}^k \mu_i \gamma_{j+l} = C_{i,j} + \mu_i \sum_{l=1}^k \gamma_{j+l}.
 \end{aligned}$$

Pero si $\beta_j := \sum_{l=0}^j \gamma_l$, entonces $\beta_{j+k} - \beta_j = \sum_{l=0}^{j+k} \gamma_l - \sum_{l=0}^j \gamma_l = \gamma_{j+1} + \dots + \gamma_{j+k} = \sum_{l=1}^{j+k} \gamma_{j+l}$.

De aquí que,

$$\mathbb{E}[C_{i,j+k}|C_{i,0}, \dots, C_{i,j}] = C_{i,j} + \mu_i \sum_{l=1}^k \gamma_{j+l} = C_{i,j} + \mu_i (\beta_{j+k} - \beta_j),$$

i.e. con μ_0, \dots, μ_I del modelo Poisson, y $\gamma_0 = \beta_0$, $\beta_j = \sum_{l=0}^j \gamma_l$ se satisface la Suposición 2.2.2. \square

Esta proposición establece que el modelo Poisson satisface las suposiciones para el método Bornhuetter & Ferguson. Por tanto, se puede usar el estimador BF para el cálculo de la reserva bajo el modelo Poisson.

Existen diferentes métodos de estimación de los parámetros (μ_i) y (γ_j) , a continuación se considerará una estimación alternativa: la estimación máximo verosímil.

La función de verosimilitud sobre el conjunto de observaciones

$$\mathcal{D}_I = \{X_{i,j}; i+j \leq I, 0 \leq j \leq J\},$$

está dada por

$$\mathcal{L}_{\mathcal{D}_I}(\mu_0, \dots, \mu_I, \gamma_0, \dots, \gamma_J) = \prod_{i+j \leq I} \left(\exp(-\mu_i \gamma_j) \frac{(\mu_i \gamma_j)^{X_{i,j}}}{(X_{i,j})!} \right).$$

Entonces, la log-verosimilitud está dada por

$$\begin{aligned}\log(\mathcal{L}_{\mathcal{D}_I}) &= \sum_{i+j \leq I} \log \left(\exp(-\mu_i \gamma_j) \frac{(\mu_i \gamma_j)^{X_{i,j}}}{(X_{i,j})!} \right) \\ &= \sum_{i+j \leq I} \left(-\mu_i \gamma_j + X_{i,j} \left(\log(\mu_i) + \log(\gamma_j) \right) - \log [(X_{i,j})!] \right) \\ &= \sum_{i+j \leq I} \left(X_{i,j} \log(\mu_i) + X_{i,j} \log(\gamma_j) - \mu_i \gamma_j - \log [(X_{i,j})!] \right).\end{aligned}$$

Se maximiza el logaritmo de la función de verosimilitud igualando a cero las $I + J + 2$ derivadas parciales con respecto a los parámetros desconocidos μ_i y γ_j , i.e.,

$$\frac{\partial}{\partial \mu_i} \log(\mathcal{L}_{\mathcal{D}_I}) = \sum_{j \leq I-i} \frac{X_{i,j}}{\mu_i} - \gamma_j, \quad \text{y} \quad \frac{\partial}{\partial \gamma_j} \log(\mathcal{L}_{\mathcal{D}_I}) = \sum_{i \leq I-j} \frac{X_{i,j}}{\gamma_j} - \mu_i.$$

Entonces para la observaciones \mathcal{D}_I , los estimadores máximo verosímiles (ML) deben satisfacer

$$\sum_{j=0}^{I-i} \hat{\mu}_i \hat{\gamma}_j = \sum_{j=0}^{I-i} X_{i,j} = C_{i,I-i}, \quad \text{y} \quad \sum_{i=0}^{I-j} \hat{\mu}_i \hat{\gamma}_j = \sum_{i=0}^{I-j} X_{i,j}, \quad (2.4.1)$$

para todo $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$ bajo la restricción de que $\sum_{j=0}^J \hat{\gamma}_j = 1$. Además, si este sistema de ecuaciones tiene solución única, se obtiene los estimadores máximo verosímiles $\hat{\mu}_i$ y $\hat{\gamma}_j$, para μ_i y γ_j , respectivamente.

Estimador 2.4.1. (*Estimador Poisson máximo verosímil*)

Los estimadores máximo verosímiles en el modelo Poisson (2.4.1) para $i + j > I$, de $\mathbb{E}[X_{i,j}]$ y $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$ están dados por

$$\hat{X}_{i,j}^{Pois} = \hat{\mathbb{E}}[X_{i,j}] = \hat{\mu}_i \hat{\gamma}_j,$$

y

$$\hat{C}_{i,J}^{Pois} = \hat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I] = C_{i,I-i} + \sum_{j=I-i+1}^J \hat{X}_{i,j}^{Pois}.$$

Observación 2.4.2.

Para la siniestralidad última bajo el modelo Poisson,

$$\begin{aligned}\widehat{C}_{i,J}^{Pois} &= C_{i,I-i} + \left(1 - \sum_{j=0}^{I-i} \widehat{\gamma}_j\right) \widehat{\mu}_i \\ &= C_{i,I-i} + \widehat{\mu}_i - \sum_{j=0}^{I-i} \widehat{\gamma}_j \widehat{\mu}_i \\ &= C_{i,I-i} + \widehat{\mu}_i - C_{i,I-i} = \widehat{\mu}_i.\end{aligned}$$

Por lo tanto, el estimador Poisson máximo verosímil tiene la misma forma que el estimador BF (2.2.2), pero los estimadores máximo verosímiles para μ_i y γ_j dependen los datos \mathcal{D}_I . ∇

2.4.1.1. Obtención del algoritmo CL a partir del modelo Poisson

En esta sección se estudiarán algunas condiciones que hacen equivalentes al modelo Poisson con el estimador CL. Entonces, se supondrá que existe una solución positiva de las ecuaciones (2.4.1).

Lema 2.4.1.

Bajo la Suposición 2.4.1, en el conjunto \mathcal{D}_I se satisface que

$$\sum_{i=0}^{I-j} C_{i,j} = \sum_{i=0}^{I-j} \widehat{\mu}_i \sum_{k=0}^j \widehat{\gamma}_k.$$

Demostración:

La prueba se hará por inducción.

Por la condición de máxima verosimilitud (2.4.1)

$$C_{i,I-i} = \sum_{j=0}^{I-i} X_{i,j} = \sum_{j=0}^{I-i} \widehat{\mu}_i \widehat{\gamma}_j.$$

Entonces para $i = 0$ y $I = J$ se satisface que $C_{0,J} = \sum_{j=0}^J X_{0,j} = \widehat{\mu}_0 \sum_{j=0}^J \widehat{\gamma}_j$.

Para el paso inductivo nótese que

$$\begin{aligned} \sum_{i=0}^{I-(j-1)} C_{i,j-1} &= \sum_{i=0}^{I-j} C_{i,j} - \sum_{i=0}^{I-j} X_{i,j} + \sum_{k=0}^{j-1} X_{I-j+1,k} \\ &= \sum_{i=0}^{I-j} \hat{\mu}_i \sum_{k=0}^j \hat{\gamma}_k - \hat{\gamma}_j \sum_{i=0}^{I-j} \hat{\mu}_i + \hat{\mu}_{I-j+1} \sum_{k=0}^{j-1} \hat{\gamma}_k. \end{aligned}$$

Entonces,

$$\sum_{i=0}^{I-(j-1)} C_{i,j-1} = \sum_{i=0}^{I-j} \hat{\mu}_i \sum_{k=0}^{j-1} \hat{\gamma}_k + \hat{\mu}_{I-j+1} \sum_{k=0}^{j-1} \hat{\gamma}_k = \sum_{i=0}^{I-(j-1)} \hat{\mu}_i \sum_{k=0}^{j-1} \hat{\gamma}_k.$$

□

Proposición 2.4.2.

El estimador CL, $\hat{C}_{i,j}^{CL} = \hat{\mathbb{E}}(C_{i,j}|\mathcal{D}_I) = C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{j-1}$ y el estimador máximo verosímil $\hat{C}_{i,J}^{Pois} = \hat{\mathbb{E}}(C_{i,J}|\mathcal{D}_I) = C_{i,I-i} + \sum_{j=I-i+1}^J \hat{X}_{i,j}^{Pois}$ bajo el modelo Poisson (2.4.1) son equivalentes, es decir,

$$\hat{C}_{i,J}^{CL} = \hat{C}_{i,J}^{Pois}.$$

Demostración:

Para $i \in \{1, 2, \dots, J\}$, el estimador Poisson 2.4.1 está dado por

$$\begin{aligned}
\widehat{C}_{i,J}^{Pois} &= C_{i,I-i} + \sum_{j=I-i+1}^J \widehat{X}_{i,j}^{Pois} \\
&= C_{i,I-i} + \sum_{j=I-i+1}^J \widehat{\mu}_i \widehat{\gamma}_j \\
&= C_{i,I-i} + \sum_{j=0}^{I-i} \widehat{\mu}_i \widehat{\gamma}_j - \sum_{j=0}^{I-i} \widehat{\mu}_i \widehat{\gamma}_j + \sum_{j=I-i+1}^J \widehat{\mu}_i \widehat{\gamma}_j \\
&= C_{i,I-i} - \sum_{j=0}^{I-i} \widehat{\mu}_i \widehat{\gamma}_j + \sum_{j=0}^J \widehat{\mu}_i \widehat{\gamma}_j \\
&= C_{i,I-i} - C_{i,I-i} + \sum_{j=0}^J \widehat{\mu}_i \widehat{\gamma}_j = \sum_{j=0}^J \widehat{\mu}_i \widehat{\gamma}_j.
\end{aligned}$$

Pero también,

$$\widehat{C}_{i,J}^{Pois} = \widehat{\mathbb{E}}(C_{i,J} | \mathcal{D}_I) = C_{i,I-i} \frac{\sum_{j=0}^J \widehat{\gamma}_j}{\sum_{j=0}^{I-i} \widehat{\gamma}_j}. \quad (2.4.2)$$

Que se puede reescribir como

$$\widehat{C}_{i,J}^{Pois} = C_{i,I-i} \frac{\sum_{j=0}^J \widehat{\gamma}_j}{\sum_{j=0}^{I-i} \widehat{\gamma}_j} = C_{i,I-i} \frac{\sum_{j=0}^{I-i+1} \widehat{\gamma}_j}{\sum_{j=0}^{I-i} \widehat{\gamma}_j} \cdots \frac{\sum_{j=0}^J \widehat{\gamma}_j}{\sum_{j=0}^{J-1} \widehat{\gamma}_j}. \quad (2.4.3)$$

Pero por el lema 2.4.1 se tiene que en el conjunto \mathcal{D}_I

$$\sum_{i=0}^{I-j} C_{i,j} = \sum_{i=0}^{I-j} \widehat{\mu}_i \sum_{k=0}^j \widehat{\gamma}_k. \quad (2.4.4)$$

Y de nuevo por la condición de máxima verosimilitud

$$\sum_{i=0}^{I-j} C_{i,j-1} = \sum_{i=0}^{I-j} (C_{i,j} - X_{i,j}) = \sum_{i=0}^{I-j} \widehat{\mu}_i \sum_{k=0}^{j-1} \widehat{\gamma}_k. \quad (2.4.5)$$

Entonces, por las ecuaciones (2.4.4) y (2.4.5) para $j \leq J$

$$\frac{\sum_{k=0}^j \hat{\gamma}_k}{j-1} = \frac{\sum_{i=0}^{I-j} C_{i,j}}{I-j} = \hat{f}_{j-1}.$$

$$\sum_{k=0}^j \hat{\gamma}_k = \sum_{i=0}^{I-j} C_{i,j-1}$$

Ahora por la ecuación (2.4.3)

$$\begin{aligned} \hat{C}_{i,J}^{Pois} &= C_{i,I-i} \frac{\sum_{k=0}^{I-(I-i+1)} C_{k,I-i+1}}{I-(I-i+1)} \cdots \frac{\sum_{k=0}^{I-J} C_{k,J}}{I-J} \\ &= C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{j-1} = \hat{C}_{i,J}^{CL}. \end{aligned}$$

□

Este resultado muestra que el estimador máximo verosímil del modelo Poisson lleva a los mismos factores *age-to-age* f_j del modelo Chain-Ladder de distribución libre.

Corolario 2.4.1.

Bajo la Suposición 2.4.1, para $j \in \{0, 1, \dots, J\}$ se satisface que

$$\sum_{k=0}^j \hat{\gamma}_k = \beta_j^{(CL)} = \prod_{k=j}^{J-1} \frac{1}{\hat{f}_k}.$$

Demostración:

Por la proposición anterior y la ecuación (2.4.2)

$$C_{i,I-i} \frac{\sum_{j=0}^J \hat{\gamma}_j}{I-i} = \hat{C}_{i,J}^{Pois} = \hat{C}_{i,J}^{CL} = C_{i,I-i} \hat{f}_{I-i} \cdots \hat{f}_{J-1}.$$

Pero como $\sum_{j=0}^J \widehat{\gamma}_j = 1$, entonces

$$1 = \sum_{j=0}^{I-i} \widehat{\gamma}_j \prod_{I-i}^{J-1} \widehat{f}_j = \sum_{j=0}^{I-i} \widehat{\gamma}_j \left(\beta_{I-i}^{(CL)} \right)^{-1}.$$

Por lo tanto, $\beta_j^{(CL)} = \sum_{k=0}^j \widehat{\gamma}_k$.

□

El corolario anterior establece que el método CL de distribución libre y el modelo Poisson tengan el mismo patrón de desarrollo, $\left\{ \beta_j^{(CL)} \right\}_j$ y por tanto la misma reserva BF en caso de que se utilice el mismo patrón de desarrollo como estimado para $\left\{ \beta_j \right\}_j$. Es decir, si se utiliza el patrón de desarrollo $\left\{ \beta_j^{(CL)} \right\}_j$ para el método BF y el modelo Poisson, sólo se tendrán diferencias en la elección de la siniestralidad última esperada μ_i , obteniéndose

$$\widehat{C}_{i,J}^{Pois} = C_{i,I-i} + \left(1 - \beta_{I-i}^{(CL)} \right) \widehat{\mu}_i,$$

donde $\widehat{\mu}_i$ es el estimador máximo verosímil del modelo Poisson en la ecuación (2.4.1)

Observación 2.4.3.

Para encontrar los estimadores máximo verosímiles $\widehat{\mu}_i$ y $\widehat{\gamma}_j$ se tiene que resolver el sistema de ecuaciones en (2.4.1). Se puede encontrar dichas soluciones sin directamente resolver el sistema. Esto se hace con ayuda de los factores \widehat{f}_k del modelo CL a partir del corolario anterior a partir de las expresiones

$$\widehat{\gamma}_j = \beta_j^{(CL)} - \beta_{j-1}^{(CL)} = \prod_{k=j}^{J-1} \frac{1}{\widehat{f}_k} \left(1 - \frac{1}{\widehat{f}_{j-1}} \right)$$

y

$$\widehat{\mu}_i = \frac{\sum_{j=0}^{I-i} X_{i,j}}{\sum_{j=0}^{I-i} \widehat{\gamma}_j} = \frac{C_{i,I-i}}{\beta_{I-i}^{(CL)}} = \widehat{C}_{i,J}^{CL}.$$

▽

El modelo CL de distribución libre se puede aplicar a conjuntos de datos más generales de los que supone el modelo Poisson. Estrictamente hablando, el modelo Poisson sólo se

aplica si todas las reclamaciones/pagos incrementales $X_{i,j}$ son positivas y enteras. El modelo CL de distribución libre se puede aplicar aún con incrementos negativos siempre y cuando los acumulados permanezcan positivos.

2.4.2. Modelo log-normal para reclamaciones acumuladas

El modelo considera las reclamaciones acumuladas y supone que los factores CL individuales, $F_{i,j} = \frac{C_{i,j}}{C_{i,j-1}}$, siguen una distribución log-normal. Además, se obtienen diferentes estimadores sesgados e insesgados, dependiendo si el parámetro de la varianza σ^2 es conocido o no.

Suposición 2.4.2. (*Modelo Log-normal para reclamaciones acumuladas*)

- (i) Los factores de desarrollo individuales, $F_{i,j} = \frac{C_{i,j}}{C_{i,j-1}}$ (donde $C_{i,-1} := 1$), tienen distribución log-normal con parámetros ξ_j, σ_j^2 , i.e. para cualesquiera $i \in \{1, \dots, I\}, j \in \{1, \dots, J\}$

$$\eta_{i,j} = \log(F_{i,j}) \sim N(\xi_j, \sigma_j^2).$$

- (ii) Para $i \in \{1, \dots, I\}, j \in \{1, \dots, J\}$, $\eta_{i,j}$ son independientes.

Proposición 2.4.3.

Si para cualesquiera $i \in \{1, \dots, I\}, j \in \{1, \dots, J\}$

$$\eta_{i,j} = \log(F_{i,j}) \sim N(\xi_j, \sigma_j^2),$$

y $\eta_{i,j}$ son independientes, entonces

- (a) $\mathbb{E}(F_{i,j}) = \exp\{\xi_j + \frac{\sigma_j^2}{2}\}.$
 (b) $Var(F_{i,j}) = \exp\{2\xi_j + \sigma_j^2\}(e^{\sigma_j^2} - 1).$
 (c) $\mathbb{E}(C_{i,j}|C_{i,j-1}) = C_{i,j-1} \cdot \exp\{\xi_j + \frac{\sigma_j^2}{2}\}.$
 (d) $Var(C_{i,j}|C_{i,j-1}) = C_{i,j-1}^2 \cdot \exp\{2\xi_j + \sigma_j^2\}(e^{\sigma_j^2} - 1).$

Observación 2.4.4.

Según esta proposición, si se cumplen las suposiciones (2.4.2), entonces

$$\mathbb{E}(C_{i,j}|C_{i,j-1}) = f_j C_{i,j-1},$$

con $f_j = \exp\{\xi_j + \frac{\sigma_j^2}{2}\}$, i.e. se cumplen las hipótesis del modelo CL (2.3.1) ▽

Los estimadores para los parámetros ξ_j y σ_j^2 , respectivamente, están dados por

$$\begin{aligned} \widehat{\xi}_j &= \frac{1}{I-j+1} \sum_{i=0}^{I-j} \log\left(\frac{C_{i,j}}{C_{i,j-1}}\right) \\ &\text{y} \\ \widehat{\sigma}_j^2 &= \frac{1}{I-j} \sum_{i=0}^{I-j} \left(\log\left(\frac{C_{i,j}}{C_{i,j-1}}\right) - \widehat{\xi}_j\right)^2. \end{aligned}$$

Lema 2.4.2.

Si se cumple la Suposición 2.4.2, entonces

$$\widehat{\xi}_j \sim N\left(\xi_j, \frac{\sigma_j^2}{I-j+1}\right)$$

y

$$\frac{I-j}{\sigma_j^2} \widehat{\sigma}_j^2 \sim \chi_{(I-j)}^2,$$

donde $\chi_{(I-j)}^2$ es una distribución- χ^2 con $I-j$ grados de libertad. Además, $\widehat{\xi}_j$ y $\widehat{\sigma}_j^2$ son estocásticamente independientes.

Demostración

La primera fórmula de este lema, es fácil de probar a partir de la Suposición 2.4.2.

La segunda parte es un caso especial del Teorema de Cochran (Hogg y Craig (1995)). □

Observación 2.4.5.

La distribución χ^2 con $n \in \mathbb{N}$ grados de libertad es la distribución Gamma $\Gamma(\gamma, c)$ con $\gamma = \frac{n}{2}$ y $c = \frac{1}{2}$. ▽

Considérese las variables aleatorias

$$Z_{i,j} = \log(C_{i,j}).$$

Entonces,

$$Z_{i,J} = Z_{i,I-i} + \sum_{j=I-i+1}^J \log \left(\frac{C_{i,j}}{C_{i,j-1}} \right) = Z_{i,I-i} + \sum_{j=I-i+1}^J \eta_{i,j}.$$

Se obtiene del lema (2.4.2)

$$\mathbb{E}[Z_{i,J}|\mathcal{D}_I] = Z_{i,I-i} + \sum_{j=I-i+1}^J \xi_j.$$

Por lo tanto, el estimador está dado por

$$\widehat{Z}_{i,J} = \widehat{\mathbb{E}}[Z_{i,J}|\mathcal{D}_I] = Z_{i,I-i} + \sum_{j=I-i+1}^J \widehat{\xi}_j.$$

Lema 2.4.3.

(a) $\mathbb{E} \left[\widehat{Z}_{i,J} | Z_{i,I-i} \right] \mathbb{E}[Z_{i,J}|\mathcal{D}_I] = \mathbb{E}[Z_{i,J}|Z_{i,I-i}].$

(b) $\mathbb{E} \left[\widehat{Z}_{i,J} \right] = \mathbb{E}[Z_{i,J}].$

(c) $ECMP_{Z_{i,J}|\mathcal{D}_I} \left(\widehat{Z}_{i,J} \right) = \sum_{j=I-i+1}^J \sigma_j^2 + \left(\sum_{j=I-i+1}^J \left(\widehat{\xi}_j - \xi_j \right) \right)^2.$

(d) $ECMP_{Z_{i,J}} \left(\widehat{Z}_{i,J} \right) = ECMP_{Z_{i,J}|Z_{i,I-i}} \left(\widehat{Z}_{i,J} \right) = \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right).$

Demostración

(a)

$$\mathbb{E} \left[\widehat{Z}_{i,J} - Z_{i,J} | Z_{i,I-i} \right] = \mathbb{E} \left[\sum_{j=I-i+1}^J \widehat{\xi}_j \right] - \mathbb{E} \left[\sum_{j=I-i+1}^J \eta_{i,j} \right] = 0.$$

(b) Es inmediato de (a).

(c) Si se cumplen las suposiciones (2.4.2), el ECMP condicional se tiene

$$\begin{aligned}
 ECMP_{Z_{i,J}|\mathcal{D}}(\hat{Z}_{i,J}) &= \mathbb{E} \left[\left(\hat{Z}_{i,J} - Z_{i,J} \right)^2 | \mathcal{D}_I \right] \\
 &= \mathbb{E} \left[\left(\sum_{j=I-i+1}^J \hat{\xi}_j - \sum_{j=I-i+1}^J \eta_{i,j} \right)^2 | \mathcal{D}_I \right] \\
 &= \text{Var} \left(\sum_{j=I-i+1}^J \eta_{i,j} \right) + \left(\sum_{j=I-i+1}^J (\hat{\xi}_j - \xi_j) \right)^2 \\
 &= \sum_{j=I-i+1}^J \sigma_j^2 + \left(\sum_{j=I-i+1}^J (\hat{\xi}_j - \xi_j) \right)^2.
 \end{aligned}$$

(d) Nótese que $\hat{Z}_{i,J} - Z_{i,J}$ es independiente de $Z_{i,I-i}$. Utilizando el lema (2.4.2) se tiene que

$$ECMP_{Z_{i,J}|Z_{i,I-i}}(\hat{Z}_{i,J}) = ECMP_{Z_{i,J}}(\hat{Z}_{i,J}) = \mathbb{E} \left[ECMP_{Z_{i,J}|\mathcal{D}_I}(\hat{Z}_{i,J}) \right] \quad (2.4.6)$$

$$= \text{Var} \left(\sum_{j=I-i+1}^J \eta_{i,j} \right) + \text{Var} \left(\sum_{j=I-i+1}^J \hat{\xi}_j \right) \quad (2.4.7)$$

$$= \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right). \quad (2.4.8)$$

□

2.4.2.1. Varianza σ_j^2 conocida

Supóngase que las varianzas $\sigma_0^2, \dots, \sigma_J^2$ son conocidas. Entonces

$$\begin{aligned}
 \mathbb{E}[C_{i,J}|C_{i,I-i}] &= \mathbb{E}[\exp\{Z_{i,J}\}|C_{i,I-i}] \\
 &= \exp\{Z_{i,I-i}\} \exp \left\{ \sum_{j=I-i+1}^J \xi_j + \sum_{j=I-i+1}^J \frac{\sigma_j^2}{2} \right\} \\
 &= C_{i,I-i} \exp \left\{ \sum_{j=I-i+1}^J \xi_j + \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \right\} \quad (2.4.9)
 \end{aligned}$$

y

$$\begin{aligned}\mathbb{E} \left[\exp \left\{ \widehat{Z}_{i,J} \right\} \mid C_{i,I-i} \right] &= \exp \{ Z_{i,I-i} \} \mathbb{E} \left[\exp \left\{ \sum_{j=I-i+1}^J \widehat{\xi}_j \right\} \right] \\ &= C_{i,I-i} \exp \left\{ \sum_{j=I-i+1}^J \xi_j + \frac{1}{2} \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\}. \quad (2.4.10)\end{aligned}$$

Debido a las diferentes varianzas condicionales de $Z_{i,J}$ y $\widehat{Z}_{i,J}$ se obtienen diferentes valores esperados para la distribución log-normal.

Estimador 2.4.2. (*Estimador insesgado con σ_j^2 conocido*)

Si se cumplen la Suposición 2.4.2, entonces para $i \in \{1, \dots, I\}$

$$\widehat{C}_{i,J}^{LN} = \widehat{\mathbb{E}}[C_{i,J} | \mathcal{D}_I] = \exp \left\{ \widehat{Z}_{i,J} + \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\}. \quad (2.4.11)$$

Lema 2.4.4. (*Insensgamiento para σ_j^2 conocida*)

Bajo las Suposición 2.4.2 se tiene

- (a) $\mathbb{E} \left[\widehat{C}_{i,J}^{LN} \mid C_{i,I-i} \right] = \mathbb{E}[C_{i,J} | C_{i,I-i}] = \mathbb{E}[C_{i,J} | \mathcal{D}_I]$.
- (b) $\mathbb{E} \left[\widehat{C}_{i,J}^{LN} \right] = \mathbb{E}[C_{i,J}]$.
- (c)

$$\begin{aligned}ECMP_{C_{i,J} | C_{i,I-i}} \left(\widehat{C}_{i,J}^{LN} \right) &= \left(\exp \left\{ \sum_{j=I-i+1}^J \sigma_j^2 \right\} \right. \\ &\quad \left. + \exp \left\{ \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} - 2 \right) (\mathbb{E}[C_{i,J} | C_{i,I-i}])^2\end{aligned}$$

Demostración

(a)

$$\begin{aligned}\mathbb{E} \left[\widehat{C}_{i,J}^{LN} \middle| C_{i,I-i} \right] &= \mathbb{E} \left[\exp \left\{ \widehat{Z}_{i,J} \right\} \middle| C_{i,I-i} \right] \exp \left\{ \frac{1}{2} \sum_{j=I-i+1}^J \left(\sigma_j^2 - \frac{\sigma_j^2}{I-j+1} \right) \right\} \\ &= \mathbb{E}[C_{i,J} | C_{i,I-i}] = \mathbb{E}[C_{i,J} | \mathcal{D}_I].\end{aligned}$$

(b) Es inmediato de (a).

(c) Usando la no correlación

$$\mathbb{E} \left[C_{i,J} \widehat{C}_{i,J}^{LN} \middle| C_{i,I-i} \right] = \mathbb{E}[C_{i,J} | C_{i,I-i}] \mathbb{E} \left[\widehat{C}_{i,J}^{LN} \middle| C_{i,I-i} \right].$$

Entonces se obtiene el ECMP

$$\begin{aligned}ECMP_{C_{i,J} | C_{i,I-i}} \left(\widehat{C}_{i,J}^{LN} \right) &= \mathbb{E} \left[\left(C_{i,J} - \widehat{C}_{i,J}^{LN} \right)^2 \middle| C_{i,I-i} \right] \\ &= \text{Var} \left(C_{i,J} - \widehat{C}_{i,J}^{LN} \middle| C_{i,I-i} \right) \\ &= \text{Var}(C_{i,J} | C_{i,I-i}) + \text{Var} \left(\widehat{C}_{i,J}^{LN} \middle| C_{i,I-i} \right) \\ &= C_{i,I-i}^2 \text{Var} \left(\exp \left\{ \sum_{j=I-i+1}^J \eta_{i,j} \right\} \right) \\ &+ C_{i,I-i}^2 \exp \left\{ \sum_{j=I-i+1}^J \sigma_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\} \\ &\times \text{Var} \left(\exp \left\{ \sum_{j=I-i+1}^J \widehat{\xi}_j \right\} \right). \quad (2.4.12)\end{aligned}$$

Por lo tanto, esto implica que

$$\begin{aligned}\mathbb{E} \left[\left(C_{i,J} - \widehat{C}_{i,J}^{LN} \right)^2 \middle| C_{i,I-i} \right] &= C_{i,I-i}^2 \exp \left\{ \sum_{j=I-i+1}^J (2\xi_j + \sigma_j^2) \right\} \\ &\times \left[\exp \left(\sum_{j=I-i+1}^J \sigma_j^2 \right) + \exp \left(\sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right) - 2 \right].\end{aligned}$$

□

Observación 2.4.6.

El Lema 2.4.4 calcula la aproximación no condicional del parámetro error. Entonces el ECMP está dado por

$$\begin{aligned} ECMP_{C_{i,J}|\mathcal{D}_I}(\widehat{C}_{i,J}^{LN}) &= \mathbb{E}\left[\left(C_{i,J} - \widehat{C}_{i,J}^{LN}\right)^2 \middle| \mathcal{D}_I\right] \\ &= C_{i,I-i}^2 \text{Var}\left(\exp\left\{\sum_{j=I-i+1}^J \eta_{i,j}\right\}\right) + \left(\mathbb{E}[C_{i,J}|C_{i,I-i}] - \widehat{C}_{i,J}^{LN}\right)^2. \end{aligned} \quad (2.4.13)$$

Por lo tanto, del inciso (c) del Lema 2.4.4 expresado en la siguiente manera

$$\begin{aligned} ECMP_{C_{i,J}|C_{i,I-i}}(\widehat{C}_{i,J}^{LN}) &= \mathbb{E}[C_{i,J}|C_{i,I-i}]^2 \\ &\quad \times \left(\exp\left\{\sum_{j=I-i+1}^J \sigma_j^2\right\} - 1 + \exp\left\{\sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1}\right\} - 1\right), \end{aligned}$$

se conforma por

$$\exp\left\{\sum_{j=I-i+1}^J \sigma_j^2\right\} - 1,$$

que corresponde al proceso de varianza, y

$$\exp\left\{\sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1}\right\} - 1,$$

la estimación del parámetro error.

Estimador 2.4.3. (*Adición de Años de Accidente*)

El ECMP para reclamaciones agregadas se estima como

$$\begin{aligned} \widehat{ECMP}_{\sum_i C_{i,J}}\left(\sum_{i=1}^I \widehat{C}_{i,J}^{LN}\right) &= \sum_{i=1}^I ECMP_{\sum_i C_{i,J}|C_{i,I-i}}(\widehat{C}_{i,J}^{LN}) \\ &+ 2 \sum_{1 \leq i < k \leq I} \mathbb{E}[C_{i,J}|C_{i,I-i}] \times \mathbb{E}[C_{k,J}|C_{k,I-k}] \left(\exp\left\{\sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1}\right\} - 1\right). \end{aligned} \quad (2.4.14)$$

El estimador $\widehat{\xi}_j$ es el mismo para diferentes año de accidente.

Notación 2.4.1.

El uso de “|·” en el subíndice de $\widehat{ECMP}_{\sum_i C_{i,J}}$ ($\sum_{i=1}^I \widehat{C}_{i,J}^{LN}$), significa que se condiciona sobre la última diagonal para medir el volumen de $C_{i,I-i}$, pero éste también es considerado una variable aleatoria en el estimador de ξ_{I-i} . ∇

Estimador 2.4.4. (*Predictor central, σ_j^2 conocido*)

Si se cumplen las suposiciones (2.4.2), entonces para cualesquiera $i \in \{1, \dots, I\}$ con $\widehat{C}_{i,J}^{LN} = \exp \left\{ \widehat{Z}_{i,J} + \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\}$

$$\begin{aligned} \widehat{C}_{i,J}^{pc} = \widehat{\mathbb{E}}[C_{i,J} | \mathcal{D}_I] &= \exp \left\{ \widehat{Z}_{i,J} + \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right) \right\} \\ &= \widehat{C}_{i,J}^{LN} \exp \left\{ \sum_{j=I-i+1}^J \left(\frac{\sigma_j^2}{I-j+1} \right) \right\} > \widehat{C}_{i,J}^{LN}. \end{aligned}$$

Lema 2.4.5. (*Predictor central, σ_j^2 conocido*)

Bajo las suposiciones (2.4.2)

(a) $\widehat{C}_{i,J}^{pc}$ satisface $\mathbb{E} \left[\frac{C_{i,J}}{\widehat{C}_{i,J}^{pc}} \middle| C_{i,I-i} \right] = 1$.

(b) El ECMP está dado por

$$\begin{aligned} ECMP_{C_{i,J} | C_{i,I-i}}(\widehat{C}_{i,J}^{pc}) &= \mathbb{E}[C_{i,J} | C_{i,I-i}]^2 \left(\exp \left\{ \sum_{j=I-i+1}^J \frac{3\sigma_j^2}{I-j+1} \right\} \right. \\ &\quad \left. + \exp \left(\sum_{j=I-i+1}^J \sigma_j^2 \right) - 2 \exp \left\{ \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} \right) \\ &> ECMP_{C_{i,J} | C_{i,I-i}}(\widehat{C}_{i,J}^{LN}). \end{aligned}$$

Demostración:

- (a) A partir de, $Z_{i,j} = \log(C_{i,j})$, $Z_{i,J} = Z_{i,I-i} + \sum_{j=I-i+1}^J \eta_{i,j}$, $\widehat{Z}_{i,J} = Z_{i,I-i} + \sum_{j=I-i+1}^J \widehat{\xi}_j$ y las ecuaciones (2.4.11) y (2.4.4)

$$\begin{aligned} \mathbb{E} \left[\frac{C_{i,J}}{\widehat{C}_{i,J}^{pc}} \middle| C_{i,I-i} \right] &= \mathbb{E} \left[\exp \left\{ \sum_{j=I-i+1}^J \eta_{i,j} - \sum_{j=I-i+1}^J \widehat{\xi}_j - \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right) \right\} \right] \\ &= \exp \left\{ -\frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right) \right\} \prod_{j=I-i+1}^J \mathbb{E}[\exp(\eta_{i,j})] \\ &\quad \cdot \prod_{j=I-i+1}^J \mathbb{E}[\exp(-\widehat{\xi}_{i,j})]. \end{aligned}$$

Si se considera, $\mathbb{E}[\exp(\eta_{i,j})] = \exp \left\{ \xi_j + \frac{\sigma_j^2}{2} \right\}$ y $\mathbb{E}[\exp(-\widehat{\xi}_{i,j})] = \exp \left\{ -\xi_j + \frac{1}{2} \frac{\sigma_j^2}{I-j+1} \right\}$, entonces

$$\mathbb{E} \left[\frac{C_{i,J}}{\widehat{C}_{i,J}^{pc}} \middle| C_{i,I-i} \right] = \exp \left\{ -\frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \left(1 + \frac{1}{I-j+1} \right) - \xi_j - \sigma_j^2 + \xi_j - \frac{\sigma_j^2}{I-j+1} \right\} = 1.$$

- (b) Dado que $\widehat{C}_{i,J}^{pc}$ no es un estimador insesgado, el ECMP se estima como

$$\begin{aligned} ECMP_{C_{i,J}|C_{i,I-i}}(\widehat{C}_{i,J}^{pc}) &= \mathbb{E} \left[\left(\widehat{C}_{i,J}^{pc} - C_{i,J} \right)^2 \middle| C_{i,I-i} \right] \\ &= C_{i,I-i}^2 \text{Var} \left(\exp \left\{ \sum_{j=I-i+1}^J \eta_{i,j} \right\} \right) + \mathbb{E} \left[\left(\mathbb{E}[C_{i,J}|C_{i,I-i}] - \widehat{C}_{i,J}^{pc} \right)^2 \middle| C_{i,I-i} \right], \end{aligned} \tag{2.4.15}$$

donde

$$\mathbb{E} \left[\widehat{C}_{i,J}^{pc} \middle| C_{i,I-i} \right] = \mathbb{E}[C_{i,J}|C_{i,I-i}] \exp \left\{ \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\}.$$

Así, el segundo término de (2.4.15) es igual a

$$\mathbb{E}[C_{i,J}|C_{i,I-i}]^2 \left(1 - \exp \left\{ \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} \right)^2 + \text{Var} \left(\widehat{C}_{i,J}^{pc} \middle| C_{i,I-i} \right).$$

Usando (2.4.4), entonces

$$\begin{aligned}
ECMP_{C_{i,J}|C_{i,I-i}}(\widehat{C}_{i,J}^{pc}) &= C_{i,I-i}^2 Var \left(\exp \left\{ \sum_{j=I-i+1}^J \eta_{i,j} \right\} \right) \\
&+ \mathbb{E}[C_{i,J}|C_{i,I-i}]^2 \left(1 - \exp \left\{ \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} \right)^2 \\
&+ \exp \left\{ 2 \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} Var \left(\widehat{C}_{i,J}^{pc} | C_{i,I-i} \right) \\
&> Var(C_{i,J}|C_{i,I-i}) + Var \left(\widehat{C}_{i,J}^{pc} | C_{i,I-i} \right) \\
&= ECMP_{C_{i,J}|C_{i,I-i}}(\widehat{C}_{i,J}^{LN}). \tag{2.4.16}
\end{aligned}$$

□

2.4.2.2. Varianza desconocida

En general, la varianza σ_j^2 se necesita estimar. Supóngase que $W \sim \chi_{(p)}^2$, i.e., $W \sim \Gamma\left(\frac{p}{2}, \frac{1}{2}\right)$ con función generadora de momentos dada por

$$\mathbb{E} \left[e^{(t\hat{\sigma}_j^2)} \right] = \left(1 - \frac{2t\sigma_j^2}{I-j} \right)^{-\frac{I-j}{2}}. \tag{2.4.17}$$

Si se elige t tal que

$$\left(1 - \frac{2t\sigma_j^2}{I-j} \right)^{-\frac{I-j}{2}} = \exp \left\{ \frac{1}{2}\sigma_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\}.$$

Es decir,

$$t = \frac{I-j}{2\sigma_j^2} \left(1 - \exp \left\{ \frac{\sigma_j^2}{I-j} \left(1 - \frac{1}{I-j+1} \right) \right\} \right).$$

Por lo tanto, se define

$$\begin{aligned}
t_j(\sigma_j^2) &= \frac{I-j}{\sigma_j^2} \left(1 - \exp \left\{ \frac{\sigma_j^2}{I-j} \left(1 - \frac{1}{I-j+1} \right) \right\} \right) \\
&= \frac{I-j}{\sigma_j^2} \left(1 - \exp \left\{ \frac{-\sigma_j^2}{I-j+1} \right\} \right). \tag{2.4.18}
\end{aligned}$$

Estimador 2.4.5. (*Estimador para σ_j^2 desconocida*)

Bajo las suposiciones (2.4.2), para cualesquiera $i \in \{1, \dots, I\}$

$$\widehat{C}_{i,J}^{LN\sigma,2} = \widehat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I] = \exp \left\{ \widehat{Z}_{i,J} + \frac{1}{2} \sum_{j=I-i+1}^J \widehat{\sigma}_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\}. \quad (2.4.19)$$

A partir de la independencia de $\widehat{\xi}_j$ y $\widehat{\sigma}_j^2$ y (2.4.9-2.4.10)

$$\begin{aligned} \mathbb{E} \left[\widehat{C}_{i,J}^{LN\sigma,2} \middle| C_{i,I-i} \right] &= C_{i,I-i} \exp \left\{ \sum_{j=I-i+1}^J \xi_j + \frac{1}{2} \sum_{j=I-i+1}^J \frac{\sigma_j^2}{I-j+1} \right\} \\ &\quad \times \mathbb{E} \left[\exp \left\{ \frac{1}{2} \sum_{j=I-i+1}^J \widehat{\sigma}_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\} \middle| C_{i,I-i} \right] \\ &= \mathbb{E}[C_{i,J}|C_{i,I-i}] \\ &\quad \times \mathbb{E} \left[\exp \left\{ \frac{1}{2} \sum_{j=I-i+1}^J (\widehat{\sigma}_j^2 - \sigma_j^2) \left(1 - \frac{1}{I-j+1} \right) \right\} \right]. \end{aligned}$$

Usando (2.4.17) y que $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n} \right)^n = e^x$, entonces

$$\begin{aligned} &\lim_{I \rightarrow \infty} \mathbb{E} \left[\exp \left\{ \frac{1}{2} \sum_{j=I-i+1}^J (\widehat{\sigma}_j^2 - \sigma_j^2) \left(1 - \frac{1}{I-j+1} \right) \right\} \right] \\ &= \lim_{I \rightarrow \infty} \left[\left(1 - \frac{\sigma_j^2}{I-j+1} \right)^{\frac{(I-j)}{2}} \exp \left\{ -\frac{\sigma_j^2(I-j)}{2(I-j+1)} \right\} \right] = \exp \left(\frac{\sigma_j^2}{2} \right) \exp \left(\frac{-\sigma_j^2}{2} \right) = 1. \end{aligned}$$

Por lo tanto, se obtiene la aproximación

$$\mathbb{E} \left[\widehat{C}_{i,J}^{LN\sigma,2} \middle| C_{i,I-i} \right] \approx \mathbb{E}[C_{i,J}|C_{i,I-i}]. \quad (2.4.20)$$

2.4.2.3. Intervalo de confianza modelo log-normal para reclamaciones acumuladas

Si los factores de desarrollo individuales, $F_{i,j} = \frac{C_{i,j}}{C_{i,j-1}}$ tienen una distribución log-normal, entonces, $\log(F_{i,j}) = \eta_{i,j} \sim N(\xi_j, \sigma_j^2)$. Se define el intervalo al $(1 - \alpha) \times 100$ de

confianza como

$$\mathbb{P}(-c < \eta_{i,j} < c) = 1 - \alpha,$$

donde α , es el nivel de significancia. Entonces, la variable aleatoria

$$Y = \frac{\eta_{i,j} - \xi_j}{\sqrt{\sigma_j^2}} \sim N(0, 1).$$

Por lo tanto,

$$\begin{aligned} -Z_{\frac{\alpha}{2}} &< \frac{\eta_{i,j} - \xi_j}{\sqrt{\sigma_j^2}} < Z_{\frac{\alpha}{2}} \\ -Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2} &< \eta_{i,j} - \xi_j < Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2} \\ \xi_j - Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2} &< \eta_{i,j} < \xi_j + Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2}. \end{aligned}$$

Y el intervalo al $(1 - \alpha) \times 100$ de confianza para $\eta_{i,j}$ es

$$\left(\xi_j - Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2}, \xi_j + Z_{\frac{\alpha}{2}}\sqrt{\sigma_j^2} \right).$$

En la sección anterior se obtuvieron los estimadores $\hat{\xi}_j, \hat{\sigma}_j^2$, para ξ_j, σ_j^2 respectivamente, y a partir de estos su distribución.

Ahora se dará un intervalo al $(1 - \alpha) \times 100$ de confianza para los parámetros de este modelo.

Si $\hat{\xi}_j \sim N\left(\xi_j, \frac{\sigma_j^2}{I-j+1}\right)$, se define a la variable aleatoria $X = \frac{\xi_j - \hat{\xi}_j}{\sqrt{\frac{\sigma_j^2}{I-j+1}}} \sim N(0, 1)$.

Entonces,

$$\begin{aligned} -Z_{\frac{\alpha}{2}} &< \frac{\xi_j - \hat{\xi}_j}{\sqrt{\frac{\sigma_j^2}{I-j+1}}} < Z_{\frac{\alpha}{2}} \\ -Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_j^2}{I-j+1}} &< \xi_j - \hat{\xi}_j < Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_j^2}{I-j+1}} \\ \hat{\xi}_j - Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_j^2}{I-j+1}} &< \xi_j < \hat{\xi}_j + Z_{\frac{\alpha}{2}}\sqrt{\frac{\sigma_j^2}{I-j+1}}. \end{aligned}$$

Por tanto, el intervalo al $(1 - \alpha) \times 100$ confianza para ξ_j es

$$\left(\widehat{\xi}_j - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_j^2}{I-j+1}}, \widehat{\xi}_j + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_j^2}{I-j+1}} \right).$$

Siguiendo el mismo procedimiento para σ_j^2 , a partir de $\frac{I-j}{\sigma_j^2} \widehat{\sigma_j^2} \sim \chi_{(I-j)}^2$, entonces

$$\begin{aligned} -\chi_{I-j}^2 &< \frac{(I-j)\widehat{\sigma_j^2}}{\sigma_j^2} < \chi_{I-j}^2 \\ -\frac{(I-j)\widehat{\sigma_j^2}}{\chi_{I-j}^2} &< \sigma_j^2 < \frac{(I-j)\widehat{\sigma_j^2}}{\chi_{I-j}^2}. \end{aligned}$$

Por tanto, el intervalo al $(1 - \alpha) \times 100$ confianza para σ_j^2 es

$$\left(-\frac{(I-j)\widehat{\sigma_j^2}}{\chi_{I-j}^2}, \frac{(I-j)\widehat{\sigma_j^2}}{\chi_{I-j}^2} \right).$$

2.4.3. Reclamaciones incrementales

Ahora, se considerarán modelos para las reclamaciones incrementales $X_{i,j}$ (pagos incrementales, cambio en las reclamaciones incurridas ó número de reclamaciones en el periodo j para el año de accidente i).

2.4.3.1. Modelo Poisson de sobredispersión

Se supondrá que existen parámetros μ_0, \dots, μ_I y $\gamma_0, \dots, \gamma_J$ tales que $X_{i,j}$ son variables independientes con distribución Poisson tales que, para $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$ y $\sum_{j=0}^J \gamma_j = 1$

$$x_{i,j} = \mathbb{E}[X_{i,j}] = \mu_i \gamma_j. \quad (2.4.21)$$

Con esta suposición también

$$x_{i,j} = \text{Var}(X_{i,j}) = \mu_i \gamma_j. \quad (2.4.22)$$

Observación 2.4.7.

En este caso, la varianza tiene la misma escala que la esperanza, i.e., una función de varianza lineal. ∇

Los parámetros se estimarán por máxima verosimilitud y el estimador para la siniestralidad última será

$$\widehat{C}_{i,J}^{Pois} = \mathbb{E}[C_{i,j} | \mathcal{D}_I] = C_{i,I-i} + \sum_{j=I-i+1}^J \widehat{X}_{i,j}^{Pois},$$

con $\widehat{X}_{i,j}^{Pois} = \widehat{\mathbb{E}}[X_{i,j}] = \widehat{\mu}_i \widehat{\gamma}_j$.

El modelo Poisson se puede generalizar a partir de un parámetro de dispersión adicional ϕ ; con esto se obtiene el modelo Poisson con sobredispersión, i.e., para $\phi > 0$ se supondrá que

$$\mathbb{E}[X_{i,j}] = x_{i,j} \quad y \quad Var(X_{i,j}) = \phi x_{i,j}.$$

2.4.3.2. Intervalo de confianza del modelo Poisson para las reclamaciones incrementales, $X_{i,j}$

Bajo las suposición distribucional Poisson con media $\mu_i \gamma_j$ para las reclamaciones incrementales, $X_{i,j}$, se define a la variable aleatoria $Y = \frac{X_{i,j} - \mu_i \gamma_j}{\sqrt{\mu_i \gamma_j}} \sim N(0, 1)$. Entonces

$$\begin{aligned} -Z_{\frac{\alpha}{2}} &< \frac{X_{i,j} - \mu_i \gamma_j}{\sqrt{\mu_i \gamma_j}} < Z_{\frac{\alpha}{2}} \\ -Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j} &< X_{i,j} - \mu_i \gamma_j < Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j} \\ \mu_i \gamma_j - Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j} &< X_{i,j} < \mu_i \gamma_j + Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j}. \end{aligned}$$

Por tanto, el intervalo al $(1 - \alpha) \times 100$ confianza para $X_{i,j}$ es

$$\left(\mu_i \gamma_j - Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j}, \mu_i \gamma_j + Z_{\frac{\alpha}{2}} \sqrt{\mu_i \gamma_j} \right).$$

2.4.3.3. Modelo binomial-negativo

El modelo binomial-negativo se puede expresar ya sea para reclamaciones incrementales como para las acumuladas. La distribución binomial-negativa es la que se obtiene a partir de la distribución binomial suponiendo que el parámetro Poisson sigue una distribución Gamma.

Supóngase que $X_{i,j}$ satisface la Suposición 2.4.1 con $x_{i,j} = \mathbb{E}[X_{i,j}] = \mu_i \gamma_j$. Por tanto, para $j \in \{0, \dots, J\}$

$$\mathbb{E}[C_{i,j}] = \mu_i \beta_j \quad \text{con} \quad \beta_j = \sum_{k=0}^j \gamma_k.$$

Entonces, si $f_{j-1} = \frac{\beta_j}{\beta_{j-1}}$ (i.e., f_j son los factores CL para valores esperados)

$$\mathbb{E}[X_{i,j}] = \frac{\mathbb{E}[C_{i,j-1}]}{\beta_{j-1}} \gamma_j = \mathbb{E}[C_{i,j-1}] (f_{j-1} - 1),$$

y

$$\mathbb{E}[C_{i,j}] = \mathbb{E}[C_{i,j-1}] \frac{\beta_j}{\beta_{j-1}} = \mathbb{E}[C_{i,j-1}] f_{j-1}.$$

Observación 2.4.8.

