



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

MODELO PARA EL CÁLCULO DE RESERVAS ACTUARIALES

REPORTE DE TRABAJO PROFESIONAL

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A:

LUIS FERNANDO GARCÍA CABRERA



TUTOR:

ACT. RICARDO VILLEGAS AZCORRA

2018



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.

1. Datos del alumno

García

Cabrera

Luis Fernando

55 22 96 39 03

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

308312740

2. Datos del tutor

Act.

Ricardo

Villegas Azcorra

3. Datos del sinodal 1

M. en I.

Fernando Eleazar

Vanegas Chávez

4. Datos del sinodal 2

Act.

Carlos Lorenzo

Ramírez Vargas

5. Datos del sinodal 3

Act.

José Luis

Botello Ramírez

6. Datos del sinodal 4

Act.

Roberto

Carrillo Lárraga

7. Datos del trabajo escrito

Modelo para el cálculo de reservas actuariales
SONR, RRC y Mejora planteada al modelo

38 p

2019

Índice

1.	Introducción y objetivos	4
1.1	Introducción	4
1.2	Objetivos	4
2.	Notación:	5
2.1	Siniestros.....	5
2.2	Primas	6
3.	Valuación de Reserva para Siniestros Ocurridos y No Reportados (SONR)	6
3.1	Definición e importancia.....	6
3.2	Metodología de cálculo y recursos empleados.....	7
3.2.1	Chain Ladder	9
3.2.2	Bootstrap paramétrico.....	10
3.2.3	Cape-Cod (CC).....	16
3.2.3	Importes Recuperables de Reaseguro	22
4.	Reserva de Riesgos en Curso (RRC)	23
4.1	Definición e importancia.....	23
4.2	Metodología de cálculo y recursos empleados.....	23
4.2.1	Índice de siniestralidad última	23
4.2.2	Mejor Estimación de la Reserva de Riesgos en Curso	25
4.2.3	Importes Recuperables de Reaseguro	26
5.	Mejora planteada al modelo mediante una “Cópula Gaussiana”	27
6.	Resultados Obtenidos	32
6.1	Prueba de BackTesting.....	32
7.	Conclusiones:.....	37
8.	Bibliografía.....	38

1. Introducción y objetivos

1.1 Introducción

En este reporte de trabajo profesional explicaré como se implementó el cálculo de reservas actuariales en una compañía de seguros especializada en salud, el trabajo se resume a mostrar la metodología, una posible mejora, así como los algoritmos y consultas principales que permitieron automatizar de manera eficaz el cálculo de las mismas.

El cálculo de las reservas actuariales es una parte fundamental en el funcionamiento de una compañía de seguros, pues es la estimación de las obligaciones futuras que tendrá la aseguradora, por lo que las reservas son indispensables para asegurar la solvencia de una compañía de seguros; sin embargo es muy importante que esta estimación no sea de más conservadora pues esto supondría que los socios de la compañía tendrían que desembolsar más capital del necesario para poder cubrir esas obligaciones.

El cálculo de la reservas se debe llevar con una minuciosa aplicación de modelos estadísticos y matemáticos para la correcta medición del riesgo por el peligro que supondría no llevar a cabo de manera prudente los cálculos necesarios, por lo que su cálculo resulta de mucha importancia y requiere de un profesionalista con amplios conocimientos estadística, probabilidad, programación, etc. Como lo es un actuario.

1.2 Objetivos

Se mostrará la metodología que se implementó en una empresa de seguros especializada en salud para calcular la reserva de siniestros ocurridos y no reportados (SONR) y la reserva de riesgos en curso (RRC) mediante la metodología de Bootstrap Paramétrico la cual hace uso de simulaciones de una distribución estocásticas.

Se describirá lo que es cada una de estas reservas, su importancia, y como se lleva a cabo su cálculo.

Se informarán los supuestos bajo los cuales se lleva a cabo el cálculo de estas, y como se llegó a estos supuestos, se mostrará cómo se verifica que estas estimaciones sean correctas mediante la prueba de "Back-Testing" en donde cada año se verifica que la siniestralidad que se pronosticó un año anterior no haya sido superada por la siniestralidad que en realidad se presentó.

Se describirán los algoritmos más importantes mediante los cuales implementé el cálculo de las reservas en "Matlab" y como se logró con ciertos comandos hacer el

cálculo mediante operaciones vectoriales para optimizar el tiempo y el uso de memoria de las computadoras donde se realizan los cálculos.

También presentaré una posible mejora al modelo mediante el uso de cópulas Gaussianas, por medio de la cual se podría medir la correlación de los siniestros entre cada año de desarrollo.

2. Notación:

2.1 Siniestros

Para efectos del análisis de los siniestros, se disponen en un arreglo matricial llamado *triángulo de desarrollo*, el cual consiste en acomodar el monto de siniestros $X_{i,j}$ de la siguiente manera:

Período de Inicio de Vigencia	Período de Desarrollo					
	0	1	...	j	...	J
0	$X_{0,0}$	$X_{0,1}$...	$X_{0,j}$...	$X_{0,J}$
1	$X_{1,0}$	$X_{1,1}$...	$X_{1,j}$...	
...		
i	$X_{i,0}$	$X_{i,1}$...			
...				
I	$X_{I,0}$					

Donde:

i = Período de inicio de vigencia de la póliza (eje vertical), donde denotaremos a I como el último período de suscripción observado.

j = Período de desarrollo del siniestro (eje horizontal), donde se denota como J el último período de desarrollo.