Nótese que se aplicará un algoritmo CL a las reclamaciones esperadas. ∇

Entonces,

$$X_{i,j} \sim \text{Poisson} \left(\mathbb{E}[C_{i,j-1}] \left(\frac{\beta_j}{\beta_{j-1}} - 1 \right) \right).$$

El objetivo es entender este modelo estocástico. Primero, se utilizará una filosofía Bayesiana para reemplazar a la media real $\mathbb{E}[C_{i,j-1}]$ por una variable aleatoria

$$\theta_{i,j-1} \sim \text{Gamma}(c_{i,j-1}, 1). \quad (2.4.23)$$

Entonces, se $X_{i,j}$ tiene una distribución binomial-negativa

$$\begin{aligned} x_{i,j} &= \mathbb{E}[X_{i,j}] \\ &= \mathbb{E}[\mathbb{E}(X_{i,j}|\theta_{i,j-1})] \\ &= \mathbb{E}[\theta_{i,j-1}](f_{j-1} - 1) \\ &= c_{i,j-1}(f_{j-1} - 1), \end{aligned}$$

y

$$\begin{aligned} \text{Var}(X_{i,j}) &= \mathbb{E}[\text{Var}(X_{i,j}|\theta_{i,j-1})] + \text{Var}(\mathbb{E}[X_{i,j}|\theta_{i,j-1}]) \\ &= \mathbb{E}[\theta_{i,j-1}](f_{j-1} - 1) + \text{Var}(\theta_{i,j-1})(f_{j-1} - 1)^2 \\ &= c_{i,j-1}(f_{j-1} - 1)f_{j-1} \\ &= x_{i,j}f_{j-1}. \end{aligned} \quad (2.4.24)$$

Por lo tanto, para $c_{i,j-1}$ fijo se puede interpretar a este modelo como un Poisson sobredisperso con parámetro de dispersión f_{j-1} .

Ahora, se reemplazará $c_{i,j-1}$ por una serie de tiempo, suponiendo que la distribución de $X_{i,j}$ depende de $C_{i,j-1} = c_{i,j-1}$. Entonces

Suposición 2.4.3. (*Modelo Binomial-Negativo, Verrall (2000)*)

- (I) Las reclamaciones de diferente año de accidente i son independientes.
- (II) Existen factores $f_0, \dots, f_{J-1} \geq 1$ tales que para $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$, los incrementos $X_{i,j}$ condicionados a $C_{0,i}, \dots, C_{i,j-1}$ tienen una distribución binomial-negativa con

$$\begin{aligned}\mathbb{E}[X_{i,j}|C_{0,i}, \dots, C_{i,j-1}] &= C_{i,j-1}(f_{j-1} - 1), \\ \text{Var}(X_{i,j}|C_{0,i}, \dots, C_{i,j-1}) &= C_{i,j-1}(f_{j-1} - 1)f_{j-1}.\end{aligned}$$

El modelo binomial-negativo satisface la Suposición 2.3.1

$$\begin{aligned}\mathbb{E}[C_{i,j}|C_{0,i}, \dots, C_{i,j-1}] &= \mathbb{E}[X_{i,j} + C_{i,j-1}|C_{0,i}, \dots, C_{i,j-1}] \\ &= f_{j-1}C_{i,j-1},\end{aligned}$$

y la función varianza

$$\begin{aligned}\text{Var}(C_{i,j}|C_{0,i}, \dots, C_{i,j-1}) &= \text{Var}(X_{i,j} + C_{i,j-1}|C_{0,i}, \dots, C_{i,j-1}) \\ &= C_{i,j-1}(f_{j-1} - 1)f_{j-1}.\end{aligned}$$

Es decir, se puede aplicar la teoría CL con el parámetro de varianza $\sigma_{j-1}^2 = (f_{j-1} - 1)f_{j-1}$.

Se tiene que suponer $f_j > 1$, de otro modo, la función varianza no tendrá sentido. Sin embargo, en aplicaciones prácticas puede ocurrir que $f_j < 1$ (por ejemplo si los ajustadores sobre-estiman la reclamación promedio). También es posible que para los pagos acumulados, los factores CL son más pequeños que 1 (que se puede deber a subrogaciones y deducibles).

Se puede aproximar esta distribución discreta binomial-negativa mediante la distribución Gaussiana con media $C_{i,j-1}(f_{j-1} - 1)$ y varianza $\phi_{i,j-1}C_{i,j-1}$, de esta manera, se evita el problema de obtener valores negativos para $X_{i,j}$.

2.4.3.4. Modelo log-normal

Ahora, se supondrá que las reclamaciones incrementales $X_{i,j}$ tienen distribución log-normal y son independientes, i.e.,

Suposición 2.4.4. (*Modelo Log-normal para Reclamaciones Incrementales*)

Para cualesquiera $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$, existen parámetros $m_{i,j}$ y σ^2 tales que

$$\log(X_{i,j}) \sim N(m_{i,j}, \sigma^2).$$

En este caso, $x_{i,j} = \mathbb{E}[X_{i,j}] = \exp\left\{m_{i,j} + \frac{1}{2}\sigma^2\right\}$ y $\text{Var}(X_{i,j}) = x_{i,j}^2(e^{\sigma^2} - 1)$.

Se puede demostrar que la información transformada $\log(X_{i,j})$ tiene una estructura aditiva. Además, se supondrá una estructura adicional para $m_{i,j}$ como $m_{i,j} = c + a_i + b_j$ con

$$\sum_{j=0}^J e^{b_j} = 1.$$

Se pueden estimar los parámetros por máxima verosimilitud. Además,

$$x_{i,j} = \exp\left\{c + \frac{1}{2}\sigma^2\right\} e^{a_i} e^{b_j} = \mu_i \gamma_j,$$

con

$$\mu_i = \exp\left\{c + \frac{1}{2}\sigma^2\right\} e^{a_i} \quad \text{y} \quad \gamma_j = e^{b_j}.$$

De esta forma, se tiene una función varianza y dispersión $(e^{\sigma^2} - 1)$.

2.4.3.5. Intervalo de confianza del modelo log-normal para las reclamaciones incrementales, $X_{i,j}$

Si las reclamaciones incrementales $X_{i,j}$ tienen distribución log-normal y son independientes, i.e., $\log(X_{i,j}) \sim \mathcal{N}(m_{i,j}, \sigma^2)$. Entonces, la variable aleatoria $W = \frac{\log(X_{i,j}) - m_{i,j}}{\sqrt{\sigma^2}} \sim N(0, 1)$.

Así,

$$\begin{aligned} -Z_{\frac{\alpha}{2}} &< \frac{\log(X_{i,j}) - m_{i,j}}{\sqrt{\sigma^2}} < Z_{\frac{\alpha}{2}} \\ -Z_{\frac{\alpha}{2}} \sqrt{\sigma^2} &< \log(X_{i,j}) - m_{i,j} < Z_{\frac{\alpha}{2}} \sqrt{\sigma^2} \\ m_{i,j} - Z_{\frac{\alpha}{2}} \sqrt{\sigma^2} &< \log(X_{i,j}) < m_{i,j} + Z_{\frac{\alpha}{2}} \sqrt{\sigma^2}. \end{aligned}$$

Por tanto, el intervalo al $(1 - \alpha) \times 100$ confianza para $\log(X_{i,j})$ es

$$\left(m_{i,j} - Z_{\frac{\alpha}{2}} \sqrt{\sigma^2}, m_{i,j} + Z_{\frac{\alpha}{2}} \sqrt{\sigma^2}\right).$$

2.4.3.6. Modelo Gamma

En la expresión $Var(X_{i,j}) = x_{i,j}^2(e^{\sigma^2} - 1)$ del modelo log-normal, se obtuvo que la función varianza corresponde al modelo Gamma. Mack (1991) definió directamente al modelo Gamma en su obtención para una función de varianza del tipo cuadrático.

Suposición 2.4.5. (*Modelo Gamma*)

Existen $r_{i,j} \in \mathbb{R}$ y variables aleatoria independientes $X_{i,j}^{(k)}$ con distribución Gamma de media $m_{i,j}$ y coeficiente variacional \sqrt{v} , tal que para cualesquiera $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$

$$X_{i,j} = \sum_{k=0}^{r_{i,j}} X_{i,j}^{(k)}.$$

Si $X_{i,j}^{(k)}$ tiene distribución Gamma con parámetros $a_{i,j}$ y $b_{i,j}$ ($X_{i,j}^{(k)} \sim Gamma(a_{i,j}, b_{i,j})$) entonces su media es

$$m_{i,j} = \mathbb{E}[X_{i,j}^{(k)}] = \frac{a_{i,j}}{b_{i,j}},$$

y su varianza está dada por

$$Var(X_{i,j}^{(k)}) = \frac{a_{i,j}}{b_{i,j}^2}.$$

Por tanto, $\sqrt{v} = Vco(X_{i,j}^{(k)}) = \sqrt{a_{i,j}}$, i.e., $X_{i,j}^{(k)} \sim Gamma\left(v, \frac{v}{m_{i,j}}\right)$.

Entonces, en virtud de la Suposición 2.4.5

$$X_{i,j}^{(k)} \sim Gamma\left(v r_{i,j}, \frac{v}{m_{i,j}}\right).$$

Con $x_{i,j} = \mathbb{E}[X_{i,j}] = r_{i,j} m_{i,j}$ y $Var(X_{i,j}) = \frac{r_{i,j}}{v} m_{i,j}^2 = \frac{x_{i,j}^2}{r_{i,j} v}$ (que es una función cuadrática).

Suposición 2.4.6. (*Modelo Gamma, Estructura Multiplicativa*)

Además de la Suposición 2.4.5, supóngase que existen parámetros $\mu_0, \dots, \mu_I > 0$ y

$\gamma_0, \dots, \gamma_J$ tales que $\sum_{j=0}^J \gamma_j = 1$ y

$$x_{i,j} = \mathbb{E}[X_{i,j}] = r_{i,j} \cdot m_{i,j} = \mu_i \cdot \gamma_j. \quad (2.4.25)$$

La función de verosimilitud está dado por

$$L_{\mathcal{D}_I}(\mu_0, \dots, \mu_I, \gamma_0, \dots, \gamma_J) = \prod_{i+j \leq I} \frac{\left(\frac{vr_{i,j}}{\mu_i \gamma_j}\right)^{vr_{i,j}}}{\Gamma(vr_{i,j})} X_{i,j}^{vr_{i,j}-1} \exp\left\{-\frac{vr_{i,j}}{\mu_i \gamma_j} X_{i,j}\right\}. \quad (2.4.26)$$

Se maximizará esta función con $\mathcal{D}_I = \{X_{i,j} : i + j \leq I, j \leq J\}$ haciendo I+J+2 derivadas parciales con respecto a μ_i y γ_j , e igualando a cero se obtienen los estimados para cualesquiera $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$

$$\hat{\mu}_i = \frac{\sum_{j=0}^{I-i} r_{ij} \frac{X_{i,j}}{\hat{\gamma}_j}}{\sum_{j=0}^{I-i} r_{ij}} \quad \text{y} \quad \hat{\gamma}_j = \frac{\sum_{i=0}^{I-j} r_{ij} \frac{X_{i,j}}{\hat{\mu}_i}}{\sum_{i=0}^{I-j} r_{ij}}, \quad (2.4.27)$$

bajo la restricción de que $\sum \hat{\gamma}_j = 1$.

Nótese que estos estimados son promedios ponderados de las observaciones, donde los pesos son $r_{i,j}$.

Estimador 2.4.6. (*Estimador Máximo Verosímil Gamma*)

El estimador máximo verosímil en el modelo Gamma (2.4.6) está dado por

$$\begin{aligned} \hat{X}_{i,j}^{Ga} &= \hat{\mathbb{E}}[X_{i,j}] = \hat{\mu}_i \hat{\gamma}_j, \\ \hat{C}_{i,J}^{Ga} &= \hat{\mathbb{E}}[C_{i,J} | \mathcal{D}_I] = C_{i,I-i} + \sum_{j=I-i+1}^J \hat{X}_{i,j}^{Ga}, \end{aligned}$$

para $i + j > I$.

Para estimar los parámetros μ_i y γ_j de este modelo se necesita información adicional $r_{i,j}, i + j \leq I$. Si $X_{i,j}^{(k)}$ denota a los pagos de reclamaciones individuales, entonces $r_{i,j}$ corresponde al número de pagos de reclamaciones individuales. Además, se puede tratar de estimar el pago promedio $m_{i,j}$ para $i + j > I$ en la estructura multiplicativa estimando el número futuro de pagos, $r_{i,j}$ para $i + j > I$.

Capítulo 3

Modelos Chain-Ladder Bayesianos

3.1. Modelos Bayesianos

El paradigma Bayesiano para métodos de reservas incorpora información previa y junto con las observaciones se obtiene una estimación para la siniestralidad última.

3.1.1. El método de Benktander-Hovinen y el modelo Cape-Cod

Para comenzar, se revisaran dos métodos para el cálculo de la reserva que no se consideran modelos Bayesianos pero están encaminados a ideas Bayesianas subyacentes.

3.1.1.1. Método de Benktander-Hovinen

Este método que permite calcular/estimar el monto total de las pérdidas, fue desarrollado de manera independiente por Benktander (1976) y Hovinen (1981).

Donde, para un año de accidente fijo $i \geq 1$, supóngase que se tiene un estimado a *priori* μ_i , para $\mathbb{E}[C_{i,j}]$; y un patrón de desarrollo de las reclamaciones $\{\beta_j\}_{j=0}^J$ por tanto $\mathbb{E}[C_{i,j}] = \mu_i \beta_j$ es conocido.

Como el método BF “ignora” la observaciones $C_{i,I-i}$ de la última diagonal observada y el método CL ignora el estimado a *priori* μ_i , entonces se puede considerar una combinación de los dos métodos. Para $\kappa \in [0, 1]$ se definirá la siguiente mezcla de credibilidad

$$u_i(\kappa) = \kappa \widehat{C}_{i,J}^{CL} + (1 - \kappa) \mu_i, \quad i \in \{0, \dots, I\}. \quad (3.1.1)$$

donde $\widehat{C}_{i,J}^{CL}$ es el estimador CL para la siniestralidad última y μ_i el estimador a *priori*.

El parámetro κ debe ser mayor a $C_{i,j}$ conforme el periodo j se incrementa, pues se está obteniendo más información acerca de $C_{i,j}$ conforme j se incrementa. Benktander (1976)

propuso que $\kappa = \beta_{I-i}$, con el que se obtiene el siguiente estimador.

Estimador 3.1.1. (*Estimador Benktander-Hovinen (BH)*)

El estimador de Benktander-Hovinen se define como

$$\widehat{C}_{i,J}^{BH} = C_{i,I-i} + (1 - \beta_{I-i}) \left[\beta_{I-i} \cdot \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \right]. \quad (3.1.2)$$

Se puede identificar al patrón de desarrollo de reclamaciones $\{\beta_j\}_{j=0}^J$ con los factores CL, $\{f_j\}_{j=0}^{J-1}$. Esto se logra si se utiliza la Suposición 2.2.3 del método BF, i.e.,

$$\beta_j = \prod_{k=j}^{J-1} f_k^{-1}. \quad (3.1.3)$$

Ya que el patrón de desarrollo β_j es conocido, también son conocidos los factores CL (a partir de la ecuación (3.1.3)), entonces para $j \in \{0, \dots, J-1\}$,

$$\widehat{f}_j = f_j.$$

Por lo tanto, el estimador BH en la ecuación (3.1.2) se puede reescribir como

$$\begin{aligned} \widehat{C}_{i,J}^{BH} &= C_{i,I-i} + \beta_{I-i} \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \\ &\quad - \beta_{I-i} \left[\beta_{I-i} \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \right] \\ &= (C_{i,I-i} + (1 - \beta_{I-i}) \mu_i) + \beta_{I-i} \widehat{C}_{i,J}^{CL} \\ &\quad - \beta_{I-i} [C_{i,I-i} + (1 - \beta_{I-i}) \mu_i] \\ &= \widehat{C}^{BF} + \beta_{I-i} \widehat{C}_{i,J}^{CL} - \beta_{I-i} \widehat{C}^{BF} \\ &= \beta_{I-i} \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \widehat{C}_{i,J}^{BF} \\ &= C_{i,I-i} + (1 - \beta_{I-i}) \widehat{C}_{i,J}^{BF}, \text{ pues } \widehat{C}_{i,J}^{CL} = \frac{C_{i,I-i}}{\beta_{I-i}}. \end{aligned} \quad (3.1.4)$$

Observación 3.1.1.

La ecuación (3.1.4) muestra que el estimador BH se puede ver como un estimador BF “iterado” utilizando al estimador BF como un nuevo estimador a *priori* en el sentido de que $\widehat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \beta_{I-i}) \mu_i$ i.e. se usó $\widehat{C}_{i,J}^{BF}$ en vez de μ_i . ∇

Proposición 3.1.1.

Si se conoce el patrón de desarrollo de reclamaciones $\{\beta_j\}_{j=0}^J$, además de la identificación de (3.1.3), entonces

$$\widehat{C}_{i,J}^{BH} = u_i(1 - (1 - \beta_{I-i})^2), \quad i \in \{1, \dots, I\}$$

donde, la función $u_i(\kappa) = \kappa \widehat{C}_{i,J}^{CL} + (1 - \kappa)\mu_i$.

Demostración:

$$\begin{aligned} \widehat{C}_{i,J}^{BH} &= u_i(1 - (1 - \beta_{I-i})^2) \\ &= [1 - (1 - \beta_{I-i})^2] \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^2 \mu_i \\ &= \beta_{I-i} \widehat{C}_{i,J}^{CL} + (\beta_{I-i} - \beta_{I-i}^2) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^2 \mu_i \\ &= C_{i,I-i} + (\beta_{I-i} - \beta_{I-i}^2) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^2 \mu_i \\ &= C_{i,I-i} + \beta_{I-i} \cdot \widehat{C}_{i,J}^{CL} - \beta_{I-i}^2 \cdot \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^2 \mu_i \\ &= C_{i,I-i} + (1 - \beta_{I-i}) \left[\beta_{I-i} \cdot \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \right]. \end{aligned}$$

□

Observación 3.1.2.

La Proposición 3.1.1 muestra que β_{I-i} no es un punto fijo de la iteración ya que se tiene que evaluar el estimador BH en $1 - (1 - \beta_{I-i})^2$. ▽

Teorema 3.1.1.

Si se conoce el patrón de desarrollo de reclamaciones, $\{\beta_j\}_{j=0}^J$, $\beta_{I-i} > 0$ y $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$, entonces

$$\lim_{m \rightarrow \infty} \widehat{C}^{(m)} = \widehat{C}_{i,J}^{CL},$$

donde, $\widehat{C}^{(0)} = \mu_i$ y

$$\widehat{C}^{(m+1)} = C_{i,I-i} + (1 - \beta_{I-i}) \widehat{C}^{(m)}, \quad m \in \mathbb{N}.$$

Demostración:

Nótese que por la proposición anterior para $m \in \mathbb{N}_+$

$$\widehat{C}^{(m)} = (1 - (1 - \beta_{I-i})^m) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^m \mu_i. \quad (3.1.5)$$

Inductivamente es cierto para $m = 1, 2$,

$$\begin{aligned} \widehat{C}^{(1)} &= [1 - (1 - \beta_{I-i})] \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \\ &= \beta_{I-i} \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i}) \mu_i \\ &= C_{i,I-i} + (1 - \beta_{I-i}) \mu_i = \widehat{C}_{i,J}^{BF}. \end{aligned}$$

Entonces, para la estimación $m + 1$

$$\begin{aligned} \widehat{C}^{(m+1)} &= C_{i,I-i} + (1 - \beta_{I-i}) \widehat{C}^{(m)} \\ &= C_{i,I-i} + (1 - \beta_{I-i}) \left[(1 - (1 - \beta_{I-i})^m) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^m \mu_i \right] \\ &= \beta_{I-i} \widehat{C}_{i,J}^{CL} + ((1 - \beta_{I-i}) - (1 - \beta_{I-i})^{m+1}) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^{m+1} \mu_i \\ &= (1 - (1 - \beta_{I-i})^{m+1}) \widehat{C}_{i,J}^{CL} + (1 - \beta_{I-i})^{m+1} \mu_i. \end{aligned}$$

□

Finalmente, como $\lim_{m \rightarrow \infty} (1 - \beta_{I-i})^m = 0$, entonces

$$\lim_{m \rightarrow \infty} \widehat{C}^{(m)} = (1 - 0) \widehat{C}_{i,J}^{CL} + 0 \cdot \mu_i = \widehat{C}_{i,J}^{CL}.$$

3.1.1.2. Modelo Cape-Cod

Una de las principales deficiencias del modelo CL es que la siniestralidad última depende de la última observación de la diagonal. Si esta última observación es un *outlier*, este *outlier* se proyecta a la siniestralidad última (a través de los factores *age-to-age*, \hat{f}_j). Además, para ramos de cola larga (como responsabilidad civil), las primeras observaciones no son representativas. Una posibilidad es “suavizar” los *outliers* en la última diagonal combinando los métodos BF y CL (como lo hace el método BH). Otra posibilidad es hacer más robusta dicha diagonal, esto se hace a partir del método Cape-Cod (Bühlmann, 1983).

Suposición 3.1.1. (*Método Cape-Cod*)

- (I) Las reclamaciones acumuladas $C_{i,j}$ de diferente año de accidente son independientes.
- (II) Existen parámetros $\pi_0, \dots, \pi_I, \kappa > 0$ y un patrón de desarrollo $\{\beta_j\}_{j=0}^J$ con $\beta_J = 1$ tales que para todo $i \in \{0, \dots, I\}$

$$\mathbb{E}[C_{i,j}] = \kappa \pi_i \beta_j.$$

Las suposiciones del modelo Cape-Cod son parecidas a las de la Suposición 2.2.3 haciendo $\mu_i = \kappa \pi_i$. Entonces, bajo las suposiciones del modelo Cape-Cop, los parámetros π_i se pueden interpretar como la prima recibida en el año de accidente i y κ es el *loss ratio*¹ promedio.

Se supondrá que κ es independiente del año de accidente i y π_i con respecto a κ es el mismo para todos los años de accidente. Si $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$ se puede estimar κ para cada año de accidente utilizando el estimador CL de siniestralidad última como

$$\widehat{\kappa}_i := \frac{\widehat{C}_{i,J}^{CL}}{\pi_i} = \frac{C_{i,I-i}}{\left(\prod_{j=I-i}^{J-1} f_j^{-1} \right) \pi_i} = \frac{C_{i,I-i}}{\beta_{I-i} \pi_i}.$$

Se puede demostrar que $\widehat{\kappa}_i$ es un estimador insesgado de κ pues

$$\begin{aligned} \mathbb{E}[\widehat{\kappa}_i] &= \mathbb{E}\left[\frac{\widehat{C}_{i,J}^{CL}}{\pi_i}\right] = \frac{1}{\pi_i} \mathbb{E}[\widehat{C}_{i,J}^{CL}] \\ &= \frac{1}{\pi_i} \cdot \frac{1}{\beta_{I-i}} \mathbb{E}[C_{i,I-i}] \\ &= \frac{1}{\pi_i} \cdot \frac{1}{\beta_{I-i}} \kappa \pi_i \beta_{I-i} = \kappa, \end{aligned}$$

y además

$$Var(\widehat{\kappa}_i) = \frac{1}{\pi_i^2 \beta_{I-i}^2} Var(C_{i,I-i}).$$

¹Proporción de las pérdidas incurridas respecto a las prima ganada

Entonces, un *loss ratio* más robusto se puede estimar como un promedio ponderado a partir de

$$\begin{aligned}\widehat{\kappa}^{CC} &:= \frac{\sum_{i=0}^I \beta_{I-i} \pi_i \widehat{\kappa}_i}{\sum_{k=0}^I \beta_{I-k} \pi_k} = \frac{\sum_{i=0}^I \beta_{I-i} \pi_i \frac{\widehat{C}_{i,J}^{CL}}{\pi_i}}{\sum_{k=0}^I \beta_{I-k} \pi_k} \\ &= \frac{\sum_{i=0}^I \beta_{I-i} \cdot \widehat{C}_{i,J}^{CL}}{\sum_{k=0}^I \beta_{I-k} \pi_k} = \frac{\sum_{i=0}^I C_{i,I-i}}{\sum_{i=0}^I \beta_{I-i} \pi_i}.\end{aligned}$$

A continuación se demostrará que este estimador es insesgado para κ .

Proposición 3.1.2.

$\mathbb{E}[\widehat{\kappa}^{CC}] = \kappa$, i.e., $\widehat{\kappa}^{CC}$ es un estimador insesgado de κ .

Demostración:

$$\mathbb{E}[\widehat{\kappa}^{CC}] = \frac{1}{\sum_{i=0}^I \beta_{I-i} \pi_i} \sum_{i=0}^I \mathbb{E}[C_{i,I-i}] = \frac{1}{\sum_{i=0}^I \beta_{I-i} \pi_i} \sum_{i=0}^I \kappa \cdot \pi_i \beta_{I-i} = \kappa.$$

□

Entonces, un valor “robusto” para $C_{i,I-i}$ está dado por

$$\widehat{C}_{i,I-i}^{CC} = \widehat{\kappa}^{CC} \pi_i \beta_{I-i}.$$

Entonces el estimador para la siniestralidad última

Estimador 3.1.2. (*Estimador Cape-Cod*)

Se define al estimador *Cape-Cod* como

$$\widehat{C}_{i,J}^{CC} = C_{i,I-i} - \widehat{C}_{i,I-i}^{CC} + \prod_{j=I-i}^{J-1} f_j \widehat{C}_{i,I-i}^{CC}, \quad i \in \{0, \dots, I\}. \quad (3.1.6)$$

Proposición 3.1.3.

Si se cumple $\hat{\kappa}^{CC} = \frac{\sum_{i=0}^I C_{i,I-i}}{\sum_{i=0}^I \beta_{I-i} \pi_i}$ y $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$, entonces el estimador $\hat{C}_{i,J}^{CC} - C_{i,I-i}$ es insesgado para $\mathbb{E}[C_{i,J} - C_{i,I-i}] = \kappa \pi_i (1 - \beta_{I-i})$.

Demostración:

Nótese que

$$\mathbb{E} \left[\hat{C}_{i,I-i}^{CC} \right] = \mathbb{E} \left[\hat{\kappa}^{CC} \right] \pi_i \beta_{I-i} = \kappa \pi_i \beta_{I-i} = \mathbb{E}[C_{i,I-i}].$$

Además, con $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$,

$$\begin{aligned} \hat{C}_{i,J}^{CC} - C_{i,I-i} &= \hat{C}_{i,I-i}^{CC} \left(\prod_{j=I-i}^{J-1} f_j - 1 \right) \\ &= \hat{\kappa}^{CC} \pi_i \prod_{k=I-i}^{J-1} f_k^{-1} \left(\prod_{j=I-i}^{J-1} f_j - 1 \right) \\ &= \hat{\kappa}^{CC} \pi_i (1 - \beta_{I-i}). \end{aligned}$$

En el método Cape-Cod, el modelo CL se aplica a la diagonal “robustecida”, $\hat{C}_{i,I-i}^{CC}$, pero para estimar la siniestralidad última, se necesita agregar la diferencia entre la observación original de la diagonal, $C_{i,I-i}$, y el valor robustecido. Si se reescribe el estimador Cape-Cod se obtiene

$$\hat{C}_{i,J}^{CC} = C_{i,I-i} + (1 - \beta_{I-i}) \hat{\kappa}^{CC} \pi_i,$$

que es un estimador del tipo BF con un estimador *a priori* $\hat{\kappa}^{CC} \cdot \pi_i$.

3.1.2. Métodos basados en credibilidad

Ya se hicieron mezclas de estimadores, de aquí que el siguiente paso natural es encontrar mezclas óptimas de estimadores CL y BF, por supuesto con un criterio de optimalidad dado.

3.1.2.1. Minimización de funciones de pérdida cuadráticas

Definición 3.1.1. (*Reclamaciones pendientes*)

Sea $i \in \{1, \dots, I\}$. Se define las reclamaciones pendientes para el año de accidente i como

$$R_i := R_{i,I-i} := C_{i,J} - C_{i,I-i}.$$

Recordando que $C_{i,j}$ son las reclamaciones acumuladas, entonces R_i son las pérdidas pendientes en el año i .

Bajo las suposiciones de que los factores CL son conocidos y que $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$ (Suposición 2.2.3) la reserva CL y BF, respectivamente, están dadas por

$$\widehat{R}_i^{CL} = \widehat{C}_{i,J}^{CL} - C_{i,I-i} = C_{i,I-i} \left(\prod_{j=I-i}^{J-1} f_j - 1 \right)$$

y

$$\widehat{R}_i^{BF} = \widehat{C}_{i,J}^{BF} - C_{i,I-i} = (1 - \beta_{I-i})\mu_i.$$

Si se mezclan los métodos CL y BF entonces se obtiene la mezcla de credibilidad

$$t \cdot \widehat{C}_{i,J}^{CL} + (1-t)\widehat{C}_{i,J}^{BF}, \quad t \in [0, 1],$$

y se tiene las siguientes expresiones para las reservas

$$\begin{aligned}
\widehat{R}_i(t) &= t\widehat{R}_i^{CL} + (1-t)\widehat{R}_i^{BF} \\
&= t\left(\widehat{C}_{i,J}^{CL} - C_{i,I-i}\right) + (1-t)\left(\widehat{C}_{i,J}^{BF} - C_{i,I-i}\right) \\
&= t\left(\widehat{C}_{i,J}^{CL} - \widehat{C}_{i,J}^{BF}\right) - t \cdot C_{i,I-i} + \widehat{C}_{i,J}^{BF} - C_{i,I-i} + t \cdot C_{i,I-i} \\
&= \widehat{C}_{i,J}^{BF} - C_{i,I-i} + t\left(\widehat{C}_{i,J}^{CL} - \widehat{C}_{i,J}^{BF}\right) \\
&= t\widehat{C}_{i,J}^{CL} + \widehat{C}_{i,J}^{BF}(1-t) - C_{i,I-i} \\
&= t\widehat{C}_{i,J}^{CL} + [C_{i,I-i} + (1-\beta_{I-i})\mu_i](1-t) - C_{i,I-i} \\
&= t\widehat{C}_{i,J}^{CL} + C_{i,I-i} - tC_{i,I-i} + (1-\beta_{I-i})\mu_i(1-t) - C_{i,I-i} \\
&= t\left(\widehat{C}_{i,J}^{CL} - C_{i,I-i}\right) + (1-\beta_{I-i})\mu_i(1-t) \\
&= t\left(\widehat{C}_{i,J}^{CL} - \widehat{C}_{i,J}^{CL}\beta_{I-i}\right) + (1-\beta_{I-i})\mu_i(1-t) \\
&= (1-\beta_{I-i})\left(t\widehat{C}_{i,J}^{CL} + (1-t)\mu_i\right) \\
&= (1-\beta_{I-i})u_i(t), \quad \text{donde } u_i(t) = t\widehat{C}_{i,J}^{CL} + (1-t)\mu_i.
\end{aligned}$$

Por lo tanto,

$$\widehat{R}_i(t) = (1-\beta_{I-i})u_i(t).$$

La idea de esta sección es encontrar t óptima y dicha optimalidad se definirá en términos de la minimización de la función de pérdida cuadrática.

Entonces el objetivo es minimizar el ECMP para los estimadores de la reserva, $\widehat{R}_i(t)$ (que estiman a la variable aleatoria R_i), pero para minimizar

$$ECMP_{R_i}\left(\widehat{R}_i(t)\right) = \mathbb{E}\left[\left(R_i - \widehat{R}_i(t)\right)^2\right],$$

primero se necesita suponer un modelo estocástico.

Suposición 3.1.2.

- (I) Las reclamaciones acumuladas $C_{i,j}$ para diferentes años de accidente i son independientes.
- (II) Existe una suposición $\{\beta_j\}_{j=0}^J$ tal que $\beta_J = 1$ y para todo $j \in \{0, \dots, J\}$.

$$\mathbb{E}[C_{i,j}] = \beta_J \mathbb{E}[C_{i,J}].$$

- (III) Existen variables aleatorias U_1, \dots, U_I tales que $C_{i,I-i}, C_{i,J} \perp U_i$ y $\mathbb{E}[U_i] = \mathbb{E}[C_{i,J}]$.

Observación 3.1.3.

La Suposición 3.1.2 es muy parecida a la Suposición 2.2.3 si se toma $U_i = \mu_i > 0$ determinista. Entonces, ya no se estará suponiendo un valor determinista *a priori* para la media, ahora se tendrán estimadores insesgados U_i para dicha media. Con esto, se rescata la noción de que también hay incertidumbre al seleccionar la verdadera media $\mathbb{E}[C_{i,J}]$.

El modelo CL satisface la Suposición 3.1.2 pero no necesariamente si se cumple la Suposición 2.2.3 se cumplirá el modelo CL.

Si se supone que $\beta_j = \prod_{k=j}^{J-1} f_k^{-1}$ entonces los estimadores CL y BF están dados por

$$\widehat{C}_{i,J}^{CL} = \frac{C_{i,I-i}}{\beta_{I-i}} \quad \text{y} \quad \widehat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \beta_{I-i})U_i. \quad (3.1.7)$$

Entonces, las reservas de credibilidad ponderadas están dadas por

$$\widehat{R}_i(t) = (1 - \beta_{I-i}) \left(t \widehat{C}_{i,J}^{CL} + (1 - t)U_i \right). \quad (3.1.8)$$

De aquí lo que se desea minimizar es

$$ECMP_{R_i} \left(\widehat{R}_i(t) \right) = \mathbb{E} \left[\left(R_i - \widehat{R}_i(t) \right)^2 \right]. \quad (3.1.9)$$

Si se supone que se cumple el modelo CL, no se puede comparar directamente la estimación del ECMP con la estimación del ECMP que se hizo anteriormente. Para la obtención de una expresión del ECMP para el método CL se supuso que los factores f_j son desconocidos.

Si se supone que estos factores son conocidos entonces

$$\begin{aligned}
 ECMP_{C_{i,J}}(\widehat{C}_{i,J}^{CL}) &= \mathbb{E} \left[\mathbb{E} \left[\left(C_{i,J} - \widehat{C}_{i,J}^{CL} \right)^2 \middle| \mathcal{D}_I \right] \right] \\
 &= \mathbb{E} \left[ECMP_{C_{i,J}|\mathcal{D}_I}(\widehat{C}_{i,J}^{CL}) \right] \\
 &= \mathbb{E} [Var(C_{i,J}|\mathcal{D}_I)] \\
 &= \mathbb{E}[Var(C_{i,J}|\mathcal{D}_I)] - Var(\mathbb{E}[C_{i,J}|\mathcal{D}_I]) \\
 &= Var(C_{i,J}) - Var(C_{i,I-i}) \prod_{j=I-i}^{J-1} f_j^2. \quad \nabla
 \end{aligned}$$

Proposición 3.1.4.

Si se estima $\widehat{R}_i(t) = (1 - \beta_{I-i})(t\widehat{C}_{i,J}^{CL} + (1-t)U_i)$ bajo la Suposición 3.1.2, $\widehat{C}_{i,J}^{CL} = \frac{C_{i,I-i}}{\beta_{I-i}}$, $\widehat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \beta_{I-i})U_i$, y además $\mathbb{E}[\widehat{R}_i(t)] = \mathbb{E}[R_i]$.

Entonces,

$$ECMP_{R_i}(\widehat{R}_i(t)) = Var(R_i) + Var(\widehat{R}_i(t)) - 2Cov(R_i, \widehat{R}_i(t)).$$

Demostración:

$$\begin{aligned}
 ECMP_{R_i}(\widehat{R}_i(t)) &= Var(R_i) + \mathbb{E} \left[\left(\mathbb{E}[R_i] - \widehat{R}_i(t) \right)^2 \right] + 2\mathbb{E} \left[(R_i - \mathbb{E}[R_i]) \left(\mathbb{E}[R_i] - \widehat{R}_i(t) \right) \right] \\
 &= Var(R_i) + Var(\widehat{R}_i(t)) - 2Cov(R_i, \widehat{R}_i(t)).
 \end{aligned}$$

□

Teorema 3.1.2.

Bajo la Suposición 3.1.2, si $\widehat{C}_{i,J}^{CL} = \frac{C_{i,I-i}}{\beta_{I-i}}$ y $\widehat{C}_{i,J}^{BF} = C_{i,I-i} + (1 - \beta_{I-i})U_i$, entonces el factor de credibilidad óptimo, t_i^* , que minimiza el $ECMP_{R_i}(\widehat{R}_i(t)) = \mathbb{E} \left[\left(R_i - \widehat{R}_i(t) \right)^2 \right]$ está dado por

$$t_i^* = \frac{\beta_{I-i}}{1 - \beta_{I-i}} \cdot \frac{Cov(C_{i,I-i}, R_i) + \beta_{I-i}(1 - \beta_{I-i})Var(U_i)}{Var(C_{i,I-i}) + \beta_{I-i}^2 Var(U_i)}.$$

Demostración:

Si se tiene que

$$\mathbb{E} \left[\left(\widehat{R}_i(t_i) - R_i \right)^2 \right] = t_i^2 \mathbb{E} \left[\left(\widehat{R}_i^{CL} - \widehat{R}_i^{BF} \right)^2 \right] + \mathbb{E} \left[\left(R_i - \widehat{R}_i^{BF} \right)^2 \right] - 2t_i \mathbb{E} \left[\left(\widehat{R}_i^{CL} - \widehat{R}_i^{BF} \right) \left(R_i - \widehat{R}_i^{BF} \right) \right].$$

Entonces, el t_i óptimo está dado por

$$\begin{aligned} t_i^* &= \frac{\mathbb{E} \left[\left(\widehat{R}_i^{CL} - \widehat{R}_i^{BF} \right) \left(R_i - \widehat{R}_i^{BF} \right) \right]}{\mathbb{E} \left[\left(\widehat{R}_i^{CL} - \widehat{R}_i^{BF} \right)^2 \right]} \\ &= \frac{\mathbb{E} \left[\left(\left(\frac{1}{\beta_{I-i}} - 1 \right) C_{i,I-i} - (1 - \beta_{I-i}) U_i \right) \left(R_i - (1 - \beta_{I-i}) U_i \right) \right]}{\mathbb{E} \left[\left(\left(\frac{1}{\beta_{I-i}} - 1 \right) C_{i,I-i} - (1 - \beta_{I-i}) U_i \right)^2 \right]} \\ &= \frac{\beta_{I-i}}{1 - \beta_{I-i}} \frac{\mathbb{E} \left[\left(C_{i,I-i} - \beta_{I-i} U_i \right) \left(R_i - (1 - \beta_{I-i}) U_i \right) \right]}{\mathbb{E} \left[\left(C_{i,I-i} - \beta_{I-i} U_i \right)^2 \right]}. \end{aligned}$$

Dado que $\mathbb{E} [\beta_{I-i} U_i] = \mathbb{E} [C_{i,I-i}]$ y $\mathbb{E} [U_i] = \mathbb{E} [C_{i,J}]$, entonces

$$\begin{aligned} t_i^* &= \frac{\beta_{I-i}}{1 - \beta_{I-i}} \frac{\text{Cov}(C_{i,I-i} - \beta_{I-i} U_i, R_i - (1 - \beta_{I-i}) U_i)}{\text{Var}(C_{i,I-i} - \beta_{I-i} U_i)} \\ &= \frac{\beta_{I-i}}{1 - \beta_{I-i}} \frac{\text{Cov}(C_{i,I-i}, R_i) + \beta_{I-i} (1 - \beta_{I-i}) \text{Var}(U_i)}{\text{Var}(C_{i,I-i}) + \beta_{I-i}^2 \text{Var}(U_i)}. \end{aligned}$$

□

Observación 3.1.4.

No se consideraron errores de estimación en el patrón de desarrollo de reclamaciones $\{\beta_j\}$ y f_j . En este sentido el Teorema 3.1.2 establece que los pesos de credibilidad óptima toman en cuenta el proceso de varianza y la incertidumbre en estimador, U_i . Entonces para calcular explícitamente t_i^* , es necesario especificar un modelo estocástico explícito.

3.1.2.2. Ejemplos distribucionales para la reserva de reclamaciones basados en credibilidad

Suposición 3.1.3. (*Mack 2000*)

- (i) Las reclamaciones acumuladas $C_{i,j}$ de diferentes años de accidente i son independientes.
- (ii) Existe una sucesión $\{\beta_j\}_{j=0}^J$ con $\beta_J = 1$ y una función $\alpha^2(\cdot)$ tales que para todo $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$
 - $\mathbb{E}[C_{i,j}|C_{i,J}] = \beta_j C_{i,J},$
 - $Var[C_{i,j}|C_{i,J}] = \beta_j(1 - \beta_j)\alpha^2(C_{i,J}).$

Observación 3.1.5.

- I) Este modelo se puede considerar como una aproximación Bayesiana de la siniestralidad última $C_{i,J}$.
- II) La Suposición 3.1.3 cumple la Suposición 2.2.3 con $\mu_i = \mathbb{E}[C_{i,J}]$ y $C_{i,j}$ satisface la Suposición 3.1.2. Esto no ocurre con el modelo CL.
- III) La varianza condicional tiende a cero cuando $\beta_j \rightarrow 1$, i.e., si las reclamaciones pendientes esperadas son bajas, las incertidumbre también.

▽

Lema 3.1.1.

$$Var(C_{i,I-i}) = \beta_{I-i}(1 - \beta_{I-i})\mathbb{E}[\alpha^2(C_{i,J})] + \beta_{I-i}^2 Var(C_{i,J}).$$

Demostración:

$$\begin{aligned} Var(C_{i,I-i}) &= \mathbb{E}[Var(C_{i,I-i}|C_{i,J})] + Var[\mathbb{E}(C_{i,I-i}|C_{i,J})] \\ &= \beta_{I-i}(1 - \beta_{I-i})\mathbb{E}[\alpha^2(C_{i,J})] + \beta_{I-i}^2 Var(C_{i,J}). \end{aligned}$$

□

Lema 3.1.2.

$$\text{Cov}(C_{i,I-i}, C_{i,J}) = \beta_{I-i} \text{Var}(C_{i,J}).$$

Demostración:

$$\begin{aligned} \text{Cov}(C_{i,I-i}, C_{i,J}) &= \mathbb{E}[\text{Cov}(C_{i,I-i}, C_{i,J} | C_{i,J})] + \text{Cov}[\mathbb{E}(C_{i,I-i} | C_{i,J}), \mathbb{E}(C_{i,J} | C_{i,J})] \\ &= 0 + \text{Cov}(\beta_{I-i} C_{i,J}, C_{i,J}) \\ &= \beta_{I-i} \text{Var}(C_{i,J}). \end{aligned}$$

□

Lema 3.1.3.

$$\text{Cov}(C_{i,I-i}, C_{i,J} - C_{i,I-i}) = \beta_{I-i} \text{Var}(C_{i,J}) - \text{Var}(C_{i,I-i}).$$

Demostración:

$$\begin{aligned} \text{Cov}(C_{i,I-i}, C_{i,J} - C_{i,I-i}) &= \text{Cov}(C_{i,I-i}, C_{i,J}) - \text{Var}(C_{i,I-i}) \\ &= \beta_{I-i} \text{Var}(C_{i,J}) - \text{Var}(C_{i,I-i}). \end{aligned}$$

□

Corolario 3.1.1. (*Factor de Credibilidad Óptimo*)

Si se satisface la Suposición 3.1.3, U_i independiente de $C_{i,I-i}$ y $C_{i,J}$, y $E[U_i] = E[C_{i,J}]$, entonces para $i \in \{0, \dots, I\}$ el factor de credibilidad óptimo que minimiza

$ECMP_{R_i}(\hat{R}_i(t)) = \mathbb{E} \left[\left(R_i - \hat{R}_i(t) \right)^2 \right]$ está dado por

$$t_i^* = \frac{\beta_{I-i}}{\beta_{I-i} + t_i} \quad \text{con} \quad t_i = \frac{\mathbb{E}[\alpha^2(C_{i,J})]}{\text{Var}(U_i) + \text{Var}(C_{i,J}) - \mathbb{E}[\alpha^2(C_{i,J})]},$$

$i \in \{1, \dots, I\}$.

Demostración:

Del Teorema 3.1.2 se tiene

$$t_i^* = \frac{\beta_{I-i}}{1 - \beta_{I-i}} \frac{\text{Cov}(C_{i,I-i}, C_{i,J} - C_{i,I-i}) + \beta_{I-i}(1 - \beta_{I-i})\text{Var}(U_i)}{\text{Var}(C_{i,I-i}) + \beta_{I-i}^2 \text{Var}(U_i)}.$$

Y de los Lemas 3.1.1, 3.1.2 y 3.1.3 se obtiene

$$\begin{aligned} t_i^* &= \frac{\beta_{I-i}}{1 - \beta_{I-i}} \frac{\beta_{I-i}\text{Var}(C_{i,J}) - \text{Var}(C_{i,I-i}) + \beta_{I-i}(1 - \beta_{I-i})\text{Var}(U_i)}{\text{Var}(C_{i,I-i}) + \beta_{I-i}^2 \text{Var}(U_i)} \\ &= \frac{\text{Var}(C_{i,J}) - \mathbb{E}[\alpha^2(C_{i,J})] + \text{Var}(U_i)}{(\beta_{I-i}^{-1} - 1)\mathbb{E}[\alpha^2(C_{i,J})] + \text{Var}(C_{i,J}) + \text{Var}(U_i)} \\ &= \frac{\text{Var}(C_{i,J}) - \mathbb{E}[\alpha^2(C_{i,J})] + \text{Var}(U_i)}{\beta_{I-i}^{-1}\mathbb{E}[\alpha^2(C_{i,J})] + \text{Var}(C_{i,J}) - \mathbb{E}[\alpha^2(C_{i,J})] + \text{Var}(U_i)}. \end{aligned}$$

□

Corolario 3.1.2. (*ECMPs*)

Si se satisfacen la Suposición 3.1.3, U_i es independiente de $C_{i,I-i}$ y de $C_{i,J}$ y $\mathbb{E}[U_i] = \mathbb{E}[C_{i,J}] =$ entonces para $i \in \{1, \dots, I\}$

- (i) $ECMP_{R_i}(\widehat{R}_i(t)) = \mathbb{E}[\alpha^2(C_{i,J})] \left(\frac{t^2}{\beta_{I-i}} + \frac{1}{1-\beta_{I-i}} + \frac{(1-t)^2}{t_i} \right) (1 - \beta_{I-i})^2.$
- (ii) $ECMP_{R_i}(\widehat{R}_i(0)) = \mathbb{E}[\alpha^2(C_{i,J})] \left(\frac{1}{1-\beta_{I-i}} + \frac{1}{t_i} \right) (1 - \beta_{I-i})^2.$
- (iii) $ECMP_{R_i}(\widehat{R}_i(1)) = \mathbb{E}[\alpha^2(C_{i,J})] \left(\frac{1}{\beta_{I-i}} + \frac{1}{1-\beta_{I-i}} \right) (1 - \beta_{I-i})^2.$
- (iv) $ECMP_{R_i}(\widehat{R}_i(t_i^*)) = \mathbb{E}[\alpha^2(C_{i,J})] \left(\frac{1}{\beta_{I-i} + t_i} + \frac{1}{1-\beta_{I-i}} \right) (1 - \beta_{I-i})^2.$

Observación 3.1.6.

La reserva $\widehat{R}_i(0)$ corresponde a la reserva \widehat{R}_i^{BF} y $\widehat{R}_i(1)$ corresponde a la reserva \widehat{R}_i^{CL} . Sin embargo, $ECMP_{R_i}(\widehat{R}_i(1))$ y $ECMP_{R_i}(\widehat{R}_i^{CL})$ no son comparables, ya que utilizan un modelo completamente diferente.

Nótese del Corolario 3.1.2 que, es mejor el estimador BF que el estimador CL si se cumple (3.1.3), en términos de ECMP si se cumple que

$$t_i > \beta_{I-i}. \quad (3.1.10)$$

Entonces, para años donde las pérdidas esperadas β_{I-i} son pequeñas, se toma el estimador BF, en otro caso se toma el estimador CL.

Ejemplo 3.1.1.

Supóngase que $\frac{C_{i,j}}{C_{i,J}} \Big| C_{i,J} \sim \text{Beta}(\alpha_i \beta_j, \alpha_i(1 - \beta_j))$. Entonces para $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$

$$\mathbb{E}[C_{i,j}|C_{i,J}] = C_{i,J} \mathbb{E} \left[\frac{C_{i,j}}{C_{i,J}} \Big| C_{i,J} \right] = \beta_j C_{i,J},$$

$$\text{Var}[C_{i,j}|C_{i,J}] = C_{i,J}^2 \text{Var} \left[\frac{C_{i,j}}{C_{i,J}} \Big| C_{i,J} \right] = \beta_j(1 - \beta_j) \frac{C_{i,J}^2}{1 + \alpha_i}.$$

▽

3.1.2.3. Modelo log-normal/log-normal

Ahora además de suponer (3.1.3), se harán algunas suposiciones distribucionales adicionales sobre $C_{i,J}$ y $C_{i,j}|C_{i,J}$.

En el Corolario 3.1.1 se pidió que $\mathbb{E}[U_i] = \mathbb{E}[C_{i,J}]$, donde U_i era un estimador a *priori*. Ahora, no se dará alguna hipótesis sobre la distribución de U_i ahora se harán suposiciones con respecto a $C_{i,J}$. Dichas suposiciones distribucionales deben ser tales que se pueda determinar la distribución exacta de $C_{i,j}|C_{i,J}$ a partir del Teorema de Bayes. Entonces, el mejor estimador para $\mathbb{E}[C_{i,j}|C_{i,I-i}]$ es una mezcla (de credibilidad) entre la observación $C_{i,I-i}$ y la media a *priori* $\mathbb{E}[C_{i,J}]$.

Suposición 3.1.4. (Modelo Log-normal/Log-normal)

- (I) Las reclamaciones acumuladas $C_{i,j}$ de diferentes años de accidente i son independientes.
- (II) $C_{i,J} \sim \text{log-normal}(\mu^{(i)}, \sigma_i^2)$, $i \in \{0, \dots, I\}$.
- (III) $C_{i,j}|C_{i,J}$ tiene distribución log-normal con parámetros $v_j = v_j(C_{i,J})$, $\tau_j^2 = \tau_j^2(C_{i,J})$, $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$.

Con la Suposición 3.1.4

$$\mu_i = \mathbb{E}[C_{i,J}] = \exp \left\{ \mu^{(i)} + \frac{1}{2} \sigma_i^2 \right\}. \quad (3.1.11)$$

Además, si $\{C_{i,j}\}_{j=0}^J$ también cumple la Suposición 3.1.3, entonces

$$\mathbb{E}[C_{i,j}|C_{i,J}] = \exp \left\{ v_j + \frac{1}{2} \tau_j^2 \right\} \stackrel{!}{=} \beta_j C_{i,J},$$

$$\text{Var}[C_{i,j}|C_{i,J}] = \exp\{2v_j + \tau_j^2\} (\exp\{\tau_j^2\} - 1) \stackrel{!}{=} \beta_j(1 - \beta_j)\alpha^2(C_{i,J}).$$

Entonces, se está escogiendo parámetros de tal forma que

$$\tau_j^2 = \tau_j^2(C_{i,J}) = \log\left(1 + \frac{1 - \beta_j}{\beta_j} \frac{\alpha^2(C_{i,J})}{C_{i,J}^2}\right), \quad (3.1.12)$$

$$v_j = v_j(C_{i,J}) = \log(\beta_j C_{i,J}) - \frac{1}{2} \log\left(1 + \frac{1 - \beta_j}{\beta_j} \frac{\alpha^2(C_{i,J})}{C_{i,J}^2}\right). \quad (3.1.13)$$

La densidad conjunta de $(C_{i,j}, C_{i,J})$ está dada por

$$\begin{aligned} f_{C_{i,j}, C_{i,J}}(x, y) &= f_{C_{i,j}|C_{i,J}}(x|y) f_{C_{i,J}}(y) \\ &= \frac{1}{(2\pi)^{1/2} \tau_j(y)} \frac{1}{x} \exp\left\{-\frac{1}{2} \left(\frac{\log(x) - v_j(y)}{\tau_j(y)}\right)^2\right\} \\ &\quad \times \frac{1}{(2\pi)^{1/2} \sigma_i y} \exp\left\{-\frac{1}{2} \left(\frac{\log(y) - \mu^{(i)}}{\sigma_i}\right)^2\right\} \\ &= \frac{1}{(2\pi) \sigma_i \tau_j(y)} \frac{1}{xy} \exp\left\{-\frac{1}{2} \left(\frac{\log(x) - v_j(y)}{\tau_j(y)}\right)^2 - \frac{1}{2} \left(\frac{\log(y) - \mu^{(i)}}{\sigma_i}\right)^2\right\}. \end{aligned} \quad (3.1.14)$$

Lema 3.1.4.

Si se cumple la Suposición 3.1.4 y 3.1.3 con $\alpha^2(\lambda) = a^2 \lambda^2$ para alguna $a \in \mathbb{R}$, entonces

$$\tau_j^2(\lambda) = \tau_j^2 = \log\left(1 + \frac{1 - \beta_j}{\beta_j} a^2\right), \quad (3.1.15)$$

$$v_j(c) = \log(\lambda) + \log(\beta_j) - \frac{1}{2} \tau_j^2. \quad (3.1.16)$$

Además, $C_{i,J}|C_{i,j}$ tiene distribución log-normal con parámetros

$$\mu_{post(i,j)} = \left(1 - \frac{\tau_j^2}{\sigma_i^2 + \tau_j^2}\right) \left(\frac{1}{2} \tau_j^2 + \log\left(\frac{C_{i,j}}{\beta_j}\right)\right) + \frac{\tau_j^2}{\sigma_i^2 + \tau_j^2} \mu^{(i)},$$

$$\sigma_{post(i,j)}^2 = \frac{\tau_j^2}{\sigma_i^2 + \tau_j^2} \sigma_i^2.$$

3.2. Métodos *Bootstrap*

Hasta el momento se centró la atención en la estimación y predicción de la siniestralidad última (esperada), $C_{i,j}$, $i \in \{0, \dots, I\}$ y los errores cuadráticos medios correspondientes. Es decir, se estimaron el primer y segundo momento de las obligaciones pendientes de cumplir y sus predictores. Sin embargo, no se pudieron calcular dichos momentos en forma cerrada. Además, generalmente se tiene interés en la distribución de probabilidad completa para calcular el VaR o algunas cantidades relacionadas con solvencia.

Una forma de “resolver” este problema es estimar el primer y segundo momento y hacer algún supuesto distribucional con respecto a las reservas totales y sus predictores, i.e., se “evita” el problema de encontrar distribuciones marginales, estructuras de dependencia y convoluciones con simplemente hacer una suposición con respecto a la distribución “resultante”.

El método *bootstrap* es una herramienta que permite obtener información con respecto a la distribución agregada de muestras (individuales) de datos. Este método fue definido por Efron (1979). De manera muy simplista se puede describir al *bootstrap* como “simular de un modelo estimado”, i.e., se define una estructura de modelo necesario para re-muestrear observaciones de los datos. Es decir, la idea detrás del *bootstrapping* es hacer un re-muestreo de los datos. El re-muestreo viene de los mismos datos (a partir de una estructura adecuada en el modelo) y a partir de esta estructura re-muestrear nuevos subconjuntos de datos del conjunto de información observada.

3.2.1. *Bootstrap* no-paramétrico de Efron

Considérese n variables (o vectores) aleatorias independientes e idénticamente distribuidas

$$Z_1, \dots, Z_n \stackrel{iid}{\sim} F,$$

donde F es una distribución desconocida.

Supóngase que se desea estimar un parámetro $h(F)$ de la distribución F . El parámetro $h(F)$ por ejemplo puede ser la media, la varianza, el VaR o el *Expected Shortfall* de la variable aleatoria Z_1 .

Si se tuviese una función conocida g (de la muestra aleatoria Z_1, \dots, Z_n) que estima $h(F)$, i.e.,

$$\hat{\theta}_n = g(Z_1, \dots, Z_n),$$

es un estimador de $h(F)$, como en la teoría clásica estadística, la idea es estudiar la distribución de $\hat{\theta}_n$.

Si se considera la distribución F , se podrían obtener simulaciones de esta distribución y con estas simulaciones se obtendrá un nuevo valor del estimador $\hat{\theta}_n$. Repitiendo este procedimiento (un número significativo de veces) se podrá obtener la distribución empírica de $\hat{\theta}_n$.

Sin embargo, como no se conoce explícitamente F , no se tiene un mecanismo explícito para generar simulaciones de F . Entonces, una de las opciones más naturales es utilizar la distribución empírica, \hat{F}_n , para generar observaciones. La distribución empírica \hat{F}_n da un peso de $\frac{1}{n}$ a cada observación Z_i y a partir de ésta se generan nuevos datos

$$Z_1^*, \dots, Z_n^* \stackrel{iid}{\sim} \hat{F}_n. \quad (3.2.1)$$

El nuevo vector simulado (Z_1^*, \dots, Z_n^*) se le conoce como muestra *bootstrap*. Entonces, para la muestra *bootstrap* se puede calcular un nuevo valor para el estimador $\hat{\theta}_n$, $\hat{\theta}_n^* = g(Z_1^*, \dots, Z_n^*)$. Repitiendo esta idea muchas veces, se obtendrá una distribución empírica F_n^* , para $\hat{\theta}_n^*$.

Ejemplo 3.2.1.

Considérese una muestra aleatoria $Z_1, \dots, Z_n \stackrel{iid}{\sim} \mathcal{N}(\theta, 1)$. La idea es estimar θ (la media de la distribución, i.e., $\mathbb{E}[Z_i] = \theta$). Si se elige al estimador $\hat{\theta} = g(Z_1, \dots, Z_n) = \frac{1}{n} \sum_{i=1}^n Z_i$ para las observaciones Z_1, \dots, Z_n definase a la distribución empírica

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{[Z_i, \infty)}^{(x)},$$

simule observaciones de esta distribución empírica

$$Z_1^*, \dots, Z_n^* \stackrel{iid}{\sim} \hat{F}_n,$$

y posteriormente se calcula $\hat{\theta}_n^* = g(Z_1^*, \dots, Z_n^*) = \frac{1}{n} \sum_{i=1}^n Z_i^*$. Para finalmente obtener una distribución empírica de $\hat{\theta}_n^*$, repitiendo el re-muestreo de \hat{F}_n muchas veces. ∇

Definición 3.2.1. (*Distribución Bootstrap*)

La distribución *bootstrap*, F_n^* , para $\hat{\theta}_n^*$ es la distribución condicional que es inducida por las observaciones independientes simuladas $Z_1^*, \dots, Z_n^* \stackrel{iid}{\sim} \hat{F}_n$ dados los datos originales Z_1, \dots, Z_n , i.e. F_n^* es una distribución condicional dados los datos originales Z_1, \dots, Z_n .

Observación 3.2.1.

- El hecho de que se condicione con respecto a los datos originales Z_1, \dots, Z_n permite que se pueda “fijar” a la distribución de re-muestreo \widehat{F}_n .
- Si la distribución empírica \widehat{F}_n es “cercana” a la distribución F original, entonces la distribución *bootstrap* F_n^* para $\widehat{\theta}_n^*$, también es “cercana” a la verdadera distribución del estimador $\widehat{\theta}_n$; por supuesto se tiene que definir formalmente el concepto de “cercanía”.
- Con este mecanismo sólo se está reproduciendo información ya contenida en las observaciones Z_1, \dots, Z_n , i.e., no está generando nueva información. Esto significa que toda la información *bootstrap* está condicionada a Z_1, \dots, Z_n . Es decir, la hipótesis subyacente es que las observaciones Z_1, \dots, Z_n sean suficientemente informativas para capturar las principales propiedades del modelo. ∇

Definición 3.2.2.

Se dice que el *bootstrap* es consistente para $\widehat{\theta}_n$ con respecto a la sucesión $\{a_n\}_{n \in \mathbb{N}}$ si para todo $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \left(\mathbb{P} \left[a_n \left(\widehat{\theta}_n - h(F) \right) \leq x \right] - \mathbb{P} \left[a_n \left(\widehat{\theta}_n^* - \widehat{\theta}_n \right) \leq x \right] \right) = 0.$$

Ejemplo 3.2.2.

En el ejemplo anterior ($Z_1, \dots, Z_n \stackrel{iid}{\sim} N(\theta, 1)$), si se escoge $a_n = \sqrt{n}$, entonces por el Teorema de Límite Central

$$a_n \left(\widehat{\theta}_n - \mathbb{E}[Z_1] \right) \underset{n \rightarrow \infty}{\overset{d}{\sim}} N(0, 1),$$

y entonces la consistencia *bootstrap* significa que

$$a_n \left(\widehat{\theta}_n^* - \widehat{\theta}_n \right) \underset{n \rightarrow \infty}{\overset{d}{\sim}} N(0, 1).$$

En general, es difícil demostrar consistencia. Generalmente se cumple cuando las variables aleatorias Z_1, \dots, Z_n son independientes e idénticamente distribuidas y si la distribución límite de $\widehat{\theta}_n$ es Gaussiana. También es común que se pida la condición técnica

$$\frac{\mathbb{E}_{F_n^*} \left[\widehat{\theta}_n^* \right] - \widehat{\theta}_n}{\mathbb{E} \left[\widehat{\theta}_n \right] - h(F)} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1 \quad \text{y} \quad \frac{\text{Var}_{F_n^*} \left(\widehat{\theta}_n^* \right)}{\text{Var} \left(\widehat{\theta}_n \right)} \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 1.$$

 ∇

3.2.1.1. Intervalos de confianza

Cuando el *bootstrap* es consistente con respecto a $\{a_n\}$ se pueden construir intervalos de confianza. Si q_α es el α -cuantil de $a_n \left(\hat{\theta}_n - h(F) \right)$ y \hat{q}_α es el α -cuantil de $a_n \left(\hat{\theta}_n^* - \hat{\theta}_n \right)$, dadas las observaciones Z_1, \dots, Z_n . Entonces el intervalo de confianza con $(1 - \alpha)100\%$ de confianza para el parámetro $h(F)$ está dado por

$$\left(\hat{\theta}_n - \frac{q_{\alpha/2}}{a_n}, \hat{\theta}_n - \frac{q_{1-\alpha/2}}{a_n} \right),$$

y entonces, dados Z_1, \dots, Z_n , el intervalo de confianza *bootstrap* es

$$\left(\hat{\theta}_n - \frac{\hat{q}_{\alpha/2}}{a_n}, \hat{\theta}_n - \frac{\hat{q}_{1-\alpha/2}}{a_n} \right).$$

Sin embargo, este intervalo no es exacto, pues utilizando diferentes niveles de *bootstrap* (*bootstrap* doble, triple, ...) su puede mejorar dicho intervalo de confianza.

3.2.2. *Bootstrap* paramétrico

El *bootstrap* no-paramétrico de Efron se puede ver como un mecanismo en el que se simula de la distribución empírica \hat{F}_n , i.e. simular de un modelo no-paramétrico donde las observaciones son independientes e idénticamente distribuidas con distribución desconocida F . Si además se supone que

$$Z_1, \dots, Z_n \stackrel{iid}{\sim} \hat{F}_\theta, \quad (3.2.2)$$

donde F_θ denota una distribución conocida pero que depende de un parámetro desconocido θ , entonces se dirá que se está haciendo *bootstrap* paramétrico.

Con el fin de re-muestrear la muestra *bootstrap*, primero se estima el parámetro desconocido θ a partir de $\hat{\theta}$ (ya sea por máxima verosimilitud, mínimos cuadrados u otro mecanismo de estimación). Posteriormente, el *bootstrap* paramétrico usa la distribución *bootstrap*

$$Z_1^*, \dots, Z_n^* \stackrel{iid}{\sim} \hat{F}_{\hat{\theta}} = F_{\hat{\theta}}, \quad (3.2.3)$$

en vez de la distribución empírica (3.2.1). Todos los pasos del *bootstrap* se hacen igual que el *bootstrap* no-paramétrico de Efron.

El *bootstrap* de Efron generalmente se aplica en los modelos estocásticos para el cálculo de la reserva, si inicialmente se supuso un modelo distribucional entonces se aplica el *bootstrap* paramétrico.

3.2.3. Modelo log-normal para reclamaciones acumuladas

Se considerarán la Suposición 2.4.2, los factores de desarrollo son independientes y tienen distribución log-normal, i.e.

$$\eta_{i,j} = \log \left(\frac{C_{i,j}}{C_{i,j-1}} \right) \sim N(\xi_j, \sigma_j^2).$$

Con el fin de utilizar el método *bootstrap* que se describió anteriormente, se hará $\theta_j = (\xi_j, \sigma_j^2)$ (los parámetros desconocidos de la distribución de $\eta_{i,j}$).

Dado \mathcal{D}_I , se tiene el interés en la distribución de

$$h(F) := \sum_{i+j>I} \mathbb{E}[X_{i,j} | \mathcal{D}_I], \quad (3.2.4)$$

ya que es la reserva de pérdidas esperadas abierta al tiempo I . Usando la expresión (2.4.9) se tiene que

$$h(F) = \sum_{i=1}^I C_{i,I-i} \left(\exp \left\{ \sum_{j=I-i+1}^J \xi_j + \frac{1}{2} \sum_{j=I-i+1}^J \sigma_j^2 \right\} - 1 \right).$$

Si ξ_j y σ_j^2 son conocidos entonces esta cantidad es constante (dada la información \mathcal{D}_I). Esto significa que para parámetros conocidos, no hay incertidumbre en la estimación de las reclamaciones abiertas pendientes.

Evidentemente, como en la mayoría de los casos, los parámetros son desconocidos y por tanto se deben estimar; dejando a su paso incertidumbre en las reclamaciones esperadas.

Una propuesta de estimadores para ξ_j, σ_j^2 está dada por

$$\hat{\xi}_j = \frac{1}{I-j+1} \sum_{i=0}^{I-j} \log \left(\frac{C_{i,j}}{C_{i,j-1}} \right) = \frac{1}{I-j+1} \sum_{i=0}^{I-j} \eta_{i,j}, \quad (3.2.5)$$

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j} \log \left(\frac{C_{i,j}}{C_{i,j-1}} - \hat{\xi}_j \right)^2 = \frac{1}{I-j} \sum_{i=0}^{I-j} \left(\eta_{i,j} - \hat{\xi}_j \right)^2, \quad (3.2.6)$$

que son los estimadores “tradicionales” para la media y la varianza Gaussianos.

Entonces, un estimador para $h(F)$ (dada la información \mathcal{D}_I), se puede construir como

$$\begin{aligned}
g(\mathcal{D}_I) &:= \sum_{i=1}^I \widehat{C}_{i,J}^{LN} - C_{i,I-i} \\
&= \sum_{i=1}^I C_{i,I-i} \left(\exp \left\{ \sum_{j=I-i+1}^J \widehat{\xi}_j + \frac{1}{2} \sum_{j=I-i+1}^J \widehat{\sigma}_j^2 \left(1 - \frac{1}{I-j+1} \right) \right\} - 1 \right).
\end{aligned} \tag{3.2.7}$$

El objetivo es estudiar la distribución del estimador $g(\mathcal{D}_I)$. Hasta el momento no se asegura que $g(\mathcal{D}_I)$ sea un estimador insesgado para $h(F)$, sólo se hace una suposición asintótica (2.4.20).

A partir de la suposición distribucional lo ideal será la aplicación del método *bootstrap* paramétrico. La tarea ahora consiste en la estimación de los parámetros desconocidos ξ_j y σ_j^2 a partir de las observaciones \mathcal{D}_I .

El método *bootstrap* paramétrico genera observaciones nuevas e independientes

$$\eta_{i,j}^* \sim N \left(\widehat{\xi}_j, \sigma_j^2 \right). \tag{3.2.8}$$

Con las observaciones *bootstrapping* los estimadores para ξ_j y σ_j^2 están dados como

$$\widehat{\xi}^* = \frac{1}{I-j+1} \sum_{i=0}^{I-j} \eta_{i,j}^*, \tag{3.2.9}$$

$$\widehat{\sigma}^{2*} = \frac{1}{I-j} \sum_{i=0}^{I-j} \left(\eta_{i,j}^* - \widehat{\xi}_j^* \right)^2. \tag{3.2.10}$$

Entonces la reserva *bootstrap* está dada como

$$g^*(\mathcal{D}_I) = \sum_{i=1}^I C_{i,I-i} \left(\exp \left\{ \sum_{j=I-i+1}^J \widehat{\xi}_j^* + \frac{1}{2} \sum_{j=I-i+1}^J \widehat{\sigma}^{2*} \left(1 - \frac{1}{I-j+1} \right) \right\} - 1 \right). \tag{3.2.11}$$

Si se repite el procedimiento “varias” veces se obtiene la distribución empírica de $g^*(\mathcal{D}_I)$ condicionada a \mathcal{D}_I .

Una vez que se propuso este estimador, ahora se estudiará el sesgo de $g^*(\mathcal{D}_I) - g(\mathcal{D}_I)$, así como la volatilidad de $g^*(\mathcal{D}_I)$. Debido a la \mathcal{D}_I -medibilidad de $g(\mathcal{D}_I)$, la siguiente expresión se puede descomponer como

$$\mathbb{E}_{\widehat{\theta}(\mathcal{D}_I)}^* \left[(g^*(\mathcal{D}_I) - g(\mathcal{D}_I))^2 \right] = \text{Var}_{\widehat{\theta}(\mathcal{D}_I)}^* (g^*(\mathcal{D}_I)) + \left(\mathbb{E}_{\widehat{\theta}(\mathcal{D}_I)}^* [g^*(\mathcal{D}_I)] - g(\mathcal{D}_I) \right)^2,$$

donde $\mathbb{P}_{\hat{\theta}(\mathcal{D}_I)}^*$ denota la distribución de probabilidad condicional a \mathcal{D}_I y obtenida de (3.2.8). Esto significa que los verdaderos parámetros son $\hat{\xi}_j$ y $\hat{\sigma}_j^2$. Bajo esta suposición se analizarán las posibles fluctuaciones al momento de volver a re-muestrear las observaciones $\eta_{i,j}^*$.

3.2.4. Modelos lineales generalizados

Suposición 3.2.1. (*Familia de Dispersión Exponencial*)

1. Las reclamaciones incrementales $X_{i,j}$ de diferente año de accidente i y diferente año de desarrollo j son independientes
2. La función de densidad de los incrementos $X_{i,j}$ está dada por

$$f(x; \theta_{i,j}, \phi_{i,j}, \omega_{i,j}) = a\left(x, \frac{\phi_{i,j}}{\omega_{i,j}}\right) \exp\left\{\frac{x\theta_{i,j} - b(\theta_{i,j})}{\frac{\phi_{i,j}}{\omega_{i,j}}}\right\}, \quad (3.2.12)$$

donde $b(\cdot) \in \mathbb{R}$ es una función con segunda derivada del parámetro $\theta_{i,j}$ tal que $(b')^{-1}(\cdot)$ existe, $\phi > 0$ es el parámetro de dispersión, $\omega_{i,j} > 0$ constante conocida y $a(\cdot, \cdot) \in \mathbb{R}$ una función normalizada.

A partir de la Suposición (3.2.1) se tiene que

$$\mathbb{E}[X_{i,j}] = x_{i,j} \quad \text{y} \quad \text{Var}(X_{i,j}) = \frac{\phi_{i,j}}{\omega_{i,j}} V(x_{i,j}),$$

donde $V(\cdot)$ es una función de la varianza adecuada. Además, si se supone una estructura multiplicativa para las reclamaciones esperadas, entonces $x_{i,j} = \mu_i \gamma_j$ (la estimación de μ_i y γ_j se obtiene por máxima verosimilitud).

Como ya se estudió, la estimación del error cuadrático medio de predicción a través de la matriz con información de Fisher junto con las suposiciones asintóticas y la estimación del error son una tarea difícil sobretodo si no se conoce la distribución, esto permite la aplicación de una técnica *bootstrap* para encontrar otro estimador.

Si se utiliza el método *bootstrap* no-paramétrico de Efron es necesario encontrar residuales idénticamente distribuidos con los que se pueda construir la distribución empírica \hat{F}_n , suponiendo que $\phi = \frac{\phi_{i,j}}{\omega_{i,j}}$ es constante.

En la teoría de modelos lineales generalizados (GLM) existen diferentes definiciones para residuales (Pearson, desviación, residuales Anscombe). Si se eligen los residuales de Pearson de forma que

$$R_{i,j}^{(P)}(x_{i,j}) = \frac{X_{i,j} - x_{i,j}}{\sqrt{V(x_{i,j})}}, \quad (3.2.13)$$

con media cero y varianza ϕ . Entonces, se puede definir una distribución *bootstrap* para $R_{i,j}^{(P)}(x_{i,j})$. Por tanto, para cualesquiera $i + j \leq I$ se define a la distribución *bootstrap* de $\hat{F}_{\mathcal{D}_I}$ como

$$Z_{i,j} = \frac{X_{i,j} - \hat{x}_{i,j}}{\sqrt{V(x_{i,j})}}.$$

Así, si se re-muestran los residuales independientes e idénticamente distribuidos un número considerable de veces, entonces

$$Z_{i,j} \sim \hat{F}_{\mathcal{D}_I}.$$

Y por tanto, las observaciones *bootstrap* de $X_{i,j}$ quedan definidas como

$$X_{i,j}^* = \hat{x}_{i,j} + \sqrt{V(\hat{x}_{i,j})} Z_{i,j}^*. \quad (3.2.14)$$

Las observaciones *bootstrap* $X_{i,j}^*$ se refieren al triángulo de la reserva de reclamaciones con $\mathcal{D}_I^* = \{X_{i,j}^*; i + j \leq I\}$. Usando métodos GLM, se puede calcular los estimadores *bootstrap* μ_i^* , γ_j^* y $\hat{x}_{i,j}^*$ a partir de las observaciones *bootstrap* $X_{i,j}^*; i + j \leq I$. Entonces se puede obtener la reserva de reclamaciones *bootstrap*, $\hat{X}_{i,j}^{*EDF}; i + j > I$. Además, repitiendo el *bootstrapping* se obtiene la distribución *bootstrap* de la reserva de reclamaciones, condicional a \mathcal{D}_I .

3.2.5. Método *bootstrap* Chain-Ladder

Para estimar el ECMP se utilizan dos métodos de muestreo, condicional y no-condicional.

A partir de estas aproximaciones de re-muestreo, el modelo CL de distribución libre está dado por

$$C_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \varepsilon_{i,j+1}.$$

Supóngase que, dado \mathcal{B}_0 , los $\varepsilon_{i,j}$ son independientes e idénticamente distribuidos. Entonces los factores individuales de desarrollo se puede reescribir como

$$F_{i,j+1} = \frac{C_{i,j+1}}{C_{i,j}} = f_j + \frac{\sigma_j \varepsilon_{i,j+1}}{\sqrt{C_{i,j}}},$$

con σ_j conocida.

Como ya se dijo, para aplicar el método *bootstrap* se necesitan residuales apropiados que permitan la construcción de la función empírica \widehat{F}_n , a partir de la cual se obtienen las observaciones *bootstrap*. Es obvio que los mejores candidatos son $\varepsilon_{i,j}$ (por su independencia y distribución idéntica). Considérese para cualesquiera $i \in \{0, \dots, I\}$, $j \in \{1, \dots, J\}$ con $i + j \leq I$, los residuales

$$\tilde{\varepsilon}_{i,j} = \frac{F_{i,j} - \widehat{f}_{j-1}}{\sigma_{j-1}} \sqrt{C_{i,j-1}}, \quad (3.2.15)$$

donde $\widehat{f}_j = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}}{\sum_{i=0}^{I-j-1} C_{i,j}}$, son los estimadores CL.

Observación 3.2.2.

$\tilde{\varepsilon}_{i,j}$ son observables para determinada σ_j pero no para f_j desconocido. Y además,

$$\mathbb{E}[\tilde{\varepsilon}_{i,j} | \mathcal{B}_{j-1}] = 0 \quad \text{y} \quad \text{Var}(\tilde{\varepsilon}_{i,j} | \mathcal{B}_{j-1}) = 1 - \frac{C_{i,j-1}}{I-j},$$

$$\sum_{i=0}^{I-j-1} C_{i,j-1}$$

con $\frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}} < 1$. ▽

Esto significa que se debería ajustar el residual observado, $\tilde{\varepsilon}_{i,j}$, de tal forma que se tenga una mejor estimación del error (la distribución empírica tiene una varianza “muy” pequeña).

Nótese que

$$\begin{aligned} \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}} \tilde{\varepsilon}_{i,j+1} &= \sum_{i=0}^{I-j-1} \sqrt{C_{i,j}} \left(\frac{F_{i,j+1} - \widehat{f}_j}{\sigma_j} \right) \sqrt{C_{i,j}} = \frac{1}{\sigma_j} \sum_{i=0}^{I-j-1} C_{i,j} (F_{i,j+1} - \widehat{f}_j) \\ &= \frac{1}{\sigma_j} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \widehat{f}_j \right) = \frac{1}{\sigma_j} \left[\sum_{i=0}^{I-j-1} C_{i,j+1} - \sum_{i=0}^{I-j-1} C_{i,j} \cdot \widehat{f}_j \right] \\ &= \frac{1}{\sigma_j} \left[\sum_{i=0}^{I-j-1} C_{i,j+1} - \sum_{i=0}^{I-j-1} C_{i,j+1} \right] = 0. \end{aligned}$$

Esto implica que la distribución *bootstrap* de los residuales $\tilde{\varepsilon}_{i,j+1}$ no es centrada, i.e. la media es distinta de cero.

Para el ajuste del residual así como la varianza de la distribución *bootstrap*, se define

$$Z_{i,j} = \left(\frac{1}{\sqrt{1 - \frac{C_{i,j-1}}{\sum_{i=0}^{I-j} C_{i,j-1}}}} \right) \frac{F_{i,j} - \hat{f}_{j-1}}{\sigma_{j-1}} \sqrt{C_{i,j-1}}.$$

Los residuales $\{Z_{i,j}; i + j \leq I\}$ definen a la distribución *bootstrap* \hat{F}_{D_I} . Entonces, remuestreando los residuos de manera independiente

$$Z_{i,j}^* \stackrel{iid}{\sim} \hat{F}_{D_I}, \quad (3.2.16)$$

se generan las observaciones *bootstrap* $F_{i,j}^*$, $i + j \leq I$.

3.2.5.1. Estimación de error no-condicional

Para la aproximación no-condicional se generan completamente nuevos triángulos.

Supóngase un año de accidente i fijo, si se define a $C_{i,0}^* = C_{i,0}$, entonces para $j \geq 1$

$$C_{i,j}^* = \hat{f}_{j-1} C_{i,j-1}^* + \hat{\sigma}_{j-1} \sqrt{C_{i,j-1}^*} Z_{i,j}^*, \quad (3.2.17)$$

bajo la aproximación no-condicional los factores *bootstrap* de desarrollo individuales están dados como

$$\Gamma_{i,j+1}^* = \frac{C_{i,j+1}^*}{C_{i,j}^*} = \hat{f}_j + \frac{\hat{\sigma}_j}{\sqrt{C_{i,j}^*}} Z_{i,j+1}^*,$$

por lo tanto, los factores de desarrollo *bootstrapping* son

$$\hat{f}_j^* = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^*}{\sum_{i=0}^{I-j-1} C_{i,j}^*} = \sum_{i=0}^{I-j-1} \frac{C_{i,j}^*}{\sum_{k=0}^{I-j-1} C_{k,j}^*} \Gamma_{i,j+1}^*. \quad (3.2.18)$$

Entonces, el estimador CL para las reclamaciones últimas $C_{i,J}$ bajo la aproximación no-condicional están dados por

$$\hat{C}_{i,J}^* = C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j^*.$$

3.2.5.2. Estimación de error condicional

Para el modelo condicional, sólo se generan nuevas observaciones para el siguiente paso en la serie de tiempo, i.e., la distribución de esta aproximación siempre es condicional a \mathcal{D}_I . De esta forma, los factores *bootstrap* de desarrollo están dados por

$$\hat{f}_j^* = \sum_{i=0}^{I-j-1} \frac{C_{i,j}}{\sum_{k=0}^{I-j-1} C_{k,j}} \Gamma_{i,j+1}^*, \quad (3.2.19)$$

con

$$\Gamma_{i,j+1}^* = \hat{f}_j + \frac{\hat{\sigma}_j}{\sqrt{C_{i,j}^*}} Z_{i,j+1}^*. \quad (3.2.20)$$

Entonces, el estimador CL para la siniestralidad última $C_{i,J}$ en la aproximación condicional están dados por

$$\hat{C}_{i,J}^* = C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j^*.$$

3.2.6. Suposiciones matemáticas sobre métodos *bootstrap***Suposición 3.2.2.** (*Residuales Gaussianos*)

- (I) Las reclamaciones acumuladas de diferentes años de accidente i son independientes.
- (II) Existen constantes $f_j, \sigma_j > 0$ y variables aleatorias $\varepsilon_{i,j+1}$ Gaussianas estándar independientes e idénticamente distribuidas, tales que para todo $i \in \{0, \dots, I\}$ y $j \in \{0, \dots, J-1\}$

$$C_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \varepsilon_{i,j+1}.$$

Nótese que la suposición (3.2.2) no es un modelo bien definido, ya que teóricamente se podrían obtener reclamaciones acumuladas negativas.

Suposición 3.2.3. (*Suposición General*)

Supóngase que σ_j es conocida. Si los factores f_j no son conocidos, entonces el error $\varepsilon_{i,j}$ de los factores no es observable. Sólo los siguientes términos del error se pueden observar

$$\begin{aligned}\tilde{\varepsilon}_{i,j+1} &= \frac{C_{i,j+1} - \hat{f}_j C_{i,j}}{\sigma_j \sqrt{C_{i,j}}} = \varepsilon_{i,j+1} + \frac{(f_j - \hat{f}_j) C_{i,j}}{\sigma_j \sqrt{C_{i,j}}} \\ &= \varepsilon_{i,j+1} - \frac{\sqrt{C_{i,j}}}{I-j-1} \sum_{k=0}^{I-j-1} \sqrt{C_{k,j}} \varepsilon_{k,j+1}.\end{aligned}\quad (3.2.21)$$

Nótese que los términos anteriores no son independientes. El primer objetivo será encontrar transformaciones $A^{(j)}$ de dimensión \mathcal{B}_j tales que hagan que los términos error observables transformados sean i.i.d. Con la suposición Gaussiana, se puede utilizar una transformación lineal, i.e., en el modelo Gaussiano las variables aleatorias no-correlacionadas son independientes.

Re-escalando los términos error observables como

$$\begin{aligned}\Delta_{i,j+1} &= \sqrt{C_{i,j}} \tilde{\varepsilon}_{i,j+1} = \frac{C_{i,j+1} - \hat{f}_j C_{i,j}}{\sigma_j} \\ &= \sqrt{C_{i,j}} \varepsilon_{i,j+1} - \frac{C_{i,j}}{I-j-1} \sum_{k=0}^{I-j-1} \sqrt{C_{k,j}} \varepsilon_{k,j+1}.\end{aligned}\quad (3.2.22)$$

Considérese el vector $\Delta_{j+1} = (\Delta_{0,j+1}, \dots, \Delta_{I-j-1,j+1})' \in \mathbb{R}^{I-j}$. De la ecuación (3.2.22) inmediatamente se observa que los términos error re-escalados generan un hiperplano H de dimensión $(I-j-1)$ en el espacio \mathbb{R}^{I-j} , pues

$$(1, \dots, 1) \Delta_{j+1} = \sum_{i=0}^{I-j-1} \Delta_{i,j+1} = 0.\quad (3.2.23)$$

Entonces, para una nueva base $\{\Delta_{i,j+1}^*\}_{i=1}^{I-j-1}$ cubren el hiperplano H de dimensión $(I-j-1)$ tal que $\Delta_{i,j+1}^*$ y $\Delta_{k,j+1}^*$ son ortogonales para $i \neq k$. En notación matricial $B^{(j)} \in \mathbb{R}^{(I-j)(I-j)}$ es \mathcal{B}_j -medible con

$$B^{(j)} \Delta_{j+1} = (0, \Delta_{1,j+1}^*, \dots, \Delta_{I-j-1,j+1}^*)' \in \mathbb{R}^{I-j},\quad (3.2.24)$$

tal que para cualesquiera $i \neq k$

$$\text{Cov}(\Delta_{i,j+1}^*, \Delta_{k,j+1}^* | \mathcal{B}_j) = 0. \quad (3.2.25)$$

La matriz \mathcal{B}_j -medible es definida como

$$B^{(j)} = \left\{ \begin{array}{cccccccc} 1 & 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & -\frac{C_{0,j}}{C_{1,j}} & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & -\frac{C_{0,j} + C_{1,j}}{C_{2,j}} & 0 & \dots & 0 & 0 \\ \vdots & & & & & & \vdots \\ 1 & 1 & 1 & 1 & \dots & -\frac{\sum_{i=0}^{I-j-3} C_{i,j}}{C_{I-j-2,j}} & 0 \\ 1 & 1 & 1 & 1 & \dots & 1 & -\frac{\sum_{i=0}^{I-j-2} C_{i,j}}{C_{I-j-1,j}} \end{array} \right\}. \quad (3.2.26)$$

Lema 3.2.1. La matriz $B^{(j)}$ definida en (3.2.26), satisface (3.2.24) y (3.2.25) con

$$\tau_{k,j}^2 = \text{Var}(\Delta_{k,j+1}^* | \mathcal{B}_j) = \sum_{i=0}^{k-1} C_{i,j} \left(1 + \sum_{m=0}^{k-1} \frac{C_{m,j}}{C_{k,j}} \right).$$

Corolario 3.2.1.

Bajo la Suposición 3.2.2. Para $j \in \{0, \dots, J-2\}$ se define la matriz \mathcal{B}_j -medible como

$$A^{(j)} = \begin{pmatrix} 1 & & & 0 \\ & \tau_{1,j}^{-2} & & \\ & & \ddots & \\ 0 & & & \tau_{I-j-1,j}^{-2} \end{pmatrix} B^{(j)},$$

y al vector aleatorio

$$\delta_{j+1} = (0, \delta_{1,j+1}, \dots, \delta_{I-j-1,j+1})' = A^{(j)} \Delta_{j+1} \in \mathbb{R}^{I-j}.$$

El conjunto $\tilde{\mathcal{D}}_I = \{\delta_{i,j}; i \in \{1, \dots, I\}, j \in \{1, \dots, J-1\}, i+j \leq I\}$ genera una familia de variables aleatorias independientes e idénticamente distribuidas, con distribución Gaussiana estándar.

3.3. Métodos Bayesianos exactos

En secciones anteriores se estudiaron modelos para el cálculo de la reserva con ideas Bayesianas subyacentes.

Los métodos Bayesianos para la estimación de la reserva de reclamaciones son aquellos que combinan suposiciones a *priori* con observaciones del triángulo de reclamaciones \mathcal{D}_I . Esta relación se puede expresar; por ejemplo, mediante la distribución a *priori* de la siniestralidad última. Entonces esta distribución se asocia con la función de verosimilitud a través del teorema de Bayes y así, obtener expresiones para la distribución a *posteriori* de esta distribución inicial, y poder calcular $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$, que se conoce como estimador Bayesiano.

Se dice que un método Bayesiano es exacto cuando el estimador Bayesiano es óptimo, en el sentido de que minimiza el error cuadrático medio de predicción en la clase $L^2_{C_{i,J}}(\mathcal{D}_I)$ de los estimadores de $C_{i,J}$, es decir,

$$\mathbb{E}[C_{i,J}|\mathcal{D}_I] = \underset{Y \in L^2_{C_{i,J}}(\mathcal{D}_I)}{\operatorname{argmín}} \mathbb{E} \left[(C_{i,J} - Y)^2 | \mathcal{D}_I \right].$$

Para éste, el error cuadrático medio de predicción condicional satisface que

$$ECMP_{C_{i,J}|\mathcal{D}_I} \left(\mathbb{E}[C_{i,J}|\mathcal{D}_I] \right) = \operatorname{Var}(C_{i,J}|\mathcal{D}_I).$$

Es claro que si se tienen parámetros no conocidos en el modelo, entonces no se puede calcular explícitamente $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$. Entonces, esos parámetros son remplazados por sus respectivos estimadores \mathcal{D}_I -medibles. Por tanto, se obtiene el estimador $\widehat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I]$ de $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$ (que también es predictor de $C_{i,J}|\mathcal{D}_I$), y esto implica que el error cuadrático medio de predicción condicional se obtiene como

$$ECMP_{C_{i,J}|\mathcal{D}_I} \left(\widehat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I] \right) = \operatorname{Var}(C_{i,J}|\mathcal{D}_I) + \left(\widehat{\mathbb{E}}[C_{i,J}|\mathcal{D}_I] - \mathbb{E}[C_{i,J}|\mathcal{D}_I] \right)^2,$$

entonces se tiene una situación similar al modelo CL.

En un modelo Bayesiano los parámetros desconocidos también se deberían estimar con métodos Bayesianos minimizando una función de pérdida.

Observación 3.3.1.

Para tarificación y *pricing* de contratos de seguros, se suele utilizar técnicas Bayesianas ya que incorporan información externa. Para la reserva de reclamaciones los métodos Bayesianos son menos usados. ∇

3.3.1. Modelo Poisson de sobredispersión con distribución Gamma a priori

Se considerará una variable latente Θ_i . Condicionado a Θ_i , se hará una suposición distribucional para las reclamaciones incrementales $X_{i,j}$ y las reclamaciones acumuladas $C_{i,j}$ en la que Θ_i describe la característica del riesgo por año accidente i , i.e. si un año es “bueno” o “malo”. Entonces, $C_{i,J}$ es una variable aleatoria con parámetros que dependen de Θ_i .

Suposición 3.3.1. (*Modelo Poisson de sobredispersión con distribución Gamma a priori*)

Para cualesquiera $i \in \{0, \dots, I\}$ y $j \in \{0, \dots, J\}$, supóngase que existen variables aleatorias Θ_i y $Z_{i,j}$, y constantes $\phi_i > 0$, $\gamma_0, \dots, \gamma_J > 0$ con $\sum_{j=0}^J \gamma_j = 1$, tales que

- (I) Condicionado a Θ_i , las variables aleatorias $Z_{i,j}$ son independientes con distribución Poisson, y además las variables incrementales $X_{i,j} = \phi_i Z_{i,j}$ satisfacen

$$\mathbb{E}[X_{i,j}|\Theta_i] = \Theta_i \gamma_j \quad \text{y} \quad \text{Var}(X_{i,j}|\Theta_i) = \phi_i \Theta_i \gamma_j. \quad (3.3.1)$$

- (II) Las parejas $(\Theta_i, (X_{i,0}, \dots, X_{i,J}))$ son independientes con $\Theta_i \sim \text{Gamma}(a_i, b_i)$.

Observación 3.3.2.

El parámetro ϕ_i denota la sobredispersión. Si $\phi_i = 1$ entonces se obtiene la distribución Poisson donde la varianza varía de forma lineal con la media. ∇

Dado Θ_i , la esperanza y varianza de $Z_{i,j}$ se obtienen como

$$\mathbb{E}[Z_{i,j}|\Theta_i] = \text{Var}(Z_{i,j}|\Theta_i) = \frac{\Theta_i \gamma_j}{\phi_i}, \quad (3.3.2)$$

entonces la esperanza a priori para las reclamaciones incrementales está dada como

$$\mathbb{E}[X_{i,j}] = \mathbb{E}[\mathbb{E}[X_{i,j}|\Theta_i]] = \gamma_j \mathbb{E}[\Theta_i] = \gamma_j \frac{a_i}{b_i}.$$

Así, la siniestralidad última se obtiene como

$$C_{i,J} = \phi_i \sum_{j=0}^J Z_{i,j}.$$

Por tanto, condicionado a Θ_i

$$\frac{C_{i,J}}{\phi_i} \sim \text{Poisson}(\Theta_i/\phi_i) \quad \text{y} \quad \mathbb{E}[C_{i,J}|\Theta_i] = \Theta_i.$$

Esto significa que Θ_i juega el papel de la siniestralidad última esperada por año accidente i . La teoría Bayesiana muestra como se debería combinar la esperanza a *priori* $\mathbb{E}[C_{i,J}] = a_i/b_i$ con las observaciones \mathcal{D}_I .

Observación 3.3.3.

Si se considera a las variables aleatorias $X_{i,j}$ como el número de siniestros reportados, se podría tener problemas con este modelo al suponer que son no-negativas. Sin embargo, si $X_{i,j}$ denota; por ejemplo, los pagos incrementales entonces se podría tener valores negativos en las observaciones. ∇

Se supondrá que el patrón de desarrollo de reclamaciones γ_j es conocido.

Proposición 3.3.1.

Bajo la Suposición 3.3.1, la distribución predictiva de Θ_i dado $(X_{i,0}, \dots, X_{i,j})$, es una distribución Gamma con parámetros

$$\begin{aligned} a_{i,j}^{post} &= a_i + \frac{C_{i,j}}{\phi_i}, \\ b_{i,j}^{post} &= b_i + \sum_{k=0}^j \frac{\gamma_k}{\phi_i} = b_i + \frac{\beta_j}{\phi_i}, \end{aligned}$$

donde $\beta_j = \sum_{k=0}^j \gamma_k$.

Demostración:

A partir de la ecuación (3.3.2), se obtiene la densidad condicional de $(X_{i,0}, \dots, X_{i,j})$ dado Θ_i , i.e.

$$f_{X_{i,0}, \dots, X_{i,j} | \Theta_i}(x_0, \dots, x_j | \theta) = \prod_{k=0}^j \exp \left\{ -\theta \frac{\gamma_k}{\phi_i} \right\} \frac{\left(\frac{\theta \gamma_k}{\phi_i} \right)^{x_k / \phi_i}}{x_k / \phi_i!}.$$

Por tanto, la distribución conjunta de $(X_{i,0}, \dots, X_{i,j})$ y Θ_i está dada como

$$\begin{aligned} f_{\Theta_i, X_{i,0}, \dots, X_{i,j}}(\theta, x_0, \dots, x_j) &= f_{X_{i,0}, \dots, X_{i,j} | \Theta_i}(x_0, \dots, x_j | \theta) f_{\Theta_i}(\theta) \\ &= \prod_{k=0}^j \exp\left\{-\theta \frac{\gamma_k}{\phi_i}\right\} \frac{\left(\frac{\theta \gamma_k}{\phi_i}\right)^{x_k / \phi_i}}{x_k / \phi_i!} \cdot \frac{b_i^{a_i}}{\Gamma(a_i)} \theta^{a_i-1} \exp(-b_i \theta) \\ &= \prod_{k=0}^j \exp\left\{-\theta \left(\frac{\gamma_k}{\phi_i} + b_i\right)\right\} \theta^{a_i + \frac{x_k}{\phi_i} - 1} \frac{\left(\frac{\gamma_k}{\phi_i}\right)^{x_k / \phi_i}}{x_k / \phi_i!} \cdot \frac{b_i^{a_i}}{\Gamma(a_i)}. \end{aligned}$$

De aquí se observa que la distribución predictiva de Θ_i dado $(X_{i,0}, \dots, X_{i,j})$ es otra una distribución Gamma con parámetros

$$a_{i,j}^{post} = a_i + \frac{C_{i,j}}{\phi_i} \quad \text{y} \quad b_{i,j}^{post} = b_i + \sum_{k=0}^j \frac{\gamma_k}{\phi_i}.$$

□

Dado que los años accidente son independientes, para calcular la distribución predictiva de Θ_i es suficiente sólo considerar las observaciones $(X_{i,0}, \dots, X_{i,j})$ para un año accidente i .

Usando la Proposición 3.3.1 se obtiene la esperanza predictiva, i.e.

$$\begin{aligned} \mathbb{E}[\Theta_i | \mathcal{D}_{\mathcal{I}}] &= \frac{a_{i,I-i}^{post}}{b_{i,I-i}^{post}} = \frac{a_i + C_{i,I-i}/\phi_i}{b_i + \beta_{i,I-i}/\phi_i} \\ &= \frac{a_i}{b_i + \beta_{i,I-i}/\phi_i} \cdot \frac{b_i}{b_i} + \frac{\frac{\beta_{I-i}}{\beta_{I-i}} \frac{C_{i,I-i}}{\phi_i}}{b_i + \beta_{i,I-i}/\phi_i} \\ &= \frac{b_i}{b_i + \beta_{i,I-i}/\phi_i} \cdot \frac{a_i}{b_i} + \frac{\frac{\beta_{I-i}}{\phi_i} + b_i - b_i}{b_i + \beta_{i,I-i}/\phi_i} \cdot \frac{C_{i,I-i}}{\beta_{I-i}} \\ &= \frac{b_i}{b_i + \beta_{I-i}/\phi_i} \cdot \frac{a_i}{b_i} + \left(1 - \frac{b_i}{b_i + \beta_{I-i}/\phi_i}\right) \frac{C_{i,I-i}}{\beta_{I-i}}, \end{aligned} \quad (3.3.3)$$

que se puede interpretar como un promedio ponderado de credibilidad entre la esperanza a priori $\mathbb{E}[\Theta_i]$ y la observación $\frac{C_{i,I-i}}{\beta_{I-i}}$.

También se puede dar una expresión para la distribución predictiva de $(C_{i,J} - C_{i,I-i})/\phi_i$ dado $\mathcal{D}_{\mathcal{I}}$, entonces para cualesquiera $k \in \{0, 1, \dots\}$

$$\begin{aligned}
 & \mathbb{P} \left[\frac{(C_{i,J} - C_{i,I-i})}{\phi_i} = k | \mathcal{D}_I \right] \\
 &= \int_{\mathbb{R}_+} \exp \left(-(1 - \beta_{I-i}) \frac{\theta}{\phi_i} \right) \frac{\left((1 - \beta_{I-i}) \frac{\theta}{\phi_i} \right)^k}{k!} \frac{(b_{i,I-i}^{post})^{a_{i,I-i}^{post}}}{\Gamma(a_{i,I-i}^{post})} \theta^{a_{i,I-i}^{post}-1} \exp \left(-b_{i,I-i}^{post} \theta \right) d\theta \\
 &= \frac{(b_{i,I-i}^{post})^{a_{i,I-i}^{post}} ((1 - \beta_{I-i})/\phi_i)^k}{\Gamma(a_{i,I-i}^{post}) k!} \underbrace{\int_{\mathbb{R}_+} \theta^{k+a_{i,I-i}^{post}-1} \exp \left(-(b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i) \theta \right) d\theta}_{\propto \Gamma(k+a_{i,I-i}^{post}, b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i)} \\
 &= \frac{(b_{i,I-i}^{post})^{a_{i,I-i}^{post}} ((1 - \beta_{I-i})/\phi_i)^k}{\Gamma(a_{i,I-i}^{post}) k!} \cdot \frac{\Gamma(k + a_{i,I-i}^{post})}{(b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i)^{k+a_{i,I-i}^{post}}} \\
 &= \frac{\Gamma(k + a_{i,I-i}^{post})}{k! \Gamma(a_{i,I-i}^{post})} \left(\frac{b_{i,I-i}^{post}}{b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i} \right)^{a_{i,I-i}^{post}} \left(\frac{(1 - \beta_{I-i})/\phi_i}{b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i} \right)^k \\
 &= \binom{k + a_{i,I-i}^{post}}{k} \left(\frac{b_{i,I-i}^{post}}{b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i} \right)^{a_{i,I-i}^{post}} \left(\frac{(1 - \beta_{I-i})/\phi_i}{b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i} \right)^k,
 \end{aligned}$$

que es una distribución binomial/negativa con parámetros $r = a_{i,I-i}^{post}$ y $p = b_{i,I-i}^{post} + (1 - \beta_{I-i})/\phi_i$.

La importancia de este resultado es que se puede calcular explícitamente la distribución de las pérdidas pendientes condicionada a $C_{i,I-i}$. La desventaja de este modelo es que no se puede obtener estimados del error cuadrático medio pero si se podrían calcular estimados del VaR y otras cantidades relacionadas con el riesgo.

Además, este modelo proporciona una representación explícita de la distribución para la siniestralidad última $C_{i,J}$ dado \mathcal{D}_I y permite cuantificar su riesgo calculando el VaR.