Cada período está compuesto por la estructura de los últimos 12 meses partiendo del mes de valuación. Es decir, cada período es una ventana móvil de 12 meses que buscan reconocer el retraso en el reporte de los siniestros en un lapso de un año¹.

$X_{i,j}$ denota todos los siniestros que se registraron con j períodos de retraso de pólizas que iniciaron su vigencia en el período i , también puede ser visto como los siniestros provenientes del período de inicio de vigencia i , con j períodos de retraso entre la ocurrencia y el registro del siniestro.

¹ En el mes de diciembre cada período corresponde a un año calendario.

Los triángulos suelen analizarse de manera acumulada definiendo la siguiente variable:

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

Estas variables generan dos triángulos: las observaciones $C_{i,j}$, $i+j \leq I$ contenidas en el triángulo superior y las estimaciones $C_{i,j}$, $i+j > I$ contenidas en el triángulo inferior.

2.2 Primas

Para efectos de análisis y cálculo, las primas se agrupan por período de inicio de vigencia de las pólizas.

Período de Inicio de Vigencia	Prima Neta
0	π_0
1	π_1
...	...
<i>i</i>	π_i
...	...
<i>I</i>	π_I

Donde:

π_i = Primas de tarifa, de pólizas que iniciaron vigencia en el período i (Prima Neta).

i = Período de inicio de vigencia de la póliza, donde denotaremos a I como el último período observado.

De forma análoga a la construcción del triángulo, cada período está compuesto por la estructura de los últimos 12 meses partiendo del mes de valuación.

3. Valuación de Reserva para Siniestros Ocurridos y No Reportados (SONR)

3.1 Definición e importancia

Esta reserva se constituye para los siniestros tardíos, es decir, los siniestros ya ocurridos pero que aún no han sido reportados a la aseguradora, por consiguiente, desconocidos al asegurador directo y al reasegurador.

Motivos para la ocurrencia de "siniestros SONR":

- avisos tardíos
- tramitación retrasada.
- reconocimiento tardío.
- sentencias judiciales.

Por otro parte, cabe señalar que las reservas SONR son críticas para la solvencia a largo plazo proporcionando las siguientes ventajas estratégicas y gerenciales a las aseguradoras:

- 1) Menos volatilidad en resultados.
- 2) Mejor desempeño financiero.
- 3) Mejores decisiones de capital

3.2 Metodología de cálculo y recursos empleados.

El primer paso para poder calcular la reserva es ordenar la información de siniestros que poseemos en el arreglo triangular descrito anteriormente.

El arreglo triangular se completa de la siguiente manera: En el eje vertical los siniestros se acomodan de acuerdo a la fecha de inicio de vigencia de la póliza.

De manera horizontal el desfase está dado por el retraso entre la fecha de ocurrido y la fecha de reporte.

Supongamos que se realiza la valuación el 31-Dec-2017; en la Figura 3.1 sobre la diagonal marcada en azul están solo siniestros que han sido reportados durante 2017, por ejemplo en el año 2016 retraso 1, están los siniestros de pólizas suscritas durante 2016 y con 1 año de desfase entre la fecha de ocurrido y fecha de reporte, es decir, siniestros que ocurrieron durante 2016 y se reportaron durante 2017.

De esta manera lo que lograremos será completar nuestro arreglo matricial con la información que poseemos, a esta información le llamaré el “triángulo superior”

	Año	Período de Inicio de Vigencia	Retraso				
			0	1	2	3	4
Inicio de vigencia de la póliza	2008	0	9,909	7,630	5,951	1,904	267
	2009	1	9,437	8,871	7,452	3,130	908
	2010	2	7,316	6,584	3,358	2,283	297
	2011	3	6,362	6,234	5,860	3,633	1,817
	2012	4	4,970	4,572	4,435	1,464	732
	2013	5	5,847	5,730	5,444	2,014	363
	2014	6	4,301	3,183	2,514	1,760	
	2015	7	4,893	3,768	2,637		
	2016	8	5,451	4,906			
	2017	9	6,454				

Figura 3.1 Arreglo de siniestros en triángulo de desarrollo

Para poder asignar a cada siniestro su período de inicio de vigencia y período de retraso, se utiliza la siguiente instrucción SQL (Structured Query Language):

```
SELECT Year(inivigpol) -2017 + Iif(Month(inivigpol)>12,11,10) + '|'
+ Year(fecreg) + Iif(Month(fecreg)>12,1,0) -
Year(fecocu)+ Iif(Month(fecocu)> 12,1,0) as Llave,
```

El campo “Llave” tendrá en los primeros 2 dígitos el periodo de inicio de vigencia de la póliza, seguido de un pipe “|”, y en los últimos 2 dígitos, el periodo de retraso.

Una vez que cada siniestro ha sido catalogado según su período de inicio de vigencia y de desarrollo, se agrupan y acomodan en el arreglo triangular mediante el siguiente algoritmo (Pseudocódigo):

```
for i = 1:10
    for j=0:5
        xIndice = (base_llave{1} == i+'|'+j);
        trianguloDeDesarrollo(i,j) = sum(base_siniestros(xIndice));
    end
end
```

Las primas se acomodan en un arreglo matricial de 10x1 (Figura 3.2) de acuerdo a la fecha de inicio de vigencia de la póliza.

Inicio de vigencia de la póliza	Período de Inicio de Vigencia	Prima Neta
2008	0	65,784
2009	1	72,514
2010	2	79,040
2011	3	78,250
2012	4	67,295
2013	5	76,043
2014	6	81,366
2015	7	78,112
2016	8	83,579
2017	9	81,908

Figura 3.2 Arreglo de Primas.

Para asignar a cada póliza su periodo de inicio de vigencia, se usa la siguiente instrucción:

```
SELECT YEAR(inivigpol) -2017 + IIF(MONTH(inivigpol)>12,11,10) as Llave ,
```

Después podremos agrupar las primas de acuerdo a su período mediante el siguiente algoritmo (Pseudocódigo):

```
for i = 1:10
    xIndice = (base_llave{1} == i);
    Primas(i) = sum(base_primas(xIndice));
end
```

Una vez que tenemos la información de primas y siniestros en los arreglos definidos, procedemos a realizar el cálculo que nos permitirá estimar los siniestros que ya ocurrieron pero aún no se reportan, los cuales de acuerdo al acomodo del arreglo matricial de siniestros definido anteriormente serán colocados en el “el triángulo inferior”.

Primero se mostrará la definición del método “Chain Ladder” el cual nos permitirá calcular los factores de desarrollo y la definición del “Bootsrap paramétrico”, el cual nos permitirá remuestrear dichos factores; después de su definición, explicaré como son utilizados dichos métodos.

3.2.1 Chain Ladder

Definición obtenida de la nota técnica (inciso 4 y 5 de la bibliografía).

Los supuestos del modelo son los siguientes:

- Los siniestros acumulados $C_{i,j}$ de diferentes períodos i son independientes.
- Existen factores de desarrollo $f_0, \dots, f_{j-1} > 0$ tal que para toda $0 \leq i \leq I$ y para toda $1 \leq j \leq J$ tenemos que

$$E[C_{i,j} | C_{i,0}, \dots, C_{i,j-1}] = E[C_{i,j} | C_{i,j-1}] = \hat{C}_{i,j}^{CL} = f_{j-1} C_{i,j-1} \quad (1)$$

- Los factores de desarrollo f_j son el objetivo central del modelo. Para el método Chain Ladder (CL), los factores de desarrollo se estiman como el promedio ponderado del incremento de la siguiente manera.

$$\hat{f}_j = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \frac{C_{i,j+1}}{C_{i,j}} \quad \text{para } j = 0, 1, \dots, J-1 \quad (2)$$

Donde el modelo clásico Chain Ladder (CL) define $\hat{f}_j = 1$ suponiendo que la siniestralidad está totalmente desarrollada en el período de observación.

$$\text{Siniestros Estimados Chain Ladder (CL)} = \sum_{i=0}^I [\hat{C}_{i,j}^{CL} - C_{i,j-i}] \quad (3)$$

A continuación se mostrará la definición del método Bootstrap paramétrico, con el cual podremos simular los factores de desarrollo (2) y después mostraré como pudimos ajustar nuestros datos.

3.2.2 Bootstrap paramétrico

Definición obtenida de la nota técnica (inciso 4 y 5 de la bibliografía).

Se conoce que F pertenece a una familia paramétrica de distribuciones y se estiman sus parámetros a partir de la muestra. Se generan muestras a partir de F usando los parámetros estimados.

Supongamos que tenemos más información sobre la distribución de Z_i , es decir,

$$Z_1, \dots, Z_n \text{ i. i. d. } \sim^{(d)} F_\theta$$

donde F_θ denota un tipo de distribución conocida con parámetro θ desconocido.

Para obtener las muestras con reemplazamiento a partir de los datos iniciales, primero se estiman los parámetros desconocidos (Máxima Verosimilitud MLE, por sus siglas en inglés). Las muestras generadas de tamaño n se obtienen a partir de la estimación paramétrica de F_θ , \hat{F}_θ .

$$Z_1^*, \dots, Z_n^* \text{ i. i. d. } \sim^{(d)} \hat{F}_\theta = F_{\hat{\theta}}$$

Una vez obtenidas las muestras siguiendo este patrón, se calculan m réplicas del estadístico a estudiar y a partir de ellas, se calcula aquello que nos interese, por ejemplo, valor esperado, desviación típica, etc.

A partir del modelo descrito por Hertig² (1985) el cual supone que los factores de desarrollo individuales CL $F_{i,j} = C_{i,j+1}/C_{i,j}$ $i+j \leq I$ (factores contenidos en el triángulo superior) se distribuyen log-normal con parámetros ξ_j y σ_j^2 ,

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

Donde,

$$\hat{\xi}_j = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \log\left(\frac{C_{i,j+1}}{C_{i,j}}\right) \text{ para } j = 0, 1, \dots, J-1 \quad (4)$$

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \left(\log\left(\frac{C_{i,j+1}}{C_{i,j}}\right) - \hat{\xi}_j \right)^2 \text{ para } j = 0, 1, \dots, J-1 \quad (5)$$

A continuación podemos ver como se ajustan nuestros datos a una distribución Log Normal.

² Hertig, J. (1985). A statistical approach to the IBNR-reserves in marine insurance. *Astin Bulletin* 15/2, 171–183.