A partir de la ecuación (3.3.1) y de la independencia de $X_{i,j}$ dado Θ_i se obtiene

$$\begin{aligned}
 \mathbb{E}[C_{i,J} | \mathcal{D}_I] &= \mathbb{E} \left[\mathbb{E}[C_{i,J} | \Theta_i, \mathcal{D}_I] | \mathcal{D}_I \right] \\
 &= C_{i,I-i} + \mathbb{E} \left[\mathbb{E} \left[\sum_{j=I-i+1}^J X_{i,j} \middle| \Theta_i \right] \middle| \mathcal{D}_I \right] \\
 &= C_{i,I-i} + (1 - \beta_{I-i}) \mathbb{E}[\Theta_i | \mathcal{D}_I].
 \end{aligned} \tag{3.3.4}$$

Entonces con (3.3.3) y (3.3.4), se obtiene el siguiente estimador:

Estimador 3.3.1. (*Estimador Poisson-Gamma*)

Bajo la Suposición 3.3.1 se obtiene el siguiente estimador para siniestralidad última $\mathbb{E}[C_{i,J}|\mathcal{D}_I]$, para cualquier $i \in \{0, \dots, I\}$

$$\widehat{C}_{i,J}^{PoiGa} = C_{i,I-i} + (1 - \beta_{I-i}) \left[\frac{b_i}{b_i + (\beta_{I-i}/\phi_i)} \frac{a_i}{b_i} + \left(1 - \frac{b_i}{b_i + (\beta_{I-i}/\phi_i)} \right) \frac{C_{i,I-i}}{\beta_{I-i}} \right].$$

El error cuadrático medio de predicción condicional del estimador está dado por

$$\begin{aligned} ECMP_{C_{i,J}|\mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) &= \mathbb{E} \left[\left(C_{i,J} - \widehat{C}_{i,J}^{PoiGa} \right)^2 \middle| \mathcal{D}_I \right] \\ &= \mathbb{E} \left[\left(\sum_{j=I-i+1}^J X_{i,j} - (1 - \beta_{I-i}) \mathbb{E}[\Theta_i | \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right] \\ &= \mathbb{E} \left[\left(\sum_{j=I-i+1}^J \left(X_{i,j} - \gamma_j \mathbb{E}[\Theta_i | \mathcal{D}_I] \right) \right)^2 \middle| \mathcal{D}_I \right]. \end{aligned}$$

Entonces, para el estimador Bayesiano se cumple

$$\mathbb{E}[X_{i,j}|\mathcal{D}_I] = \mathbb{E} \left[\mathbb{E}[X_{i,j}|\Theta_i, \mathcal{D}_I] \middle| \mathcal{D}_I \right] = \mathbb{E} \left[\mathbb{E}[X_{i,j}|\Theta_i] \middle| \mathcal{D}_I \right] = \gamma \mathbb{E}[\Theta_i | \mathcal{D}_I],$$

para $j > I - i$.

Por lo tanto, con este resultado se tiene que

$$ECMP_{C_{i,J}|\mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) = Var \left(\sum_{j=I-i+1}^J X_{i,j} \middle| \mathcal{D}_I \right).$$

Esta expresión se puede calcular con ayuda de la distribución Binomial-negativa. Utili-

zando la independencia de $X_{i,j}$ dado Θ_i y la ecuación (3.3.1) entonces se obtiene

$$\begin{aligned}
 \text{Var} \left(\sum_{j=I-i+1}^J X_{ij} \middle| \mathcal{D}_I \right) &= \mathbb{E} \left[\text{Var} \left(\sum_{j=I-i+1}^J X_{i,j} \middle| \Theta_i \right) \middle| \mathcal{D}_I \right] + \text{Var} \left[\mathbb{E} \left(\sum_{j=I-i+1}^J X_{i,j} \middle| \Theta_i \right) \middle| \mathcal{D}_I \right] \\
 &= \mathbb{E} \left[\sum_{j=I-i+1}^J \phi_i \Theta_i \gamma_j \middle| \mathcal{D}_I \right] + \text{Var} \left(\sum_{j=I-i+1}^J \Theta_i \gamma_j \middle| \mathcal{D}_I \right) \\
 &= \phi_i (1 - \beta_{I-i}) \mathbb{E}[\Theta_i | \mathcal{D}_I] + (1 - \beta_{I-i})^2 \text{Var}(\Theta_i | \mathcal{D}_I).
 \end{aligned}$$

Proposición 3.3.2.

Bajo la Suposición 3.3.1, el error cuadrático medio de predicción para un año accidente i , $i \in \{1, \dots, I\}$ está dado por

$$ECMP_{C_{i,J} | \mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) = \phi_i (1 - \beta_{I-i}) \frac{a_{i,I-i}^{post}}{b_{i,I-i}^{post}} + (1 - \beta_{I-i})^2 \frac{a_{i,I-i}^{post}}{(b_{i,I-i}^{post})^2}.$$

Observación 3.3.4.

Nótese que a_i , b_i , ϕ_i y γ_j son conocidos. Si necesitarán ser estimados se obtiene un término adicional en el cálculo del ECMP, correspondiente al error de estimación del parámetro. ∇

El error cuadrático medio de predicción no-condicional se calcula como

$$\begin{aligned}
 ECMP_{C_{i,J} | \mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) &= \mathbb{E} \left[ECMP_{C_{i,J} | \mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) \right] \\
 &= \phi_i (1 - \beta_{I-i}) \frac{\mathbb{E}[a_{i,I-i}^{post}]}{b_{i,I-i}^{post}} + (1 - \beta_{I-i})^2 \frac{\mathbb{E}[a_{i,I-i}^{post}]}{(b_{i,I-i}^{post})^2},
 \end{aligned}$$

y usando $C_{i,I-i} = \beta_{I-i} \left(\frac{a_i}{b_i} \right)$ se obtiene

$$ECMP_{C_{i,J} | \mathcal{D}_I} \left(\widehat{C}_{i,J}^{PoiGa} \right) = \phi_i (1 - \beta_{I-i}) \frac{a_i}{b_i} \cdot \frac{1 + \phi_i b_i}{\phi_i b_i + \beta_{I-i}}.$$

3.4. Credibilidad Bayesiana para el modelo Chain-Ladder

En las secciones anteriores ya se estudiaron algunos estimadores tipo Chain-Ladder que se “combinan” para formar nuevos estimadores; de hecho es la idea básica de los métodos basados en credibilidad. Es decir, se retomará la idea de combinar información siniestral (los datos históricos) y algún otro tipo de información que ayude a “mejorar” las estimaciones (por ejemplo, índices de siniestralidad de cierto ramo, alguna índice construido por un experto o el regulador, algún factor de ajuste corporativo, información de algún ramo “similar” al que se está estudiando, etc).

En esta sección se desarrollará un estimador Bayesiano y un estimador de credibilidad (“diferentes” a otros que se han definido en secciones anteriores) con esta idea Bayesiana de incorporar información adicional.

Además, se demostrará que este estimador de credibilidad es un estimador Bayesiano exacto en el caso de la familia exponencial con sus *a priori*s naturales conjugadas; y como en todo el desarrollo de este trabajo, se relacionará con el modelo Chain-Ladder clásico, mostrando que el error cuadrático medio de estimación es diferente al del Chain-Ladder clásico.

Notación 3.4.1.

Se denotará por \mathbf{C}_j a la j -ésima columna del triángulo de reclamaciones acumuladas, es decir, $\mathbf{C}_j = (C_{0,j}, C_{1,j}, \dots, C_{I-j,j})^\top$, $j \in \{0, 1, \dots, J\}$. Además, para $k \in \{0, \dots, I-j\}$ defínase la suma truncada

$$S_j^{[k]} := \sum_{i=0}^k C_{i,j}.$$

▽

Como ya se observó en el capítulo 2, la suposición básica detrás del método Chain-Ladder se que los vectores columna $\{\mathbf{C}_j\}_{j=0}^J$ son “proporcionales” (salvo variaciones aleatorias), i.e.

$$C_{i,j+1} \simeq f_j \cdot C_{i,j},$$

para algunas constante adecuadas f_j (que se les llamó factores de desarrollo) y el pronóstico/estimación Chain-Ladder está dado por

$$C_{i,k}^{CL} = C_{i,I-i} \prod_{j=I-i}^{k-1} \hat{f}_j,$$

donde

$$\hat{f}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}.$$

Sólo para tener referencia y hacer una comparación más fluida, se recordarán las suposiciones de Mack para definir un modelo estocástico subyacente para el método Chain-Ladder.

Suposición 3.4.1.

1. Las variables aleatorias $C_{i,j}$ son independientes para diferentes años accidentes $i \in \{0, 1, \dots, I\}$.
2. Existen constantes $f_j > 0$ y $\sigma_j^2 > 0$ tales que para cualesquiera $i \in \{0, 1, \dots, I\}$, $j \in \{0, 1, \dots, J-1\}$ tales que

$$\begin{aligned}\mathbb{E}[C_{i,j+1} | C_{i,j}] &= f_j C_{i,j}, \\ \text{Var}[C_{i,j+1} | C_{i,j}] &= \sigma_j^2 C_{i,j}.\end{aligned}$$

Como antes, \mathcal{B}_j será el conjunto $\{C_{i,k} : i+k \leq I, k \leq j\}$, es decir $\mathcal{B}_j = \{\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_j\} \subseteq \mathcal{D}_I$.

Para el desarrollo de esta metodología se hará un cambio de las variables aleatorias $C_{i,j}$ por las variables aleatorias $Y_{i,j}$ (factores de desarrollo individuales) que se definen como

$$Y_{i,j} := \frac{C_{i,j+1}}{C_{i,j}}.$$

Notación 3.4.2.

Para $j \in \{0, 1, \dots, J-1\}$, se denotará al vector columna de factores de desarrollo individuales como

$$\mathbf{Y}_j := (Y_{0,j}, \dots, Y_{I-j-1,j})^\top,$$

y como

$$\mathbf{y}_j = (y_{0,j}, \dots, y_{I-j-1,j})^\top,$$

a una realización de \mathbf{Y}_j . ▽

Entonces la Suposición 3.4.1 del modelo Chain-Ladder clásico es equivalente a

$$\begin{aligned}\mathbb{E}[Y_{i,j} | C_{i,j}] &= f_j, \\ \text{Var}[Y_{i,j} | C_{i,j}] &= \frac{\sigma_j^2}{C_{i,j}}.\end{aligned}$$

En el modelo de Mack sólo se consideran los datos individuales de un triángulo específico. Ahora se estudiará, a través de información *a priori* o de información de un portafolio con riesgos “similares”, si existe la posibilidad de desarrollar un patrón de desarrollo de la reclamaciones.

3.4.1. Chain-Ladder Bayesiano

En esta perspectiva Bayesiana, se supondrá que los factores de desarrollo desconocidos $f_j, j \in \{0, 1, \dots, J-1\}$ son realizaciones independientes de las variables aleatorias independientes F_j .

Notación 3.4.3.

Se denotará por $\mathbf{F} = (F_0, F_1, \dots, F_{J-1})^\top$ al vector de F_j 's y por $\mathbf{f} = (f_0, f_1, \dots, f_{J-1})^\top$ a una realización de \mathbf{F} . ▽

El modelo Chain-Ladder Bayesiano supone que, dado \mathbf{F} , la Suposición 2.3.1 del modelo CL se cumple.

Suposición 3.4.2. (*Chain-Ladder Bayesiano*)

1. Dado \mathbf{F} , las variables aleatorias $C_{i,j}$ de diferentes años accidente i son independientes.
2. Dados \mathbf{F} y \mathcal{B}_j , la distribución condicional de $F_{i,j}$ depende sólo de $C_{i,j}$ y además se cumple que

$$\mathbb{E}[Y_{i,j} | \mathbf{F}, \mathcal{B}_j] = F_j, \quad (3.4.1)$$

$$\text{Var}[Y_{i,j} | \mathbf{F}, \mathcal{B}_j] = \frac{\sigma_j^2(F_j)}{C_{i,j}}. \quad (3.4.2)$$

3. Las variables aleatorias F_0, F_1, \dots, F_{J-1} son independientes.

Si se satisface la Suposición 3.4.2, la esperanza condicional de $Y_{i,j}$ dados \mathbf{F} y \mathcal{B}_j , depende sólo del factor Chain-Ladder (aleatorio) desconocido, F_j , y no de los factores Chain-Ladder F_k de otro periodo de desarrollo $k \neq j$.

Observación 3.4.1.

- En la ecuación (3.4.2), $C_{i,j}$ tiene el papel del un ponderador, i.e. la varianza condicional de $Y_{i,j}$ dados \mathbf{F} y \mathcal{B}_j , es inversamente proporcional a $C_{i,j}$ y el denominador depende de F_j .

- La distribución no condicional de \mathbf{F} no depende de \mathcal{D}_I . A la distribución de los F_j 's se le conoce como función estructural.

▽

Si se define $\widehat{F}_j := \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}$ (que es el estimador del factor Chain-Ladder f_j clásico), entonces se satisface la Suposición 3.4.2, entonces

$$\mathbb{E} \left[\widehat{F}_{i,j} \mid \mathbf{F}, \mathcal{B}_j \right] = F_j, \quad (3.4.3)$$

$$\text{Var} \left[\widehat{F}_{i,j} \mid \mathbf{F}, \mathcal{B}_j \right] = \frac{\sigma_j^2(F_j)}{S_j^{[I-j-1]}}. \quad (3.4.4)$$

Este planteamiento tiene algunas analogías con el modelo de Mack, por ejemplo

- Dado \mathbf{F} , $\{C_{i,j} : j \in \{0, \dots, J\}\}$ tiene la propiedad de Markov, i.e. la distribución condicional de $C_{i,j+1} \mid C_{i,0}, \dots, C_{i,j}$ sólo depende de la última observación $C_{i,j}$. Esta es una suposición más fuerte que la de Mack ya que en esa sólo se hizo la afirmación para el primer y el segundo momentos condicionales, no sobre la función de distribución en sí misma.
- Dado \mathbf{F} , $\{Y_{i,j} : j \in \{0, \dots, J-1\}\}$ son no-correlacionados y además para $i \neq k$, $Y_{i,j}$ y $Y_{k,l}$ son independientes. Esta suposición es más fuerte que la de Mack ya que en esa sólo se cumplía la no correlación.

El objetivo es encontrar mejores predictores de $C_{i,j}$, dadas las observaciones \mathcal{D}_I , para $i + j > I$.

Teorema 3.4.1.

Bajo la Suposición 3.4.2, dadas las observaciones \mathcal{D}_I , las variables aleatorias F_0, F_1, \dots, F_{J-1} son independientes con distribución a *posteriori* dada por

$$K \prod_{j=0}^{J-1} dU(f_j \mid \mathcal{D}_I),$$

donde K es una constante de normalización y $U(f_j)$ es el conjunto de funciones de distribución de las variables aleatorias F_j .

Demostración:

Para simplificar la notación, para $j \in \{0, \dots, J-1\}$, se denotará como $U(f_j)$ al conjunto de funciones de distribución de las variables aleatorias F_j . También se denotará por $F_{f_i}(\cdot)$ y $F_{\mathbf{f}}(\cdot)$ a las distribuciones condicionales dados $F_j = f_j$ (o $\mathbf{F} = \mathbf{f}$), respectivamente. Por ejemplo, $F_{f_j}(y_{i,j} | \mathcal{B}_j)$ es la función de distribución condicional de $Y_{i,j}$ dados $F_j = f_j$ y \mathcal{B}_j .

A partir de la Suposición 3.4.2 se tiene que

$$dF_{\mathbf{f}}(\mathbf{y}_{0,j}, \mathbf{y}_{1,j}, \dots, \mathbf{y}_{J-1} | \mathcal{B}_0) = \prod_{j=0}^{J-1} \prod_{i=0}^{I-j-1} dF_{f_j}(y_{i,j} | C_{i,j}),$$

donde, $C_{i,j} = y_{i,j-1} C_{i,j-1}$ para $j \geq 1$. Entonces, la función de distribución conjunta a posteriori de F dado \mathcal{D}_I está dada por

$$\begin{aligned} dU(f_0, \dots, f_{J-1} | \mathcal{D}_I) &\propto \prod_{j=0}^{J-1} \left\{ \prod_{i=0}^{I-j-1} dF_{f_j}(y_{i,j} | C_{i,j}) dU(f_j) \right\} \\ &\propto \prod_{j=0}^{J-1} dU(f_j | \mathcal{D}_I). \end{aligned}$$

□

Observación 3.4.2.

La distribución condicional de F_j , dado \mathcal{D}_I , sólo depende de \mathbf{Y}_j y \mathbf{C}_j , donde $C_{i,j}$, $i \in \{0, \dots, I-j\}$, son las ponderaciones ya que $\text{Var}[Y_{i,j} | \mathbf{F}, \mathcal{B}_j] = \frac{\sigma_j^2(F_j)}{C_{i,j}}$. De hecho, las variables aleatorias $Y_{i,j}$ son las únicas en el Y -trapezoide que contienen información acerca de F_j . ▽

Ahora, se obtendrá el estimador de siniestralidad última Bayesiano. Para esto, primero se definirá un criterio de optimalidad.

Definición 3.4.1.

Se Z una variable aleatoria. Sean \hat{Z}, \tilde{Z} estimadores de Z . Se dice que \hat{Z} es mejor o igual que \tilde{Z} si

$$\mathbb{E} \left[\left(\hat{Z} - Z \right)^2 \right] \leq \mathbb{E} \left[\left(\tilde{Z} - Z \right)^2 \right].$$

Observación 3.4.3.

La definición 3.4.1 establece que se usará a la pérdida cuadrática esperada como criterio de optimalidad. ∇

Lema 3.4.1.

Sea Z una variable aleatoria y \mathbf{X} un vector aleatorio de observaciones. Entonces, el mejor estimador de Z es

$$Z^{Bayes} := \mathbb{E}[Z|\mathbf{X}].$$

Además, Z^{Bayes} minimiza la pérdida cuadrática esperada condicional, i.e.

$$Z^{Bayes} = \arg \min_{\hat{Z}} \mathbb{E} \left[\left(\hat{Z} - Z \right)^2 \middle| \mathbf{X} \right]. \quad (3.4.5)$$

Definición 3.4.2.

Sea $\hat{C}_{i,J}$ un estimador de la siniestralidad última $C_{i,J}$ dadas las observaciones \mathcal{D}_I . Se define al error cuadrático medio condicional de $\hat{C}_{i,J}$ como

$$ECMP(\hat{C}_{i,J}) = \mathbb{E} \left[\left(\hat{C}_{i,J} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right]. \quad (3.4.6)$$

Como antes, sea \hat{R}_i la reserva estimada, es decir

$$\hat{R}_i = \hat{C}_{i,J} - C_{i,I-i}.$$

Nótese que

$$ECMP(\hat{C}_{i,J}) = ECMP(\hat{R}_i) = \mathbb{E} \left[\left(\hat{R}_i - R_i \right)^2 \middle| \mathcal{D}_I \right].$$

Entonces, como en la ecuación (3.4.5),

$$C_{i,J}^{Bayes} := \mathbb{E}[C_{i,J} | \mathcal{D}_I], \quad (3.4.7)$$

y este es el mejor estimador que minimiza el error cuadrático medio condicional (3.4.6).

Teorema 3.4.2.

Si se cumple la Suposición 3.4.2, entonces

$$C_{i,J}^{Bayes} = C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes}, \quad (3.4.8)$$

donde F_j^{Bayes} denota al estimador Bayesiano de F_j .

Demostración:

Por la independencia *a posteriori* de F_j dado \mathcal{D}_I , según el Teorema 3.4.1, se tiene que para $j \in \{0, \dots, J-1\}$, $Y_{i,j}$ son también condicionalmente no-correlacionados. Entonces,

$$\begin{aligned} C_{i,J}^{Bayes} &= \mathbb{E} \left[C_{i,I-i} \prod_{j=I-i}^{J-1} Y_{i,j} \middle| \mathcal{D}_I \right] \\ &= C_{i,I-i} \prod_{j=I-i}^{J-1} \mathbb{E}[Y_{i,j} \mid \mathcal{D}_I] \\ &= C_{i,I-i} \prod_{j=I-i}^{J-1} \mathbb{E}[\mathbb{E}(Y_{i,j} \mid \mathbf{F}, \mathcal{D}_I) \mid \mathcal{D}_I] \\ &= C_{i,I-i} \prod_{j=I-i}^{J-1} \mathbb{E}[F_j \mid \mathcal{D}_I] \\ &= C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes}. \end{aligned}$$

□

Notación 3.4.4.

La reserva estimada de Bayes correspondiente está dada por

$$R_i^{Bayes} = C_{i,J}^{Bayes} - C_{i,I-i}.$$

▽

El siguiente paso natural, es querer obtener una fórmula para el error cuadrático medio de R_i^{Bayes} , que es el mismo que el error cuadrático medio de $C_{i,J}^{Bayes}$. Gracias al hecho de que $\mathbb{E}[(X - a)^2] = \text{Var}(X) + (\mathbb{E}[X] - a)^2$ se tiene que

$$\begin{aligned}
 ECMP(C_{i,J}^{Bayes}) &= \mathbb{E} \left[(C_{i,J}^{Bayes} - C_{i,J})^2 \middle| \mathcal{D}_I \right] \\
 &= \mathbb{E} [Var(C_{i,J} | \mathbf{F}, \mathcal{D}_I)] + \mathbb{E} \left[\left(C_{i,J}^{Bayes} - \mathbb{E}[C_{i,J} | \mathbf{F}, \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right].
 \end{aligned}$$

En el modelo de Mack, el primer término correspondía al proceso de error y el segundo al error de estimación; pero aquí esto no es muy claro. Como \mathbf{F} es un vector aleatorio, el primer término es una especie de “promedio” del error y el segundo una especie del “promedio” de la estimación del error. Además,

$$\begin{aligned}
 Var[C_{i,J} | \mathbf{F}, \mathcal{D}_I] &= \mathbb{E}[Var(C_{i,J} | \mathbf{F}, \mathcal{D}_I, C_{i,J-1}) | \mathbf{F}, \mathcal{D}_I] + Var(\mathbb{E}[C_{i,J} | \mathbf{F}, \mathcal{D}_I, C_{i,J-1}] | \mathbf{F}, \mathcal{D}_I) \\
 &= C_{i,I-i} \sigma^2(F_{J-1}) \prod_{j=I-i}^{J-2} F_j + F_{J-1}^2 Var[C_{i,J-1} | \mathbf{F}, \mathcal{D}_I].
 \end{aligned}$$

Iterativamente se obtiene

$$Var(C_{i,J-1} | \mathbf{F}, \mathcal{D}_I) = C_{i,I-1} \sum_{k=I-i}^{J-1} F_{I-i} \cdots \cdots F_{k-1} \sigma^2(F_k) \cdot F_{k+1}^2 \cdots \cdots F_{J-1}^2. \quad (3.4.9)$$

Observación 3.4.4.

La expresión (3.4.9) es la misma que en el modelo de Mack, que no es relevante, ya que condicionalmente a \mathbf{F} , las suposiciones del modelo de Mack se cumplen. Con la ecuación (3.4.9) y dado que los F_j 's son independientes, dado \mathcal{D}_I , se obtiene el “promedio” del error como

$$\mathbb{E}[Var(C_{i,J} | \mathbf{F}, \mathcal{D}_I)] = C_{i,I-1} \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes} \cdot \mathbb{E}[\sigma^2(F_k) | \mathcal{D}_I] \cdot \prod_{n=k+1}^{J-1} \mathbb{E}[F_n^2 | \mathcal{D}_I] \right\}. \quad (3.4.10)$$

Además, el “promedio” del error de estimación de $C_{i,J}^{Bayes}$ está dado por

$$\begin{aligned}
 \mathbb{E} \left[\left(C_{i,J}^{Bayes} - \mathbb{E}[C_{i,J} | \mathbf{F}, \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right] &= C_{i,I-i}^2 \cdot \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} \mathbb{E}[F_j | \mathcal{D}_I] - \prod_{j=I-i}^{J-1} F_j \right)^2 \middle| \mathcal{D}_I \right] \\
 &= C_{i,I-i}^2 Var \left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right). \quad (3.4.11)
 \end{aligned}$$

De las ecuaciones (3.4.10) y (3.4.11) se sigue inmediatamente el siguiente resultado.

Teorema 3.4.3.

El error cuadrático medio condicional de la reserva de Bayes para el año de accidente i está dado por

$$\begin{aligned} ECMP \left(R_i^{Bayes} \right) &= \mathbb{E} \left[\left(C_{i,J}^{Bayes} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\ &= C_{i,I-i} \Gamma_{I-i} + C_{I-i}^2 \Delta_{I-i}^B, \end{aligned}$$

donde

$$\begin{aligned} \Gamma_{I-i} &= \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes} \cdot \mathbb{E}[\sigma^2(F_k) \mid \mathcal{D}_I] \cdot \prod_{n=k+1}^{J-1} \mathbb{E}[F_n^2 \mid \mathcal{D}_I] \right\}, \quad (3.4.12) \\ \Delta_{I-i}^B &= Var \left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right). \end{aligned}$$

3.4.2. Credibilidad para el modelo Chain-Ladder

Ya se estableció que, en la construcción Bayesianas, el mejor estimador de la siniestralidad última es

$$C_{i,J}^{Bayes} = C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes}. \quad (3.4.13)$$

Sin embargo, para calcular F_j^{Bayes} se necesita conocer la distribución de los F_j^{Bayes} , así como la distribución condicional de los $C_{i,j}$'s dado \mathbf{F} . Estas distribuciones generalmente son desconocidas. Una de las ventajas de la teoría de credibilidad es que sólo se necesita conocer el primer y segundo momentos. Se está suponiendo que el primer y segundo momentos existen y son finitos para todas las variables aleatorias. Dado un portafolio de riesgos similares, se pueden estimar esos momentos de los datos observados.

Observación 3.4.5.

Cuando se utiliza el término “credibilidad para el modelo Chain-Ladder” se está haciendo referencia a que se reemplazará F_j^{Bayes} en la ecuación (3.4.13) por los estimadores F_j^{Cred} . ▽

Definición 3.4.3.

Con base en teoría de credibilidad, se define al predictor de siniestralidad última $C_{i,J}$ dado \mathcal{D}_I como

$$C_{i,J}^{(Cred)} := C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Cred}.$$

Observación 3.4.6.

Cuando se utiliza el superíndice “(Cred)” se le está llamando estimador basado en credibilidad y éste no es un estimador de credibilidad ya que por definición un estimador de credibilidad es una función lineal de las observaciones. Sin embargo, dada la estructura multiplicativa de la metodología Chain-Ladder, no tendrían sentido estimadores lineales de $C_{i,J}$. ∇

La reserva estimada correspondiente para este modelo es

$$R_i^{(Cred)} = C_{i,J}^{(Cred)} - C_{i,I-i}.$$

Los estimadores de credibilidad basados en alguna estadística \mathbf{X} son los mejores estimadores que son una función lineal de las entradas de \mathbf{X} . Para la estimación de F_j se basará en las observaciones $Y_{i,j}$ con $i \in \{0, \dots, I-j-1\}$.

Definición 3.4.4.

Se define el estimador de credibilidad como

$$F_j^{Cred} = \underset{\{\hat{F}_j : \hat{F}_j = a_0^{(j)} + \sum_{i=0}^{I-j-1} a_i^{(j)} Y_{i,j}\}}{\text{arg mín}} \mathbb{E} \left[\left(\hat{F}_j - F_j \right)^2 \middle| \mathcal{B}_j \right].$$

En otras palabras, F_j^{Cred} se define como el mejor estimador en el conjunto

$$\left\{ \hat{F}_j : \hat{F}_j = a_0^{(j)} + \sum_{i=0}^{I-j-1} a_i^{(j)} Y_{i,j} \right\},$$

que minimiza a $\mathbb{E} \left[\left(\hat{F}_j - F_j \right)^2 \middle| \mathcal{B}_j \right] = \mathbb{E} \left[\left(\hat{F}_j - F_j \right)^2 \middle| \mathcal{C}_j \right]$.

Teorema 3.4.4. (*Estimador de credibilidad*)

- (i) Los estimadores de credibilidad de los factores Chain-Ladder F_j (desconocidos) son

$$F_j^{(Cred)} = \alpha_j \widehat{F}_j + (1 - \alpha) f_j, \quad (3.4.14)$$

donde

$$\begin{aligned} \widehat{F}_j &= \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}, \\ \alpha_j &= \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]} + \frac{\sigma_j^2}{\tau_j^2}}, \\ f_j &= \mathbb{E}(F_j), \\ \sigma_j^2 &= \mathbb{E}[\sigma_j^2(F_j)], \\ \tau_j^2 &= \text{Var}(F_j). \end{aligned}$$

- (ii) El error cuadrático medio de $F_j^{(Cred)}$ es

$$\mathbb{E} \left[\left(F_j^{(Cred)} - F_j \right)^2 \middle| \mathcal{B}_j \right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} = (1 - \alpha_j) \tau_j^2.$$

Demostración:

Condicionado a \mathcal{B}_j , las variables aleatorias $Y_{i,j}$ ($i \in \{0, \dots, I - j - 1\}$) cumplen las suposiciones del modelo de Bühlmann & Straub (1970). De hecho, la ecuación (3.4.14) se conoce como el estimador de credibilidad de Bühlmann & Straub (1970). \square

La cantidad \widehat{F}_j es el estimado del factor de desarrollo \widehat{f}_j en el modelo Chain-Ladder clásico, i.e. $F_j^{(Cred)} = \alpha_j \widehat{F}_j + (1 - \alpha) f_j$ es un promedio de credibilidad ponderado entre el estimador Chain-Ladder clásico \widehat{F}_j y el valor esperado a priori f_j .

El estimador de credibilidad (3.4.14) depende de los parámetros estructurales f_j, σ_j^2 y τ_j^2 . Estos estimadores se pueden estimar a partir de los datos usando métodos de estimación estándar.

Para el error cuadrático medio se obtiene,

$$\begin{aligned}
 ECMP\left(R_i^{(Cred)}\right) &= ECMP\left(C_{i,J}^{(Cred)}\right) \\
 &= \mathbb{E}\left[\left(C_{i,J}^{(Cred)} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \\
 &= \mathbb{E}[Var[C_{i,J} | \mathbf{F}, \mathcal{D}_I] | \mathcal{D}_I] + \mathbb{E}\left[\left(C_{i,J}^{(Cred)} - \mathbb{E}[C_{i,J} | \mathbf{F}, \mathcal{D}_I]\right)^2 \middle| \mathcal{D}_I\right].
 \end{aligned}$$

El primer término, es el proceso de error “promedio” y se mantiene sin cambios, es el mismo como en el Teorema 3.4.3. El error de estimación “promedio” se obtiene como

$$\mathbb{E}\left[\left(C_{i,J}^{(Cred)} - \mathbb{E}[C_{i,J} | \mathbf{F}, \mathcal{D}_I]\right)^2 \middle| \mathcal{D}_I\right] = C_{i,I-i}^2 \mathbb{E}\left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right].$$

Por lo tanto,

$$ECMP\left(R_i^{(Cred)}\right) = C_{i,I-i} \Gamma_{I-i} + C_{i,I-i}^2 \Delta_{I-i}^C,$$

donde

$$\Gamma_{I-i} = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes} \cdot \mathbb{E}[\sigma^2(F_k) | \mathcal{D}_I] \cdot \prod_{n=k+1}^{J-1} \mathbb{E}[F_n^2 | \mathcal{D}_I] \right\}, \quad (3.4.15)$$

$$\Delta_{I-i}^C = \mathbb{E}\left[\left(\prod_{j=I-i}^{J-1} F_{i,J}^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right]. \quad (3.4.16)$$

Para encontrar un estimador del error cuadrático medio se hacen las siguientes aproximaciones en las ecuaciones (3.4.15) y (3.4.16).

$$\begin{aligned}
 F_j^{Bayes} &\simeq F_j^{Cred} \\
 \mathbb{E}\left[\left(F_j - F_j^{Bayes}\right)^2 \middle| \mathcal{D}_I\right] &\simeq \mathbb{E}\left[\left(F_j - F_j^{Cred}\right)^2 \middle| \mathcal{B}_I\right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}}.
 \end{aligned}$$

Entonces se obtiene

$$\begin{aligned}
 \mathbb{E}[F_j^2 | \mathcal{D}_I] &= \mathbb{E}\left[\left(F_j - F_j^{Bayes}\right)^2 \middle| \mathcal{D}_I\right]^2 + \left(F_j^{Bayes}\right)^2 \\
 &\simeq \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} + \left(F_j^{Cred}\right)^2,
 \end{aligned}$$

y también

$$\begin{aligned}
\mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j \right)^2 \middle| \mathcal{D}_I \right] &\simeq \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} F_j^{Bayes} - \prod_{j=I-i}^{J-1} F_j \right)^2 \middle| \mathcal{D}_I \right] \\
&= \text{Var} \left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right) \\
&= \prod_{j=I-i}^{J-1} \mathbb{E}[F_j^2 \mid \mathcal{D}_I] - \prod_{j=I-i}^{J-1} (\mathbb{E}[F_j \mid \mathcal{D}_I])^2 \\
&\simeq \prod_{j=I-i}^{J-1} \left(\alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} + (F_j^{Cred})^2 \right) - \prod_{j=I-i}^{J-1} (F_j^{Cred})^2.
\end{aligned}$$

Así, se tiene la siguiente aproximación para el error cuadrático medio de $R_i^{(Cred)}$.

Teorema 3.4.5.

El error cuadrático medio condicional de $R_i^{(Cred)}$ está dado por

$$ECMP \left(R_i^{(Cred)} \right) \simeq C_{i,I-i} \Gamma_{I-i}^* + C_{i,I-i}^2 \Delta_{I-i}^*,$$

donde

$$\Gamma_{I-i}^* = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Cred} \sigma_k^2 \prod_{n=k+1}^{J-1} \left((F_n^{Cred})^2 + \alpha_n \frac{\sigma_n^2}{S_n^{[I-n-1]}} \right) \right\} \quad (3.4.17)$$

$$\Delta_{I-i}^* = \prod_{j=I-i}^{J-1} \left((F_j^{Cred})^2 + \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} (F_j^{Cred})^2. \quad (3.4.18)$$

Observación 3.4.7.

Si se reemplazan en (3.4.17) y (3.4.18) los estimadores apropiados de σ_j^2 y τ_j^2 en α_j , se obtiene el estimador del error cuadrático medio de $R_i^{(Cred)}$. ∇

Denótese al total de la reserva de credibilidad como

$$R^{(Cred)} = \sum_{i=1}^I R_i^{(Cred)},$$

donde $R_i^{(Cred)} = 0$ para las i 's tales que $I - i \geq J$. Entonces,

$$\begin{aligned} ECMP(R^{(Cred)}) &= \mathbb{E} \left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\ &= \mathbb{E} \left[Var \left(\sum_i C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I \right) \middle| \mathcal{D}_I \right] \\ &\quad + \mathbb{E} \left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i \mathbb{E}[C_{i,J} \mid \mathbf{F}, \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right]. \end{aligned}$$

Por la independencia condicional de los años accidente se obtiene que

$$\begin{aligned} \mathbb{E} \left[Var \left(\sum_i C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I \right) \middle| \mathcal{D}_I \right] &= \sum_i \mathbb{E}[Var(C_{i,J} \mid \mathbf{F}, \mathcal{D}_I) \mid \mathcal{D}_I] \\ &= \sum_i C_{i,I-i} \Gamma_{I-i}, \end{aligned}$$

donde Γ_{I-i} está dada en la ecuación (3.4.12). Nótese que el proceso error promedio de $R^{(Cred)}$ es la suma de los errores de $R_i^{(Cred)}$.

El segundo sumando, es la estimación del error promedio, y se obtiene como

$$\begin{aligned} &\mathbb{E} \left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i \mathbb{E}[C_{i,J} \mid \mathbf{F}, \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right] \\ &= \sum_{i=0}^I \mathbb{E} \left[\left(C_{i,J}^{(Cred)} - \mathbb{E}[C_{i,J} \mid \mathbf{F}, \mathcal{D}_I] \right)^2 \middle| \mathcal{D}_I \right] \\ &+ 2 \sum_{i=0}^I \sum_{k=i+1}^I \mathbb{E} \left[\left(C_{i,J}^{(Cred)} - \mathbb{E}[C_{i,J} \mid \mathbf{F}, \mathcal{D}_I] \right) \left(C_{k,J}^{(Cred)} - \mathbb{E}[C_{k,J} \mid \mathbf{F}, \mathcal{D}_I] \right) \middle| \mathcal{D}_I \right]. \end{aligned}$$

De nuevo con la aproximación

$$\begin{aligned} F_j^{Bayes} &\simeq F_j^{Cred}, \\ \mathbb{E} \left[\left(F_j - F_j^{Bayes} \right)^2 \middle| \mathcal{D}_I \right] &\simeq \mathbb{E} \left[\left(F_j - F_j^{Cred} \right)^2 \middle| \mathcal{B}_I \right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}}, \end{aligned}$$

se obtiene el segundo término

$$\begin{aligned}
& \mathbb{E} \left[\left(C_{i,J}^{(Cred)} - \mathbb{E}[C_{i,J} | F, \mathcal{D}_I] \right) \left(C_{k,J}^{(Cred)} - \mathbb{E}[C_{k,J} | F, \mathcal{D}_I] \right) \middle| \mathcal{D}_I \right] \\
&= C_{i,I-i} C_{k,I-k} \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j \right) \left(\prod_{l=I-k}^{J-1} F_l^{Cred} - \prod_{l=I-k}^{J-1} F_l \right) \middle| \mathcal{D}_I \right] \\
&\simeq C_{i,I-i} C_{k,I-k} \mathbb{E} \left[\left(\prod_{j=I-i}^{J-1} F_j^{Bayes} - \prod_{j=I-i}^{J-1} F_j \right) \left(\prod_{l=I-k}^{J-1} F_l^{Bayes} - \prod_{l=I-k}^{J-1} F_l \right) \middle| \mathcal{D}_I \right] \\
&= C_{i,I-i} C_{k,I-k} Cov \left(\prod_{j=I-i}^{J-1} F_j, \prod_{l=I-k}^{J-1} F_l \middle| \mathcal{D}_I \right) \\
&= C_{i,I-i} C_{k,I-k} \prod_{j=I-k}^{I-i-1} F_j^{Bayes} Var \left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right) \\
&\simeq C_{i,I-i} C_{k,I-k}^{(Cred)} \Delta_{I-i}^*,
\end{aligned}$$

donde Δ_{I-i}^* está dado en (3.4.16) y se utilizó el hecho de que

$$\begin{aligned}
& Cov \left(\prod_{j=I-i}^{J-1} F_j, \prod_{l=I-k}^{J-1} F_l \middle| \mathcal{D}_I \right) \\
&= \mathbb{E} \left[Cov \left(\prod_{j=I-i}^{J-1} F_j, \prod_{l=I-k}^{J-1} F_l \middle| \mathcal{D}_I, F_{I-k}, \dots, F_{I-i-1} \right) \middle| \mathcal{D}_I \right] \\
&\quad + Cov \left(\mathbb{E} \left[\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I, F_{I-k}, \dots, F_{I-i-1} \right], \mathbb{E} \left[\prod_{j=I-k}^{J-1} F_j \middle| \mathcal{D}_I, F_{I-k}, \dots, F_{I-i-1} \right] \middle| \mathcal{D}_I \right) \\
&= \mathbb{E} \left[\prod_{l=I-k}^{I-i-1} F_l \cdot Var \left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right) \middle| \mathcal{D}_I \right] + 0.
\end{aligned}$$

De esta forma, se obtiene el siguiente resultado.

Proposición 3.4.1.

Se puede aproximar el error cuadrático medio condicional de $R^{(Cred)}$ como

$$ECMP\left(R^{(Cred)}\right) \simeq \sum_i ECMP\left(R_i^{(Cred)}\right) + 2 \sum_{i=0}^I \sum_{k=i+1}^I C_{i,I-i} C_{k,I-i}^{(Cred)},$$

donde $ECMP\left(R_i^{(Cred)}\right)$ y Δ_{i-i}^* están dados en el Teorema 3.4.5.

3.4.3. Credibilidad exacta para el modelo Chain-Ladder

El caso en el que el estimador de Bayes sea del “tipo de credibilidad” entonces se dirá que hay credibilidad exacta. En esta parte se considerará una clase de modelos Chain-Ladder en la que ocurre esta condición i.e. $F_j^{Bayes} = F_j^{Cred}$.

La suposición fundamental para esta clase de modelos es que, condicionalmente a \mathbf{F} y \mathcal{B}_j , las variables aleatorias $Y_{0,j}, \dots, Y_{I,j}$ son independientes con una distribución que pertenece a la familia de dispersión exponencial de un parámetro, y que la distribución a *priori* de F_j pertenece a la familia de conjugadas naturales a *priori*.

La familia de dispersión exponencial generalmente se parametriza por lo que se conoce como parámetro canónico. Por lo tanto, en vez de \mathbf{F} con realizaciones \mathbf{f} se considerará el vector Θ de parámetros canónicos Θ_j y realizaciones ϑ . A continuación se verá que los dos parámetros están relacionados entre sí.

Definición 3.4.5.

Se dice que una distribución es de dispersión exponencial, si se puede expresar como

$$dF(x) = \exp\left[\frac{x\vartheta - b(\vartheta)}{\varphi/\omega} + c(x, \varphi/\omega)\right] d\nu(x), x \in A \subset \mathbb{R}, \quad (3.4.19)$$

donde

$\nu(\cdot)$ es la medida de Lebesgue en \mathbb{R} o una medida de conteo,

$\varphi \in \mathbb{R}^+$ es el parámetro de dispersión,

$\omega \in \mathbb{R}^+$ es un ponderador,

$b(\vartheta)$ es una función con segunda derivada cuya primera derivada $b'(\vartheta)$ tiene una única inversa.

Proposición 3.4.2.

Si X tiene una función de distribución de distribución exponencial entonces

$$(i) \quad \mu_X = \mathbb{E}[X] = b'(\vartheta). \quad (3.4.20)$$

$$(ii) \quad \sigma_X^2 = \text{Var}(X) = \frac{\varphi}{\omega} b''(\vartheta).$$

Si se saca la inversa en la ecuación (3.4.20), se obtiene

$$\vartheta = (b')^{-1}(\mu_X) =: h(\mu_X),$$

donde a $h(\cdot)$ se le conoce como función canónica de *link*.

Observación 3.4.8.

La varianza también se puede expresar como una función de la media, i.e.

$$\text{Var}(X) = \frac{\varphi}{\omega} b''(h(\mu_X)) = \frac{\varphi}{\omega} V(\mu_X),$$

donde a $V(\cdot)$ se le conoce como función de varianza. ∇

Definición 3.4.6.

La familia de distribuciones que se definió en (3.4.3) se le conoce como familia exponencial de un parámetro y se denota por

$$\mathcal{F}_{\text{exp}} = \{\mathcal{F}_{\vartheta} : \vartheta \in M\},$$

donde M es el espacio de parámetros canónicos (conjunto de posibles valores de ϑ). Si se especifica la forma de $b(\cdot)$ y $c(\cdot, \cdot)$ entonces se denota por $\mathcal{F}_{\text{exp}}^{b,c}$.

Una vez que se dieron estas definiciones, se supondrá que, condicionalmente a \mathbf{F} y \mathcal{B}_j , las variables aleatorias $Y_{0,j}, \dots, Y_{I,j}$ son independientes con una distribución que pertenece a la familia exponencial de un parámetro $\mathcal{F}_{\text{exp}}^{b,c}$.

Los siguientes resultados son válidos bajo esta suposición general. Sin embargo, no todas las familias exponenciales son adecuados para el problema de *reserving* ya que las variables aleatorias $Y_{i,j}$ son no-negativas. Por lo tanto, sólo las distribuciones que tienen soporte en \mathbb{R}^+ son adecuadas para el modelo Chain-Ladder.

Una subclase de la familia exponencial que tiene esta propiedad es la que considera los modelos de Tweedie.

Definición 3.4.7.

Los modelos de Tweedie son una subclase de la familia exponencial de un parámetro con función de varianza dada por

$$V(\mu) = \mu^p,$$

Observación 3.4.9.

La familia de los modelos de Tweedie incluye las siguientes distribuciones.

- Para $p = 0$ se tiene la distribución Normal.
- Para $p = 1$ se tiene la distribución Poisson con sobredispersión.
- Para $p \in (1, 2)$ se tiene la distribución Poisson compuesta donde los montos de reclamaciones tienen distribución Gamma con parámetro $\gamma = \frac{2-p}{p-1}$.
- Para $p = 2$ se tiene la distribución Gamma.

Suposición 3.4.3. (*Familia exponencial y conjugadas a priori*)

1. Condicionado a $\Theta = \vartheta$ y \mathcal{B}_j , las variables aleatorias $Y_{o,j}, \dots, Y_{I,j}$ son independientes con distribución de la forma exponencial dada en la definición 3.4.3, con funciones $b(\cdot)$ y $c(\cdot, \cdot)$, parámetro de dispersión φ_j y ponderador $\omega_{i,j} = C_{i,j}$.
2. $\Theta_0, \dots, \Theta_{J-1}$ son independientes con densidades (con respecto a la medida de Lebesgue)

$$u_j(\vartheta) = \exp \left[\frac{f_j^{(0)}\vartheta - b(\vartheta)}{\eta_j^2} + d(f_j^{(0)}, \eta_j^2) \right], \quad (3.4.21)$$

donde $f_j^{(0)}$ y η_j^2 son hiperparámetros y $\exp [d(f_j^{(0)}, \eta_j^2)]$ es un factor de normalización.

A partir del inciso 1. de la Suposición 3.4.3 se deduce que

$$f_{\vartheta_j}(y_{i,j} | \mathcal{B}_j) = \exp \left\{ \frac{y_{i,j}\vartheta_j - b(\vartheta_j)}{\varphi_j/C_{i,j}} \right\} a(y_{i,j}, \varphi_j/C_{i,j}),$$

y por tanto

$$\begin{aligned} \mathbb{E}[Y_{i,j} | \Theta_j] &= b'(\Theta_j) = F_j, \\ \text{Var}(Y_{i,j} | \Theta_j, C_{i,j}) &= b''(\Theta_j) \frac{\varphi_j}{C_{i,j}} = V(F_j) \frac{\varphi_j}{C_{i,j}}. \end{aligned}$$

Por lo tanto, dado F_j , la Suposición 3.4.1 del modelo Chain-Ladder clásico de Mack se cumple. La diferencia con los supuestos de Mack, es que ahora se hacen suposiciones específicas sobre la distribución condicional y no sólo en la primer y segundo momento condicional. Por ejemplo, el modelo Tweedie con $p = 1$, supone que $Y_{i,j}$ tiene distribución condicional Poisson con sobre-dispersión. Sin embargo, esta hipótesis condicional no es comparable con el CL clásico.

Además, las distribuciones de $\Theta_0, \dots, \Theta_{J-1}$ con densidad de la forma (3.4.21) pertenecen a la familia U_{exp}^b de las distribuciones de la conjugada natural a *priori* de $\mathcal{F}_{\text{exp}}^{b,c}$ que está dada por

$$\begin{aligned} U_{\text{exp}}^b &= \{u_\gamma(\vartheta) : \gamma = (x_0, \eta^2) \in \mathbb{R} \times \mathbb{R}^+\}, \\ u_\gamma(\vartheta) &= \exp \left[\frac{x_0\vartheta - b(\vartheta)}{\eta^2} + d(x_0, \eta^2) \right], \vartheta \in M. \end{aligned}$$

Observación 3.4.10.

El modelo se formula en términos de los factores de desarrollo individuales de $Y_{i,j}$. Sin embargo, también se podría estructurar en términos de $C_{i,j}$. En ese caso, dado ϑ , se obtendría una serie de tiempo del modelo estudiado en Murphy (1994), Barnett-Zehnwirth (2000) y Buchwalder *et al.* (2006) ▽

Teorema 3.4.6.

Bajo la Suposición 3.4.3 y si la región M es tal que $u_j(\vartheta) = 0$ frontera de M entonces se cumple que

1.

$$\mathbb{E}[F_j] = f_j^{(0)}.$$

2.

$$F_j^{Bayes} = F_j^{Cred} = \alpha_j \widehat{F}_j + (1 - \alpha) f_j^{(0)},$$

donde

$$\widehat{F}_j = \frac{S^{[I-j-1]_j}}{S_j^{[I-j-1]} + \frac{\sigma_j^2}{\tau_j^2}},$$

$$\alpha_j = \frac{S^{[I-j-1]_j}}{S_j^{I-j-1} + \frac{\sigma_j^2}{\tau_j^2}},$$

$$\sigma_j^2 = \varphi \mathbb{E}[b''(\Theta_j)] = \varphi_j \mathbb{E}[V(F_j)],$$

$$\tau_j^2 = \text{Var}[b'(\Theta_j)] = \text{Var}(F_j).$$

Nótese que F_j^{Bayes} en el Teorema 3.4.6 coincide con la estimación del Teorema 3.4.4.

Teorema 3.4.7.

Bajo la Suposición 3.4.3 y si la región M es tal que $u_j(\vartheta) = u'_j(\vartheta) = 0$ frontera de M entonces se cumple que

1.
$$\mathbb{E} \left[\left(F_j^{Bayes} - F_j \right)^2 \middle| \mathcal{B}_j \right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} = (1 - \alpha_j) \tau_j^2. \quad (3.4.22)$$

2. El error cuadrático medio condicional de la reserva R_i^{Bayes} de año de accidente i está dado por

$$\begin{aligned} ECMP \left(R_i^{(Bayes)} \right) &= \mathbb{E} \left[\left(C_{i,I-i}^{(Cred)} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\ &\simeq \mathbb{E} \left[\left(C_{i,I-i}^{(Cred)} - C_{i,J} \right)^2 \middle| \mathcal{B}_{I-i} \right] \\ &= C_{i,I-i}^2 \Gamma_{I-i}^* + C_{i,I-i}^2 \Delta_{I-i}^*, \end{aligned} \quad (3.4.23)$$

donde

$$\begin{aligned} \Gamma_{I-i}^* &= \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Cred} \cdot \sigma_k^2 \prod_{n=k+1}^{J-1} \left(\left(F_n^{Cred} \right)^2 \right) + \alpha_n \frac{\sigma_n^2}{S_n^{[I-n-1]}} \right\}, \\ \Delta_{I-i}^* &= C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \left(\left(F_j^{(Cred)} \right)^2 + \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \left(F_j^{Cred} \right)^2 \right]. \end{aligned}$$

3. El error cuadrático medio de la reserva total $R^{Bayes} = \sum_i R_i^{Bayes}$ es

$$ECM(R^{(Bayes)}) \simeq \sum_i ECM \left(R_i^{(Bayes)} \right) + 2 \sum_{i=0}^I \sum_{k=i+1}^I C_{i,I-i} C_{k,I-i}^{(Bayes)} \Delta_{I-i} +$$

Demostración:

Ya que F_j^{Bayes} es un estimador de credibilidad se sigue del Teorema 3.4.4, que (3.4.22) se cumple, si todas las variables aleatorias tienen varianza finita. Entonces, a partir de (3.4.21)

se obtiene

$$\begin{aligned} u'(\vartheta) &= \frac{1}{\eta^2} \left(f_j^{(0)} - b'(\vartheta) \right) u_j(\vartheta), \\ u''(\vartheta) &= \frac{1}{\eta^4} \left(f_j^{(0)} - b'(\vartheta) \right)^2 u_j(\vartheta) + \frac{1}{\eta^2} b''(\vartheta) u_j(\vartheta). \end{aligned}$$

Como $u'_j(\vartheta)$ es 0 en la frontera de M , se tiene que

$$\begin{aligned} 0 &= \frac{1}{\eta^4} \int_M \left(f_j^{(0)} - b'(\vartheta) \right)^2 u_j(\vartheta) d\vartheta + \frac{1}{\eta^2} \int_M b''(\vartheta) u_j(\vartheta) d\vartheta \\ &= \frac{1}{\eta^4} \text{Var}(F_j) + \frac{1}{\eta^2} \varphi_j \mathbb{E}[V(F_j)]. \end{aligned}$$

Por lo tanto, F_j tiene varianza finita.