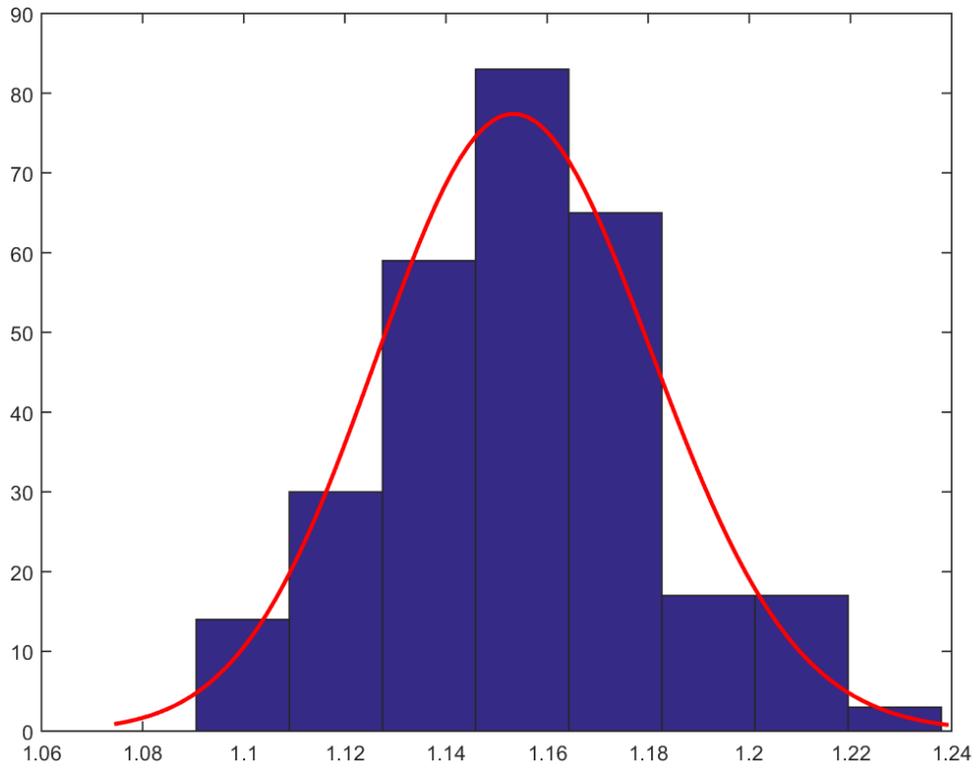


Figura 3.3 Histograma de los factores de desarrollo O ajustándole una función de densidad log normal. En la cual se puede observar que la distribución se ajusta de buena manera a nuestra información.

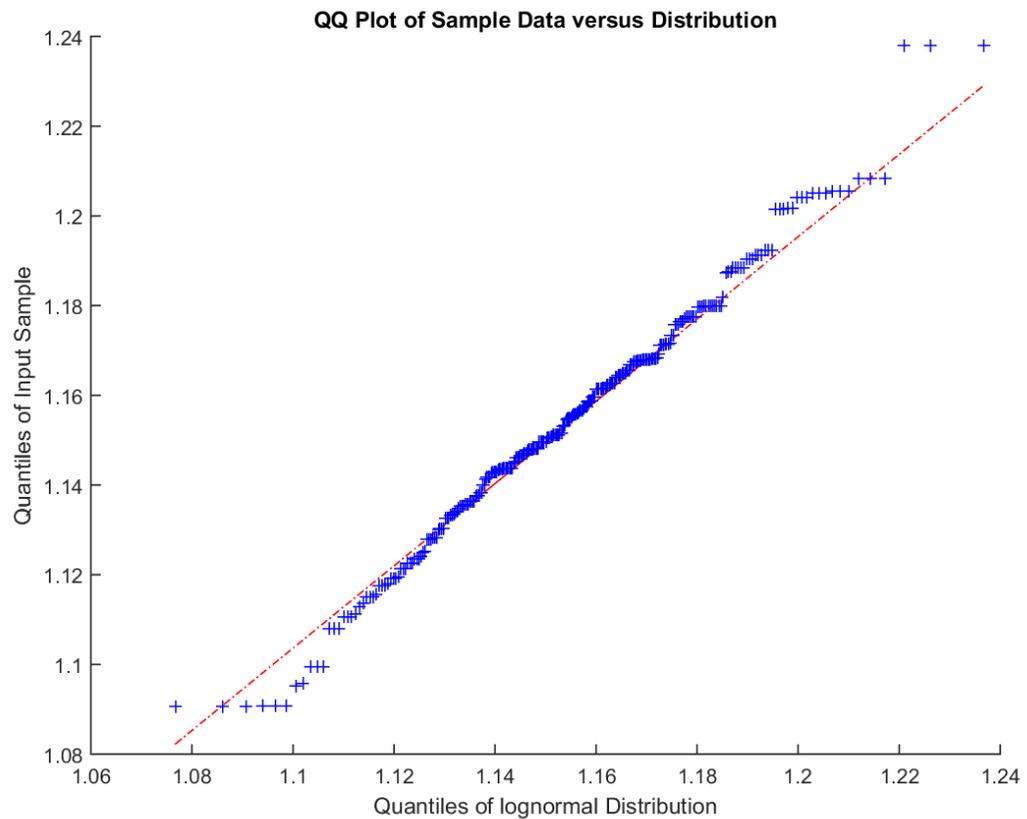


Figura 3.4 Gráfica cuantil-cuantil (quantile-quantile), en la que se muestran los cuantiles teóricos de la distribución log-normal, contra los cuantiles de nuestra muestra de datos, podemos observar que los cuantiles de la muestra de datos, son similares a los cuantiles teóricos de la distribución log-normal.

Aplicamos el método de bootstrap paramétrico. Para lo cual se estiman los parámetros desconocidos ξ_j y σ_j^2 a partir de los datos observados. La aplicación del método bootstrap paramétrico, consiste entonces en generar nuevas observaciones independientes

$$\eta_{i,j}^* \sim^{(d)} N(\hat{\xi}_j, \hat{\sigma}_j^2)$$

A continuación mostraré con nuestros siniestros de ejemplo (Figura 3.1) como se calculan los factores de desarrollo y cómo se remuestran estos con el modelo “Bootstrap Paramétrico”; el primer paso consiste en formar un triángulo acumulando los siniestros de la siguiente manera (lo llamaremos “triángulo acumulado”):

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

	Período de Inicio de Vigencia	Retraso Acumulado				
		0	1	2	3	4
Inicio de vigencia de la póliza	0	9,909	17,539	23,491	25,395	25,662
	1	9,437	18,308	25,760	28,890	29,798
	2	7,316	13,900	17,258	19,541	19,838
	3	6,362	12,596	18,456	22,089	23,906
	4	4,970	9,542	13,977	15,441	16,173
	5	5,847	11,577	17,021	19,035	19,397
	6	4,301	7,484	9,998	11,758	
	7	4,893	8,661	11,298		
	8	5,451	10,357			
	9	6,454				

Figura 3.5 Triángulo acumulado

Ahora podremos calcular los factores de desarrollo:

$$F_{i,j} = C_{i,j+1} / C_{i,j}$$

Fac. Desarrollo	0	1	2	3
0	1.77000	1.33932	1.08107	1.01050
1	1.94000	1.40701	1.12149	1.03142
2	1.90000	1.24158	1.13231	1.01519
3	1.98000	1.46525	1.19686	1.08224
4	1.92000	1.46479	1.10471	1.04739
5	1.98000	1.47020	1.11833	1.01905
6	1.74000	1.33598	1.17604	
7	1.77000	1.30452		
8	1.90000			

Figura 3.6 Triángulo de factores de desarrollo

Los factores de desarrollo nos indican cuanto crece o decrece la siniestralidad entre cada año de retraso, por ejemplo la entrada (0,0) nos dice que los siniestros de pólizas con 1 año de retraso que iniciaron vigencia en el año 0, representan el 77% de los siniestros de la pólizas que empezaron vigencia en el año 0 y tuvieron 0 años de retraso.

Una vez que tenemos los factores de desarrollo procederemos a obtener su logaritmo natural para poder estimar los parámetros $\hat{\xi}_j$ y $\hat{\sigma}_j^2$ De acuerdo a (4) y (5).

Fac Log Normal	0	1	2	3
0	0.57098	0.29216	0.07795	0.01044
1	0.66269	0.34147	0.11466	0.03093
2	0.64185	0.21638	0.12426	0.01508
3	0.68310	0.38203	0.17971	0.07903
4	0.65233	0.38171	0.09958	0.04630
5	0.68310	0.38540	0.11184	0.01887
6	0.55389	0.28966	0.16215	
7	0.57098	0.26583		
8	0.64185			
Mu	0.6289732	0.3193316	0.1243084	0.0334434
sigma	0.0474193	0.0589021	0.0327533	0.0235551

Figura 3.7 Triangulo con el logaritmo natural de factores de desarrollo, que nos permite estimar $\hat{\xi}_j$ y $\hat{\sigma}_j^2$

Ahora que tenemos los parámetros de nuestra distribución log-normal simularemos 20,000 factores de desarrollo para cada año de desarrollo y cada año de vigencia de la póliza, asumiendo que los factores de desarrollo son idénticamente distribuidos por cada año de desarrollo.

Para poder optimizar el tiempo de cálculo de las reservas se optó por hacer los cálculos de manera vectorial, es decir obtener todas las simulaciones al mismo tiempo. El comando “lognrnd” nos permite obtener en Matlab todas las simulaciones de manera vectorial:

`lognrnd(mu, sigma, 20000, 1);`

Por lo que ahora tendremos una matriz de factores de desarrollo simulados en R3 de [9,4,20000].

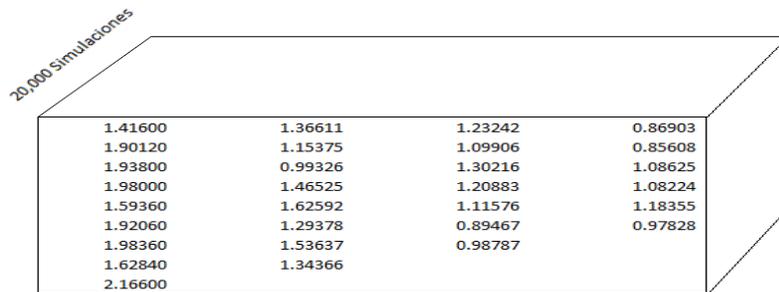


Figura 3.8 Matriz de factores de desarrollo simulada.

Una vez que tenemos nuestros factores de desarrollo simulados, obtendremos la matriz acumulada simulada, A partir de nuestra ecuación (1).

$$\hat{C}_{i,j}^{CL} = f_{j-1}^{sim} C_{i,j-1} \quad (1)$$

A partir de este momento todas nuestras matrices serán en R^3 , sin embargo para poder ejemplificar de buena manera el cálculo solo mostraré la primera simulación.

Por lo que utilizando (1) podremos obtener nuestros 20,000 triángulos superiores simulados

Simulado/Acumulado	0	1	2	3	4
0	9,909	18,604	25,639	29,045	30,045
1	9,437	17,727	24,457	27,705	28,652
2	7,316	13,739	18,935	21,458	22,195
3	6,362	11,938	16,462	18,651	19,292
4	4,970	9,336	12,867	14,578	15,075
5	5,847	10,979	15,130	17,144	17,732
6	4,301	8,078	11,136	12,614	
7	4,893	9,193	12,678		
8	5,451	10,237			
9	6,454				

Figura 3.9 Primera simulación de los triángulos superiores simulados.