La demostración del inciso 2. es la misma que el Teorema 3.4.5 y la demostración del inciso 3. es la misma que la que se usó en la proposición 3.4.1. \square

Para los modelos de Tweedie si $p \geq 0$ la condición y el resultado del Teorema 3.4.7 se cumplen para $p = 0$ (distribución Normal) y para $1 < p < 2$ (Poisson compuesto). Para $p = 2$ (distribución Gamma) sólo se cumple para $\eta^2 < 1$. En este caso la conjugada a *priori* natural $u_j(\vartheta)$ es también una distribución Gamma y $\eta^2 < 1$ significa que el parámetro γ de esta distribución Gamma es menor que uno.

3.4.3.1. Relación con el modelo Chain-Ladder clásico

En esta sección se considera la misma familia exponencial, pero ahora se estudiará el caso en el que se tiene una distribución a *priori* no-informativa, i.e se tiene el caso limite en el que $\tau_j^2 \rightarrow \infty$.

Teorema 3.4.8.

Bajo la Suposición 3.4.3 y si $u_j(\vartheta) = 0$ en la frontera de M , si $\tau_j^2 \rightarrow \infty$, entonces

$$\begin{aligned} F_j^{Bayes} &= \hat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}} = \hat{f}_j, \\ R_i^{Bayes} &= R_i^{CL}. \end{aligned} \tag{3.4.24}$$

Demostración:

Los pesos de credibilidad $\alpha_j \rightarrow 1$ cuando $\tau_j^2 \rightarrow \infty$. El resultado se deduce inmediatamente del Teorema(3.4.6). \square

Los \hat{f}_j 's son los estimadores del modelo Chain-Ladder clásico. Por lo tanto, los pronósticos/estimaciones Chain-Ladder Bayesianos son los mismos que los pronósticos/estimaciones Chain-Ladder clásico y las reservas también son las mismas.

El siguiente resultado muestra como se puede estimar el error cuadrático medio de predicción en el caso límite de distribuciones *a priori* no informativas, que en este caso es el error cuadrático de predicción para el Chain-Ladder clásico. Así el siguiente resultado da otra visión de la estimación del error cuadrático medio Chain-Ladder y sugiere un estimador diferente a los encontrados en este parte del trabajo.

A partir del Teorema 3.4.7 y el Teorema 3.4.8 se deduce inmediatamente el siguiente resultado.

Teorema 3.4.9.

Bajo la Suposición 3.4.3 y si $u_j(\vartheta) = 0 = u'_j(\vartheta)$ en la frontera de M y $\tau_j^2 \rightarrow \infty$, entonces el error cuadrático medio condicional Chain-Ladder de la reserva R_i^{CL} se puede estimar como

$$\widehat{ECMC}(R_i^{CL}) = C_{i,I-i} \hat{\Gamma}_{I-i} + C_{I-i}^2 \hat{\Delta}_{I-i},$$

donde

$$\Gamma_{I-i}^* = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} \hat{f}_m \cdot \hat{\sigma}_k^2 \prod_{n=k+1}^{J-1} \left(\hat{f}_n^2 + \frac{\hat{\sigma}_n^2}{S_n^{[I-n-1]}} \right) \right\},$$

$$\Delta_{I-i}^* = \prod_{j=I-i}^{J-1} \left(\hat{f}_j^2 + \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \hat{f}_j^2,$$

con $\hat{\sigma}_j^2$ estimadores adecuados para σ_j^2 .

Para los modelos Tweedie, el resultado es válido para $p = 0$ y para $p \in (1, 2)$, pero no para $p = 2$ ya que η^2 tiene que ser menor que uno. En particular, en el caso $p \in (1, 2)$ (Poisson compuesto) parece más adecuado y un buen modelo condicional para el desarrollo de las reclamaciones Chain-Ladder.

Estos estimadores se deben comparar con los sugeridos por Mack (1993) y Buchwalder

et al. (2006). El estimador $\widehat{\Gamma}_{I-i}$ del “promedio” del error es un poco mayor gracias al término $\widehat{\sigma}_n^2/S_n^{[I-n-1]}$. El estimador $\widehat{\Delta}_{I-i}$ del proceso de error “promedio” es el mismo que el de remuestreo condicional de Buchwalder *et al.* (2006) pero diferente a Mack (1993).

Finalmente, a partir del Teorema 3.4.8, el Teorema 3.4.9 y la proposición 3.4.1 se obtiene el siguiente resultado.

Proposición 3.4.3.

El error cuadrático medio de la reserva total $R^{CL} = \sum_i R_i^{CL}$ se puede estimar como

$$\widehat{ECM}(R^{CL}) = \sum_i \widehat{ECM}(R_i^{CL}) + 2 \sum_{i=0}^I \sum_{k=i+1}^I C_{i,I-i} \widehat{C}_{k,I-i} \widehat{\Delta}_{I-i}$$

donde

$$\widehat{C}_{k,I-i} = C_{k,I-k} \prod_{j=I-k}^{I-i-1} \widehat{f}_j,$$

donde $\widehat{ECM}(R^{CL})$ y $\widehat{\Delta}_{I-i}$ están dados en el Teorema 3.4.9, \widehat{f}_j son factores de desarrollo Chain-Ladder dados en (3.4.24).

3.5. Metodología Bayesiana de cálculo aproximado (ABC)

Siguiendo con las ideas de las secciones anteriores de este capítulo, la intención de este capítulo es desarrollar “otro” modelo Bayesiano Chain-Ladder de distribución libre utilizando una metodología Bayesiana de cálculo aproximado (*Approximate Bayesian Computation* o ABC)). También se estimarán cuantiles de las correspondientes reservas y se le compararán con los estimados de la reserva que se obtiene desde las perspectivas clásicas (Capítulo 2) y de credibilidad (Sección 3.4). Este desarrollo utilizará teoría de cadenas de Markov Monte-Carlo (MCMC), el ABC y un *bootstrap* Bayesiano.

Esta metodología ABC se desarrolla en un contexto de “distribución libre” pues no se harán suposiciones parametrales, de aquí que no se pueda evaluar la verosimilitud puntualmente ni simular directamente de un modelo de verosimilitud. Con la ayuda del *bootstrap* se generarán muestras de verosimilitud poco manejables analíticamente sin hacer suposiciones distribucionales. A partir de estas muestras se obtendrá una distribución empírica en el modelo CL y la distribución predictiva de las reclamaciones pendientes (condicionado a las reclamaciones observadas). Finalmente se darán estimados predictivos Bayesianos, el Valor en Riesgo (VaR) y el error cuadrático medio de predicción.

Se utilizará la formulación de en series de tiempo del modelo CL y con éste se hará *bootstrap* de las reclamaciones. Una característica importante de este modelo es que proporciona una justificación de algoritmo “determinista” Chain-Ladder que se estudió en el Capítulo 2 de manera mecánica.

La metodología ABC se requiere ya que no se puede evaluar puntualmente la verosimilitud (pues no la hay) y “evita” dichas evaluaciones simulando de alguna otra verosimilitud. Finalmente, se utilizará un algoritmo MCMC es una distribución proveniente de un *bootstrap* (analíticamente poco manejable) para obtener muestras de la distribución posterior.

3.5.1. Modelo Chain-Ladder Bayesiano de distribución libre

Como en las secciones anteriores, se dará una suposición que especifique la versión del modelo CL que se utilizará.

Suposición 3.5.1.

1. Considerese los factores CL (aleatorios), $\mathbf{F} = (F_0, \dots, F_{J-1})$ y los parámetros de desviación estándar (también aleatorios) $\Xi = (\Xi_0, \dots, \Xi_{J-1})$. Se supondrá independencia entre todos estos parámetros, i.e. la densidad *a priori* de (\mathbf{F}, Ξ) está dada por

$$\pi(\mathbf{f}, \boldsymbol{\sigma}) = \prod_{j=0}^{J-1} \pi(f_j) \pi(\sigma_j),$$

donde $\pi(f_j)$ denota la densidad de F_j y $\pi(\sigma_j)$ denota la densidad de Ξ_j .

2. Dados $\mathbf{F} = \mathbf{f} = (f_0, \dots, f_{J-1})$ y $\Xi = \boldsymbol{\sigma} = (\sigma_0, \dots, \sigma_{J-1})$ se satisface que
- (a) Las reclamaciones acumuladas $C_{i,j}$ son independientes para años accidente i diferentes.
 - (b) Para cualesquiera i, j , las reclamaciones acumuladas satisfacen

$$C_{i,j+1} = f_j C_{i,j} + \sigma_j \sqrt{C_{i,j}} \epsilon_{i,j+1},$$

donde si se condiciona respecto a \mathcal{B}_0 , los residuales $\epsilon_{i,j}$ son independientes e idénticamente distribuidas tales que

$$\mathbb{E}[\epsilon_{i,j} | \mathcal{B}_0, \mathbf{F}, \Xi] = 0 \quad \text{y} \quad \text{Var}[\epsilon_{i,j} | \mathcal{B}_0, \mathbf{F}, \Xi] = 1$$

Además,

$$\mathbb{P}[C_{i,j} > 0 | \mathcal{B}_0, \mathbf{F}, \Xi] = 1.$$

Observación 3.5.1.

El inciso (b) del numeral 2 de la Suposición 3.5.1 sobre los residuales garantiza que las reclamaciones acumuladas $C_{i,j}$ son positivas $\mathbb{P}(\cdot | \mathcal{B}_0, \mathbf{F}, \Xi)$ -casi seguramente. ∇

Proposición 3.5.1.

Bajo la Suposición 3.5.1. Dado \mathcal{D}_I , las variables aleatorias $(F_0, \Xi_0) \dots (F_{J-1}, \Xi_{J-1})$ son independientes. Por tanto, la distribución *a posteriori*, para (\mathbf{F}, Ξ) dado \mathcal{D}_I , está dada por

$$\pi(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I) = \prod_{j=0}^{J-1} \pi(f_j, \sigma_j | \mathcal{D}_I).$$

Para hacer un análisis Bayesiano completo, se harán algunas suposiciones con respecto a las distribuciones *a priori* de (\mathbf{F}, Ξ) .

Suposición 3.5.2.

Además de la Suposición 3.5.1, supóngase que para cualquier $j \in \{0, \dots, J-1\}$, el modelo *a priori* está dado por

1. $F_j \sim \text{Gamma}(\alpha_j, \beta_j)$, donde $\text{Gamma}(\alpha_j, \beta_j)$ es una distribución Gamma tal que $\mathbb{E}(F_j) = \alpha_j \beta_j = f_j^{(CL)}$.
2. $\Xi_j^2 \sim \text{IG}(a_j, b_j)$, donde $\text{IG}(a_j, b_j)$ es la distribución gamma inversa tal que $\mathbb{E}(\Xi_j^2) = \frac{b_j}{a_j - 1} = \hat{\sigma}_j^{2(CL)}$.

La función de verosimilitud del modelo es poco manejable, i.e., no se puede escribir una densidad bajo este análisis en el modelo CL. Para el modelo Bayesiano sólo se supone una distribución *a priori* para los parámetros (\mathbf{F}, Ξ) , pero no para las reclamaciones acumuladas, $C_{i,j}$. A pesar de que se hace una suposición distribucional sobre los parámetros, el modelo se conoce como de "distribución libre porque no la supone sobre las reclamaciones acumuladas. En este sentido, como sólo se están haciendo suposiciones distribucionales con respecto a los parámetros, no se puede hacer una análisis tradicional con el Teorema de Bayes para calcular las distribuciones posteriores.

Observación 3.5.2.

Las *a priori*s que se seleccionan en la Suposición 3.5.2 tienen varianzas grandes; esto enfatiza las diferencias entre la especificación de las distribuciones *a priori* y hacer alguna hipótesis distribucional en la verosimilitud. Estas ideas pareciesen contradictorias, pero no lo son. ▽

Las *a priori*s que se seleccionaron garantizan cierta congruencia con los parámetros del modelo CL clásico. En particular, estas *a priori*s garantizan que los parámetros f_j, σ_j sean estrictamente positivos.

A continuación se recordarán estimadores clásicos y Bayesianos para la estructura CL, incluyendo los factores de desarrollo y los parámetros de varianza.

En el modelo clásico, los factores CL se denotarán por $\hat{f}_j^{(CL)}$ y los parámetros de la varianza se estiman como

$$\hat{\sigma}_j^{2(CL)} = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j} \left(\frac{C_{i,j+1}}{C_{i,j}} - \hat{f}_j^{(CL)} \right)^2.$$

Notése que sólo está bien definido para $j < I-1$. Sin embargo, en la literatura existe una gran discusión sobre la estimación de los parámetros de la cola, por lo que se tomará la estimación de Mack (1993) para el estimador del último parámetro de varianza

$$\hat{\sigma}_{J-1}^{2(CL)} := \min \left\{ \frac{\hat{\sigma}_{J-2}^{4(CL)}}{\hat{\sigma}_{J-3}^{2(CL)}}, \hat{\sigma}_{J-3}^{2(CL)}, \hat{\sigma}_{J-2}^{2(CL)} \right\}.$$

Desde un perspectiva Bayesiana se calcula la distribución posterior de los parámetros, dado \mathcal{D}_I , que se denota como $\pi(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I)$.

A partir del procedimiento *bootstrap* MCMC-ABC se obtendrán muestras de la distribución *a posteriori* del modelo CL Bayesiano, de esta forma se podrán considerar esas muestras para el estimador puntual CL.

Los dos estimadores más comunes en el análisis Bayesiano que corresponden al de moda posterior (MAP) y media posterior (MMSE) y están dados por

$$\left(\hat{f}_j^{(MAP)}, \hat{\sigma}_j^{(MAP)} \right) = \arg \max_{f_j, \sigma_j} (\pi(f_j, \sigma_j | \mathcal{D}_I)),$$

y

$$\begin{aligned} \hat{f}_j^{(MMSE)} &= \int f_j \pi(f_j | \mathcal{D}_I) df_j = \mathbb{E}[f_j | \mathcal{D}_I], \\ \hat{\sigma}_j^{(MMSE)} &= \int \sigma_j \pi(\sigma_j | \mathcal{D}_I) d\sigma_j = \mathbb{E}[\sigma_j | \mathcal{D}_I]. \end{aligned}$$

Observación 3.5.3.

Si f_j y σ_j no son independientes, entonces los estimadores MAP se obtendrán a través de la maximización de la distribución conjunta. Sin embargo, en la práctica generalmente

se trabaja con estimadores marginales. ▽

Además, para distribuciones a *priori* difusas que se seleccionaron se satisface que

$$\hat{f}_j^{(MMSE)} \approx \hat{f}_j^{(CL)}.$$

Por lo tanto, a partir de la Proposición 3.5.1 se obtiene la siguiente aproximación

$$\begin{aligned} \mathbb{E}[C_{i,J}|\mathcal{D}_I] &= \mathbb{E}[\mathbb{E}[C_{i,J}|\mathcal{D}_I, \mathbf{F}, \boldsymbol{\Xi}]|\mathcal{D}_I] = C_{i,I-i} \mathbb{E} \left[\prod_{j=I-i}^{J-1} f_j \middle| \mathcal{D}_I \right] \\ &= C_{i,I-i} \prod_{j=I-i}^{J-1} \mathbb{E}[f_j|\mathcal{D}_I] = C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j^{(MMSE)} \\ &\approx C_{i,I-i} \prod_{j=I-i}^{J-1} \hat{f}_j^{(CL)} = \hat{C}_{i,J}, \end{aligned} \quad (3.5.1)$$

donde la aproximación del último renglón cumple la igualdad si $\pi(f_j)$ tiende a infinito.

Este es el argumento necesario para justificar porque el modelo CL Bayesiano se puede usar para los predictores CL.

3.5.2. Distribución predictiva completa y Valor en Riesgo

Además de las muestras obtenidas para el modelo CL paramétrico, vía bootstrap MCMC-ABC, se obtendrá la distribución predictiva de las reclamaciones de dos formas. La primera es la distribución de predicción completa de las reclamaciones que se obtiene al integrar posterior asociada al modelo paramétrico Bayesiano DFCL que se estimó empíricamente, i.e.

$$\pi(\mathcal{D}_I^c|\mathcal{D}_I) = \int \int \pi(\mathcal{D}_I^c|\mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I) d\mathbf{f} d\boldsymbol{\sigma}.$$

La segunda aproximación utiliza uno de los estimadores puntuales Bayesianos para los parámetros tal como es el MMSE para obtener $\pi(\mathcal{D}_I^c|\hat{\mathbf{f}}^{(MMSE)}, \hat{\boldsymbol{\sigma}}^{(MMSE)})$. De forma alternativa, se puede considerar la versión Rao-Blackwellizada de la distribución predictiva Bayesiana de las reclamaciones

$$\pi(\mathcal{D}_I^c|\hat{\mathbf{f}}^{(MMSE)}, \mathcal{D}_I) = \int \pi(\mathcal{D}_I^c|\hat{\mathbf{f}}^{(MMSE)}, \boldsymbol{\sigma}) \pi(\boldsymbol{\sigma}|\hat{\mathbf{f}}^{(MMSE)}, \mathcal{D}_I) d\boldsymbol{\sigma}.$$

Estos resultado se pueden aplicar para estimar diferentes medidas de riesgo tales como el Valor en Riesgo, Valor en Riesgo Condicional, *Expected Shortfall*, etc. En particular, dado un nivel de confianza del 95 % se puede calcular el VaR correspondiente como

$$VaR_{0.95}(C_{i,J} - \mathbb{E}[C_{i,J}|\mathcal{D}_I]|\mathcal{D}_I) = \text{mín}\{x \in \mathbb{R} : \mathbb{P}[C_{i,J} - \mathbb{E}[C_{i,J}|\mathcal{D}_I] > x|\mathcal{D}_I] \leq 0.05\}.$$

3.5.3. *Bootstrap* y el error cuadrático medio de predicción

Supóngase que ya se calculó el predictor Bayesiano o el predictor CL dado en (3.5.1). Entonces, lo siguiente es determinar la desviación de $C_{i,J}$ alrededor de estos predictores. Si sólo se considera el segundo momento, entonces se quiere un estimador del error cuadrático medio de predicción (ECMP) como en los métodos anteriores; éste generalmente se puede estimar de manera analítica. Sin embargo, si se desea otra “medida de dispersión” como el Valor en Riesgo (VaR) éste sólo se puede determinar numéricamente.

Como ya se vio a lo largo de este trabajo, un método numérico tradicional para simular este tipo de medidas es el *bootstrap*. Para fijar ideas y dar la notación que se utilizará, se hará una breve revisión de los conceptos de *bootstrap* que ya se estudiaron, es decir, se presentará el algoritmo *bootstrap* en su versión frecuentista. Es decir, se supondrá que los factores Chain-Ladder, $\mathbf{F} = \mathbf{f}$, y los parámetros de la desviación estándar $\mathbf{\Xi} = \boldsymbol{\sigma}$ en la Suposición 3.5.1 son constantes desconocidas. Entonces, el *bootstrap* genera datos simulados, que se denotarán como D_I^* que permitirán analizar las variaciones de $\hat{f}^{(CL)}$ y $\hat{\sigma}^{2(CL)}$.

3.5.3.1. *Bootstrap* no-paramétrico condicional

El algoritmo recursivo para realizar el *bootstrap* no-paramétrico y generar muestras de siniestralidad está dado por los siguientes pasos:

1. Para $i + j \leq I$, $j > 0$, calcular los residuales estimados $\hat{\varepsilon}_{i,j}$, condicionado a los estimadores $\hat{f}_{0:j-1}^{(CL)}$, $\hat{\sigma}_{0:j-1}^{2(CL)}$ y los datos observados \mathcal{D}_I

$$\hat{\varepsilon}_{i,j} = \hat{\varepsilon}_{i,j}(\hat{f}_{j-1}^{(CL)}, \hat{\sigma}_{j-1}^{2(CL)}) = \frac{C_{i,j} - \hat{f}_{j-1}^{(CL)} C_{i,j-1}}{\hat{\sigma}_{j-1}^{2(CL)} \sqrt{C_{i,j-1}}}.$$

2. Con estos residuales $\{\hat{\varepsilon}_{i,j}\}_{i+j \leq I}$ generar una distribución *bootstrap* empírica $\hat{F}_{\mathcal{D}_I}$
3. Simular de manera independiente $\tilde{\varepsilon}_{i,j}^* \sim \hat{F}_{\mathcal{D}_I}$.
4. Generar observaciones *bootstrap* (remuestreando)

$$C_{i,j}^* = \hat{f}_{j-1}^{(CL)} C_{i,j-1} + \hat{\sigma}_{j-1}^{(CL)} \sqrt{C_{i,j-1}} \tilde{\varepsilon}_{i,j}^*,$$

de donde se obtiene $\mathcal{D}_I^* = \mathcal{D}_I^*(\widehat{f}^{(CL)}, \widehat{\sigma}^{(CL)})$.

Nótese que para la versión *bootstrap* no-condicional se debería generar

$$C_{i,j}^* = \widehat{f}_{j-1}^{(CL)} C_{i,j-1}^* + \widehat{\sigma}^{(CL)} \sqrt{C_{i,j-1}^*} \widehat{\varepsilon}_{i,j}^*.$$

5. Calcular los parámetros Chain-Ladder “remuestreados”, \widehat{f}_j^* y $\widehat{\sigma}_j^{2*}$

$$\widehat{f}_j^* = \frac{\sum_{i=0}^{I-j-1} C_{i,j+1}^*}{\sum_{i=0}^{I-j-1} C_{i,j}^*} \quad \text{y} \quad \widehat{\sigma}_j^{2*} = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} C_{i,j}^* \left(\frac{C_{i,j+1}^*}{C_{i,j}^*} - \widehat{f}_j^* \right)^2.$$

6. Repetir los pasos 3-5 hasta obtener la distribución empírica de las muestras *bootstrap* $\widehat{C}_{i,J}^*$, $\widehat{f}_{i,J}^*$ y $\widehat{\sigma}_{i,J}^*$. Estos últimos son los estimadores de los parámetros desconocidos $C_{i,J}$, $f_{i,J}$, $\sigma_{i,J}$, respectivamente.

Este método *bootstrap* clásico no-paramétrico se puede ver como un método frecuentista; esto significa que no se considera la incertidumbre en los parámetros por la elección de una distribución a *priori* adecuada. Por tanto, se usará un estimador puntual para los parámetros desconocido y después las posibles fluctuaciones de este estimador puntual.

Sin embargo, este *bootstrap* no-paramétrico subestima la incertidumbre “verdadera”; esto viene del hecho de que los residuales estimados $\widehat{\varepsilon}_{i,j}$ que en general tienen varianza menor que 1, i.e. los residuales estimados no tienen la escala adecuada. Por tanto, los mecanismos frecuentistas usan diferentes escalamientos para “corregir” este hecho. En esta última parte de este capítulo, se usará un *bootstrap* Bayesiano con un algoritmo MCMC-ABC para obtener empíricamente la distribución del modelo CL Bayesiano.

3.5.3.2. Estimación *bootstrap* frecuentista

Se considerará el ECMP condicional dado por

$$\begin{aligned} ECMP_{C_{i,j}|\mathcal{D}_I}(\widehat{C}_{i,j}) &= \mathbb{E} \left[(C_{i,J} - \widehat{C}_{i,J})^2 | \mathcal{D}_I \right] \\ &= \text{Var}(C_{i,J} | \mathcal{D}_I) + \left(\mathbb{E}[C_{i,J} | \mathcal{D}_I] - \widehat{C}_{i,J} \right)^2. \end{aligned}$$

Ya se dijo que el primer término se conoce como proceso de varianza condicional y el segundo término se conoce como incertidumbre de estimación o error de estimación. En el modelo frecuentista (i.e., para $\mathbf{F} = \mathbf{f}$ y $\mathbf{\Xi} = \boldsymbol{\sigma}$) estos términos se calculan como

$$\text{Var}(C_{i,J} | \mathcal{D}_I) = \left(\mathbb{E}[C_{i,J} | C_{i,I-i}] \right)^2 \sum_{j=I-i}^{J-1} \frac{\sigma_j^2 / f_j^2}{\mathbb{E}[C_{i,j} | C_{i,I-i}]} = C_{i,I-i} \Gamma_{I-i}, \quad (3.5.2)$$

y

$$\left(\mathbb{E}[C_{i,J}|\mathcal{D}_I] - \widehat{C}_{i,J}\right)^2 = C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} f_j - \prod_{j=I-i}^{J-1} \widehat{f}_j^{(CL)} \right)^2 = C_{i,I-i}^2 \Delta_{I-i}. \quad (3.5.3)$$

Entonces, el proceso de varianza (3.5.2) se estima reemplazando los parámetros por sus estimadores i.e.

$$\widehat{Var}(C_{i,J}|\mathcal{D}_I) = \left(\widehat{C}_{i,J}\right)^2 \sum_{j=I-i}^{J-1} \frac{\widehat{\sigma}_j^{2(CL)}/\widehat{f}_j^{2(CL)}}{\widehat{C}_{i,j}} = C_{i,I-i} \widehat{\Gamma}_{I-i}^{freq}.$$

La estimación del parámetro error es más complicada por lo que es necesario usar el algoritmo *bootstrap*. Supóngase que se generan T muestras bootstrapping, i.e., $\widehat{f}_j^{*(1)}, \dots, \widehat{f}_j^{*(T)}$. Entonces el error de estimación se estima como la varianza de la muestra por el producto de las estimaciones *bootstrap* de los parámetros CL $\widehat{f}_j^{*(1)}, \dots, \widehat{f}_j^{*(T)}$, i.e., $C_{i,I-i}^2 \widehat{\Delta}_{I-i}^{freq}$.

3.5.3.3. Estimación Bayesiana del error cuadrático medio de predicción

En el modelo Bayesiano (i.e., considerando una distribución *a priori* para los parámetros desconocidos \mathbf{F} y Ξ) se puede descomponer ECMP condicional como

$$\begin{aligned} ECMP_{C_{i,J}|\mathcal{D}_I}(\mathbb{E}[C_{i,J}|\mathcal{D}_I]) &= Var(C_{i,J}|\mathcal{D}_I) \\ &= \mathbb{E}\left[Var(C_{i,J}|\mathcal{D}_I, f_j, \sigma_j)\right] + Var(\mathbb{E}[C_{i,J}|\mathcal{D}_I, f_j, \sigma_j]|\mathcal{D}_I). \end{aligned}$$

El proceso de varianza promedio está dado por

$$\begin{aligned} \mathbb{E}[Var(C_{i,J}|\mathcal{D}_I, f_j, \sigma_j)] &= C_{i,I-i} \sum_{j=I-i}^{J-1} \mathbb{E}\left[\prod_{m=I-i}^{j-1} f_m \sigma^2 \prod_{n=j+1}^{J-1} f_n^2 \middle| \mathcal{D}_I\right] \\ &= C_{i,I-i} \sum_{j=I-i}^{J-1} \prod_{m=I-i}^{j-1} \mathbb{E}[F_m|\mathcal{D}_I] \mathbb{E}[\Xi_j^2|\mathcal{D}_I] \prod_{n=j+1}^{J-1} \mathbb{E}[F_n^2|\mathcal{D}_I] = C_{i,I-i} \widehat{\Gamma}_{I-i}^{Bayes}, \end{aligned} \quad (3.5.4)$$

donde se usó que f_j, σ_j son independientes (a *posteriori*).

El error de estimación está dado por

$$Var\left(\mathbb{E}[C_{i,J}|\mathcal{D}_I, f_j, \sigma_j]|\mathcal{D}_I\right) = C_{i,I-i}^2 \left(\prod_{j=I-i}^{J-1} f_j \middle| \mathcal{D}_I \right) = C_{i,I-i}^2 \widehat{\Delta}_{I-i}^{Bayes}.$$

Usando de nuevo la independencia a *posteriori* de los parámetros, el último término se puede escribir como

$$C_{i,I-i}^2 \hat{\Delta}_{i,I-i}^{Bayes} = C_{i,I-i}^2 \left[\prod_{j=I-i}^{J-1} \mathbb{E}[f_j^2 | \mathcal{D}_I] - \prod_{j=I-i}^{J-1} \mathbb{E}[f_j | \mathcal{D}_I]^2 \right]. \quad (3.5.5)$$

Para calcular los dos términos en (3.5.4) y (3.5.5) se necesita calcular la distribución a *posteriori* de (\mathbf{F}, Ξ) , dado \mathcal{D}_I . Como no se tiene un modelo distribucional completo, no se puede escribir la función de verosimilitud, lo que permitiría soluciones analíticas o simulaciones MCMC. Por lo tanto, se usará una estructura ABC que permitiría la simulación en un entorno de distribución libre usando muestras *bootstrap* adecuadas y una métrica de distancia.

3.5.3.4. Estimador de credibilidad para el error cuadrático medio de predicción

Como se mencionó antes, también se puede considerar estimadores de credibilidad (que son los estimadores de varianza mínima que son combinación lineal de las observaciones) ya que sólo se tiene interés en los segundos momentos (i.e. el ECMP condicional). Para distribuciones a *priori* difusivas se obtiene la aproximación

$$\widehat{ECMP}_{C_{i,j} | \mathcal{D}_I} = C_{i,I-i} \hat{\Gamma}_{I-i}^{cred} + C_{i,I-i}^2 \hat{\Delta}_{I-i}^{cred},$$

donde

$$\begin{aligned} \hat{\Gamma}_{I-i}^{cred} &= \sum_{j=I-i}^{J-1} \left\{ \prod_{m=I-i}^{j-1} \hat{f}_m^{(CL)} \hat{\sigma}_j^{2(CL)} \prod_{n=j+1}^{J-1} \left(\left(\hat{f}_n^{(CL)} \right)^2 + \frac{\hat{\sigma}_n^{2(CL)}}{\sum_{i=0}^{I-n-1} C_{i,n}} \right) \right\}, \\ \hat{\Delta}_{I-i}^{cred} &= \prod_{j=I-i}^{J-1} \left(\left(\hat{f}_j^{(CL)} \right)^2 + \frac{\hat{\sigma}_j^{2(CL)}}{\sum_{i=0}^{I-j-1} C_{i,j}} \right) - \prod_{j=I-i}^{J-1} \left(\hat{f}_m^{(CL)} \right)^2. \end{aligned}$$

3.6. Metodología ABC (*Approximate Bayesian Computation*) para verosimilitudes complejas

Para estimar numéricamente los parámetros, predecir la siniestralidad y las medidas de incertidumbre como el ECMP, la metodología Bayesiana requiere simular datos de la distribución a *posteriori* de los parámetros del modelo Chain-Ladder de distribución libre. Sin embargo, obtener muestras $\left\{ \mathbf{f}^{(t)}, \boldsymbol{\sigma}_j^{2(t)} \right\}_{t=1}^T$ (que son realizaciones de un vector aleatorio con una distribución a *posteriori* $\pi(f_j, \sigma_j | \mathcal{D}_I)$) en el modelo CL es difícil ya que la verosimilitud no tiene una forma analítica cerrada y por tanto es difícil de manejar. Es por esto que las aproximaciones numéricas como los métodos MCMC no se pueden usar

directamente, ya que dichos algoritmos recursivos requieren de la evaluación repetida de la función de verosimilitud en cada etapa del muestreo de la cadena de Markov.

Una manera de evitar estas complicaciones es haciendo algún tipo de hipótesis distribucional en el momento de definir la función de verosimilitud. Por supuesto, tales suposiciones alteran el modelo CL. Una primera alternativa (la tradicional) es hacer una suposición Gaussiana para la verosimilitud. Un problema con esta suposición es que hace muy rígida la *skewness* a considerar.

Sin embargo, lo novedoso que tiene este método es que no se hacen este tipo de suposiciones, sólo se trabajará con el modelo de distribución libre utilizando una metodología de cálculo aproximado Bayesiano (ABC) para “facilitar” la simulación en la distribución a *posteriori* compleja.

Hay una dificultad en el modelo CL que no se presenta generalmente cuando se trabaja con una metodología ABC. En una metodología ABC no se pueden hacer evaluaciones puntuales en la verosimilitud, pero, condicionado a los parámetros, sí se puede simular fácilmente. Sin embargo, en el modelo CL sólo se tienen condiciones con respecto a los dos primeros momentos. Es decir, no se puede evaluar la verosimilitud directamente ni se pueden obtener simulaciones fácilmente de la verosimilitud. Es por esto que se le asociará un *bootstrap* Bayesiano a la metodología ABC.

En esta sección se dará una descripción del modelo ABC que sirve para trabajar con casos donde la verosimilitud es difícil de obtener y hacer evaluaciones puntuales tampoco es posible.

El método ABC considera una distribución a *posteriori* compleja, que se denota por $\pi(f_j, \sigma_j | \mathcal{D}_I)$ a través de un modelo aumentado

$$\pi(\mathbf{f}, \boldsymbol{\sigma}, \mathcal{D}_I^*, \mathcal{D}_I) = \pi(\mathcal{D}_I | \mathcal{D}_I^*, \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}),$$

donde \mathcal{D}_I^* es un vector auxiliar en el mismo espacio que \mathcal{D}_I . En este modelo Bayesiano aumentado, la función $\pi(\mathcal{D}_I | \mathcal{D}_I^*, \mathbf{f}, \boldsymbol{\sigma})$ es una ponderación de la posterior compleja. En general, se supondrá que $\pi(\mathcal{D}_I | \mathcal{D}_I^*, \mathbf{f}, \boldsymbol{\sigma}) = g(\mathcal{D}_I | \mathcal{D}_I^*)$ (i.e. es una suposición de modelo jerárquico).

El mecanismo que permite que el modelo ABC pueda evitar la evaluación en la verosimilitud compleja, consiste en reemplazar esta evaluación con datos simulados de la verosimilitud. Es decir, dada una realización de los parámetros del modelo, se genera un conjunto de datos sintéticos \mathcal{D}_I^* y se compara con los datos originales. La idea fundamental para hacer dichas simulaciones que usará el ABC, es generar simulaciones del modelo CL a través de un *bootstrap*.

Entonces, se obtendrá una estadística “resumen”, $S(\mathcal{D}_I^*)$, a partir de esta información y se comparará con una estadística “resumen”, $S(\mathcal{D}_I)$, de los datos observados a partir de una distancia $\rho(S(\mathcal{D}_I^*), S(\mathcal{D}_I))$. Finalmente, se asignará un peso a esos parámetros de acuerdo a la función de ponderación $g(\mathcal{D}_I|\mathcal{D}_I^*)$ que podría tener más peso cuando la distancia entre $S(\mathcal{D}_I^*)$ y $S(\mathcal{D}_I)$ sea pequeña.

Por ejemplo, supóngase que la ponderación está dada por

$$g(\mathcal{D}_I|\mathcal{D}_I^*) \propto \begin{cases} 1 & \text{si } \rho(S(\mathcal{D}_I^*), S(\mathcal{D}_I)) \leq \epsilon, \\ 0 & \text{en otro caso,} \end{cases}$$

entonces se está “premiando a la estadística “resumen”, $S(\mathcal{D}_I^*)$ con respecto a la estadística “resumen”, $S(\mathcal{D}_I)$, con una tolerancia de ϵ si se mide con la métrica ρ .

Por tanto, en un contexto ABC, una aproximación para la distribución marginal posterior (objetivo), $\pi(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I)$ está dada por

$$\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I, \epsilon) \propto \int \pi(\mathcal{D}_I|\mathcal{D}_I^*)\pi(\mathcal{D}_I^*|\mathbf{f}, \boldsymbol{\sigma})\pi(\mathbf{f}, \boldsymbol{\sigma})d\mathcal{D}_I^*.$$

Para formular un algoritmo MCMC se supondrá que la distribución estacionaria de la cadena de Markov está dada por $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}, \mathcal{D}_I^*|\mathcal{D}_I, \epsilon)$ y la correspondiente distribución objetivo para la distribución marginal $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I, \epsilon)$ se obtiene integrando numéricamente.

Nótese que la distribución marginal de predicción, $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I, \epsilon)$ converge a $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I)$ si $\epsilon \rightarrow 0$, $S(\mathcal{D}_I)$ sea una estadística suficiente y que la función de ponderación converja puntualmente a $S(\mathcal{D}_I)$. De esta manera, se obtendría $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I)$, la posterior “verdadera” que es analíticamente poco manejable.

3.6.1. Justificación del algoritmo MCMC-ABC

Dadas las observaciones \mathcal{D}_I , se quiere obtener muestras de $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma}|\mathcal{D}_I)$ con función de verosimilitud compleja. Supóngase que $S(\mathcal{D}_I)$ son los datos originales o un “resumen” de los mismos, por ejemplo, una estadística suficiente para el modelo del que se está suponiendo que \mathcal{D}_I es una realización. Así, dados los parámetros $(\mathbf{f}, \boldsymbol{\sigma})$ se pueden generar, a partir del modelo CL y *bootstrap*, un conjunto de datos simulados \mathcal{D}_I^* .

Considérese la función $g(\mathcal{D}_I^*, \mathcal{D}_I) = \mathbb{I}_{\{\rho(S(\mathcal{D}_I^*), S(\mathcal{D}_I)) < \epsilon\}}$ para un nivel de tolerancia $\epsilon > 0$ y una métrica $\rho(\cdot, \cdot)$ dados.

Entonces, el modelo Bayesiano CL se puede establecer a través de

$$\begin{aligned}
 \pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I, \epsilon) &= \frac{\int g(\mathcal{D}_I | \mathcal{D}_I^*) \pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}) d\mathcal{D}_I^*}{\int g(\mathcal{D}_I | \mathcal{D}_I^*) \pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}) d\mathcal{D}_I^* d\mathbf{f} d\boldsymbol{\sigma}} \\
 &= \frac{\pi(\mathbf{f}, \boldsymbol{\sigma}) \mathbb{E}[g(\mathcal{D}_I | \mathcal{D}_I^*) | \mathbf{f}, \boldsymbol{\sigma}]}{\mathbb{E}[g(\mathcal{D}_I | \mathcal{D}_I^*)]}. \tag{3.6.1}
 \end{aligned}$$

El siguiente paso consiste en aproximar el numerador de la expresión (3.6.1) a través de la distribución empírica, i.e.

$$\pi(\mathbf{f}, \boldsymbol{\sigma}) \mathbb{E}[g(\mathcal{D}_I | \mathcal{D}_I^*) | \mathbf{f}, \boldsymbol{\sigma}] \approx \pi(\mathbf{f}, \boldsymbol{\sigma}) \frac{1}{L} \sum_{l=1}^L g\left(\mathcal{D}_I | \mathcal{D}_I^{*,(l)}(\mathbf{f}, \boldsymbol{\sigma})\right),$$

donde, $\mathcal{D}_I^{*,(l)}(\mathbf{f}, \boldsymbol{\sigma})$ son simulaciones independientes de $\pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma})$.

Observación 3.6.1.

Se necesita considerar la expresión $\mathbb{E}[g(X|y)]$ en el denominador. En general, esta expresión no se puede calcular analíticamente; sin embargo, dado que se utilizó un algoritmo MCMC, el denominador es cancelado en alguna etapa de este desarrollo. ∇

Por lo tanto, para obtener simulaciones de $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I, \epsilon)$ se tiene que usar

$$\begin{aligned}
 \pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I, \epsilon) &\approx \frac{\int g(\mathcal{D}_I | \mathcal{D}_I^*) \pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}) d\mathcal{D}_I^*}{\int g(\mathcal{D}_I | \mathcal{D}_I^*) \pi(\mathcal{D}_I^* | \mathbf{f}, \boldsymbol{\sigma}) \pi(\mathbf{f}, \boldsymbol{\sigma}) d\mathcal{D}_I^* d\mathbf{f} d\boldsymbol{\sigma}} \\
 &\propto \pi(\mathbf{f}, \boldsymbol{\sigma}) \mathbb{E}[g(\mathcal{D}_I | \mathcal{D}_I^*) | \mathbf{f}, \boldsymbol{\sigma}] \tag{3.6.2} \\
 &\approx \pi(\mathbf{f}) \pi(\boldsymbol{\sigma}) \frac{1}{L} \sum_{l=1}^L g\left(\mathcal{D}_I | \mathcal{D}_I^{*,(l)}(\mathbf{f}, \boldsymbol{\sigma})\right).
 \end{aligned}$$

Observación 3.6.2.

Generalmente, se toma $L = 1$ para reducir el proceso computacional aunque esto disminuye la probabilidad de converger a la distribución estacionaria y por supuesto, se pueden utilizar otras funciones de decisión $g(\cdot | \cdot)$ y otras métricas $\rho(\cdot, \cdot)$. ∇

Se puede resumir esta propuesta para el caso en el que se tenga un verosimilitud compleja/poco manejable de la siguiente manera:

1. Se utiliza una metodología ABC para trabajar con la verosimilitud y por tanto, también con la posterior poco manejable.
 2. Para usar la metodología ABC, se necesita simular muestras del modelo CL dadas realizaciones de los parámetros. Dichas observaciones se obtienen de un *bootstrap*.
 3. Se utiliza un algoritmo MCMC (que no requiere el cálculo de constantes de normalización) sobre la distribución objetivo $\pi_{ABC}(\mathbf{f}, \boldsymbol{\sigma} | \mathcal{D}_I, \epsilon)$.
-

Capítulo 4

Generalizaciones del modelo Chain-Ladder de Mack

4.1. Introducción

La reserva IBNR para un portafolio se puede calcular con base en el triángulo de pérdidas pagadas o el triángulo de pérdidas incurridas. El problema que surge es que las proyecciones que se obtienen en cada triángulo son diferentes e incluso para un año accidente el triángulo de pérdidas pagadas puede producir una peor estimación de la siniestralidad última pero existe la posibilidad de que el siguiente año la situación sea inversa con el triángulo de reclamaciones incurridas.

4.2. Modelo Munich Chain Ladder

4.2.1. Introducción al método Munich Chain Ladder

El método más común para obtener la reserva IBNR es el Chain-Ladder, ya sea con triángulos de pérdidas pagadas o con el de pérdidas incurridas, estimada independientemente. Entonces se realizan dos aplicaciones del algoritmo Chain-Ladder y se le conoce como Chain-Ladder separado (SCL, *Separate Chain-Ladder*).

En esta sección se analiza el problema de estimación del método Chain-Ladder y se describe una posible solución: el método Munich Chain-Ladder (MCL). Exactamente, el método muestra la correlación entre el triángulo de pérdidas pagadas e incurridas que aplicando el método Chain-Ladder por separado, se ignoraría. Por tal motivo, el método Munich Chain-Ladder toma ventaja de esto juntando las pérdidas pagadas e incurridas pasadas para obtener una proyección futura. Es decir, el método Munich Chain Ladder combina las

pérdidas pagadas (P) con las pérdidas incurridas (I) a partir de los *ratios* (P/I). Aquí el *ratio* (P/I) se usa para el cociente entre las pérdidas pagadas y pérdidas incurridas, i.e. la proporción de las pérdidas incurridas que se han pagado hasta cierto punto del tiempo.

Se empezará estudiando el desarrollo de los *ratios* (P/I) utilizando el método SCL.

4.2.1.1. El problema (P/I) en cálculos del Chain-Ladder separado

Primero se revisará cómo los *ratios* (P/I) ejercen influencia en el grado de concurrencia de las proyecciones pagadas e incurridas generalmente a partir del método SCL.

Considérese los dos cuadrángulos que se obtienen aplicando el método SCL a los triángulos pagados e incurridos de un portafolio arbitrario.

Notación 4.2.1.

Para $i \in \{0, \dots, I\}, j \in \{0, \dots, J\}$,

- $P_{i,j}$ serán las pérdidas pagadas totales del i -ésimo año de accidente después de j periodos de desarrollo.
- $I_{i,j}$ serán las pérdidas incurridas totales del i -ésimo año de accidente después de j periodos de desarrollo. ▽

Observación 4.2.1.

Se seguirá considerando $I = J$. ▽

Los valores $P_{i,j}, I_{i,j}$ están dados para $j \in \{0, \dots, I - i\}$ y se proyectan para $j \in \{I + 1 - i, \dots, J\}$. Es decir, los números $P_{i,j}, I_{i,j}$ son las entradas de la hipotenusa de los triángulos correspondientes, i.e., los totales acumulados por año de accidente.

También se considerará el *ratio* $(P/I)_{i,j}$

$$(P/I)_{i,j} := \frac{P_{i,j}}{I_{i,j}},$$

del i -ésimo año de accidente y periodo de desarrollo j y el promedio $(P/I)_j$ de todos los años accidente en el periodo j

$$(P/I)_j := \frac{\sum_{i=0}^I P_{i,j}}{\sum_{i=0}^I I_{i,j}} = \frac{1}{\sum_{i=0}^I I_{i,j}} \sum_{i=0}^I I_{i,j} (P/I)_{i,j},$$

que es el promedio de los *ratios* $(P/I)_{i,j}$ en el periodo j ponderado con cantidades incurridas.

Además,

- f_j^P , es el factor de desarrollo promedio para pérdidas pagadas en el periodo de desarrollo j al $j + 1$ del método Chain-Ladder,

$$f_j^P = \frac{\sum_{i=0}^{I-j-1} P_{i,j+1}}{\sum_{i=0}^{I-j-1} P_{i,j}},$$

- f_j^I , es el factor de desarrollo promedio para pérdidas incurridas en el periodo de desarrollo j al $j + 1$ del método Chain-Ladder,

$$f_j^I = \frac{\sum_{i=0}^{I-j-1} I_{i,j+1}}{\sum_{i=0}^{I-j-1} I_{i,j}}.$$

Para $j \geq I - i$, las cantidades proyectadas $P_{i,j+1}$, $I_{i,j+1}$ son

$$P_{i,j+1} = P_{i,j} \cdot f_j^P \quad \text{y} \quad I_{i,j+1} = I_{i,j} \cdot f_j^I.$$

Entonces para $j > I - i$, los *ratios* del “futuro” $(P/I)_{i,j}$ son

$$(P/I)_{i,j} = \frac{P_{i,j}}{I_{i,j}} = \frac{P_{i,I-i} \cdot f_{I-i}^P \cdots f_{j-1}^P}{I_{i,I-i} \cdot f_{I-i}^I \cdots f_{j-1}^I}. \quad (4.2.1)$$

Para los factores de desarrollo de las pérdidas pagadas, se obtiene la ecuación

$$\begin{aligned} f_j^P \sum_{i=0}^I P_{i,j} &= f_j^P \left(\sum_{i=0}^{I-j-1} P_{i,j} + \sum_{i=I-j}^I P_{i,j} \right) \\ &= \frac{\sum_{i=0}^{I-j-1} P_{i,j+1}}{\sum_{i=0}^{I-j-1} P_{i,j}} \sum_{i=0}^{I-j-1} P_{i,j} + \sum_{i=I-j}^I f_j^P \cdot P_{i,j} \\ &= \sum_{i=0}^{I-j-1} P_{i,j+1} + \sum_{i=I-j}^I P_{i,j+1} = \sum_{i=0}^I P_{i,j+1}. \end{aligned}$$

A partir de esto y de las correspondientes expresiones de los factores de desarrollo para los incurridos se obtienen las siguientes expresiones

$$f_j^P = \frac{\sum_{i=0}^I P_{i,j+1}}{\sum_{i=0}^I P_{i,j}} \quad \text{y} \quad f_j^I = \frac{\sum_{i=0}^I I_{i,j+1}}{\sum_{i=0}^I I_{i,j}}.$$

Sustituyendo en la ecuación (4.2.1) para los *ratios* futuros se obtiene

$$\begin{aligned} (P/I)_{i,J} &= \frac{P_{i,I-i} \frac{\sum_{i=0}^I P_{i,I-i+1}}{\sum_{i=0}^I P_{i,I-i}} \dots \frac{\sum_{i=0}^I P_{i,J}}{\sum_{i=0}^I P_{i,J-1}}}{I_{i,I-i} \frac{\sum_{i=0}^I I_{i,I-i+1}}{\sum_{i=0}^I I_{i,I-i}} \dots \frac{\sum_{i=0}^I I_{i,J}}{\sum_{i=0}^I I_{i,J-1}}} \\ &= \frac{P_{i,I-i} \frac{\sum_{i=0}^I P_{i,J}}{\sum_{i=0}^I P_{i,I-i}}}{I_{i,I-i} \frac{\sum_{i=0}^I I_{i,J}}{\sum_{i=0}^I I_{i,I-i}}}. \end{aligned}$$

Reacomodando los términos, se obtiene

$$\frac{(P/I)_{i,J}}{(P/I)_J} = \frac{(P/I)_{i,I-i}}{(P/I)_{I-i}}. \quad (4.2.2)$$

Se puede interpretar esta ecuación de la siguiente manera:

Para cada año de accidente, el *ratio* de un valor proyectado $(P/I)_{i,J}$ para su correspondiente promedio es el mismo que el *ratio* del valor actual de $(P/I)_{i,I-i}$ para su correspondiente promedio. Por tanto, este *ratio* permanece constante en proyecciones basadas en los cálculos del Chain-Ladder separado. Entonces, para un año de accidente con un *ratio* actual $(P/I)_{i,I-i}$ por encima o por debajo del promedio también tendrá un *ratio* proyectado por encima o por debajo del promedio al final del cuadrángulo, i.e. en periodo de desarrollo J .

4.2.1.2. Correlación entre datos pagados e incurridos

La ecuación que se desarrolló en la sección anterior, muestra que el método en el que se desarrolla dos veces el Chain-Ladder, uno para las pérdidas pagadas y otro para las pérdidas incurridas tiene una “debilidad” sistemática y por lo tanto es común que genere proyecciones no realistas que contradicen la experiencia pasada.

En observaciones pasadas, un *ratio* (P/I) relativamente bajo fue seguido de factores de desarrollo relativamente altos para las pérdidas pagadas o factores de desarrollo relativamente bajos para las pérdidas incurridas o bien ambos. Para un *ratio* relativamente alto (P/I) , la situación es a la inversa.

Toda esta discusión sugiere que no se debe usar el mismo promedio de los factores de desarrollo Chain-Ladder para todos los años accidente, como en el caso del método SCL; se debe considerar la siguiente regla basada en la experiencia pasada:

Dependiendo de si el *ratio* (P/I) está por debajo o arriba del promedio, se debe usar un factor de desarrollo de las observaciones pagadas que esté por arriba o abajo de ese promedio y/o un factor de desarrollo de las incurridas que esté por arriba o abajo del promedio,

respectivamente.

El modelo Munich Chain-Ladder se basa en el modelo Chain-Ladder definido por Mack. Es un modelo conjunto para las pérdidas pagadas e incurridas que considera la dependencia de los factores de desarrollo en los *ratios* (P/I) (o (I/P)).

4.2.2. El método Chain-Ladder de Mack

Primero se dará alguna notación y las suposiciones del modelo Chain-Ladder, que ya se vio a lo largo de todo el trabajo, adaptando algunas hipótesis y la notación que haga más explícita la relación entre reclamaciones pagadas e incurridas.

Notación 4.2.2.

Sea $i \in \{0, \dots, I\}$ el número de años accidente y $j \in \{1, \dots, J\}$ el periodo de desarrollo. $P_i = (P_{i,j})$ es el proceso de reclamaciones pagadas en el año accidente i e $I_i = (I_{i,j})$ es el proceso de reclamaciones incurridas en el año i . Por tanto, la variable aleatoria $P_{i,j}$ denota las pérdidas pagadas para el año de accidente i después de j años de desarrollo e $I_{i,j}$ denota las pérdidas incurridas para el año de accidente i después de j años de desarrollo. Los procesos P_i e I_i describen el desarrollo de las pérdidas pagadas e incurridas del año de accidente i a través de los años de desarrollo.

Además, $\mathcal{P}_i(s) = \{P_{i,1}, \dots, P_{i,s}\}$ es el desarrollo de lo pagado del año accidente i hasta el final de año de desarrollo s . Análogamente, $\mathcal{I}_i(s) = \{I_{i,1}, \dots, I_{i,s}\}$ es el desarrollo de lo incurrido del año accidente i hasta el año de desarrollo s (incluyendo este último).

Suposición 4.2.1.

Las suposiciones del modelo Chain-Ladder para reclamaciones pagadas son:

1. Para $j \in J$, existe un factor de desarrollo $f_j^P > 0$ tal que para todo $i \in \{1, \dots, I\}$

$$\mathbb{E} \left[\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j) \right] = f_j^P. \quad (4.2.3)$$

2. Para $j \in J$, existe una constante de proporcionalidad $\sigma_j^P \geq 0$ tal que para todo $i \in \{1, \dots, I\}$

$$\text{Var} \left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j) \right) = \frac{(\sigma_j^P)^2}{P_{i,j}}. \quad (4.2.4)$$

3. Los años de accidente son independientes, i.e., los procesos

$$\{P_{1,j}\}, \dots, \{P_{I,j}\}, \quad (4.2.5)$$

son estocásticamente independientes.

Las suposiciones del modelo Chain-Ladder para reclamaciones incurridas son:

4. Para $j \in J$, existe un factor de desarrollo $f_j^I > 0$ tal que para todo $i \in \{1, \dots, I\}$

$$\mathbb{E} \left[\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j) \right] = f_j^I. \quad (4.2.6)$$

5. Para $j \in J$, existe una constante de proporcionalidad $\sigma_j^I \geq 0$ tal que para todo $i \in \{1, \dots, I\}$

$$\text{Var} \left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j) \right) = \frac{(\sigma_j^I)^2}{I_{i,j}}. \quad (4.2.7)$$

6. Los diferentes años de accidente son independientes, i.e., los procesos

$$\{I_{1,j}\}, \dots, \{I_{I,j}\}, \quad (4.2.8)$$

son estocásticamente independientes.

Las suposiciones anteriores se diseñan para proyectar su correspondiente triángulo pero

no hay relaciones entre los procesos pagados e incurridos. La esperanza condicional describe la mejor posibilidad de pronóstico de $P_{i,j+1}$ si sólo se conoce el proceso de pagos de este año de accidente hasta el periodo j .

Análogamente, para el proceso de los incurridos. Sin embargo, ambos triángulos quieren hacer proyecciones en términos de las esperanzas condicionales

$$\mathbb{E} \left[\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{B}_i(j) \right] \quad \text{y} \quad \mathbb{E} \left[\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{B}_i(j) \right],$$

donde $\mathcal{B}_i(j) = \{P_{i,1}, \dots, P_{i,j}, I_{i,1}, \dots, I_{i,j}\}$ es la información de desarrollo de los procesos hasta el final del periodo de desarrollo j .

4.2.3. El modelo Munich Chain-Ladder

Para desarrollar este método se tiene que suponer que los conjuntos

$$\{P_{1,j}, I_{1,j}\}, \dots, \{P_{I,j}, I_{I,j}\}, \quad (4.2.9)$$

son independientes.

Notación 4.2.3.

Sea,

$$Q_i := \frac{P_i}{I_i} = \left(\frac{P_{i,j}}{I_{i,j}} \right),$$

que denota al proceso (P/I) .

Definición 4.2.1. (*Desviación Estándar y Residual Condicional*)

Sea X una variable aleatoria y \mathcal{C} una condición.

(I) Se define a la desviación estándar condicional de X dado \mathcal{C} como

$$\sigma(X|\mathcal{C}) := \sqrt{\text{Var}(X|\mathcal{C})}.$$

(II) Se define al residual condicional de X dado \mathcal{C} como

$$\text{Res}(X|\mathcal{C}) = \frac{X - \mathbb{E}[X|\mathcal{C}]}{\sigma(X|\mathcal{C})}.$$

Observación 4.2.2.

$$\mathbb{E}[Res(X|\mathcal{C})|\mathcal{C}] = \mathbb{E}\left[\frac{X - \mathbb{E}[X|\mathcal{C}]}{\sigma(X|\mathcal{C})} \middle| \mathcal{C}\right] = \frac{\mathbb{E}[X|\mathcal{C}] - \mathbb{E}[X|\mathcal{C}]}{\sigma(X|\mathcal{C})} = 0,$$

y

$$Var(Res(X|\mathcal{C})|\mathcal{C}) = Var\left(\frac{X - \mathbb{E}[X|\mathcal{C}]}{\sigma(X|\mathcal{C})} \middle| \mathcal{C}\right) = \frac{1}{\sigma^2(X|\mathcal{C})} Var(X|\mathcal{C}) = 1.$$

▽

Se tiene interés en las esperanzas condicionales para los factores de desarrollo de lo pagado, lo ocurrido y sus residuales

$$Res\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right) \quad \text{y} \quad Res\left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j)\right),$$

donde los procesos de desarrollo son conocidos. Comparado con el método SCL, el método MCL fórmula sus suposiciones en términos de

$$\mathbb{E}\left[Res\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right) \middle| \mathcal{B}_i(j)\right] \quad \text{y} \quad \mathbb{E}\left[Res\left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j)\right) \middle| \mathcal{B}_i(j)\right].$$

Estas suposiciones, convertirán la hipótesis de dependencia lineal de estas esperanzas condicionales sobre los residuales de los *ratios* (I/P) o (P/I),

$$Res\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right) \quad \text{y} \quad Res(Q_{i,j} \middle| \mathcal{I}_i(j)),$$

en una ecuación matemática.

Suposición 4.2.2.

1. Existe una constante λ^P tal que para todo $j \in \{1, \dots, J\}$ e $i \in \{1, \dots, I\}$

$$\mathbb{E} \left[\text{Res} \left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j) \right) \middle| \mathcal{B}_i(j) \right] = \lambda^P \cdot \text{Res} \left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right), \quad (4.2.10)$$

o equivalentemente

$$\mathbb{E} \left[\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{B}_i(j) \right] = f_j^P + \lambda^P \frac{\sigma \left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j) \right)}{\sigma \left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right)} \left(Q_{i,j}^{-1} - \mathbb{E} \left[Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right] \right). \quad (4.2.11)$$

2. Existe una constante λ^I tal que para todo $j \in \{1, \dots, J\}$ e $i \in \{1, \dots, I\}$

$$\mathbb{E} \left[\text{Res} \left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j) \right) \middle| \mathcal{B}_i(j) \right] = \lambda^I \cdot \text{Res} \left(Q_{i,j} \middle| \mathcal{I}_i(j) \right), \quad (4.2.12)$$

o equivalentemente

$$\mathbb{E} \left[\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{B}_i(j) \right] = f_j^I + \lambda^I \frac{\sigma \left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j) \right)}{\sigma \left(Q_{i,j} \middle| \mathcal{I}_i(j) \right)} \left(Q_{i,j} - \mathbb{E} \left[Q_{i,j} \middle| \mathcal{I}_i(j) \right] \right). \quad (4.2.13)$$

Los parámetros λ^P y λ^I representan las pendientes de las rectas de regresión en las correspondientes gráficas de residuales y no dependen del año de desarrollo. Generalmente, $\lambda^P, \lambda^I \geq 0$.

Las ecuaciones (4.2.11) y (4.2.13) representan las esperanzas condicionales de los factores de desarrollo como la suma del factor de desarrollo Chain-Ladder y un término de corrección que es función de ambos tipos de datos.

El modelo MCL consiste de la suposición de independencia (4.2.9) para los años de accidente, los pre-requisitos del Chain Ladder usuales en las ecuaciones (4.2.3), (4.2.4), (4.2.6) y (4.2.7) para las pérdidas pagadas e incurridas y las suposiciones (4.2.10) y (4.2.12) que describen la dependencia de los factores de desarrollo de los pagados e incurridos a los *ratios* (P/I) y (I/P), respectivamente.

4.2.3.1. Análisis de las suposiciones del modelo Munich Chain-Ladder

A continuación se analizará con más detalle el modelo MCL suponiendo que $\lambda^P, \lambda^I \geq 0$.

La esperanza condicional $\mathbb{E} \left[\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{B}_i(s) \right]$, del factor de desarrollo para las reclamaciones incurridas que se utiliza para proyectar un año de accidente i en los periodos de desarrollo de j a $j + 1$, es una función lineal y creciente con respecto a $Q_{i,j}$. Esto muestra que las observaciones se expresan en términos de las suposiciones teóricas. La ecuación (4.2.13) representa la esperanza condicional como la suma del factor de desarrollo Chain-Ladder, f_j^I , y un término de corrección que es lineal en $Q_{i,j}$. Los tres factores del término de corrección se pueden explicar de la siguiente manera:

- El factor λ^I es el coeficiente de correlación de los residuales de los factores de desarrollo y de los residuales de los *ratios* (P/I) y además no depende de i ni de j . Por lo tanto, a λ^I se le conoce como factor de correlación o parámetro de correlación. El factor λ^I generalmente está entre 0 y 1 y “mide” la dependencia de los factores de desarrollo y los *ratios* (P/I). Si hay mucha dependencia de los datos entonces $\lambda^I \approx 0$ y los factores de desarrollo proyectan las estimaciones como en el método SCL.
- El factor de desviación estándar es el cociente de las desviaciones estándar condicionales de los factores de desarrollo incurridos y el *ratio* (P/I). Mientras más grande sea la desviación estándar del factor de desarrollo, es más probable que haya una desviación significativa del promedio y más grande será el factor de corrección. Mientras más pequeña sea la desviación estándar del *ratio* (P/I), será más raro que haya desviaciones significativas del promedio y de nuevo será más grande el término de corrección.
- El término lineal $Q_{i,s} - \mathbb{E}[Q_{i,s} | \mathcal{I}_i(s)]$ incorpora al *ratio* (P/I) en la proyección.

Los *ratios* (P/I) arriba del promedio tienen un efecto de corrección del factor de desarrollo hacia arriba y los *ratios* por debajo del promedio tienen un efecto de corrección del factor de desarrollo hacia abajo.

Observación 4.2.3.

Las afirmaciones que se hicieron con respecto a los factores de desarrollo incurridos y los *ratios* (P/I) se pueden hacer para los factores de desarrollo pagados y los *ratios* (I/P). ∇

Los parámetros de correlación λ^P y λ^I representa la relación entre los triángulos de las reclamaciones pagadas y de las reclamaciones incurridas. La magnitud de estos parámetros describe cómo son afectados por las pérdidas incurridas y pagadas con respecto a otro tipo de datos y por esto dichos parámetros son importantes para el tamaño de la proyección

última. El método de residuales considera todos los factores de desarrollo juntos, i.e., hace que se tengan suficientes datos y su estimación sea relativamente estable.

Para justificar la descripción de λ^P y λ^I como parámetros de correlación se calcula

$$\begin{aligned}
Cov\left(Q_{i,j}^{-1}, \frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right) &= Cov\left(Q_{i,j}^{-1}, \mathbb{E}\left[\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{B}_i(j)\right] \middle| \mathcal{P}_i(j)\right) \\
&= Cov\left(Q_{i,j}^{-1}, f_j^P + \lambda^P \frac{\sigma\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right)}{\sigma\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right)} \cdot \left(Q_{i,j}^{-1} - \mathbb{E}\left[Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right]\right) \middle| \mathcal{P}_i(j)\right) \\
&= \lambda^P \frac{\sigma\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right)}{\sigma\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right)} Var\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right) \\
&= \lambda^P \sigma\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right) \sigma\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right).
\end{aligned}$$

Por lo tanto, el coeficiente de correlación se calcula como

$$\begin{aligned}
\lambda^P &= \frac{Cov\left(Q_{i,j}^{-1}, \frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right)}{\sigma\left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j)\right) \sigma\left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j)\right)} \\
&= Corr\left(Q_{i,s}^{-1}, \frac{P_{i,t}}{P_{i,s}} \middle| \mathcal{P}_i(s)\right).
\end{aligned}$$

Siguiendo el mismo procedimiento para las incurridas, se puede probar que

$$\lambda^I = Corr\left(Q_{i,s}, \frac{I_{i,t}}{I_{i,s}} \middle| \mathcal{I}_i(s)\right).$$

Para los coeficientes de correlación condicionales,

$$Corr\left(Res\left(Q_{i,s}^{-1} \middle| \mathcal{P}_i(s)\right), Res\left(\frac{P_{i,t}}{P_{i,s}} \middle| \mathcal{P}_i(s)\right)\right) = \lambda^P,$$

y

$$Corr\left(Res\left(Q_{i,s} \middle| \mathcal{I}_i(s)\right), Res\left(\frac{I_{i,t}}{I_{i,s}} \middle| \mathcal{I}_i(s)\right)\right) = \lambda^I.$$

Así, los parámetros λ del modelo MCL y sus coeficiente de correlación correspondientes concuerdan.

4.2.4. Estimación de parámetros

Para calcular los residuales y los factores de desarrollo esperados, se tiene que estimar cada parámetro del modelo MCL.

4.2.4.1. Parámetros del Chain-Ladder por separado

Para $j \in \{0, \dots, J-1\}$, los factores de desarrollo f_j^P y f_j^I se estiman como los parámetros usuales del Chain-Ladder, i.e.

$$\widehat{f}_j^P = \frac{1}{\sum_{i=0}^{I-j-1} P_{i,j}} \sum_{i=0}^{I-j-1} P_{i,j} \frac{P_{i,j+1}}{P_{i,j}} = \frac{\sum_{i=0}^{I-j-1} P_{i,j+1}}{\sum_{i=0}^{I-j-1} P_{i,s}},$$

y

$$\widehat{f}_j^I = \frac{1}{\sum_{i=0}^{I-j-1} I_{i,j}} \sum_{i=0}^{I-j-1} I_{i,j} \frac{I_{i,j+1}}{I_{i,j}} = \frac{\sum_{i=0}^{I-j-1} I_{i,j+1}}{\sum_{i=0}^{I-j-1} I_{i,j}}.$$

Lo mismo para la estimación de los parámetros σ , i.e.

$$(\widehat{\sigma}_j^P)^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} P_{i,j} \left(\frac{P_{i,j+1}}{P_{i,j}} - \widehat{f}_j^P \right)^2$$

y

$$(\widehat{\sigma}_j^I)^2 = \frac{1}{I-j-1} \sum_{i=0}^{I-j-1} I_{i,j} \left(\frac{I_{i,j+1}}{I_{i,j}} - \widehat{f}_j^I \right)^2,$$

exactamente, $\widehat{\sigma}_j^P = \sqrt{(\widehat{\sigma}_j^P)^2}$ y $\widehat{\sigma}_j^I = \sqrt{(\widehat{\sigma}_j^I)^2}$.

4.2.4.2. Parámetros del Munich Chain-Ladder

Para calcular los residuales condicionales de los *ratios* (P/I) e (I/P) se necesitan estimadores para las esperanzas condicionales $\mathbb{E}[Q_{i,j}|\mathcal{I}_i(j)]$ y $\mathbb{E}[Q_{i,j}^{-1}|\mathcal{P}_i(j)]$ y las desviaciones estándar condicionales $\sigma(Q_{i,j}|\mathcal{I}_i(j))$ y $\sigma(Q_{i,j}^{-1}|\mathcal{P}_i(j))$.

Se supondrá que $\mathbb{E}[Q_{i,j}|\mathcal{I}_i(j)]$ es constante (análogo a la condición (4.2.6) del Chain-Ladder para incurridos). Además, se supondrá una dependencia de la varianza condicional del *ratio* (P/I) con el volumen incurrido (análogo a la condición (4.2.7)). Mientras mayor sea el volumen incurrido mayor será la varianza del *ratio* (P/I).

Si $j \in \{0, \dots, J\}$ estas suposiciones para las esperanzas condicionales y varianzas de los *ratios* (P/I) construyen un estimador para $\mathbb{E}[Q_{i,j}|\mathcal{I}_i(j)]$, definido como

$$\widehat{q}_j := \frac{1}{\sum_{k=0}^{I-j+1} I_{k,j}} \sum_{k=0}^{I-j+1} I_{k,j} \cdot Q_{k,j} = \frac{\sum_{k=0}^{I-j+1} P_{k,j}}{\sum_{k=0}^{I-j+1} I_{k,j}},$$

que es el mismo para todos los años de accidente.

Para $j \in \{1, \dots, J-1\}$, se tiene que el estimador de $\sigma(Q_{i,j}|\mathcal{I}_i(j))$ es

$$\hat{\sigma}(Q_{i,j}|\mathcal{I}_i(j)) = \frac{\hat{\rho}_j^I}{\sqrt{I_{i,j}}},$$

donde

$$(\hat{\rho}_j^I)^2 = \frac{1}{I-j} \sum_{k=0}^{I-j+1} I_{k,j} (Q_{k,j} - \hat{q}_j)^2.$$

Observación 4.2.4.

$\hat{\rho}_j^I$ es independiente del índice i del año de accidente. ∇

De manera análoga, las suposiciones para los *ratios* (P/I) se pueden hacer para las esperanzas condicionales y varianzas de los *ratios* (I/P), entonces el estimador para $\mathbb{E} \left[Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right]$ se obtiene como

$$\hat{q}_j^{-1} = \frac{1}{\sum_{k=0}^{I-j+1} P_{k,j}} \sum_{k=0}^{I-j+1} P_{k,j} \cdot Q_{k,j}^{-1} = \frac{\sum_{k=0}^{I-j+1} I_{k,j}}{\sum_{k=0}^{I-j+1} P_{k,j}},$$

y el estimador para $\sigma \left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right)$ es

$$\hat{\sigma} \left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right) = \frac{\hat{\rho}_j^P}{\sqrt{P_{i,j}}},$$

donde,

$$(\hat{\rho}_j^P)^2 = \frac{1}{I-j} \sum_{k=0}^{I-j+1} P_{k,j} \left(Q_{k,j}^{-1} - \hat{q}_j^{-1} \right)^2.$$

Sin embargo, las esperanzas condicionales $\mathbb{E}[Q_{i,j}|\mathcal{I}_i(j)]$ y $\mathbb{E}[Q_{i,j}^{-1}|\mathcal{P}_i(j)]$ son constantes si $Q_{i,j}$ también lo es, hecho que no se presenta, por esto no se pueden hacer estas suposiciones y habrá una estructura de dependencia más complicada de las esperanzas condicionales sobre $\mathcal{I}_i(s)$ y $\mathcal{P}_i(s)$, respectivamente.

Dada una cantidad suficiente de datos, se podrá estimar $\mathbb{E}[Q_{i,j}|\mathcal{I}_i(j)]$ promediando los *ratios* (P/I) de $Q_{k,j}$ sobre los años de accidente k para los cuales $\mathcal{I}_k(j)$ es igual o similar a $\mathcal{I}_i(s)$. Dentro del modelo Chain-Ladder, “similar” podría significar que el nivel de $I_k(j)$ es muy cercano a $I_i(s)$ o que los factores de desarrollo individuales $I_{k,j}/I_{k,j-1}$ son cercanos a $I_{i,j}/I_{i,j-1}$. Por lo menos, se omitiría algún año de accidente j donde $\mathcal{I}_k(j)$ es claramente diferente de $\mathcal{I}_i(s)$. Por supuesto, estos conceptos también aplican para la estimación de

$$\mathbb{E} \left[Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right].$$

Para las varianzas condicionales, la situación es similar. Dado un conjunto de datos suficientemente grande se tendrá una estructura de dependencia más complicada para las varianzas condicionales de $Q_{i,s}$ y $Q_{i,s}^{-1}$ sobre $\mathcal{I}_i(s)$ y $\mathcal{P}_i(s)$, respectivamente.

Sin embargo, por simplicidad se mantendrán los estimadores anteriores. Este hecho se justificará siempre y cuando $\mathbb{E}[Q_{i,s} | \mathcal{I}_i(s)]$ y $\mathbb{E}[Q_{i,s}^{-1} | \mathcal{P}_i(s)]$ sean funciones constantes de $\mathcal{I}_i(s)$ y $\mathcal{P}_i(s)$, respectivamente.

Ahora se estimaran los residuales condicionales y para simplificar la notación se define:

Notación 4.2.4.

- $\widehat{Res}(P_{i,j+1})$ será el estimador de $Res \left(\frac{P_{i,j+1}}{P_{i,j}} \middle| \mathcal{P}_i(j) \right)$.
- $\widehat{Res}(I_{i,j+1})$ será el estimador de $Res \left(\frac{I_{i,j+1}}{I_{i,j}} \middle| \mathcal{I}_i(j) \right)$.
- $\widehat{Res}(Q_{i,j}^{-1})$ será el estimador de $Res \left(Q_{i,j}^{-1} \middle| \mathcal{P}_i(j) \right)$.
- $\widehat{Res}(Q_{i,j})$ será el estimador de $Res(Q_{i,j} | \mathcal{I}_i(s))$.

▽

Los estimadores para los residuales condicionales están dados por

$$\widehat{Res}(P_{i,j+1}) = \frac{\frac{P_{i,j+1}}{P_{i,j}} - \widehat{f}_j^P}{\widehat{\sigma}_j^P} \sqrt{P_{i,j}}, \quad \widehat{Res}(I_{i,j+1}) = \frac{\frac{I_{i,j+1}}{I_{i,j}} - \widehat{f}_j^I}{\widehat{\sigma}_j^I} \sqrt{I_{i,j}},$$

y

$$\widehat{Res}(Q_{i,j}^{-1}) = \frac{Q_{i,j}^{-1} - \widehat{q}_j^{-1}}{\widehat{\rho}_j^P} \sqrt{P_{i,j}}, \quad \widehat{Res}(Q_{i,j}) = \frac{Q_{i,j} - \widehat{q}_j}{\widehat{\rho}_j^I} \sqrt{I_{i,j}}.$$

Para los parámetros de correlación λ^P y λ^I , se usarán los estimadores que minimizan las distancias promedio cuadráticas de las coordenadas y de los puntos en la gráfica de los residuales de la recta de regresión al origen con pendiente λ^P (o λ^I , respectivamente).

$$\lambda^P := \frac{1}{\sum_{i,j} \widehat{Res}^2(Q_{i,j}^{-1})} \sum_{i,j} \widehat{Res}^2(Q_{i,s}^{-1}) \frac{\widehat{Res}(P_{i,j+1})}{\widehat{Res}(Q_{i,j}^{-1})} = \frac{\sum_{i,j} \widehat{Res}(Q_{i,j}^{-1}) \widehat{Res}(P_{i,j+1})}{\sum_{i,j} \widehat{Res}^2(Q_{i,j}^{-1})}$$

y

$$\lambda^I = \frac{1}{\sum_{i,j} \widehat{Res}^2(Q_{i,j})} \sum_{i,j} \widehat{Res}^2(Q_{i,j}) \frac{\widehat{Res}(I_{i,j+1})}{\widehat{Res}(Q_{i,j})} = \frac{\sum_{i,j} \widehat{Res}(Q_{i,j}) \widehat{Res}(I_{i,j+1})}{\sum_{i,j} \widehat{Res}^2(Q_{i,j})}.$$

En las sumas anteriores el índice j corre de 0 hasta $I - 2$ y el índice i va de 0 hasta $I - j$.

Si se fija un periodo de desarrollo j en las fórmulas de estimación de λ^P y λ^I y sumando sobre los i 's se obtienen estimados anuales de los parámetros λ 's. Los parámetros λ 's para cada año de desarrollo no deben presentar tendencias que contradigan las hipótesis del modelo MCL.

Finalmente, de acuerdo a las ecuaciones (4.2.10) y (4.2.12) se obtienen las fórmulas recursivas, para $j \geq I - i + 1$

$$\widehat{P}_{i,j+1} = \widehat{P}_{i,j} \left[\widehat{f}_j^P + \widehat{\lambda}^P \frac{\widehat{\sigma}_j^P}{\widehat{\rho}_j^P} \left(\frac{\widehat{I}_{i,j}}{\widehat{P}_{i,j}} - \widehat{q}_j^{-1} \right) \right]$$

y

$$\widehat{I}_{i,j+1} = \widehat{I}_{i,j} \left[\widehat{f}_j^I + \widehat{\lambda}^I \frac{\widehat{\sigma}_j^I}{\widehat{\rho}_j^I} \left(\frac{\widehat{P}_{i,j}}{\widehat{I}_{i,j}} - \widehat{q}_j \right) \right],$$

con valores iniciales $\widehat{P}_{i,j} = P_{i,j}$ y $\widehat{I}_{i,j} = I_{i,j}$.

Ejemplo 4.2.1.

El siguiente ejemplo muestra el procedimiento para obtener la reserva IBNR con el método Munich Chain-Ladder, primero se mostrará los datos observados que son triángulos correspondientes a las reclamaciones incurridas y pagadas, respectivamente, para un portafolio de riesgos de incendio (cifras en miles).

```

Código de R
R> MCLincurrred
  dev
origin  1    2    3    4    5    6    7
  1  978 2104 2134 2144 2174 2182 2174
  2 1844 2552 2466 2480 2508 2454  NA
  3 2904 4354 4698 4600 4644  NA  NA
  4 3502 5958 6070 6142  NA  NA  NA
  5 2812 4882 4852  NA  NA  NA  NA
  6 2642 4406  NA  NA  NA  NA  NA
  7 5022  NA  NA  NA  NA  NA  NA

R> MCLpaid
  dev
origin  1    2    3    4    5    6    7

```

```

1  576 1804 1970 2024 2074 2102 2131
2  866 1948 2162 2232 2284 2348  NA
3 1412 3758 4252 4416 4494  NA  NA
4 2286 5292 5724 5850  NA  NA  NA
5 1868 3778 4648  NA  NA  NA  NA
6 1442 4010  NA  NA  NA  NA  NA
7 2044  NA  NA  NA  NA  NA  NA

```

Código de R

Ahora, basado en las suposiciones anteriores, con el modelo de Mack se obtendrá el estimador para σ , i.e.,

```

R> sigmaI<-MackChainLadder(MCLincurred)$sigma
[1] 9.7273990 2.5444838 1.0040570 0.1200991 0.8603340 0.1200991

R> sigmaP<-MackChainLadder(MCLpaid)$sigma
[1] 13.4559310 3.6656420 0.4819578 0.2100029 0.4787308 0.2100029

```

Código de R

La paquetería ChainLadder proporciona una forma de obtener el método Munich Chain-Ladder, pero necesita los dos triángulos mencionados anteriormente y el estimador σ de cada uno, entonces

```

R> MCL<-MunichChainLadder(MCLpaid,MCLincurred,est.sigmaP =sigmaP,est.sigmaI = sigmaI)

MunichChainLadder(Paid = MCLpaid, Incurred = MCLincurred, est.sigmaP = sigmaP,
  est.sigmaI = sigmaI)

  Latest Paid Latest Incurred Latest P/I Ratio Ult. Paid Ult. Incurred Ult. P/I Ratio
1      2,131          2,174      0.980      2,131          2,174      0.980
2      2,348          2,454      0.957      2,385          2,443      0.976
3      4,494          4,644      0.968      4,554          4,634      0.983
4      5,850          6,142      0.952      6,070          6,182      0.982
5      4,648          4,852      0.958      4,879          4,958      0.984
6      4,010          4,406      0.910      4,599          4,672      0.984
7      2,044          5,022      0.407      7,505          7,655      0.980

Totals
      Paid Incurred P/I Ratio
Latest: 25,525 29,694 0.86
Ultimate: 32,121 32,720 0.98

```

Código de R

En esta tabla se obtienen los *ratios* de la última diagonal de las observaciones de los triángulos (*Latest*), así como la estimación de la siniestralidad última con sus respectivos *ratios* (*Ultimate*), la ultima columna corresponde al total de siniestros pagados e incurridos y la estimación de los futuros, por tanto para calcular la reserva IBNR y RBNS respectivamente, basta con restar, en cada caso, las reclamaciones estimadas con las de la última

diagonal.

Ahora se mostrará la forma de obtener los estimadores de los parámetros y coeficientes de correlación, residuales.

```

Código de R
R> MCL$IncurredResiduals
[1] 1.605020358 -1.183723533 -0.846383581 0.299450714 0.458137992 0.082352480 NA -0.078980622
[9] -1.039118036 1.565493903 0.004806834 -0.680589085 NA NA 0.221585430 0.287221014
[17] -1.415119918 0.930505026 NA NA NA 1.131348046 0.096288253 -0.843077797
[25] NA NA NA NA NA 0.731893789 -0.681418728 NA NA
[33] NA NA NA NA NA NA NA NA NA
[41] NA NA NA NA NA NA NA NA NA
[49] NA

R> MCL$PaidResiduals
[1] 1.240062341 -0.409542454 0.627731906 -0.432520662 -1.330415661 0.971284236 NA -0.454493935
[9] -0.257500458 0.003519297 -0.984526495 1.660669736 NA NA -0.178096214 0.292551156
[17] 1.248117471 -1.151042283 NA NA NA 0.845586577 0.571652630 -0.978875177
[25] NA NA NA NA NA -0.723943106 0.689859681 NA NA
[33] NA NA NA NA NA NA NA NA NA
[41] NA NA NA NA NA NA NA NA NA
[49] NA
    
```

```

Código de R
R> MCL$QResiduals
[1] 0.30871615 -0.47335565 -0.43743334 1.24543896 1.22304681 0.11895559 NA 0.10271271 -1.13141939
[10] 0.24623239 0.79537557 -1.37205649 NA NA -0.10709430 -1.31687503 -0.80498376 0.62550343
[19] NA NA NA NA -0.03308516 -1.53672981 0.69308386 NA NA NA
[28] NA 0.13749654 -1.17743245 NA NA NA NA NA NA
[37] NA NA NA NA NA NA NA NA NA NA
[46] NA NA NA NA NA

R> MCL$QinverseResiduals
[1] -0.28866099 0.49565517 0.45015637 -1.10614730 -1.07679458 -0.11554182 NA -0.10001795 1.16766593
[10] -0.23897922 -0.76095980 1.40633581 NA NA 0.10639977 1.34252842 0.80771329 -0.61487164
[19] NA NA NA NA 0.03251491 1.54674415 -0.67544873 NA NA NA
[28] NA -0.13556094 1.18814416 NA NA NA NA NA NA NA
[37] NA NA NA NA NA NA NA NA NA NA
[46] NA NA NA NA NA
    
```

```

Código de R
R> MCL$lambdaP

Call:
lm(formula = PaidResiduals ~ QinverseResiduals + 0)

Coefficients:
QinverseResiduals
0.636

R> MCL$lambdaI

Call:
lm(formula = IncurredResiduals ~ QResiduals + 0)

Coefficients:
QResiduals
0.4362
    
```

Código de R

Código de R

```
R> MCL$q.f
[1] 0.5325822 0.8488621 0.9275964 0.9450735 0.9491744 0.9598792 0.9802208
R> MCL$qinverse.f
[1] 1.877644 1.178048 1.078055 1.058119 1.053547 1.041798 1.020178
```

Código de R

Código de R

```
R> MCL$rhoP.sigma
[1] 14.9430129 4.9899464 2.1665565 1.6186098 1.7910011 0.2359799 0.1966189

R> MCL$rhoI.sigma
[1] 5.7107795 3.8192859 1.9184007 1.4606629 1.6370399 0.2219648 0.2494577
```

Código de R

4.3. Modelo Double Chain Ladder

El método Chain-Ladder opera con datos de reclamaciones acumuladas, i.e. sumas de reclamaciones individuales pagadas o incurridas. Desde un punto de vista teórico, se interpretaría como una distribución Poisson compuesta. En esta sección se presentará un método -relacionado con el CL- que se puede formular como un modelo estadístico y que explícitamente supone que los datos tienen una distribución Poisson compuesta.

En el método CL clásico una vez calculadas las reservas RBNS e IBNR no se pueden dividir las reclamaciones estimadas dentro de ellas, aquí, se mostrará que con un enfoque de regresión, incluyendo datos de conteo, se puede hacer exactamente esto de una manera muy simple y concisa. Así, con este método se podrá describir el flujo de efectivo en la reserva RBNS. Esto es de gran importancia cuando las compañías aseguradoras tengan que cumplir con los requisitos de la actual regulación de Solvencia II.

El método que se estudiará en esta sección parte de la combinación de los datos de conteo incurridos con los pagados. Ambos conjuntos de datos se representan de forma usual en un triángulo. Este método generaliza el modelo de Verrall *et. al.* (2010) y Martínez-Miranda *et. al.* (2011), donde se supone que sólo podría ocurrir un pago por reclamación y este pago se modela con una constante de severidad promedio, de modo que permite cambiar la severidad promedio en dirección al año de suscripción en el triángulo de reclamaciones pagadas.

4.3.1. Suposiciones del modelo Double Chain-Ladder

Se supondrá que se tienen dos triángulos de reclamaciones, el de los pagos agregados y el de conteos incurridos, que se definen de la siguiente manera:

Definición 4.3.1. (*Conteos incurridos agregados*)

Se define al conjunto de conteos incurridos agregados como

$$\aleph_I := \{N_{i,j} : (i, j) \in \mathcal{I}\},$$

donde $N_{i,j}$ es el número total de reclamaciones incurridas en el año i que se reportaron en el año $i + j$, i.e., j periodos de retraso del año i , donde

$$\mathcal{I} := \{(i, j) : i = 1, \dots, I, j = 0, 1, \dots, J - 1; i + j \leq I\}.$$

Definición 4.3.2. (*Pagos Agregados*)

Se define al conjunto de pagos como

$$\Delta_I := \{X_{i,j} : (i, j) \in \mathcal{I}\},$$

donde $X_{i,j}$ son los pagos totales de las reclamaciones incurridas en el año i y con j periodo de retraso del año i .

4.3.1.1. La extensión del modelo de Verrall *et. al.* (2010)

El modelo de Verrall *et. al.* (2010) y Martínez-Miranda *et. al.* (2011) se plantearon considerando tres componentes estocásticos: el retraso de liquidación, los pagos individuales y conteos reportados.

Los triángulos de conteos y pagos (\aleph_I, Δ_I) se observan de datos reales pero el modelo de micro-estructura define un nuevo triángulo no observado,

$$\aleph_I^{pag} := \{N_{i,j}^{pag} : (i, j) \in \mathcal{I}\},$$

donde $N_{i,j}^{pag}$ es el número de pagos incurridos en el año i y pagados con j años de retraso. Aquí el retraso de pago (o retraso RBNS) es una componente estocástica que se modela considerando variables no observadas $N_{i,j,l}^{pag}$, donde $N_{i,j,l}^{pag}$ es el número de pagos futuros originados de las reclamaciones reportadas $N_{i,j}$ que se pagan con l periodos de retraso.

Si d es el máximo número de periodos de retraso, entonces

$$N_{i,j}^{pag} = \sum_{l=0}^{\min\{j,d\}} N_{i,j-l,l}^{pag}. \quad (4.3.1)$$

También se definen a $Y_{i,j}^{(k)}$ como los pagos individuales que surgen de $N_{i,j}^{pag}$, $k \in \{1, \dots, N_{i,j}^{pag}\}$, $(i, j) \in \mathcal{I}$. Utilizando estas componentes, se puede estimar la reserva RBNS.

Para la reserva IBNR se necesita modelar el retraso IBNR y esto se puede hacer aplicando el modelo CL al triángulo \aleph_I .

Suposición 4.3.1.

- (1) **El retraso RBNS.** Dado $N_{i,j}$, la distribución del número de reclamaciones pagadas es multinomial, i.e. para todo $(i, j) \in \mathcal{I}$

$$(N_{i,j,0}^{pag}, \dots, N_{i,j,d}^{pag}) | N_{i,j} \sim \text{Multinom}(N_{i,j}; p_0, \dots, p_d),$$

donde $\mathbf{P} = (p_0, \dots, p_d)$ son las probabilidades de retraso tales que $\sum_{l=0}^d p_l = 1$ y $p_l \in (0, 1)$.

- (2) **Los pagos.** Los pagos individuales $\{Y_{i,j}^{(k)}\}$ son independientes con distribución F_i . Sea μ_i, σ_i^2 la media y varianza respectivamente de F_i ($i \in \{1, \dots, m\}$). Supóngase que $\mu_i = \mu\gamma_i$, donde μ es un factor de media y γ_i es la inflación en los años accidente. También, las varianzas serán $\sigma_i^2 = \sigma^2\gamma_i^2$, donde σ^2 es un factor de varianza.

- (3) **Los conteos.** Los conteos $\{N_{i,j}\}_j$ son variables aleatorias independientes con distribución Poisson, tales que $\mathbb{E}[N_{i,j}] = \alpha_i\beta_j$, $\sum_{j=0}^{m-1} \beta_j = 1$.

- (4) **Independencia.** Se supondrá que las variables $Y_{i,j}^{(k)}$ son independientes de los conteos $N_{i,j}$ y también de los retrasos RBNS e IBNR.

- (5) También se supondrá que las reclamaciones se pagan mediante un solo pago.

Los pagos agregados observados se pueden escribir como

$$X_{i,j} = \sum_{k=1}^{N_{i,j}^{pag}} Y_{i,j}^{(k)}, \quad (i, j) \in \mathcal{I}.$$

Bajo la Suposición 4.3.1

$$\begin{aligned}
\mathbb{E}[X_{i,j}|\mathfrak{N}_I] &= \mathbb{E}\left[\sum_{k=1}^{N_{ij}^{pag}} Y_{i,j}^{(k)} \middle| \mathfrak{N}_I\right] = \mathbb{E}[N_{ij}^{pag} | \mathfrak{N}_I] \mathbb{E}[Y_{i,j}^{(k)}] \\
&= \mathbb{E}\left[\sum_{l=0}^{\min\{j,d\}} N_{i,j-l}^{pag} \middle| \mathfrak{N}_I\right] \mu\gamma_i \\
&= \sum_{l=0}^{\min\{j,d\}} N_{i,j-l} \cdot p_l \cdot \mu\gamma_i.
\end{aligned} \tag{4.3.2}$$

Entonces,

$$\begin{aligned}
\mathbb{E}[X_{i,j}] &= \mathbb{E}[\mathbb{E}(X_{i,j}|\mathfrak{N}_I)] = \mathbb{E}\left[\sum_{l=0}^{\min\{j,d\}} N_{i,j-l} \cdot p_l \cdot \mu\gamma_i\right] \\
&= \alpha_i \mu\gamma_i \sum_{l=0}^{\min\{j,d\}} \beta_{j-l} \cdot p_l.
\end{aligned} \tag{4.3.3}$$

Se puede utilizar (4.3.2) o (4.3.3) para construir la reserva RBNS, aunque generalmente es más común utilizar (4.3.2) con el número de reclamaciones incurridas; excepto cuando se pueda generar exactamente los pronósticos Chain-Ladder estándar (en este caso es mejor usar 4.3.3).

Para la reserva IBNR, se tiene que usar (4.3.3) con los estimadores del número futuro de reclamaciones incurridas.

En Verrall *et. al.* (2010) se demuestra que la varianza condicional de $X_{i,j}$, bajo la Suposición 4.3.1 es aproximadamente proporcional a la media. Es decir,

$$\begin{aligned}
Var(X_{i,j}|\mathfrak{N}_I) &\approx \frac{\sigma_i^2 + \mu_i^2}{\mu_i} \mathbb{E}[X_{i,j}|\mathfrak{N}_I] \\
&= \gamma_i \frac{\sigma^2 + \mu^2}{\mu} \mathbb{E}[X_{i,j}|\mathfrak{N}_I] \\
&= \varphi_i \cdot \mathbb{E}[X_{i,j}|\mathfrak{N}_I],
\end{aligned}$$

donde $\varphi := \frac{\sigma^2 + \mu^2}{\mu}$ y $\varphi_i := \gamma_i \cdot \varphi$.

Entonces, se puede estimar los parámetros a partir de un modelo Poisson con sobre-dispersión (como en el capítulo 3).

Antes de empezar a estudiar el problema de estimación, se dará una pequeña modificación del modelo dado por (4.3.2) y (4.3.3) para que se pueda relacionar dicho modelo, con el modelo CL.

Este modelo es muy parecido pues usa un modelo Poisson con sobre-dispersión pero se reemplaza $\mathbf{P} = (p_0, \dots, p_d)$ por $\boldsymbol{\pi} = (\pi_0, \dots, \pi_{m-1})$ sin restricciones a los valores π_0, \dots, π_{m-1} . Entonces, las “medias” de las expresiones (4.3.2) y (4.3.3) se convierten en

$$\mathbb{E}[X_{i,j}|\mathfrak{N}_I] = \sum_{l=0}^j N_{i,j-l} \cdot \pi_l \cdot \mu \cdot \gamma_i, \quad (4.3.4)$$

y

$$\mathbb{E}[X_{i,j}] = \alpha_i \mu \gamma_i \sum_{l=0}^j \beta_{j-l} \cdot \pi_l. \quad (4.3.5)$$

Observación 4.3.1.

En general, se espera que los valores de los parámetros \mathbf{P} y $\boldsymbol{\pi}$ sean muy parecidos y que las predicciones provenientes de éstos, sean muy parecidas. ∇

Ahora, se considerará un modelo estocástico Poisson para aplicar un algoritmo CL a los pagos agregados Δ_I .

El modelo CL supone que $\{X_{i,j}\}_j$ son variables aleatorias independientes tales que

$$\mathbb{E}[X_{i,j}] = \tilde{\alpha}_i \tilde{\beta}_j,$$

con $\sum_{j=1}^J \tilde{\beta}_j = 1$ y análogamente, el modelo CL se puede aplicar al triángulo de los conteos incurridos, i.e.,

$$\mathbb{E}[N_{i,j}] = \alpha_i \beta_j, \quad (4.3.6)$$

con $\sum_{j=1}^J \beta_j = 1$.

4.3.2. Algoritmo Double Chain-Ladder

Se puede estimar los parámetros del modelo que se estudiaron en la Sección 4.3.1 por máxima verosimilitud; sin embargo, se aplicará el algoritmo CL a los triángulos de reclamaciones pagadas (Δ_I) y los conteos incurridos (\mathfrak{N}_I) para estimar todos los parámetros en todos los modelos.

Como su nombre lo dice, en el *Double Chain-Ladder* (DCL), se aplica dos veces la técnica CL clásica, i.e. la estimación DCL utiliza los estimadores de los parámetros CL de los triángulos de conteo y de pagos.

Notación 4.3.1. Para $i \in \{1, \dots, I\}$, $j \in \{0, 1, \dots, J-1\}$,

- $(\widehat{\alpha}_i, \widehat{\beta}_j)$, serán los estimadores del modelo CL para el proceso de conteo.
- $(\widetilde{\alpha}_i, \widetilde{\beta}_j)$, serán los estimadores CL para el proceso de pagos.

□

El algoritmo CL generará estimados de los factores $\{\widehat{f}_j\}_{j=1}^{J-1}$, que se pueden convertir en estimadores para β_j , $j \in \{0, \dots, J-1\}$ a partir de

$$\widehat{\beta}_0 = \frac{1}{\prod_{l=1}^{J-1} \widehat{f}_l}, \quad \widehat{\beta}_j = \frac{\widehat{f}_j - 1}{\prod_{l=1}^{J-1} \widehat{f}_l}, \text{ para } j \in \{1, \dots, J-1\}.$$

4.3.2.1. Estimación del retraso de reporte

Primero se considerará el modelo

$$\mathbb{E}[X_{i,j}] = \alpha_i \mu \gamma_i \sum_{l=0}^j \beta_{j-l} \cdot \pi_l,$$

y se discutirán las modificaciones para obtener los estimadores de los parámetros en

$$\mathbb{E}[X_{i,j}] = \alpha_i \mu \gamma_i \sum_{l=0}^{\min\{j,d\}} \beta_{j-l} \cdot p_l.$$

Los parámetros $\boldsymbol{\pi} = \{\pi_l\}_{l=0}^{J-1}$ se pueden estimar resolviendo el sistema lineal

$$\begin{pmatrix} \widetilde{\beta}_0 \\ \vdots \\ \widetilde{\beta}_{J-1} \end{pmatrix} = \begin{pmatrix} \beta_0 & 0 & \cdots & 0 \\ \beta_1 & \beta_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ \beta_{J-1} & \cdots & \beta_1 & \beta_0 \end{pmatrix} \begin{pmatrix} \pi_0 \\ \vdots \\ \pi_{J-1} \end{pmatrix}, \quad (4.3.7)$$

donde $\widetilde{\beta}_j, \beta_j$ se estimaron previamente. Sea $\widehat{\boldsymbol{\pi}} = (\widehat{\pi}_0, \dots, \widehat{\pi}_{J-1})$ una solución de (4.3.7).

Observación 4.3.2.

Los valores $\hat{\pi}_l$ pueden ser negativos y pueden sumar más de 1. ▽

Se estimará el máximo número de periodos de retraso contando el número de $\hat{\pi}_l \geq 0$ tales que $\sum_{l=0}^{d-1} \hat{\pi}_l < 1 \leq \sum_{l=0}^d \hat{\pi}_l$ y se estimarán los parámetros de retraso en (4.3.3) como

$$\hat{p}_l = \hat{\pi}_l, l = 0, \dots, d-1,$$

y

$$\hat{p}_d = 1 - \sum_{l=0}^{d-1} \hat{\pi}_l.$$

Aunque \hat{p}_j y $\hat{\pi}_j$ son muy parecidos (por construcción), habrá diferencias para retrasos de reporte más largos que afectarán los estimadores de las reservas.

4.3.2.2. Estimación de los parámetros de la distribución de los pagos individuales

Ahora, se estimarán la media y varianza de la distribución de los pagos y el parámetro que mide la inflación en los años de accidentes.

Primero se considerarán los parámetros de inflación $\gamma = (\gamma_1, \dots, \gamma_I)$ y μ . Como $\alpha_i \cdot \mu_i \cdot \gamma_i = \tilde{\gamma}_i$. Entonces, para $i \in \{1, \dots, I\}$

$$\hat{\gamma}_i = \frac{\hat{\alpha}_i}{\hat{\alpha}_i \hat{\mu}}. \quad (4.3.8)$$

Observación 4.3.3.

El modelo está sobre-parametrizado ya que hay demasiados parámetros de inflación. La manera más simple de asegurar la identificabilidad es hacer $\gamma_1 = 1$ y estimar $\hat{\mu}$ como

$$\hat{\mu} = \frac{\hat{\alpha}_1}{\hat{\alpha}_1}.$$

▽

Una vez que se tiene $\hat{\mu}$, los estimadores para los demás parámetros se obtienen a partir de la ecuación (4.3.8).

Para la estimación de las varianzas $\{\sigma_i^2\}_{i=1}^I$ se utilizará el estimador propuesto en Verrall *et. al.* (2010)

$$\hat{\varphi} = \frac{1}{n - (d+1)} \sum_{(i,j) \in \mathcal{I}} \frac{(X_{i,j} - \hat{X}_{i,j}^{DCL})^2}{\hat{X}_{i,j}^{DCL} \hat{\gamma}_i}, \quad (4.3.9)$$

donde $n = \frac{I(I+1)}{2}$ y $\widehat{X}_{i,j}^{DCL}$ es el estimado DCL de $\mathbb{E}[X_{i,j}|\mathfrak{R}_I]$, con

$$\widehat{X}_{i,j}^{DCL} = \sum_{l=0}^{\min\{j,d\}} N_{i,j-l} \cdot \widehat{p}_l \widehat{\mu} \widehat{\gamma}_i.$$

Entonces, el factor de varianza del pago total se puede estimar como

$$\widehat{\sigma}_i^2 = \widehat{\sigma}^2 \widehat{\gamma}_i^2, \quad i = 1, \dots, I,$$

donde $\widehat{\sigma}^2 = \widehat{\mu} \widehat{\varphi} - \widehat{\mu}^2$.

4.3.3. Estimadores DCL para las reservas RBNS e IBNR

Los parámetros estimados $\widehat{\theta}_1 = (\widehat{\mathbf{P}}, \widehat{\mu}, \widehat{\gamma})$ ó $\widehat{\theta}_2 = (\widehat{\pi}, \widehat{\mu}, \widehat{\gamma})$ se pueden utilizar para pronosticar los componentes de las reservas RBNS e IBNR.

Para la reserva RBNS, se debe condicionar al número actual de reclamaciones y usar

$$\mathbb{E}[X_{i,j}|\mathfrak{R}_I] = \sum_{l=0}^{j \wedge d} N_{i,j-l} \cdot p_l \cdot \mu \cdot \gamma_i$$

Para la reserva IBNR primero se deben construir predicciones del número de reclamaciones reportados (utilizando el algoritmo CL).