Pseudocódigo del algoritmo para obtener los triángulos superiores simulados

```

for i = (0:9)
    trianguloDeDesarrolloSimulado(:,i,1) = trianguloDeDesarrollo(i,1);
    for j = (1:4)
        trianguloDeDesarrolloSimulado(:,i,j) = trianguloDeDesarrolloSimulado(:,i,j-1)...
            .*lognrnd(xParametroLogNorm(j,1),xParametroLogNorm(j,2),20000,1);
    end
end
end

```

El siguiente paso, será calcular el Cape-Cod sobre cada uno de los triángulos superiores simulados para poder obtener lo que buscamos, “los triángulos inferiores” es decir la estimación de los siniestros ocurridos y no reportados.

Primero se mostrará la definición del “Cape-Cod” y después explicaré como se aplica en nuestro ejemplo que hemos venido trabajando.

3.2.3 Cape-Cod (CC)

Definición obtenida de la nota técnica (inciso 4 y 5 de la bibliografía).

Una de las principales deficiencias en el modelo Chain Ladder (CL) es el hecho de que la siniestralidad total depende completamente de la última observación registrada sobre la diagonal, si esta observación es un valor atípico, entonces se amplificará la desviación en la siniestralidad última esperada. Una alternativa para robustecer la estimación última es utilizar el modelo Cape-Cod (CC), el cual es un

método combinado entre los puramente basados en información (CL) y los métodos de siniestralidad a priori como el Borhuetter-Ferguson (BF).

Los supuestos del modelo son los siguientes:

- Los siniestros acumulados $C_{i,j}$ de diferentes períodos i son independientes.
- Existe un conjunto de parámetros $\pi_0^{Dev}, \dots, \pi_I^{Dev} > 0$, $\kappa > 0$ y un patrón de retraso de siniestros β_i $0 \leq i \leq I$ con $\beta_I = 1$ tales que:

$$E[C_{i,j}] = \kappa \pi_i^{Dev} \beta_j \quad (6)$$

Para toda $i = 0, \dots, I$

π_i^{Dev} Primas de tarifa devengada descontando derechos de póliza e IVA, de pólizas que iniciaron vigencia en el año i (Prima Neta).

κ Índice de siniestralidad media

β_j Factor de retraso de la reclamación

Se define β_i como el inverso de los factores acumulados de la siguiente forma:

$$\hat{\beta}_i = \begin{cases} \prod_{k=i}^{J-1} \hat{f}_k^{-1}; & \text{para } i = 0, \dots, J-1 \\ 1; & \text{para } i = J, J+1, \dots, I \end{cases}$$

Para poder llevar a cabo el cálculo κ es preciso obtener previamente los κ_i , para posteriormente calcular el índice de siniestralidad total κ robustecido, de la siguiente manera:

$$\hat{\kappa}_i = \frac{\hat{C}_{i,J}^{CL}}{\pi_i} = \frac{C_{i,\min(I-i,J)}}{\prod_{j=i}^{J-1} \hat{f}_j^{-1} \pi_i^{Dev}} = \frac{C_{i,\min(I-i,J)}}{\beta_{I-i} \pi_i^{Dev}} \quad (7)$$

Para toda $i = 0, \dots, I$.

Donde (β_{I-i}) representa el factor de siniestralidad desarrollada en cada período i .

Entonces el índice de siniestralidad total κ robustecido se calcula con el promedio ponderado de los $\hat{\kappa}_i$ con la prima y su factor de retraso β_{I-i} de la siguiente manera:

$$\hat{\kappa}^{CC} = \sum_{i=0}^I \frac{\beta_{I-i} \pi_i^{Dev}}{\sum_{k=0}^I \beta_{I-k} \pi_k^{Dev}} \hat{\kappa}_i = \frac{\sum_{i=0}^I C_{i,\min(I-i,J)}}{\sum_{i=0}^I \beta_{I-i} \pi_i^{Dev}} \quad (8)$$

De esta manera el estimador para $\hat{C}_{i,j}^{CC} - C_{i,\min(I-i,J)}$ se calcula de la siguiente forma:

$$\hat{C}_{i,J}^{CC} - C_{i,\min(I-i,J)} = \hat{\kappa}^{CC} \Pi_i (1 - \beta_{I-i}) \quad (9)$$

Al realizar el cálculo de esta manera en realidad estaríamos obteniendo la proyección de siniestralidad última esperada total en la ventana de observación. Es decir, una estimación de la reserva de obligaciones pendientes de cumplir así como la reserva de riesgos en curso al mismo tiempo. Por tal motivo, se utiliza la prima devengada (hasta la fecha de valuación) de cada período, obteniendo de esta manera, una estimación únicamente por las obligaciones pendientes de cumplir del riesgo ya devengado, como se muestra a continuación:

$$\hat{C}_{i,J}^{CC} - C_{i,\min(I-i,J)} = \hat{\kappa}^{CC} \pi_i^{Dev} (1 - \beta_{I-i}) \quad (10)$$

Donde,

π_i^{Dev} = Prima Devengada a la fecha de valuación para el período de inicio de vigencia i .

Por tanto la estimación de SONR se calcula de la siguiente manera:

$$EstimaciónSONR_i = \hat{\kappa}^{CC} \pi_i^{Dev} \sum_{k=0}^{i-1} (\beta_{I-k} - \beta_{I-k-1}) v^{i-k} * 1_{i=1,2,\dots,I} \quad (11)$$

Donde v^i es el factor de descuento con la tasa libre de riesgo.

Finalmente, el monto de la mejor estimación de Obligaciones Pendientes de Cumplir por Siniestros Ocurridos No Reportados (SONR) se calculará como la suma de las estimaciones de cada período:

$$SONR_n = \sum_{i=0}^I EstimaciónSONR_i \quad (12)$$

Con lo cual, la mejor estimación de reserva de siniestros ocurridos no reportados se calcula como el promedio de las simulaciones.

$$SONR_{BEL} = \frac{\sum_{n=1}^N SONR_n}{N} \quad (13)$$

Donde,

$SONR_n$ = Estimación de la reserva SONR con el método Cape-Cod de la simulación n con el método de bootstrap paramétrico.

N = número de simulaciones. Al menos se realizan 20,000 simulaciones, a partir de las cuales se verifica si al aumentar una simulación, el promedio del

SONR_{BEL} no varía en más del 0.01%. En caso de que la variación sea superior, se adicionarán el número de simulaciones necesarias hasta que la variación sea menor o igual a 0.01%.

Ahora procederemos a realizar el cálculo del Cape-Cod en nuestro ejemplo, recordemos que se le realizará el procedimiento Cape-Cod a cada una de las 20,000 simulaciones (aunque en este caso solo mostraremos una simulación), el primer paso es calcular los factores $\hat{\beta}_i$ los cuales se definen de la siguiente manera:

$$\hat{\beta}_i = \begin{cases} \frac{1}{F_j}; & \text{para } i = 0, \dots, J - 1 \\ 1; & \text{para } i = J, J + 1, \dots, I \end{cases}$$

Por lo que primero calcularemos los \hat{f} de acuerdo a (2) y a su vez calcularemos \hat{F} como $F_j = \prod_{k=i}^{J-1} \hat{f}_k$.

Fac. Desarrollo	0	1	2	3	
0	1.87744	1.37814	1.13286	1.03442	
1	1.87839	1.37962	1.13283	1.03417	
2	1.87794	1.37827	1.13320	1.03436	
3	1.87654	1.37900	1.13296	1.03437	
4	1.87840	1.37829	1.13293	1.03410	
5	1.87767	1.37809	1.13315	1.03429	
6	1.87824	1.37857	1.13270		
7	1.87877	1.37916			
8	1.87797				
fj=	1.8779	1.3786	1.1329	1.0343	1.0000
Fj=	3.0338	1.6155	1.1718	1.0343	1.0000

Figura 3.10 Matriz de factores de desarrollo de la primera simulación.

Con lo que al obtener el inverso de los factores \hat{F} , obtendremos los factores $\hat{\beta}_i$.

	Bj
9	1.00000
8	1.00000
7	1.00000
6	1.00000
5	1.00000
4	1.00000
3	0.96685
2	0.85339
1	0.61901
0	0.32962

Figura 3.11 Matriz de factores β de la primera simulación.

Lo que nos dicen los factores $\hat{\beta}_i$ es por ejemplo: al año 0 nos han llegado en promedio, el 32.96% de las reclamaciones, al año 1 el 61.901% y así sucesivamente hasta el año 4 en el que ya nos llegaron en promedio el 100% de las reclamaciones.

Por lo que ahora ya tenemos los elementos necesarios para calcular el índice de siniestralidad robustecido definido en (8). El cual intuitivamente no es más que los siniestros que ya nos reportaron entre la prima con la que le hicimos frente.

Inicio de vigencia de la póliza	Período de Inicio de Vigencia	Prima Neta Devengada
2008	0	65,784
2009	1	72,514
2010	2	79,040
2011	3	78,250
2012	4	67,295
2013	5	76,043
2014	6	81,366
2015	7	78,112
2016	8	73,251
2017	9	32,541

Figura 3.12 Matriz de primas.

$$\hat{\kappa}^{CC} = \frac{\sum_{i=0}^I C_{i, \min(I-i, J)}}{\sum_{i=0}^I \beta_{I-i} \pi_i^{Dev}}$$

Con lo que obtenemos:

Ind Sin Kcc
27.33%

Con lo que ya tendremos los elementos necesarios para calcular la estimación del “triángulo inferior”, es decir la estimación SONR:

$$EstimaciónSONR_i = \hat{\kappa}^{CC} \pi_i^{Dev} \sum_{k=0}^{i-1} (\beta_{I-k} - \beta_{I-k-1}) v^{i-k} * 1_{i=1,2,\dots,I} \quad (11)$$

La diferencia ($\beta_{I-k} - \beta_{I-k-1}$) lo que nos dice es cuanta siniestralidad llegará por cada año de retraso, por ejemplo para nuestra primer simulación, en el año 0 nos han llegado 32.96% de las reclamaciones, al año 1 el 61.901% por lo que entre el año 0 y el año 1 esperamos que nos lleguen el $61.901\% - 32.96\% = 28.93\%$ de la siniestralidad total esperada para ese año ($\hat{\kappa}^{CC}\pi_i^{Dev}$).

Sin Cape Cod VP

9,909	8,695	7,035	3,406	1,000
9,437	8,290	6,730	3,249	947
7,316	6,423	5,197	2,522	737
6,362	5,576	4,524	2,189	641
4,970	4,366	3,532	1,710	497
5,847	5,132	4,151	2,014	588
4,301	3,777	3,058	1,478	692
4,893	4,300	3,486	2,273	618
5,451	4,786	4,404	1,984	540
6,454	2,415	1,821	822	223

Figura 3.13 Primer simulación de la Matriz de siniestros con el triángulo inferior (en azul) generado mediante Cape-Cod.