Notación 4.3.2.

$$\begin{aligned} \mathcal{J}_1 &= \{i = 2, \dots, I; j = 0, \dots, J-1 \text{ tales que } i+j = J+1, \dots, J+d\}, \\ \mathcal{J}_2 &= \{i = 1, \dots, I; j = J, \dots, J-d \text{ tales que } i+j = J+1, \dots, J+d\}, \\ \mathcal{J}_3 &= \{i = 2, \dots, I; j = 0, \dots, J-1 \text{ tales que } i+j = J+d+1, \dots, 2J+d-1\}. \end{aligned}$$

▽

Observación 4.3.4.

El algoritmo CL sólo se puede aplicar al triángulo \mathcal{J}_1 .

▽

4.3.4. Estimación de las reclamaciones pendientes ignorando colas

Se pueden construir estimados de las reclamaciones pendientes utilizando el algoritmo CL, obteniendo $\widehat{X}_{i,j}^{CL} = \widehat{\alpha}_i \widehat{\beta}_j$ para $(i, j) \in \mathcal{J}_1$.

Hay dos posibles estimados que se denotarán por $\widehat{X}_{i,j}^{rbns(1)}$ basado en (4.3.4) y $\widehat{X}_{i,j}^{rbns(2)}$ basados en (4.3.5)

$$\widehat{X}_{i,j}^{rbns(1)} = \sum_{l=i-J+j}^j N_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i$$

y

$$\widehat{X}_{i,j}^{rbns(2)} = \sum_{l=i-J+j}^j \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i.$$

Y para la componente del IBNR

$$\widehat{X}_{i,j}^{ibnr} = \sum_{l=0}^{i-J+j-1} \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i.$$

Teorema 4.3.1.

Para $(i, j) \in \mathcal{J}_1$, defínase

$$\widehat{X}_{i,j}^{CL} = \widehat{\alpha}_i \cdot \widehat{\beta}_j,$$

$$\widehat{X}_{i,j}^{rbns(1)} = \sum_{l=i-J+j}^j N_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i,$$

$$\widehat{X}_{i,j}^{rbns(2)} = \sum_{l=i-J+j}^j \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i,$$

donde, $\widehat{\alpha}_i \cdot \widehat{\mu} \cdot \widehat{\gamma}_i = \widehat{\alpha}_i$ y $\sum_{l=0}^j \widehat{\beta}_{j-l} \cdot \widehat{\pi}_l = \widehat{\beta}_j$.

Entonces,

$$\widehat{X}_{i,j}^{CL} = \widehat{X}_{i,j}^{rbns(2)} + \widehat{X}_{i,j}^{ibnr}.$$

Demostración:

$$\begin{aligned}
\widehat{X}_{i,j}^{rbns(2)} + \widehat{X}_{i,j}^{ibnr} &= \sum_{l=i-J+j}^j \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i + \sum_{l=0}^{i-J+j-1} \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i \\
&= \sum_{l=0}^j \widehat{N}_{i,j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i \\
&= \sum_{l=0}^j \widehat{\alpha}_i \widehat{\beta}_{j-l} \widehat{\pi}_l \cdot \widehat{\mu} \cdot \widehat{\gamma}_i \\
&= \sum_{l=0}^j (\widehat{\alpha}_i \cdot \widehat{\mu} \cdot \widehat{\gamma}_i) \widehat{\beta}_{j-l} \widehat{\pi}_l \\
&= \widehat{\alpha}_i \sum_{l=0}^j \widehat{\beta}_{j-l} \widehat{\pi}_l \\
&= \widehat{\alpha}_i \widehat{\beta}_j = \widehat{X}_{i,j}^{CL}.
\end{aligned}$$

□

4.3.5. Método Double Chain-Ladder-Bornhutter&Ferguson (DCL-BF)

El método Double Chain Ladder-Bornhutter&Ferguson (DCL-BF) sigue los mismos pasos que el Double Chain-Ladder (DCL) pero incorporando un parámetro de inflación como información determinística adicional (idea completamente BF).

Es decir, se hace un ajuste utilizando parámetros de inflación que se obtienen a partir del DCL con los conteos reportados e incurridos en vez de los pagos agregados.

El método DCL-BF es un algoritmo de dos pasos:

Paso 1: *Estimación de parámetros*

Primero se estiman los parámetros del modelo DCL a partir de los triángulos Δ_I y \aleph_I ; obteniéndose $\widehat{\rho}, \widehat{\mu}, \widehat{\sigma}^2, \widehat{\gamma}_{1,0}, \dots, \widehat{\gamma}_{I,0}$.

Luego, repetir el algoritmo DCL pero con el triángulo de los pagos $\widehat{\Delta}_I = \{\widehat{X}_{i,j} : (i,j) < I\}$ donde $\widehat{X}_{i,j}$ es el incurrido correspondiente al año de accidente i y periodo de desarrollo j obteniéndose parámetros de inflación $\widehat{\gamma}_{1,1}, \dots, \widehat{\gamma}_{I,1}$.

Paso 2: *Ajuste Bornhutter-Ferguson*

Reemplazar los parámetros de inflación $\hat{\gamma}_{1,0}, \dots, \hat{\gamma}_{I,0}$ de los datos pagados y sustituirlo por información más “realista” por ejemplo $\hat{\gamma}_{1,1}, \hat{\gamma}_{2,1}, \dots, \hat{\gamma}_{I,1}$. Este paso se puede definir para un porcentaje arbitrario de ajuste BF.

A partir de los pasos 1 y 2 se tendrán los parámetros finales

$$\hat{\theta} = (\hat{\rho}_0, \dots, \hat{\rho}_d, \hat{\mu}, \hat{\sigma}^2, \hat{\gamma}_{1,1}, \hat{\gamma}_{2,1}, \dots, \hat{\gamma}_{I,1}).$$

Como ya se dijo antes, los algoritmos CL y BF son muy sencillos y populares. El método BF (Bornhuetter-Ferguson (1972)) tiene la ventaja de trabajar con *outliers* a diferencia del CL. Con este objetivo en mente, el método BF incorpora información previa de “expertos” y por lo tanto es más robusto que el método CL (que sólo utiliza la información en el triángulo correspondiente).

El método CL estima las reclamaciones pendientes para el año accidente $i > 1$ a partir de

$$\hat{R}_i^{CL} = C_{i,I-i} \left(\prod_{j=I-i+1}^{J-1} \hat{f}_j - 1 \right),$$

donde $C_{i,I-i}$ son las reclamaciones acumuladas y $\hat{f}_1, \dots, \hat{f}_{J-1}$ son los factores de desarrollo, i.e. la reserva CL depende fuertemente de las cantidades $C_{i,I-i}$, de aquí se pueden obtener predicciones que no tengan sentido si el triángulo tiene poca información o información muy volátil. El método BF “evita” esta dependencia y reemplaza las últimas reclamaciones acumuladas por un estimado externo (*a priori*). Entonces el método BF reemplaza $C_{i,I-i}$

por $\frac{\mu_i^{prior}}{\prod_{j=I-i+1}^{J-1} \hat{f}_j}$ y el estimador BF de las reclamaciones pendientes es

$$\hat{R}_i^{BF} = \frac{\mu_i^{prior}}{\prod_{j=I-i+1}^{J-1} \hat{f}_j} \left(\prod_{j=I-i+1}^{J-1} \hat{f}_j - 1 \right),$$

suponiendo un modelo Poisson $\mathbb{E}[X_{i,j}] = \tilde{\alpha}_i \tilde{\beta}_j$, entonces la relación entre la reserva CL y BF es

$$\widehat{R}_i^{CL} = \widehat{\alpha}_i \sum_{j=0}^{J-1} \widehat{\beta}_j \left(\frac{\sum_{j=I-i+1}^{J-1} \widehat{\beta}_j}{\sum_{j=0}^{J-1} \widehat{\beta}_j} \right) = C_{i,J}^{CL} \left(\frac{\sum_{j=I-i+1}^{J-1} \widehat{\beta}_j}{\sum_{j=0}^{J-1} \widehat{\beta}_j} \right)$$

y

$$\widehat{R}_i^{BF} = \mu_i^{prior} \left(\frac{\sum_{j=I-i+1}^{J-1} \widehat{\beta}_j}{\sum_{j=0}^{J-1} \widehat{\beta}_j} \right),$$

donde $C_{i,J}^{CL}$ es la estimación CL de la siniestralidad última y $\widehat{\alpha}_i, \widehat{\beta}_j$ son las estimaciones de los parámetros CL.

Si se supone que $\sum_{j=0}^{J-1} \beta_j = 1$ y $\widehat{C}_{i,J}^{CL} = \widehat{\alpha}_i$, entonces lo que el algoritmo BF reemplaza son los parámetros por renglón en el modelo Poisson.

El método DCL-BF tiene la misma lógica. El objetivo es “estabilizar” los parámetros por renglón con información adicional, en este caso, que viene de los datos incurridos.

Los renglones por parámetros en el modelo DCL son $\widetilde{\alpha}_i = \alpha_i \mu \gamma_i$ para información de lo pagado y $\widehat{\alpha}_i = \alpha_i \widehat{\mu} \widehat{\gamma}_i$ para información de lo incurrido, donde los α_i 's son los parámetros por renglón en el modelo para los conteos reportados.

Por lo tanto, si el método DCL-BF utiliza los parámetros de inflación de los datos incurridos, entonces se está reemplazando el parámetro más volátil en $\widetilde{\alpha}_i$, i.e. la inflación, por año accidente, γ_i , por el estimado que se obtiene de los triángulos $(\aleph_I, \widehat{\Delta}_I)$. Con este reemplazo las predicciones se vuelven más estables y precisas.

4.3.5.1. Pronóstico de las reservas IBNR y RBNS en el modelo DCL-BF

Los parámetros estimados $\widehat{\theta} = (\widehat{\rho}_0, \dots, \widehat{\rho}_d, \widehat{\mu}, \widehat{\sigma}^2, \widehat{\gamma}_{1,1}, \widehat{\gamma}_{2,1}, \dots, \widehat{\gamma}_{I,1})$ que se obtienen de los pasos 1 y 2 se pueden usar para obtener un pronóstico puntual de los componentes RBNS e IBNR de la reserva.

Observación 4.3.5.

1. El método CL genera sólo pronóstico sobre \mathcal{J}_1 .
2. Los métodos DCL y DCL-BF también obtienen pronósticos en la cola $\mathcal{J}_2 \cup \mathcal{J}_3$.

Como antes, se usará

$$\hat{X}_{i,j}^{rbns(1)} = \sum_{l=i-J+j}^j N_{i,j-l} \hat{\pi}_l \cdot \hat{\mu} \cdot \hat{\gamma}_i, \quad (i, j) \in \mathcal{J}_1 \cup \mathcal{J}_2 \quad (4.3.10)$$

y

$$\hat{X}_{i,j}^{ibnr} = \sum_{l=0}^{i-J+j-1} \hat{N}_{i,j-l} \hat{\pi}_l \cdot \hat{\mu} \cdot \hat{\gamma}_i \quad (i, j) \in \mathcal{J}_1 \cup \mathcal{J}_2 \cup \mathcal{J}_3. \quad (4.3.11)$$

Utilizando los datos en los triángulos observados \aleph_I, Δ_I y el modelo con el parámetro estimado $\hat{\theta}$ se puede obtener una distribución predictiva *bootstrap*.

▽

Ejemplo 4.3.1.

El método Double Chain-Ladder se pone en práctica con el paquete estadístico R. Con apoyo de la librería DCL se puede obtener una estimación de los parámetros requeridos para la aplicación del método DCL, para esto se requieren dos triángulos de observaciones, el de los pagos agregados y conteos incurridos. A continuación se presentan ambos triángulos:

Código de R

```
R> # Mostrar objeto triángulo, Ptriangulo (Pagos Agregados)
R> Ptriangulo

dev
origin    1      2      3      4      5      6      7      8      9     10
2000 451288 339519 333371 144988 93243 45511 25217 20406 31482 1729
2001 448627 512882 168467 130674 56044 33397 56071 26522 14346 NA
2002 693574 497737 202272 120753 125046 37154 27608 17864 NA NA
2003 652043 546406 244474 200896 106802 106753 63688 NA NA NA
2004 566082 503970 217838 145181 165519 91313 NA NA NA NA
2005 606606 562543 227374 153551 132743 NA NA NA NA NA
2006 536976 472525 154205 150564 NA NA NA NA NA NA
2007 554833 590880 300964 NA NA NA NA NA NA NA NA
2008 537238 701111 NA NA NA NA NA NA NA NA NA
2009 684944 NA NA NA NA NA NA NA NA NA NA

R> #Mostrar objeto triángulo, Ctriangulo (Conteos incurridos agregados)
R> Ctriangulo

dev
origin    1      2      3      4      5      6      7      8      9     10
2000 6238 831 49 7 1 1 2 1 2 3
```

```

2001 7773 1381 23 4 1 3 1 1 3 NA
2002 10306 1093 17 5 2 0 2 2 NA NA
2003 9639 995 17 6 1 5 4 NA NA NA
2004 9511 1386 39 4 6 5 NA NA NA NA
2005 10023 1342 31 16 9 NA NA NA NA NA
2006 9834 1424 59 24 NA NA NA NA NA NA
2007 10899 1503 84 NA NA NA NA NA NA NA
2008 11954 1704 NA NA NA NA NA NA NA NA
2009 10989 NA NA NA NA NA NA NA NA NA

```

Código de R

La función `dcl. estimation()` realiza la estimación de los parámetros usando el método Double Chain-Ladder, recibe como argumentos los triángulos antes mencionados, y generará una tabla sintetizada con distintas estimaciones derivadas del método DCL.

Existen funciones como `bdcl. estimation()` y `idcl. estimation()` las cuáles tienen la misma funcionalidad que la anterior pero la primera hace un ajuste Bornhuetter-Ferguson a la inflación usando los datos incurridos y la segunda corrige el parámetro de la inflación utilizando los datos incurridos para proporcionar el flujo de efectivo incurrido.

Código de R

```

#Estimación de los parámetros de DCL
R>DCL<-dcl. estimation(Ptriángulo,Ctriángulo)

      delay.par delay.prob inflation severity.mean severity.var
2000   0.3649    0.3649    1.0000    208.4910    2055848.1
2001   0.2924    0.2924    0.7562    157.6619    1175628.8
2002   0.1119    0.1119    0.7350    153.2415    1110629.4
2003   0.0839    0.0839    0.8908    185.7203    1631305.5
2004   0.0630    0.0630    0.7840    163.4627    1263728.0
2005   0.0332    0.0332    0.7791    162.4267    1247760.4
2006   0.0245    0.0245    0.6605    137.7131     896947.6
2007   0.0121    0.0121    0.7370    153.6665    1116798.1
2008   0.0158    0.0142    0.6990    145.7439    1004609.0
2009  -0.0012    0.0000    0.8198    170.9139    1381564.2

      mean.factor mean.factor.adj variance.factor
2000    208.3748         208.491         2055848

```

Código de R

La primera y segunda columna (`delay.par`, `delay.prob`) corresponde a la estimación del retraso de pago, i.e., $\hat{\pi}_i$ una solución del sistema de ecuaciones mencionado antes. La columna `inflation` estima la inflación γ_i en los pagos de siniestros por año de accidente. La columna `severity.mean` es el factor de severidad promedio μ_i , `mean.factor.adj` es el factor de severidad promedio ajustado que en este trabajo no se estudiará. La última columna (`severity.var`) estima la varianza de la severidad $\hat{\sigma}_i^2$.

Si se quisiera conocer individualmente la estimación de los parámetros así como la estimación de las reclamaciones pagadas (Δ_I) y los conteos incurridos (\aleph_I), dentro de la función

`dcl.estimacion()` se tendría que des-habilitar la opción `Tables`.

```

Código de R
R> dcl.completo<-dcl.estimacion(Ptriángulo,Ctriángulo,Tables=FALSE)
R> dcl.completo

$alpha.N
      2000      2001      2002      2003      2004      2005      2006      2007      2008      2009
7135.000  9193.866 11435.310 10676.296 10963.113 11436.877 11360.506 12518.938 13745.925 12556.030

$beta.N
[1] 0.8751970027 0.1184065569 0.0037653489 0.0009141152 0.0003287286 0.0002833802 0.0002341282
[8] 0.0001440705 0.0003062062 0.0004204625

$Nhat
      dev
origin 1      2      3      4      5      6      7      8      9     10
2000  6244.531  844.8308 26.86576  6.522212 2.345479 2.021918 1.670505 1.027943 2.184781 3.000000
2001  8046.444 1088.6140 34.61811  8.404252 3.022287 2.605360 2.152544 1.324565 2.815219 3.865676
2002 10008.149 1354.0156 43.05793 10.453190 3.759113 3.240541 2.677329 1.647491 3.501563 4.808119
2003  9343.862 1264.1435 40.19998  9.759364 3.509604 3.025451 2.499622 1.538140 3.269148 4.488982
2004  9594.883 1298.1044 41.27994 10.021548 3.603889 3.106730 2.566774 1.579462 3.356973 4.609578
2005 10009.521 1354.2013 43.06383 10.454623 3.759629 3.240985 2.677696 1.647717 3.502043 4.808778
2006  9942.681 1345.1584 42.77627 10.384811 3.734523 3.219343 2.659815 1.636714 3.478657 4.776667
2007 10956.537 1482.3244 47.13817 11.443752 4.115333 3.547620 2.931037 1.803610 3.833377 5.263744
2008 12030.392 1627.6076 51.75820 12.565359 4.518679 3.895324 3.218309 1.980383 4.209088 5.779646
2009 10989.000 1486.7163 47.27783 11.477658 4.127526 3.558131 2.939721 1.808954 3.844734 5.279340

$alpha.X
      2000      2001      2002      2003      2004      2005      2006      2007      2008      2009
1486754 1448715 1751387 1981700 1791061 1856619 1563619 1922669 2002268 2144804

$beta.X
[1] 0.319350463 0.299122798 0.133958486 0.088099318 0.065857069 0.037132653 0.025880266 0.013824193
[9] 0.015611817 0.001162936

$Xhat
      dev
origin 1      2      3      4      5      6      7      8      9     10
2000  474795.6 444722.0 199163.3 130982.0  97913.26 55207.12 38477.59 20553.17 23210.93 1729.000
2001  462647.7 433343.6 194067.6 127630.8  95408.11 53794.62 37493.12 20027.31 22617.07 1684.763
2002  559306.3 523879.8 234613.2 154296.0 115341.22 65033.65 45326.36 24211.51 27342.33 2036.751
2003  632856.8 592771.6 265465.5 174586.4 130508.95 73585.78 51286.92 27395.40 30937.94 2304.591
2004  571976.1 535747.1 239927.8 157791.2 117954.01 66506.83 46353.13 24759.97 27961.71 2082.889
2005  592912.0 555356.9 248709.8 163566.8 122271.45 68941.17 48049.78 25666.25 28985.19 2159.129
2006  499342.3 467714.0 209460.0 137753.7 102975.34 58061.31 40466.87 21615.77 24410.93 1818.389
2007  614005.2 575114.1 257557.8 169385.8 126621.33 71393.79 49759.18 26579.34 30016.35 2235.941
2008  639425.1 598923.9 268220.7 176398.4 131863.48 74349.51 51819.22 27679.73 31259.04 2328.509
2009  684944.0 641559.6 287314.6 188955.7 141250.47 79642.24 55508.09 29650.18 33484.28 2494.270

```

Código de R

Se puede graficar las estimaciones de los parámetros para cada caso: el modelo CL y B&F . Con la función `plot.dcl.par(DCL , type.inflat = 'DCL')` .

En la Figura 4.1, la primera gráfica (CL underwriting parameters) muestra el desarrollo del parámetro α_i CL que describe el comportamiento de los pagos agregados por año de accidente. La segunda gráfica (CL development parameters) es el desarrollo por periodo del factor β_j del modelo CL para pagos agregados con la propiedad de que $\sum_{j=1}^J \beta_j = 1$. La tercera gráfica (Severity inflation) es el desarrollo del factor inflación por cada año de accidente. La última gráfica (Delay parameters) muestra el desarrollo de la estimación del

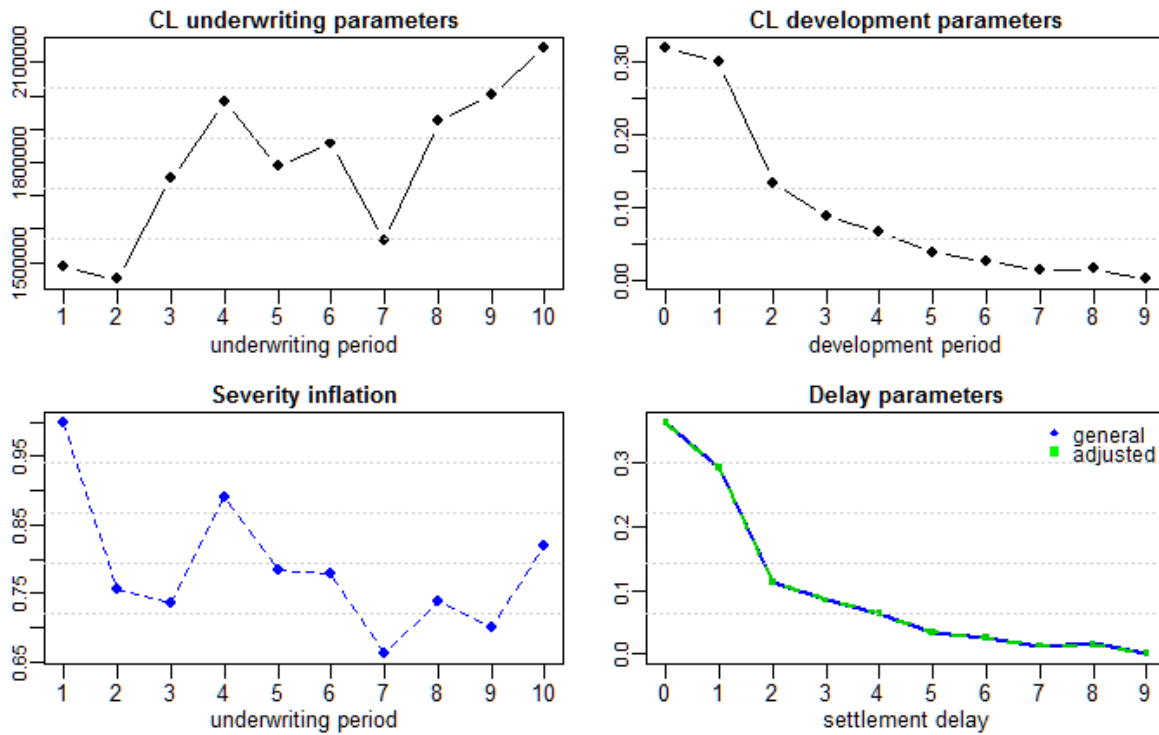


Figura 4.1: Gráfica de estimación de parámetros

parámetro de retraso con el modelo general (CL) y el ajustado (BF).

La paquetería DCL , también proporciona una predicción de la reserva IBNR y RBNS a partir de los triángulos de pagos agregados y conteos incurridos. Esto se hace a través de la función `dcl.predict()` que genera un tabla con la estimación de las reservas para obligaciones pendientes de cumplir con el método Double Chain-Ladder, que es la suma de la reserva RBNS e IBNR, i.e. la columna `total`. Adicionalmente la columna `clm` proporciona la estimación de la reserva con el método clásico Chain-Ladder.

— Código de R —

```
# Los pronósticos de flujo de efectivo por año
R> DCL_prediccion<-dcl.predict(dcl.par=DCL,Ctriángulo,num.dec=0)

Future.years  rbns  ibnr  total  clm
1             1 1260908  97168 1358076 1353858
2             2  672018  82620  754638  754180
3             3  453361  35506  488866  488612
4             4  292540  26503  319043  318043
```



```

5          5 164970 20353 185324 184611
6          6 103125 11971 115096 115023
7          7  54037  9074   63111   63145
8          8  30397  5412   35808   35813
9          9     0   5460   5460   2494
10         10    0   1119   1119    NA
11         11    0   580    580    NA
12         12    0   355    355    NA
13         13    0   211    211    NA
14         14    0   116    116    NA
15         15    0    65    65    NA
16         16    0    32    32    NA
17         17    0    13    13    NA
18         18    0     0     0    NA
19         Tot. 3031355 296558 3327913 3315779

```

Código de R

Nótese que la columna `Future.years` llega hasta el año 19, esto se debe a la estimación de las colas en el método DCL, pero esta opción se puede deshabilitar con el comando `Tail` dentro de la función y otorga sólo el estimado de los años accidente, además como en la función anterior ésta tiene la opción de dar un análisis más detallado como por ejemplo, el triángulo inferior ó estimación de las reclamaciones pagadas y conteos incurridos con el mismo comando antes mencionado (`Tables=FALSE`).

Por último, se puede hacer una estimación *bootstrap* a partir del método DCL, el cuál simula la distribución del flujo de reclamaciones y genera para cada triángulo por año calendario, la media, varianza así como los cuantiles más significativos. El método DCL *bootstrap* realiza 999 simulaciones por default pero se puede realizar la cantidad que sea considerada de acuerdo al criterio del usuario, en este caso se consideraron 10,000.

Código de R

```

# Se muestran las matrices simuladas bootstrapping usando DCL
R> DCL_boot<-dcl.boot(dcl.par=DCL,Ntriangle=Ctriángulo)

[1] "Please wait, simulating the distribution..."
[1] "Done!"
  period      rbns mean.rbns  sd.rbns  Q1.rbns  Q5.rbns  Q50.rbns  Q95.rbns  Q99.rbns
1      1 1260907.90 1257300.52 120628.87 997860.39 1068635.06 1251951.08 1462419.05 1554046.52
2      2  672017.58  668047.92  79917.07  500253.26  543894.59  664697.22  803911.93  874227.06
3      3  453360.52  451120.36  65914.32  311783.78  350283.16  446716.41  565587.80  620009.20
4      4  292539.65  290760.45  53806.04  177409.96  209660.80  287331.21  384584.46  432022.69
5      5  164970.41  163150.06  40309.54  82450.34  102475.75  159778.78  233747.07  270676.13
6      6  103125.19  100549.36  33033.12  38529.93  52477.11  97191.80  161244.62  192274.58
7      7   54037.12  52526.38  26072.60  9757.10  16978.00  48551.34  100892.79  129273.76
8      8   30396.54  29031.54  21571.14  105.07  2953.56  24495.19  70189.22  99459.79
9      9     0.00    2305.75   7802.71     0.00     0.00     0.00  15737.07  39226.06
10     10     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
11     11     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
12     12     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
13     13     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
14     14     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
15     15     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
16     16     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
17     17     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
18     18     0.00     0.00     0.00     0.00     0.00     0.00     0.00     0.00
19     Tot. 3031354.91 3014792.34 278297.14 2413415.27 2575765.28 3004562.74 3490692.44 3731463.66

  period      ibnr mean.ibnr  sd.ibnr  Q1.ibnr  Q5.ibnr  Q50.ibnr  Q95.ibnr  Q99.ibnr

```

| | | | | | | | | | |
|----|--------|------------|------------|-----------|------------|------------|------------|------------|------------|
| 1 | 1 | 97168.11 | 96715.55 | 27642.50 | 44410.32 | 56824.16 | 93829.50 | 145767.44 | 175326.27 |
| 2 | 2 | 82620.00 | 81898.56 | 25290.40 | 34655.15 | 45529.91 | 79281.81 | 127069.27 | 153597.49 |
| 3 | 3 | 35505.74 | 35116.32 | 15889.64 | 8426.15 | 13666.12 | 32749.95 | 64566.59 | 82498.19 |
| 4 | 4 | 26503.46 | 25984.33 | 13473.81 | 4820.61 | 8435.73 | 23730.51 | 51214.37 | 68247.99 |
| 5 | 5 | 20353.21 | 20024.10 | 12034.33 | 2476.86 | 5049.91 | 17722.96 | 42592.52 | 58027.18 |
| 6 | 6 | 11970.64 | 11599.02 | 9074.38 | 460.52 | 1500.23 | 9484.49 | 28846.12 | 43140.89 |
| 7 | 7 | 9074.00 | 8652.29 | 7877.64 | 114.48 | 643.61 | 6436.57 | 23758.64 | 35790.02 |
| 8 | 8 | 5411.50 | 5140.71 | 6022.78 | 2.38 | 64.03 | 3101.09 | 16821.11 | 28087.13 |
| 9 | 9 | 5459.61 | 4718.61 | 5962.56 | 0.23 | 23.22 | 2679.73 | 16412.98 | 26684.38 |
| 10 | 10 | 1119.08 | 1288.21 | 2979.28 | 0.00 | 0.00 | 88.33 | 6778.57 | 15143.68 |
| 11 | 11 | 580.33 | 520.49 | 1859.12 | 0.00 | 0.00 | 0.35 | 2916.03 | 9120.69 |
| 12 | 12 | 355.37 | 291.09 | 1339.00 | 0.00 | 0.00 | 0.00 | 1620.52 | 5968.27 |
| 13 | 13 | 210.61 | 181.11 | 1011.32 | 0.00 | 0.00 | 0.00 | 692.17 | 4626.77 |
| 14 | 14 | 116.42 | 89.88 | 646.81 | 0.00 | 0.00 | 0.00 | 166.63 | 2521.05 |
| 15 | 15 | 64.70 | 57.35 | 587.54 | 0.00 | 0.00 | 0.00 | 14.76 | 1589.64 |
| 16 | 16 | 32.12 | 21.60 | 412.65 | 0.00 | 0.00 | 0.00 | 0.00 | 268.31 |
| 17 | 17 | 12.77 | 11.03 | 221.02 | 0.00 | 0.00 | 0.00 | 0.00 | 7.88 |
| 18 | 18 | 0.00 | 0.15 | 14.42 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 19 | Tot. | 296557.68 | 292310.41 | 53909.51 | 186288.54 | 210887.21 | 288272.21 | 386644.09 | 434619.70 |
| | period | total | mean.total | sd.total | Q1.total | Q5.total | Q50.total | Q95.total | Q99.total |
| 1 | 1 | 1358076.01 | 1354016.07 | 128595.54 | 1076027.01 | 1153141.62 | 1347485.50 | 1573515.71 | 1675105.79 |
| 2 | 2 | 754637.59 | 749946.48 | 86115.86 | 567707.69 | 614230.41 | 746245.71 | 897695.17 | 970686.28 |
| 3 | 3 | 488866.26 | 486236.68 | 68828.50 | 340441.69 | 381899.36 | 481711.52 | 604653.53 | 663973.60 |
| 4 | 4 | 319043.11 | 316744.78 | 56181.44 | 200749.49 | 231238.94 | 312746.53 | 414350.66 | 467647.44 |
| 5 | 5 | 185323.62 | 183174.16 | 42245.13 | 101126.62 | 119643.88 | 179831.28 | 257534.99 | 297535.47 |
| 6 | 6 | 115095.83 | 112148.38 | 34470.09 | 47090.73 | 62050.98 | 109052.78 | 175104.99 | 207346.37 |
| 7 | 7 | 63111.11 | 61178.67 | 27062.56 | 15201.37 | 23777.43 | 57384.57 | 110712.29 | 139877.58 |
| 8 | 8 | 35808.04 | 34172.26 | 22439.30 | 2203.23 | 6395.38 | 29944.70 | 76607.83 | 105198.27 |
| 9 | 9 | 5459.61 | 7024.36 | 9832.25 | 0.52 | 42.00 | 3635.92 | 25493.62 | 47428.11 |
| 10 | 10 | 1119.08 | 1288.21 | 2979.28 | 0.00 | 0.00 | 88.33 | 6778.57 | 15143.68 |
| 11 | 11 | 580.33 | 520.49 | 1859.12 | 0.00 | 0.00 | 0.35 | 2916.03 | 9120.69 |
| 12 | 12 | 355.37 | 291.09 | 1339.00 | 0.00 | 0.00 | 0.00 | 1620.52 | 5968.27 |
| 13 | 13 | 210.61 | 181.11 | 1011.32 | 0.00 | 0.00 | 0.00 | 692.17 | 4626.77 |
| 14 | 14 | 116.42 | 89.88 | 646.81 | 0.00 | 0.00 | 0.00 | 166.63 | 2521.05 |
| 15 | 15 | 64.70 | 57.35 | 587.54 | 0.00 | 0.00 | 0.00 | 14.76 | 1589.64 |
| 16 | 16 | 32.12 | 21.60 | 412.65 | 0.00 | 0.00 | 0.00 | 0.00 | 268.31 |
| 17 | 17 | 12.77 | 11.03 | 221.02 | 0.00 | 0.00 | 0.00 | 0.00 | 7.88 |
| 18 | 18 | 0.00 | 0.15 | 14.42 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 19 | Tot. | 3327912.59 | 3307102.75 | 301357.52 | 2649495.72 | 2832084.92 | 3295960.39 | 3822522.40 | 4084974.14 |

———— Código de R ————

Se puede obtener una gráfica (Figura 4.2) de la estimación DCL *bootstrap*, en la cuál se muestra un análisis estadístico clásico, i.e., una gráfica de caja y el histograma para cada reserva por año calendario futuro.

———— Código de R ————

```
R>Plot.dclboot(DCL_boot)
```

———— Código de R ————

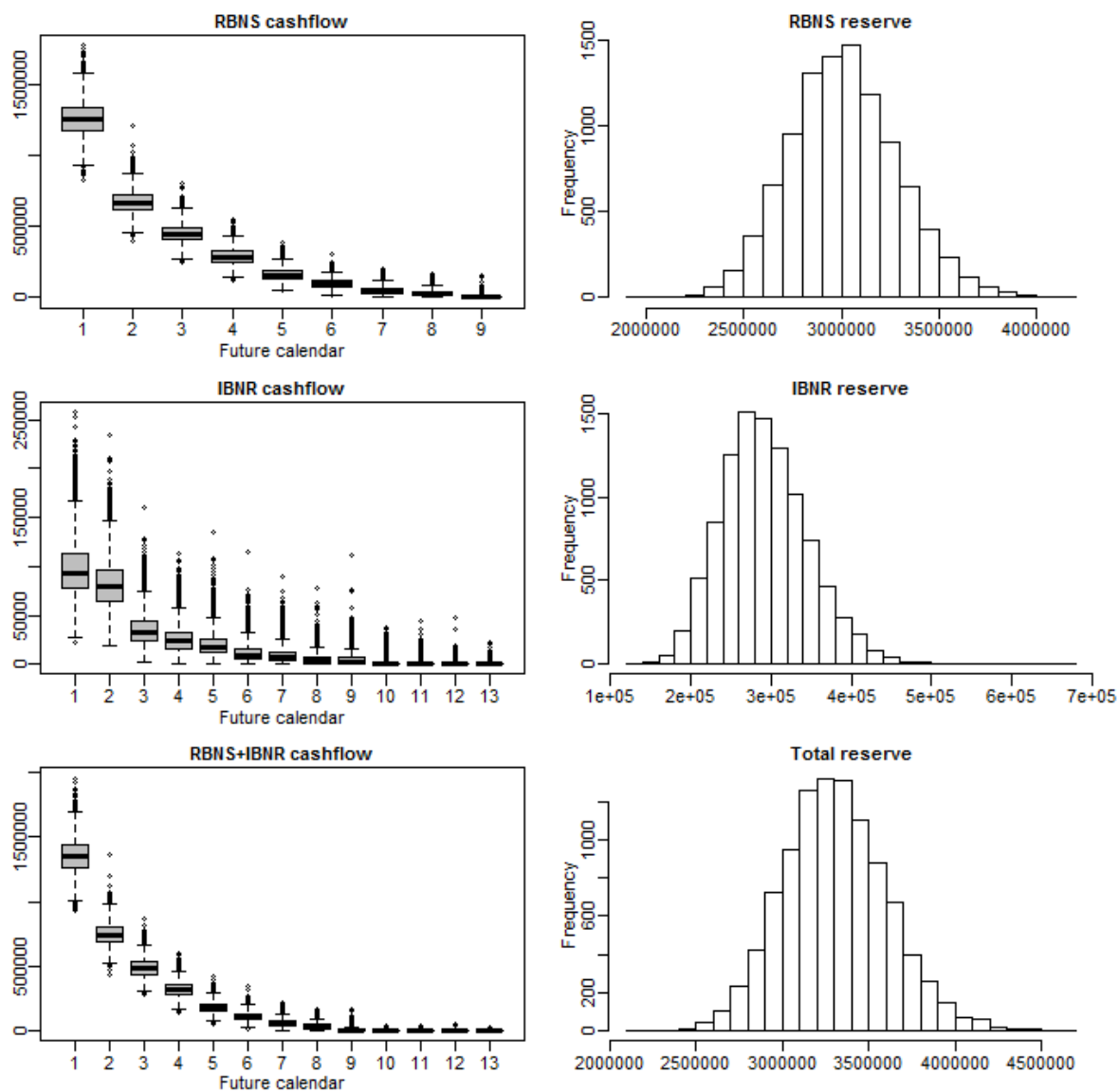


Figura 4.2: Gráfica de estimación DCL bootstrap

4.4. Modelo Chain-Ladder continuo

En esta sección, se introducirá una metodología aún más reciente, aunque muy teórica, para la estimación de la reserva de perdidas pendientes para los ramos de no-vida. En particular, este método re-formula la técnica actuarial clásica, Chain-Ladder, en un ti-

po histograma de aproximación y mejora esta técnica realizando un suavizamiento de *kernel*.

Una de las características en la mayoría de los métodos de reservas, incluyendo el Chain-Ladder, es que sus suposiciones están basadas en las reclamaciones acumuladas, donde el punto clave es notar que estos datos acumulados implican algún tipo de pre-suavizamiento. Sin embargo, los métodos para el cálculo de las reservas de reclamaciones con modelos “continuos” (paramétricos o no-paramétricos) siempre incluyen el uso de datos acumulados y por tanto se convierten en datos de tiempo discreto.

El modelo que se estudiará en esta sección es parecido al método CL clásico pero difiere en el sentido en el que no se supondrá que los datos son acumulados ya que utiliza datos recopilados en tiempo continuo. Por tal motivo se le conoce como “Chain-Ladder continuo”. Entonces se hará una aproximación al modelo Chain-Ladder clásico con ayuda de un histograma como estimador y se optimizará el nuevo modelo por medio de una técnica de suavizamiento. Además, se retomará el problema estadístico del CL como un problema de estimación de la densidad conjunta con soporte en el triángulo de reclamaciones.

El objetivo de este método es mostrar como el cálculo de la reserva con datos continuos se puede ver como una simple transición del método Chain-Ladder a un moderno método estadístico.

A partir de una técnica de suavizamiento no-paramétrico, se demostrará como el modelo clásico CL se puede considerar como la forma de un histograma encima de un triángulo, Jones (1989). El método CL original agrupa los datos y se desarrolla como un modelo multiplicativo, en este caso se simplificará el mismo modelo pero sin agrupar los datos. Cuando los datos no son agrupados, se puede trabajar como si se tuviera una densidad del triángulo.

Por otra parte, la densidad multivariada se puede estimar con métodos de suavizamiento lineales, y así ésta se puede aproximar por la densidad multiplicativa. La forma continua del Chain-Ladder se puede considerar como un modelo lineal generalizado.

4.4.1. El problema de estimación de la densidad para datos “agregados”

Primero se intentará replantear el problema de reservar en términos de un problema de estimación de la densidad multivariada y luego se desarrollará el método de *kernel* para estimar la densidad no-paramétrica.

Desde una perspectiva probabilista se puede tener mejor entendimiento de los métodos clásicos para el cálculo de la reserva, y por qué no, ser la entrada al desarrollo de mejores y modernos modelos.

La relación entre la densidad y los modelos de regresión motiva al desarrollo del método de estimación de la densidad, por ejemplo, el estimador lineal local. Esta relación se describirá primero para un conjunto unidimensional, como lo hizo Fan & Gijbels (1996).

Notación 4.4.1.

Sea X_1, \dots, X_n una muestra de variables aleatorias independientes e idénticamente distribuidas de una población \mathbf{X} con función de densidad continua, f .

Denótese:

- $a_1 < \dots < a_J$, como los puntos en el plano con la misma separación, definiendo $J - 1$ intervalos continuos;
- Para $j \in \{0, \dots, J - 1\}$, $B_j = (a_j, a_{j+1}]$ son los intervalos (continuos);
- a_1 y a_J , son los puntos extremos, que se eligen de tal forma que el intervalo $(a_1, a_J]$ esté contenido en el soporte de f ;
- Λ_J , es el rango del intervalo, donde $\Lambda_J = a_2 - a_1$;
- Para $j \in \{0, \dots, J\}$, x_j es el punto medio del intervalo, donde

$$x_j = \frac{(a_{j+1} + a_j)}{2};$$

- N_j , es el contador del intervalo, definido como el número de datos que caen dentro del intervalo B_j .

▽

Es claro que el contador N_j tiene una distribución binomial con parámetros n (tamaño del contador) y probabilidad de éxito p_j , $p_j = \int_{B_j} f(x)dx$.

Proposición 4.4.1.

Para $j \in \{0, \dots, J\}$, si $J \rightarrow \infty$ (o equivalentemente $\Lambda_J \rightarrow 0$, entonces

$$\mathbb{E} \left[\frac{N_j}{n\Lambda_J} \right] \approx f(x_j) \quad \text{y} \quad \text{Var} \left(\frac{N_j}{n\Lambda_J} \right) \approx \frac{f(x_j)}{n\Lambda_J}.$$

Observación 4.4.1.

Nótese que la estimación de la densidad puede considerarse como un modelo de regresión heterocedástico basado en los datos, $\left(x_j, \frac{N_j}{n\Lambda_J}\right)$, $j \in \{0, \dots, J\}$, que son aproximadamente independientes. Equivalentemente el modelo de regresión para el contador N_j se puede escribir como

$$N_j = r(x_j) + \varepsilon_j, \quad (4.4.1)$$

donde $r(\cdot) = n\Lambda_J f(\cdot)$ es la función de regresión.

A partir de una regresión estimada, la densidad objetivo se puede estimar como

$$\hat{r}(\cdot) = \hat{f}(\cdot)r(\cdot)n\Lambda_J.$$

▽

Para el caso bidimensional, el método Chain-Ladder se centra en la estimación de la densidad a través de la fórmula de regresión (4.4.1). Específicamente, el método CL estima de manera multiplicativa una densidad bi-dimensional, f , con soporte en el triángulo \mathcal{I} .

Los intervalos se construyen como planos de la forma $B_{i,j} = (a_{1,i}, a_{1,i+1}] \times (a_{2,j}, a_{2,j+1}]$ para $i, j \in \mathcal{I}$ con longitud constante $\Delta = a_{1,2} - a_{1,1} = a_{2,2} - a_{2,1}$. Es posible tener diferentes longitudes del intervalo en cada dimensión, pero se supondrá que es la misma para fines de exposición.

Entonces, dado el contador $N_{i,j}$ (número de datos que caen en $B_{i,j}$), el problema de regresión se puede formular como

$$N_{ij} = r(z_{i,j} + \varepsilon_{i,j}), \quad (4.4.2)$$

basado en los datos, $\{(z_{i,j}, N_{i,j}), i, j \in \mathcal{I}\}$. La covariable bidimensional $z_{i,j} = (x_i, y_j)$, se define de tal modo que x_i y y_j son los puntos medios de los intervalos $(a_{1,i}, a_{1,i+1}]$ y $(a_{2,j}, a_{2,j+1}]$, respectivamente, para $i, j \in \mathcal{I}$.

La función de regresión se relaciona con la densidad por medio de $r(\cdot, \cdot) = n\Lambda_J^2 f(\cdot, \cdot)$.

Ahora, si se supone que la función de regresión, r , es desconocida y de la forma $r(\cdot, \cdot) = r_1(\cdot)r_2(\cdot)$, este problema se puede resolver con el uso de modelos lineales generalizados.

El modelo Poisson para conteo de reclamaciones como se define en (4.3.6), se deriva de la aproximación a una distribución Poisson a partir de una distribución binomial. Por tanto, cada contador $N_{i,j} \sim \text{Binom}(n, p_{i,j})$, donde n es el número total de reclamaciones por año de accidente y $p_{i,j} = \int_{B_{i,j}} f(z)dz$. Entonces la distribución de $N_{i,j}$ se puede aproximar como

Poisson(np_{ij}). Evidentemente se tendrá una mejor estimación del modelo CL para valores más grandes de n y valores más pequeños de p_{ij} .

Ahora se cambiará la perspectiva de un modelo de regresión, el cual fue útil cuando los datos son agregados, a un modelo continuo donde la función objetivo es una densidad continua bidimensional.

Como ya se había discutido anteriormente, el método CL se puede definir como una función de regresión a partir de los datos agrupados. Aquí se estudiará como obtener una versión continua del modelo Chain-Ladder y a partir de éste, proporcionar un estimado de las reclamaciones pendientes con mejores propiedades estadísticas.

El método que se desarrollará se puede aplicar en datos agregados (como lo hace el método CL), sin embargo, su estudio sólo cubrirá el caso de reclamaciones individuales.

Se comenzará con el caso más simple, donde se describirá cómo a partir de un simple histograma se puede desarrollar una densidad de *kernel*.

4.4.2. Histograma con suavizamiento de *kernel*

Los métodos de *kernel* para la estimación de la densidad surgen de la idea intuitiva y natural del estimador histograma. La aplicación de éste en reservas depende de recordar que el método clásico Chain-Ladder se deriva de un histograma estructurado en un triángulo.

El histograma es la aproximación no-paramétrica más simple para la estimación de la función de densidad ya que separa los datos en distintos intervalos que no se duplican y construye barras con medidas definidas en proporción o número de las observaciones que caen en cada intervalo. Esta proporción da un estimado de la función de densidad de probabilidad en el punto medio del intervalo.

Se comenzará con la estimación para el caso univariado y después se extenderá al caso bivariado.

Supóngase una muestra de variables aleatorias, X_1, \dots, X_n de una población \mathbf{X} con función de densidad continua f . Considérense $J - 1$ intervalos continuos $B_j = (a_j, a_{j+1}]$ divididos en el soporte de f , con longitud Δ constante, sean x_j los puntos medios para $j \in \{0, \dots, J - 1\}$. La medida de la barra del histograma con base B_j proporciona el estimado de la función de densidad en el punto medio, x_j . Así, un estimador de la densidad f en un punto x_0 de su soporte se puede obtener como el límite del cociente entre la densidad

de un conjunto de puntos y el tamaño del conjunto, i.e.

$$\lim_{\Lambda_J \rightarrow 0} \frac{P(X \in B_j)}{\Lambda_J} = f(x) \text{ si } x \in B_j (j = 1, \dots, J-1). \quad (4.4.3)$$

A partir de esta ecuación, la expresión del estimador histograma en un punto x_0 del soporte de f está dada como

$$\hat{f}_{hist}(x_0) = \frac{\sum_{i=1}^n I\{X_i \in B_j\}}{n\Lambda_J} \text{ si } x_0 \in B_j (j = 1, \dots, J).$$

Nótese que el histograma no es una función continua, pero tiene saltos en algunos puntos del cuadro y derivada cero en cualquier otro sitio. Esto genera estimados que no son estéticamente deseables pero, para un observador sin experiencia podría generar una buena impresión. De hecho, la forma del histograma está fuertemente influenciada por la posición de los puntos medios.

Para el caso bidimensional, supóngase que Z_1, \dots, Z_n son variables aleatorias independientes e idénticamente distribuidas de una población $\mathbf{Z} = (X, Y)$ con función de densidad bivariada, f , con soporte dentro de los cuadros $B_{i,j} = (a_{1,i}, a_{1,i+1}] \times (a_{2,j}, a_{2,j+1}]$ y longitud constante $\Lambda_J = a_{1,2} - a_{1,1} = a_{2,2} - a_{2,1}$. Siguiendo el mismo procedimiento que el caso univariado, el estimador histograma definido en los puntos $z_0 = (x_0, y_0)$ del soporte de f está dado como

$$\hat{f}_{hist}(z_0) = \frac{v(z_0)}{n\Lambda_J^2},$$

donde $v(z_0)$ es el número de datos de la muestra que caen dentro del cuadro que contiene z_0 . Un estimador de la densidad de *kernel* supera el problema del estimador histograma respecto a la posición de los intervalos y proporciona un estimador suavizado de la densidad objetivo.

Definición 4.4.1. (*Densidad de kernel*)

Para un punto $z_0 = (x_0, y_0)$ en el soporte de f , se define a la densidad de *kernel* como

$$\hat{f}_h = |h|^{-1} \sum_{i=1}^n \kappa_h(z_0 - Z_i), \quad (4.4.4)$$

donde $\kappa_h(\cdot)$ es una *kernel* bidimensional y $h = (h_1, h_2)^\top \in \mathbb{R}_+^2$ es el parámetro de la amplitud de la barra con $|h| = h_1 h_2$.

Aquí, se usa la forma multiplicativa de *kernel* dada por $\kappa_h(x, y) = \kappa_{h_1}(x) \kappa_{h_2}(y)$ con $\kappa_{h_1}(x) = h_1^{-1} \kappa(x/h_1)$, $\kappa_{h_2}(y) = h_2^{-1} \kappa(y/h_2)$ y κ una función de densidad simétrica univariada.

Observación 4.4.2.

Esta estructura multiplicativa coincide con las suposiciones del método CL. ∇

Se puede ver el problema de *reserving* como un problema en la estimación de la densidad dentro de la región límite definida por el triángulo de reclamaciones. Puesto que el objetivo es predecir la densidad en todo el cuadrante que incluye el futuro en la parte inferior del triángulo, se supondrá una estructura multiplicativa y se usará el método de integración marginal por Linton&Nielsen (1995) para dar un estimador lineal multiplicativo de la densidad.

4.4.3. El estimador lineal local no estructurado en las observaciones del triángulo

Nielsen (1999) extendió el principio de la estimación lineal local a la estimación de la densidad multivariada no paramétrica con regiones límite arbitrarias.

Sea f una función de densidad bidimensional con soporte en el triángulo $\mathcal{I} = \{z = (x, y)^\top | 0 \leq x, y \leq I, x + y \leq I\}$, $I > 0$. El estimador lineal local de Nielsen se define en cada punto $z_0 = (x_0, y_0)^\top \in \mathcal{I}$ con solución $\hat{\Theta}_0$ del siguiente problema de minimización

$$\begin{pmatrix} \hat{\Theta}_0 \\ \hat{\Theta}_1 \end{pmatrix} = \arg \min \left\{ \lim_{b \rightarrow 0} \int_{\mathcal{I}} \left[\tilde{f}_b(z) - \hat{\Theta}_0 - \hat{\Theta}_1^t (z_0 - z) \right]^2 \kappa_h(z - z_0) dz \right\}, \quad (4.4.5)$$

donde $\tilde{f}_b(z) = n^{-1} (b_1, b_2)^{-1} \sum_{i=1}^n \kappa_b(z - z_0)$ es el estimador estándar de *kernel* en (4.4.4) en el punto z con parámetro de amplitud $b = (b_1, b_2)^t \in \mathbb{R}_+^2$ y κ una *kernel* bi-dimensional.

Nótese que este estimador sólo está definido en el triángulo de observaciones y por lo tanto no es apropiado para propósitos de pronóstico.

4.4.4. El estimador de la densidad lineal local

Ahora se supondrá que la densidad objetivo en el cuadro $\mathcal{S} = \{z = (x, y)^t | 0 \leq x, y \leq I\}$ es multiplicativa, i.e., $f(x, y) = f_1(x) f_2(y)$. El método de integración marginal, Linton & Nielsen (1995), se puede extender al problema de estimación de la densidad siguiendo los siguientes dos pasos:

1. A partir de los datos observados, estimar la densidad conjunta en el conjunto de observaciones \mathcal{I} con el estimador $\hat{f}_h^j(x, y)$, es tal que los estimadores lineales locales están dados por (4.4.5).

2. Supóngase que la densidad objetivo es de la forma $f(x, y) = f_1(x)f_2(y)$, entonces los estimados de $f_1(x)$ y $f_2(y)$ están dados por la minimización

$$\min_{f_1, f_2} \int_{\mathcal{I}} \left(\tilde{f}_h^{\mathcal{I}}(x, y) - f_1(x)f_2(y) \right)^2 \omega(x, y) dx dy, \quad (4.4.6)$$

donde $\omega(x, y)$ es una función ponderadora.

Si se tomará $\omega(x, y) = \left(\tilde{f}_h^{\mathcal{I}}(x, y) \right)^{-1}$, la minimización del paso 2 se puede desarrollar utilizando el siguiente algoritmo iterativo:

- (i) Considérese un estimador inicial del componente f_1 denotado por $\hat{f}_1^{(0)}$. Sea $\hat{f}^{(0)}$ el estimador no-estructurado para la densidad en \mathcal{I} definida en el paso 1.
- (ii) Usando $\hat{f}_1^{(0)}$, $f(x, y) \approx \hat{f}_1^{(0)}(x)f_2(y)$, entonces

$$\int_{\mathcal{I}_y} f(x, y) dx \approx f_2(y) \int_{\mathcal{I}_y} \hat{f}_1^{(0)}(x) dx$$

con $\mathcal{I}_y = \{x | (x, y) \in \mathcal{I}\}$.

Entonces el estimador para f_2 es

$$\hat{f}_2^{(1)}(y) = \frac{\int_{\mathcal{I}_y} \hat{f}^{(0)}(x, y) dx}{\int_{\mathcal{I}_y} \hat{f}_1^{(0)}(x) dx}.$$

- (iii) Utilizando $\hat{f}_2^{(1)}$, se obtiene el estimador de f_1 como

$$\hat{f}_1^{(1)}(x) = \frac{\int_{\mathcal{I}_x} \hat{f}^{(0)}(x, y) dy}{\int_{\mathcal{I}_x} \hat{f}_2^{(1)}(y) dy},$$

con $\mathcal{I}_x = \{y | (x, y) \in \mathcal{I}\}$.

- (iv) Repitiendo los pasos (ii)-(iii) hasta obtener el criterio de convergencia deseado.

A partir de este algoritmo se obtienen estimados en un punto del cuadro $\mathcal{S} = \{z = (x, y)^\top | 0 \leq x, y \leq I\}$. El único requisito es elegir un función de *kernel*, κ , y el parámetro de amplitud, h , de tal forma que el nivel de suavizamiento sea el adecuado. La elección de κ se puede hacer mediante argumentos prácticos o teóricos, ya que el impacto en el estimador es menor. A diferencia de la elección del parámetro de amplitud donde se afecta directamente el estimador de *kernel*.

4.4.5. Corrección multiplicativa sesgada

En esta parte, se considera una segunda mejora al estimador de *kernel* a través de técnicas de reducción de sesgo. Se sabe que los métodos de *kernel* proporcionan estimadores sesgados. Para la estimación de la reserva, el sesgo se debe corregir ya que de no ser así lleva a cálculos de la reserva erróneos con consecuencias importantes en la solvencia de las aseguradoras.

Existen varios métodos de reducción de sesgo para mejorar las estimaciones de *kernel* en los puntos de la curva. Aquí, se considerará el método de corrección de sesgo multiplicativo (MBC) (Jones, Linton & Nielsen (1995)) para la estimación de la densidad univariada. El estimador se genera en dos pasos. Primero se define un estimador de corrección de sesgo multiplicativo para la densidad en el triángulo de observaciones como un estimador MBC no-estructurado. Segundo, se aplica el método de integración marginal para generar un estimador de la densidad MBC estructurado.

Considérese el estimador lineal local no-estructurado (4.4.5). Denótese a este estimador como $\widehat{f}_{LL,h}^{\mathcal{I}}$ con soporte en \mathcal{I} , entonces se define al estimador MBC no-estructurado como

$$\widehat{f}_{MBC,h}^{\mathcal{I}}(z) = \widehat{f}_{LL,h}^{\mathcal{I}}(z)\widehat{g}_{LL,h}^{\mathcal{I}}(z), \quad (4.4.7)$$

donde $\widehat{g}_{LL,h}^{\mathcal{I}}$ es el estimador lineal local de el cociente $\frac{f(z)}{\widehat{f}_{LL,h}^{\mathcal{I}}(z)}$ obtenido a partir la minimización de la siguiente expresión en Ψ_0

$$\min_{\psi_0, \psi_1} \left\{ \lim_{b \rightarrow 0} \int_{\mathcal{I}} \left[\widetilde{f}_b(z) - (\Psi_0 - \Psi_1^t(z_0 - z))\widehat{f}_{LL,h}^{\mathcal{I}}(z) \right]^2 \kappa_h(z - z_0) dz \right\}. \quad (4.4.8)$$

Ahora, a partir del estimador definido en la ecuación (4.4.7) y utilizando un método como el descrito en la Subsección 4.4.4, se obtiene el estimador de las densidades f_1 y f_2 junto con el estimador MBC estructurado.

4.4.6. Elección del grado de suavizamiento

Es necesario hacer una elección del parámetro de amplitud a través de un criterio de optimalidad. Sin embargo, para una estimación de *kernel*, la elección teórica de la amplitud óptima no es viable, por lo tanto se necesita de una estimación razonable de este parámetro.

Para los estimadores de la densidad de *kernel* definidos anteriormente, la amplitud es un parámetro bidimensional, $h = (h_1, h_2)^\top$, que controla el grado de suavizamiento en cada dirección. Específicamente $h_1, h_2 \in (0, \infty)$, donde h_1 es el grado de suavizamiento en la dirección de año de suscripción y h_2 en dirección del desarrollo .

Existen varios métodos para la elección de la amplitud en una densidad bivariada. Uno de los más simples y comunes es el método *cross-validation*, el cual estima la amplitud óptima de una muestra de datos. En este trabajo se proponen dos candidatos con estructura multiplicativa para la estimación de la densidad a partir de una muestra de los datos observados en el triángulo, el estimador lineal local y una versión del MBC. Aquí se propone simplemente usar el *cross-validation* para encontrar un “buen” estimador no-estructurado de la densidad en el triángulo de observaciones, y a partir de éste calcular las correspondientes densidades estructuradas siguiendo el método de la Subsección 4.4.4.

Ya sea para el estimador de la densidad no-estructurado en el triángulo \mathcal{I} , el estimador lineal local $\hat{f}_{LL,h}^{\mathcal{I}}$ o el estimador MBC $\hat{f}_{MBC,h}^{\mathcal{I}}$, el valor del *cross-validation* se define como

$$CV(h) = \int_{\mathcal{I}} \hat{f}_h^{\mathcal{I}}(z)^2 dz - 2 \sum_{i=1}^n \int_{\mathcal{I}} \hat{f}_h^{\mathcal{I},[-i]}(z) d\tilde{F}_n(z),$$

donde $\hat{f}_h^{\mathcal{I},[-i]}$ es la versión *leave-one-out* de los estimadores $\hat{f}_h^{\mathcal{I}}$ y \tilde{F}_n es la función de distribución empírica de la muestra. El parámetro de amplitud *cross-validation* resulta de la minimización de $CV(h)$. La forma multiplicativa del estimador *cross-validation* se define de una manera similar.

4.4.7. Predicción/pronóstico de las reclamaciones pendientes

Como ya se mencionó, el método Chain-Ladder continuo no proporciona predicciones fuera de la estructura triangular definida. Sin embargo, en el estudio del cálculo de reservas no basta con dar estimados para las reclamaciones individuales sino se busca proporcionar una forma de predecir las reclamaciones futuras.

De acuerdo al método CL, ya se mostró que para estimar las reclamaciones pendientes se suman los valores pronosticados de las reclamaciones en el triángulo inferior, i.e. para $j \in \{1, \dots, J-1\}$

$$\hat{C}_{i,j} = \sum_{i=1}^J \hat{N}_{i,I-i+j+1} = \sum_{i=1}^J \hat{\alpha}_i \hat{\beta}_{i,I-i+j+1}.$$

Para el modelo continuo, estas predicciones se definen como

Definición 4.4.2. (*Predicción de reclamaciones*)

Para el tiempo $I + t$ con $t \in (0, I)$, se define

$$\widehat{C}_t = \tau \int_0^J \widehat{f}_1(x) \widehat{f}_2(J - x + t) dx.$$

τ es el total de riesgos expuestos en \mathcal{I} .

Las predicciones se derivan del método CL y el modelo de estimación continuo. Sólo como una extensión de la forma en la que se “agrega” la información.

Conclusión

En este trabajo se estimó la reserva de reclamaciones pendientes con diferentes metodologías, clásica (Chain-Ladder) y Bayesiana, y se quiso hacer una comparación para saber en que casos cual método es más conveniente. Al final del trabajo la postura que se tomó fue la siguiente.

Una primera impresión de los modelos mecánicos (razón, Chain-Ladder y Bornhuetter-Ferguson) fue que no resultarían “buenas” estimaciones sobre las reclamaciones futuras, al tratarse de algoritmos recursivos basados de información pasada, “promedios ponderados” como factores de desarrollo y parámetros a *priori* propuestos. Pero esta idea determinista del método Chain-Ladder fue remplazada con el modelo de Mack suponiendo una distribución libre y así proponer una forma de medir que tan buena es la estimación.

La idea de solucionar este problema se basa en calcular el error cuadrático medio de predicción y éste se conforma de dos conceptos, el proceso varianza y el error de estimación del parámetro, entonces el objetivo fue minimizar el ECMP. Como no es posible reducir el proceso de varianza ya que es la aleatoriedad de los datos, lo que se minimiza es el error de estimación. La manera de evaluar la calidad de estimación de los parámetros y predictores es a partir del primer y segundo momento.

Con el uso de metodologías Bayesianas subyacentes, en el Capítulo 3, el método Benktander-Hovinen propone estimar la siniestralidad última con mezclas de estimadores de siniestralidad última obtenidos para los métodos CL y BF para estimar la siniestralidad última, asignándoles un peso a cada uno. Cabe resaltar que lo importante de esta metodología es mostrar como la mezcla en función del peso que pondera a los estimadores converge al estimador de siniestralidad última del método CL, y motivó al estudio de modelos de credibilidad para obtener los pesos óptimos de estos estimadores.

El estudio del método de simulación *bootstrap*, en primera instancia, era adicional para obtener muestras de las observaciones y simular una distribución empírica a *priori* de los datos individuales necesaria para las hipótesis Bayesianas. Lo novedoso de este trabajo fue la combinación del método *bootstrap* con la teoría Bayesiana en el método ABC porque si

bien, ya se había dado una forma de evaluar la calidad de estimación de las predicciones (por momentos) pero no se tenían expresiones cerradas debido a la distribución libre, de este modo con el algoritmo *bootstrap* y cierta suposición distribucional a *priori* se da una forma de obtener la densidad predictiva de las reclamaciones, con la desventaja es que el uso de este método requiere de un conocimiento computacional avanzado para la estimación pero la ventaja es que puede tratar con cualquier verosimilitud aunque sea de compleja expresión.

En los capítulos 2 y 3 se estudiaron modelos que corregían el error de estimación paramétrico y no-paramétrico del modelo CL, el capítulo 4 es una extensión del método CL para corregir el error estadístico en los datos, difícil de apreciar.

El método MCL muestra la diferencia de estimación entre considerar los datos de reclamaciones pagadas e incurridas por separado o considerar un cociente entre éstas. El objetivo fue analizar la dependencia entre estos dos tipos de datos con los coeficientes de correlación entre los factores de desarrollo y los residuales de los *ratios*.

El método Double Chain-Ladder y el Chain-Ladder continuo corrigen el error estadístico que surge de considerar las reclamaciones acumuladas para la estimación de reclamaciones, en particular el DCL retoma la técnica antigua para la estimación de las reclamaciones futuras con el método Poisson para conteo de reclamaciones, donde se demuestra que coincide con los estimadores Chain-Ladder.

Después de hacer este análisis con las distintas metodologías no sería correcto mostrar una postura sobre cual es el mejor método para estimar la reserva porque depende de la situación y los datos que proporcione la aseguradora, la importancia de este trabajo es observar que la mayoría de los métodos siguen una estructura del modelo Chain-Ladder o convergen a él, y se describe donde sería factible aplicar cualquiera de los métodos.

En general, el modelo Chain-Ladder solo no es muy común utilizarlo, suele “combinarse” con otros triángulos (MCL y DCL) o incorporar información externa (Bayesianos y credibilidad) para precisar la estimación de las reclamaciones futuras.

El paquete Chain-Ladder de R proporcionó una forma de implementación de los métodos más práctica, y así obtener una estimación numérica de las reservas, además es una forma fácil de compararlas y decidir cual método es el más apropiado para el conjunto de datos considerado.

En este caso, se dio una conclusión en particular sobre los datos observados dado que como en algún momento se mencionó, las interpretaciones no siempre serán las mismas a pesar de que los datos sean los mismos.

Apéndices

Apéndice A

Aspectos generales del algoritmo Metropolis-Hastings

A.1. Cadenas de Markov

Definición A.1.1. (*Cadenas de Markov*)

Una cadena de Markov $\{X_i\}_{i=0}^t$ es un conjunto de variables aleatorias X_0, X_1, \dots, X_t tales que la probabilidad de X_t dado las variables pasadas sólo depende de la anterior, i.e.,

$$\mathbb{P}(X_t|X_0, \dots, X_{t-2}, X_{t-1}) = \mathbb{P}(X_t|X_{t-1}),$$

a esta distribución se le conoce como transición o Markov *kernel*, K .

Primero se describirá el funcionamiento del método Monte Carlo cadenas de Markov (MCMC). Dada una densidad objetivo f , se puede obtener una Markov *kernel* con distribución estacionaria f , y así se genera una cadena de Markov (X_t) a partir de K , de tal modo que la distribución límite de X_t es f y se pueden aproximar sus integrales de acuerdo al Teorema Ergódico. Robert & Casella (2010).

La dificultad de este proceso es la construcción de K ya que está asociada a una distribución arbitraria de f . Pero existen métodos para obtener dicha K que son teóricamente válidos para cualquier densidad f . El algoritmo Metropolis-Hastings y el muestreo de Gibbs (que a continuación se estudiarán), son ejemplos de estos.

A.2. Algoritmo Metropolis-Hastings

A partir de la definición anterior, el algoritmo Metropolis-Hastings (M-H) que a continuación se describirá genera variables correlacionadas de una cadena de Markov.

Una característica por la cual es importante considerar este método, es que a partir de las diferentes propiedades de convergencia de una cadena de Markov, se pueden hacer suposiciones que simplifiquen la dificultad de muestreo para ciertos casos. Pues por una parte, dado que los requisitos con los que tiene que cumplir la densidad f son mínimos, se permite hacer ajustes cuando es el caso donde se tiene muy poca información de ésta. Otra razón, es que la estructura Markoviana puede descomponer problemas de gran dimensión en un conjunto de pequeños problemas más sencillos de resolver.

Para el algoritmo M-H, dada la función objetivo f , asociada a un proceso condicional con densidad $q(y|x)$, de fácil simulación, y además q puede ser arbitraria, si cumple que para un valor fijo de X la distribución $\frac{f(y)}{q(y|x)}$ es conocida y que $q(\cdot|x)$ tiene la suficiente dispersión en el soporte de f .

La principal característica del algoritmo Metropolis-Hastings es que dada cualesquiera q , se puede construir a K tal que f es una función estacionaria.

A.2.1. El algoritmo Monte Carlo para cadenas de Markov

El algoritmo M-H relacionado con la densidad objetivo f y la densidad condicional q , genera una cadena de Markov, (X_t) , a través de la siguiente transición de *kernel*.

Algoritmo A.2.1. (*Algoritmo Metropolis-Hastings*)

Dada X_t ,

1. Genera $Y_t \sim q(y|x_t)$.

2.

$$X_{t+1} = \begin{cases} Y_t & \text{con probabilidad } \rho(x_t, Y_t) \\ x_t & \text{con probabilidad } 1 - \rho(x_t, Y_t), \end{cases}$$

donde $\rho(x, y) = \min \left\{ \frac{f(y) q(x|y)}{f(x) q(y|x)}, 1 \right\}$.

La distribución q es conocida como la distribución candidata y la probabilidad $\rho(x, y)$ como la probabilidad de aprobación del modelo M-H. Esta se puede distinguir por la tasa

de aprobación, que es el promedio de las probabilidades de aprobación de las iteraciones, i.e.,

$$\bar{\rho} = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^T \rho(X_t, Y_t) = \int \rho(x, y) f(x) q(y|x) dy dx$$

Observación A.2.1.

Esta cantidad evalúa el desarrollo del algoritmo. ∇

El Algoritmo A.2.1 satisface la condición de balance $f(x)K(y|x) = f(y)K(x|y)$.

A.2.2. El algoritmo Metropolis-Hastings independiente

Como ya se menciona antes, para el método Metropolis-Hastings, se necesita una distribución propuesta q que sólo depende del estado actual de la cadena. Si ahora se desea que la distribución q sea independiente de este estado, i.e., $q(y|x) = g(y)$, entonces se tiene un caso especial del algoritmo original.

Algoritmo A.2.2. (*Algoritmo Metropolis-Hastings independiente*)

Dada X_t ,

1. Genera $Y_t \sim g(y)$.

2.

$$X_{t+1} = \begin{cases} Y_t & \text{con probabilidad } \min \left\{ \frac{f(Y_t)g(x_t)}{f(x_t)g(Y_t)}, 1 \right\} \\ x_t & \text{e.o.c.} \end{cases}$$

Este método se puede interpretar como una simple generalización del método aceptar-rechazar en el sentido que la distribución propuesta es la misma que la densidad g como en éste método. Entonces los valores propuestos para Y_t son los mismos.

El estudio del algoritmo M-H es interesante, pero la implementación es más complicada por la dificultad de construcción de la función propuesta, ya que esta función está fuertemente influenciada por el desarrollo del algoritmo.

A.3. Muestreo de Gibbs

Ahora se extiende el algoritmo MCMC estudiando otra clase de este método llamada muestreo de Gibbs.

La aportación de este algoritmo en específico es, primero, da mayor calibración desde la densidad objetivo y, segundo, descompone problemas complejos (de grandes dimensiones en la distribución objetivo, que para una camina aleatoria de M-H sería imposible de construir) en un conjunto de problemas más sencillos, por ejemplo, una serie de densidades de menor dimensión.

A.3.1. Muestreo Gibbs de dos etapas

El procedimiento del muestreo Gibbs de dos etapas, genera una cadena de Markov a partir de una distribución conjunta $f(x, y)$.

Es decir, si dos variables X y Y tienen una distribución conjunta $f(x, y)$, con sus correspondientes densidades condicionales $f(y|x)$ y $f(x|y)$, entonces las dos etapas del muestreo de Gibbs genera una cadena de Markov (X_t, Y_t) de acuerdo a los siguientes pasos:

Algoritmo A.3.1. (*Muestreo Gibbs de dos etapas*)

Supóngase $X_0 = x_0$, para $t = 1, 2, \dots$ se genera

1. $Y_t \sim f_{Y|X}(\cdot|x_{t-1})$.
2. $X_t \sim f_{X|Y}(\cdot|y_t)$.

El Algoritmo A.3.1 es de fácil implementación siempre y cuando la simulación de las condicionales sea factible. Esto también es fácil de ver porque, si (X_t, Y_t) tienen una distribución f , entonces X_{t+1}, Y_{t+1} también, esto porque ambos pasos de iteración t -ésimos utilizan la simulación de sus condicionales originales. La convergencia de la cadena de Markov se garantiza a menos que los soportes de las condicionales no coincidan.

A.3.2. Muestreo Gibbs multi-etapas

Existe una extensión del muestreo Gibbs de dos etapas, que es el muestreo de Gibbs multi-etapas.

Supóngase que para alguna $p > 1$, la variable aleatoria $\mathbf{X} \in \mathcal{X}$ se puede escribir como $\mathbf{X} = (X_1, X_2, \dots, X_p)$, donde las X_i 's son componentes uni ó multidimensionales. Además, supóngase que se puede simular las correspondientes densidades condicionales f_1, \dots, f_p , i.e., para $i = 1, 2, \dots, p$

$$X_i|x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_p \sim f_i(x_i|x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_p).$$

Entonces el procedimiento para el muestreo de Gibbs multi-etapas está dado por la siguiente transición de X_t a X_{t+1} .

Algoritmo A.3.2. (*Muestreo de Gibbs multi-etapas*)

Para la iteración $t = 1, 2, \dots$, dado $x_1^{(t)}, \dots, x_p^{(t)}$ se genera

1. $X_1^{(t+1)} \sim f_1(x_1|x_2^{(t)}, \dots, x_p^{(t)})$.
2. $X_2^{(t+1)} \sim f_2(x_2|x_1^{(t)}, x_3^{(t)}, \dots, x_p^{(t)})$.
- ⋮
- p . $X_p^{(t+1)} \sim f_p(x_p|x_1^{(t)}, x_2^{(t)}, \dots, x_{p-1}^{(t)})$.

Las densidades f_1, \dots, f_p se conocen como condicionales completas, y una ventaja del muestreo Gibbs es que estas son las densidades utilizadas para la simulación. Por lo tanto, incluso en un problema de grandes dimensiones, todas las simulaciones se podrían convertir en un caso univariado.

Apéndice B

Familias conjugadas

Cuando se habla de inferencia Bayesiana, surge la idea de una distribución inicial (a *priori*) y una distribución final (a *posteriori*), la primera siempre es de elección arbitraria o basada en la experiencia pero esta elección no siempre resulta simple para la segunda ya que puede resultar una función no conocida, i.e., sin expresión cerrada. Sarabia & Gómez & Vázquez (2006).

Resultaría por tanto útil realizar una elección a *priori* basada en la simplicidad y conveniencia. Por tal motivo, existe una clase de distribuciones a *priori* conocidas como **familias conjugadas** que son muy considerables, en primera porque si la distribución a *priori* es de una familia conjugada, entonces la distribución a *posteriori* pertenecerá a la misma familia. En segundo lugar, las distribuciones conjugadas son lo bastante sencillas como para hacer una inferencia sin necesidad de un método computacional complejo.

Definición B.0.1. (*Familia conjugada*)

Supóngase una variable aleatoria X con distribución $f(x|\theta)$. Una familia de densidades a *priori* \mathcal{F} con parámetro θ , es conjugada para el muestreo $f(x|\theta)$ si para cualquier densidad a *priori* $\pi(\theta) \in \mathcal{F}$ se cumple que la densidad a *posteriori* $\pi(\theta|x) \propto L(x|\theta)\pi(\theta) \in \mathcal{F}$.

A continuación se ejemplificará esta definición con algunas distribuciones a *priori* pertenecientes a familias conjugadas.

Lema B.0.1. (*Muestreo Binomial*)

La familia de densidades a *priori* $Beta(\alpha, \beta)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución Binomial(m, θ).

Demostración:

Supóngase que la muestra X de tamaño n tiene distribución Binomial con parámetros m y θ , con función de verosimilitud está dada por

$$L(x|\theta) \propto \theta^{\sum_{i=1}^n x_i} (1 - \theta)^{nm - \sum_{i=1}^n x_i}.$$

Si la familia de densidades a *priori* es *beta*, entonces una distribución a *priori* de esta familia tiene la expresión

$$\pi(x|\theta) \propto \theta^{\alpha-1} (1 - \theta)^{\beta-1}.$$

Por lo tanto, dada una distribución inicial de esta familia, su distribución a *posteriori* será

$$\pi(\theta|x) \propto \theta^{\alpha+m \sum_{i=1}^n x_i - 1} (1 - \theta)^{nm - \sum_{i=1}^n x_i + \beta - 1} \sim Beta\left(\alpha + \sum_{i=1}^n x_i, nm - \sum_{i=1}^n x_i + \beta\right).$$

Lema B.0.2. (*Muestreo Binomial-negativa*)

La familia de densidades a *priori* $Beta(\alpha, \beta)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución Binomial-negativa(r, θ).

Demostración:

Supóngase que la muestra X de tamaño n tiene distribución Binomial-negativa con parámetros r y θ , con función de verosimilitud está dada por

$$L(x|\theta) \propto \theta^{nr} (1 - \theta)^{n\bar{x}}.$$

Si la familia de densidades a *priori* es *Beta*, entonces una distribución a *priori* de esta familia tiene la expresión

$$\pi(x|\theta) \propto \theta^{\alpha-1} (1 - \theta)^{\beta-1}.$$

Por lo tanto, dada una distribución *priori* de esta familia, su distribución *posteriori* será

$$\pi(\theta|x) \propto \theta^{\alpha+nr-1} (1 - \theta)^{n\bar{x}+\beta-1} \sim Beta(\alpha + nr, n\bar{x} + \beta).$$

Observación B.0.1.

En general, dada una distribución a *priori* $Beta(\alpha, \beta)$, para una muestra de tamaño n con cualquier distribución de muestreo de tipo Bernoulli, Binomial, geométrica o Binomial negativa, entonces la distribución a *posteriori* es $Beta(\alpha + x, \beta + n - x)$. ∇

Lema B.0.3. (*Muestreo Poisson*)

La familia de densidades a *priori* $Gamma(\alpha, 1/\beta)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución $Poisson(\lambda)$.

Demostración:

Supóngase que la muestra X de tamaño n tiene distribución $Poisson$ con parámetros λ , con función de verosimilitud está dada por

$$L(x|\lambda) \propto \lambda^{n\bar{x}} e^{-n\lambda},$$

donde, \bar{x} es la media muestral y una estadística suficiente, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$.

Si la familia de densidades a *priori* es *gamma*, entonces una distribución a *priori* de esta familia tiene la expresión

$$\pi(\lambda|x) \propto \lambda^{\alpha-1} e^{(-\beta\lambda)}.$$

Por lo tanto, dada una distribución a *priori* de esta familia, su distribución a *posteriori* será

$$\pi(\lambda|\bar{x}) \propto \lambda^{n\bar{x}+\alpha-1} e^{-(n+\beta)\lambda} \sim Gamma\left(n\bar{x} + \alpha, \frac{1}{(n + \beta)}\right).$$

Lema B.0.4. (*Muestreo Poisson*)

La familia de densidades a *priori* inversa Gaussiana $IG(\mu, \phi)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución $Poisson(\lambda)$.

Demostración:

Si la familia de densidades es inversa Gaussiana con parámetros $\mu, \phi > 0$, entonces una distribución a *priori* de esta familia tendrá la expresión

$$\pi(\lambda) \propto \lambda^{-\frac{3}{2}} \exp\left\{-\frac{\lambda}{2\phi} - \frac{\mu^2}{2\phi\lambda}\right\}.$$

Por lo tanto, dada una distribución a *priori* de esta familia, su distribución a *posteriori* será

$$\pi(\lambda|x) \propto \lambda^{n\bar{x}-\frac{1}{2}-1} \exp \left[-\frac{\lambda}{2\frac{\phi}{2n\phi+1}} - \frac{\left(\mu\sqrt{\frac{1}{2n\phi+1}}\right)^2}{2\lambda\frac{\phi}{2n\phi+1}} \right] \sim GIG \left(n\bar{x} - \frac{1}{2}, \mu\sqrt{\frac{1}{2n\phi+1}}, \frac{\phi}{2n\phi+1} \right).$$

que corresponde a una distribución Gamma inversa generalizada.

Lema B.0.5. (*Muestreo de datos Gamma*($\alpha, 1/\beta$), con α conocido)

La familia de densidades a *priori* para $\beta \sim \text{Gamma}(a, 1/b)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución $X \sim \text{Gamma}(\alpha, 1/\beta)$, con $\alpha > 0$ conocida.

Demostración:

Supóngase una muestra X de tamaño n donde $X|\beta \sim \text{Gamma}(\alpha, 1/\beta)$ con $\alpha > 0$ conocida, entonces la función de verosimilitud está dada por

$$L(x|\beta) \propto \beta^{n\alpha} e^{-n\beta\bar{x}}.$$

Si se elige una densidad a *priori* gamma para β , entonces la distribución a *posteriori* está dada por

$$\begin{aligned} \pi(\beta|x) &\propto \beta^{n\alpha} e^{-n\beta\bar{x}} \beta^{a-1} e^{-b\beta} \\ &\propto \beta^{n\alpha+a-1} e^{-(n\bar{x}+b)\beta} \sim \text{Gamma} \left(n\alpha + a, \frac{1}{(n\bar{x} + b)} \right). \end{aligned}$$

Lema B.0.6. (*Muestreo de datos exponenciales*)

La familia de densidades a *priori* para $\lambda \sim \text{Gamma}(\alpha, 1/\beta)$ es conjugada para $X|\lambda \sim \exp(\lambda)$.

Demostración:

Supóngase una muestra X de tamaño n donde $X|\lambda$ tiene una distribución exponencial con parámetro λ , entonces la función de verosimilitud está dada por

$$L(x|\lambda) \propto \lambda^n e^{-n\lambda\bar{x}},$$

que combinada con una distribución a *priori* $Gamma(\alpha, 1/\beta)$ con función de distribución

$$\pi(\lambda) \propto \lambda^{\alpha-1} e^{-\beta\lambda},$$

entonces la distribución a *posteriori* está dada por

$$\pi(\lambda|x) \propto \lambda^{n+\alpha-1} e^{-(n\bar{x}+\beta)\lambda}.$$

Lema B.0.7. (*Muestreo de datos uniformes*)

La familia de densidades a *priori* para $\theta \sim Pareto(\alpha, \beta)$ es conjugada para una muestra de variables aleatorias, $X = (X_1, \dots, X_n)$, con distribución $X \sim Unif(0, \theta)$.

Demostración:

Supóngase una muestra $x = (x_1, \dots, x_n)$ de tamaño n con distribución uniforme en el intervalo $(0, \theta)$, entonces $f(x_i|\theta) = \frac{1}{\theta} I_{(0,\theta)}(x_i)$, donde $I_A(s) = \begin{cases} 1, & \text{si } s \in A \\ 0, & \text{e.o.c.} \end{cases}$

La función de verosimilitud está dada por

$$L(x|\theta) \propto \theta^{-n} I_{(\underline{x}_i, \infty)}(\theta),$$

donde, $\underline{x}_i = \max(x_1, \dots, x_n)$ es una estadística suficiente.

Considerando una distribución a *priori* $Pareto(a, b)$ con función de distribución

$$\pi(\theta) \propto \theta^{a-1} I_{(b, \infty)}(\theta),$$

entonces la distribución a *posteriori* está dada por

$$\begin{aligned} \pi(\theta|x) &\propto \theta^{-n} I_{(\underline{x}_i, \infty)}(\theta) \theta^{a-1} I_{(b, \infty)}(\theta) \\ &\propto \theta^{-(a+n)-1} I_{(b_{\max}, \infty)}(\theta) \sim Pareto(a+n, b_{\max}) \end{aligned}$$

donde $b_{\max} = \max\{\underline{x}_i, b\}$.

La siguiente tabla muestra una recopilación de familias conjugadas

| Verosimilitud (X) | Distribución a <i>priori</i> (θ) | Distribución a <i>posteriori</i> (θX) |
|---|---|--|
| $X \sim \text{Poisson}(\theta)$ | $\theta \sim \text{Gamma}(a, 1/b)$ | $\text{Gamma}\left(a + n\bar{x}, \frac{1}{(b+n)}\right)$ |
| $X \sim \text{Poisson}(\theta)$ | $\theta \sim \text{IGamma}(\mu, \beta)$ | $\text{GIG}\left(n\bar{x} - \frac{1}{2}, \mu\sqrt{\frac{1}{2n\phi+1}}, \frac{\phi}{2n\phi+1}\right)$ |
| $X \sim \text{BinNeg}(r, \theta)$ | $\theta \sim \text{Beta}(a, b)$ | $\text{Beta}(a + nr, b + n\bar{x})$ |
| $X \sim \text{Bin}(m, \theta)$ | $\theta \sim \text{Beta}(a, b)$ | $\text{Beta}(a + n\bar{x}, b + mn - n\bar{x})$ |
| $X \sim \text{Gamma}(\alpha, 1/\theta)$ | $\theta \sim \text{Gamma}(a, 1/b)$ | $\text{Gamma}\left(a + n\alpha, \frac{1}{(b+n\bar{x})}\right)$ |
| $X \sim N(\theta, \sigma^2)$ | $\theta \sim N(\mu, S^2)$ | $N\left(\frac{\mu\sigma^2 + n\bar{x}S^2}{\sigma^2 + nS^2}, \frac{\sigma^2S^2}{\sigma^2 + nS^2}\right)$ |
| $X \sim \text{Unif}(0, \theta)$ | $\theta \sim \text{Pareto}(a, b)$ | $\text{Pareto}(a + n, b_M), b_M = \max\{x_i, b\}$ |

Apéndice C

Librería Chain-Ladder de R

La paquetería de R, `ChainLadder`, contiene algunos de los métodos estadísticos más comunes para el cálculo de la reserva de reclamaciones.

La instalación de la paquetería es fácil y se realiza desde el *Comprehensive R Archive Network* (CRAN).

```
_____ Código de R _____  
R> install.packages('ChainLadder')  
_____ Código de R _____
```

La instalación se realizó exitosamente si al ejecutar el comando `library(ChainLadder)` arroja el siguiente mensaje

```
_____ Código de R _____  
R> library("ChainLadder", lib.loc="~/R/win-library/3.2")  
Welcome to ChainLadder version 0.2.2  
  
Type vignette('ChainLadder', package='ChainLadder') to access  
the overall package documentation.  
  
See demo(package='ChainLadder') for a list of demos.  
  
More information is available on the ChainLadder project web-site:  
https://github.com/mages/ChainLadder  
  
To suppress this message use:  
suppressPackageStartupMessages(library(ChainLadder))  
_____ Código de R _____
```

C.1. Triángulos de datos

El triángulo de reclamaciones es donde las aseguradoras registran los siniestros para cada periodo de exposición (año de origen).

En ocasiones los datos no son guardados en la forma triangular, i.e., provienen de tablas donde los datos son recopilados en forma de lista. La paquetería ChainLadder contiene el comando `CSV-file` que puede convertir datos desde una tabla a la forma triangular y regresarla a su forma original de nuevo.

Por ejemplo, supóngase que los datos provienen de otro software, el comando `read.csv` generará una lista de todas las observaciones organizadas por año de origen, periodo de desarrollo, valor y línea de negocio (lob), además genera un resumen por cada categoría mencionada, i.e.,

Código de R

```
R> #Bases de datos
R> archivo.fuente<-file.path(system.file("Database",package="ChainLadder"),"Datos.csv")
R> datos<- read.csv(archivo.fuente)
R> #Muestra la tabla de datos
R> head(datos)
  origin dev  value lob
1  1977  1 153638 ABC
2  1978  1 178536 ABC
3  1979  1 210172 ABC
4  1980  1 211448 ABC
5  1981  1 219810 ABC
6  1982  1 205654 ABC
R> #Resumen por categoría
R> summary(datos)
  origin          dev          value          lob
Min.   : 1.0   Min.   : 1.000   Min.   : -17657   AutoLiab       :105
1st Qu.: 3.0   1st Qu.: 2.000   1st Qu.: 10324   GeneralLiab    :105
Median : 6.0   Median : 4.000   Median : 72468   M3IR5          :105
Mean   : 642.1 Mean   : 4.609   Mean   : 176632   ABC            : 66
3rd Qu.:1979.0 3rd Qu.: 7.000   3rd Qu.: 197716   CommercialAutoPaid: 55
Max.   :1991.0 Max.   :14.000   Max.   :3258646   GenIns         : 55
                (Other)          :210
```

Código de R

Se puede tener un subconjunto de los datos con los que se desea trabajar, en este caso se seleccionarán los datos ya considerados anteriormente, i.e., de la línea de negocio "GenIns".

Código de R

```
R> #Filtro de la base de datos
R> datos.genins<- subset(datos, lob in "GenIns")
R> #Muestra los datos en forma de lista del triángulo GenIns
R> head(datos.genins)
```

```

origin dev value lob
122 1 1 357848 GenIns
123 2 1 352118 GenIns
124 3 1 290507 GenIns
125 4 1 310608 GenIns
126 5 1 443160 GenIns
127 6 1 396132 GenIns

```

Código de R

Para transformar los datos de la tabla a su forma triangular se usa el comando *as.triangle*. Los argumentos que recibe esta función son la tabla de datos, el nombre de los encabezados para los años de origen y desarrollo, respectivamente y además la columna con los valores, i.e.,

Código de R

```

R> #Convertir tabla de datos a un objeto triángulo
R> genins.tri <- as.triangle(genins,origin="origin",dev="dev",value="value")
R> #Objeto triángulo
R> genins.tri
      dev
origin 1 2 3 4 5 6 7 8 9 10
1 357848 766940 610542 482940 527326 574398 146342 139950 227229 67948
2 352118 884021 933894 1183289 445745 320996 527804 266172 425046 NA
3 290507 1001799 926219 1016654 750816 146923 495992 280405 NA NA
4 310608 1108250 776189 1562400 272482 352053 206286 NA NA NA
5 443160 693190 991983 769488 504851 470639 NA NA NA NA
6 396132 937085 847498 805037 705960 NA NA NA NA NA
7 440832 847631 1131398 1063269 NA NA NA NA NA NA
8 359480 1061648 1443370 NA NA NA NA NA NA NA NA
9 376686 986608 NA NA NA NA NA NA NA NA NA NA
10 344014 NA NA NA NA NA NA NA NA NA NA

```

Código de R

Nótese que los datos se guardaron en un triángulo de observaciones incrementales.

Para regresar a la forma de tabla de nuevo a partir de este triángulo, se tiene el siguiente comando,

Código de R

```

R> #Convertir triángulo de datos a la forma tabla
R> genins.tab <- as.data.frame(genins.tri, na.rm=TRUE)
R> #Muestra de datos en forma de lista
R> head(genins.tab)
      origin dev value
1-1 1 1 357848
2-1 2 1 352118
3-1 3 1 290507
4-1 4 1 310608
5-1 5 1 443160
6-1 6 1 396132

```

Código de R

C.1.1. Importación de triángulos de otras fuentes

Cuando los datos analizados son propios, provenientes de otras fuentes donde fueron procesados.

R proporciona una forma de poder importar bases de datos almacenados en hojas de cálculo. A continuación se dará una forma de importar un archivo .csv.

El primer paso es abrir un nuevo libro de Excel y copiar el triángulo en la celda A1, con la primera columna siendo el primer año de origen de accidente y el primer renglón describe su periodo de desarrollo.

Asegúrese de que los datos no contengan encabezados y ningún tipo de formato, por ejemplo, comas para separar miles, celdas del tipo carácter, etc., y guárdese como un documento .csv(delimitado por comas).

Por ejemplo,

| | A | B | C | D | E | F | G | H | I | J | K | L |
|----|-----------|-----------|------------|------------|------------|------------|------------|------------|------------|------------|---|---|
| 1 | 5,946,975 | 9,668,212 | 10,563,929 | 10,771,690 | 10,978,394 | 11,040,518 | 11,106,331 | 11,121,181 | 11,132,310 | 11,148,124 | | |
| 2 | 6,346,756 | 9,593,162 | 10,316,383 | 10,468,180 | 10,536,004 | 10,572,608 | 10,625,360 | 10,636,546 | 10,648,192 | | | |
| 3 | 6,269,090 | 9,245,313 | 10,092,366 | 10,355,134 | 10,507,837 | 10,573,282 | 10,626,827 | 10,635,751 | | | | |
| 4 | 5,863,015 | 8,546,239 | 9,268,771 | 9,459,424 | 9,592,399 | 9,680,740 | 9,724,068 | | | | | |
| 5 | 5,778,885 | 8,524,114 | 9,178,009 | 9,451,404 | 9,681,692 | 9,786,916 | | | | | | |
| 6 | 6,184,793 | 9,013,132 | 9,585,897 | 9,830,796 | 9,935,735 | | | | | | | |
| 7 | 5,600,184 | 8,493,931 | 9,056,505 | 9,282,022 | | | | | | | | |
| 8 | 5,288,066 | 7,728,169 | 8,256,211 | | | | | | | | | |
| 9 | 5,290,793 | 7,648,729 | | | | | | | | | | |
| 10 | 5,675,568 | | | | | | | | | | | |
| 11 | | | | | | | | | | | | |

Ahora, abrir R e introducir el siguiente código que el cuál convertirá el archivo .csv en un objeto *triangle* y así ser manejado a partir de este software.

————— Código de R —————

```
R> #Comando de R
R> #Selección interactiva del archivo
R> miarchivo <- "path/to/folder/with/Triangulo.csv"
R> archivo.csv <- file.choose()

R> #Selección manual del archivo
R> #Ruta donde se almacena triángulo
R> archivo.csv <- "C:/Users/Daniel Alfaro/Tesis Alfaro/Aplicación en R/Triangulo.csv"

R> #Leer el archivo
R> tri <- read.csv(file=archivo.csv, header = FALSE)

R> #Conversión a un objeto triangle
```

```
R> tri <- as.triangle(as.matrix(tri))
R> # etiquetas de columnas
R> colnames(tri)<-c("1","2","3","4","5","6","7","8","9","10")
R>#etiquetas de renglones
R> rownames(tri)<-c("2001","2002","2003","2004","2005","2006","2007","2008","2009","2010")

R> #Objeto triangle
R> tri
      dev
origin 1      2      3      4      5      6      7      8      9      10
2001 5946975 9668212 10563929 10771690 10978394 11040518 11106331 11121181 11132310 11148124
2002 6346756 9593162 10316383 10468180 10536004 10572608 10625360 10636546 10648192      NA
2003 6269090 9245313 10092366 10355134 10507837 10573282 10626827 10635751      NA      NA
2004 5863015 8546239 9268771 9459424 9592399 9680740 9724068      NA      NA      NA
2005 5778885 8524114 9178009 9451404 9681692 9786916      NA      NA      NA      NA
2006 6184793 9013132 9585897 9830796 9935735      NA      NA      NA      NA      NA
2007 5600184 8493931 9056505 9282022      NA      NA      NA      NA      NA      NA
2008 5288066 7728169 8256211      NA      NA      NA      NA      NA      NA      NA
2009 5290793 7648729      NA      NA      NA      NA      NA      NA      NA      NA
2010 5675568      NA      NA      NA      NA      NA      NA      NA      NA      NA
```

Código de R

Si los datos provienen de una base de datos, R facilita su acceso utilizando paqueterías para SQL, por ejemplo, una conexión ODBC.

La paquetería **ChainLadder** incluye un demo donde se muestran casos cuando los datos son importados desde una base de datos ,por ejemplo MS Access,

```
R> demo(DatabaseExamples)
```

Código de R

Referencias

- Barnett, G. & Zenhwirth, B. (2000). *Best estimates for reserves*. Proc. CAS, Vol. LXXXVII, 245-321.
- Benktander, G. (1976). *An approach to credibility in calculating IBNR for casualty excess reinsurance*. The Actuarial Review, April 1976, Vol. 312,7.
- Bornhuetter, R. & Ferguson R. (1972). *The actuary and IBNR*. CAS, Vol. LIX, 181-195.
- Buchwalder, M. Bühlmann, H. Merz, M. & Wüthrich, M. (2006). *The mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy Revisited)*. Astin Bulletin, 36(2), 251-542.
- Bühlmann, H. (1983). *Estimation of IBNR reserves by the methods chain ladder, Cape Cod and complimentary loss ratio*. International Summer School 1983, unpublished.
- Bühlmann, H. & Gisler, A. (1983). *A course in Credibility Theory and its Applications*. Springer, Berlin.
- Bühlmann, H. & Straub, E. (1970). *Glaubwürdigkeit für Schadensätze*. Bulletin of the Swiss Association of Actuaries, 70:111-133.
- Casella, G. & Berger, R. (2008). *Statistical Inference*. Thomson Press.
- Charpentier, A. (2014). *Computational Actuarial Science with R*. Chapman & Hall/CRC The R Series.
- *Circular Única de Seguros*. (2016). Comisión Nacional de Seguros y Fianzas. 123-176.
- Comité de Solvencia II de Amis. (2014). *Solvencia II*. Asociación Mexicana de Instituciones de Seguros 11-29.
- Efron, B. (1979). *Bootstrap methods: another look at the jackknife*. Ann. Statist. 7/1, 1-26.

-
- England, P. & Verrall, R. (2008). *Stochastic Claims Reserving in General Insurance (with discussion)*. British Actuarial Journal 8, 443-544.
 - England, P. & Verrall, R. (2007). *Predictive distributions of outstanding liabilities in general insurance*. Ann. Actuarial Science 1/2, 221-270.
 - Hachemeister, C. & Stanard, J. (1975). *IBNR claims count estimation with static lag functions*. Astin Colloquium, Portimao, Portugal.
 - Hogg, R. & Craig, A. (1995). *Introduction to Mathematical Statistics*. Prentice Hall, Upper Saddle River, NJ.
 - Hovinen, E. (1981). *Additive and continuous IBNR*. Astin Colloquium 1981, Loen, Norway.
 - Jones, M. (1989). *Discretized and Interpolated Kernel Density Estimates*. Journal of the American Statistical Association, 84(407), 733-741.
 - Kremer, E. (1985). *Einführung in die Versicherungsmathematik*. Vandenhoeck & Ruprecht, Göttingen.
 - Kremer, E. (1982). *IBNR-claims and the two-way model of ANOVA*. Scandinavian Actuarial Journal. 47-55.
 - Kuang, D., Nielsen, B. & Nielsen, J. (2009). *Chain-Ladder as maximum likelihood revisited*. Annals of Actuarial Science 4(1), 105-121.
 - *Ley General de Instituciones y Sociedades Mutualistas de Seguros*. (2014).
 - Linton, O. & Nielsen, J. (1995). *A kernel method of estimating structured non-parametric regression based on marginal integration*. Biometrika, 82, 93-100.
 - Mack, T. (1990). *Improved estimation of IBNR claims by credibility theory*. Insurance: Math. Econom. 9/1, 51-57.
 - Mack, T. (1991). *A simple parametric model for rating automobile insurance or estimating IBNR claims reserves*. Astin Bulletin. 21/1, 93-109.
 - Mack, T. (1993). *Distribution-free calculation of the standard error of chain ladder reserve estimates*. Astin Bulletin. 23/2, 213-225.
 - Mack, T. (1994). *Measuring the variability of chain ladder reserve estimates*. CAS Forum (Spring), 101-182.
 - Mack, T. (2000). *Credible claims reserves: the Benktander method*. Astin Bulletin. 30/2, 333-347.
-

-
- Martínez-Miranda, M., Nielsen, B. Nielsen, J. & Verrall, R. (2011) *Cash flow simulation for a model of outstanding liabilities based on claim amounts and claim numbers*. ASTIN Bulletin, 41(1), 107-129.
 - Martínez-Miranda, M., Nielsen, J. & Verrall, R. (2012a) *Double Chain Ladder*. ASTIN Bulletin, 42(1), 59-76.
 - McCullag, P. & Nelder, J. (1989). *Generalized Linear Models*. 2nd Edition. Chapman & Hall, London.
 - Merz, M. & Wuthrich, M. (2015). *Modified Munich chain-ladder method*. Risks 3, no. 4, 624-646.
 - Murphy, D. (1994). *Unbiased loss development factors*. Proc. CAS, Vol. LXXXI, 154-222.
 - Nielsen, J. (1999). *Multivariate Boundary Kernels from Local Linear Estimation*. Scandinavian Actuarial Journal, 1, 93-95.
 - Peters, G., Wuthrich, M. & Shevchenko, P. (2010). *Chain Ladder Method: Bayesian Bootstrap versus Classical Bootstrap*. Insurance: Mathematics and Economics , 47 (1) pp. 36-51.
 - Renshaw, A. (1989). *Chain ladder and interactive modelling (claims reserving and GLIM)*. J. Institute Actuaries 116, 559-587.
 - Renshaw, A. & Verrall, R. (1998). *A stochastic model underlying the chain-ladder technique*. J. British Actuarial J. 4/4, 903-923.
 - Robert, C. & Casella, G. (2010). *Introducing Monte Carlo Methods with R*. Springer Verlag.
 - Sarabia, J. & Gómez, E & Vázquez, F. (2006). *Estadística Actuarial. Teoría y Aplicaciones*. Pearson Educación.
 - Verrall, R. (1989). *A state space representation of the chain ladder linear model*. J. Institute Actuaries 116, 589-609.
 - Verrall, R. (2000). *An investigation into stochastic claims reserving models and the chain-ladder technique*. Insurance: Math. Econom. 26/1, 91-99.
 - Verrall, R. (2004). *A Bayesian generalized linear model for the Bornhuetter-Ferguson method claims reserving*. North American Actuarial J. 8/3, 67-89.
 - Verrall, R. Nielsen, J. & Jessen, A. (2010). *Prediction of RBNS and IBNR claims using claim amounts and claim counts*. ASTIN Bulletin, 40(2), 871-887.
-

- Wüthrich M. & Merz M. (2008). *Stochastic Claims Reserving Methods in Insurance*.
-