Por lo que una vez que realizamos este procedimiento para las 20,000 simulaciones:

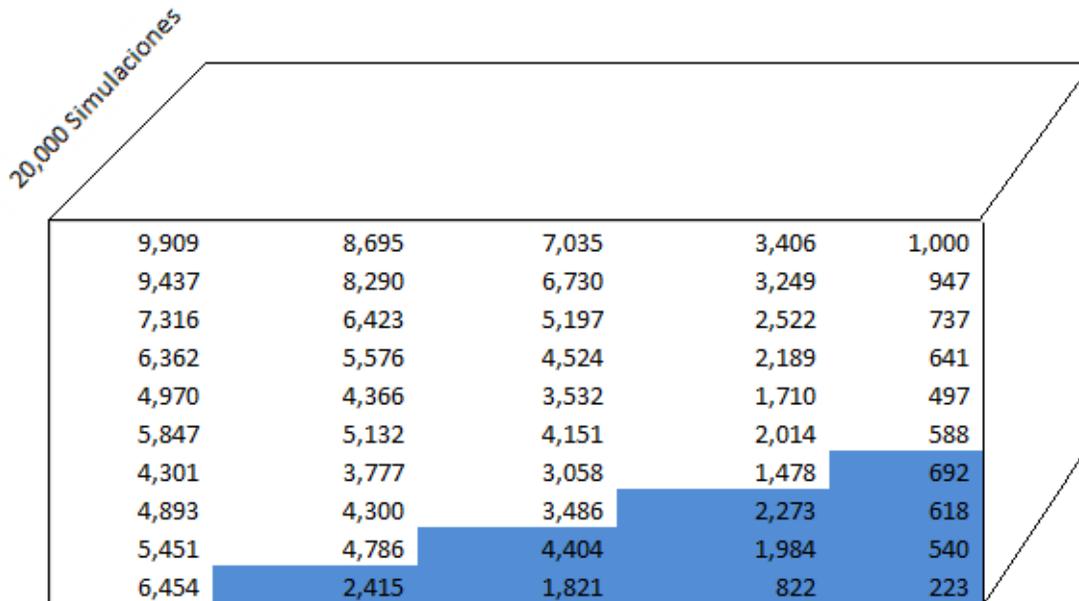


Figura 3.14 Matriz de siniestros completa con los 20,000 triángulos inferiores (marcado en azul) generados mediante Cape-Cod.

Pseudocódigo del algoritmo para la estimación SONR (11), mediante "Cape-Cod":

```

for i = (1:numColumnas-1)
  xCVP = 0
  for j = (i+1:numColumnas)
    xCVP = xCVP + 1
    EstimaciónSONR(xR-i+1,j,:) = IndSinCC(xR-i+1,:)...
    .*MatrizPrimaD(i,1)...
    .* (B(j,:) - B(j-1,:))...
    .*xValorPresente(1,xCVP);
  end
end

```

Podremos obtener la estimación $SONR_n$ de cada simulación n , como la suma de la estimación para cada año simulación:

$$SONR_n = \sum_{i=0}^I EstimaciónSONR_i \quad (12)$$

Y a su vez calcularemos el SONR total como el promedio de las $N=20,000$ simulaciones:

$$SONR_{BEL} = \frac{\sum_{n=1}^N SONR_n}{N} \quad (13)$$

3.2.3 Importes Recuperables de Reaseguro

Una vez que estimamos la siniestralidad proveniente de siniestros ocurridos y no reportados, los importes recuperables de reaseguro son la estimación de la cantidad que recuperará de los contratos de reaseguro proporcionales.

Se calculan de acuerdo a lo establecido la fracción III de disposición 5.3.5 de la CUSF, e intuitivamente no es más que la reserva SONR proveniente de cada periodo i , multiplicada por el porcentaje de cesión a reaseguro, por 1 menos la probabilidad de default (nos indica que tan probable es que la reaseguradora incumpla el pago de sus obligaciones).

$$IRR = \sum_{i=1}^I SONR_i * RC_i * FCR_i$$

Dónde:

$SONR_i$ =Reserva SONR proveniente de cada período i .

RC_i = Porcentaje de prima cedida, en el período i , en contratos de reaseguro que implican una transferencia cierta de riesgo.

FCR_i = Factor de calidad de reaseguro y es igual a $(1-PD_i^3)$ de la entidad reaseguradora al momento de la valuación).

Cuando la siniestralidad proveniente del período i , esté cubierta en contratos de Reaseguro que impliquen una Transferencia Cierta de Riesgo por más de una entidad reaseguradora, el factor de calidad de reaseguro deberá calcularse como:

$$FCR_i = \left(1 - \sum_{R=1}^l \alpha_{R,i} * PD_{R,i} \right)$$

Dónde:

l = número de entidades reaseguradoras que cubren la siniestralidad proveniente del período i .

$\alpha_{R,i}$ = es la proporción que representa la prima cedida a una entidad reaseguradora R en el período i , respecto de la prima cedida en contratos de Reaseguro que impliquen una Transferencia Cierta de Riesgo de Seguro en dicho período i .

$PD_{R,i}$ = es la probabilidad de incumplimiento⁴ de la entidad reaseguradora R .

4. Reserva de Riesgos en Curso (RRC)

4.1 Definición e importancia

La reserva de riesgos en curso tiene como propósito 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.

4.2 Metodología de cálculo y recursos empleados

4.2.1 Índice de siniestralidad última

La Siniestralidad Última se calculará a partir de la metodología de triángulo de desarrollo de siniestros ocurridos mediante el método Cape-Cod estocástico. El

³ La probabilidad de incumplimiento PD_i es la que corresponde a la Institución de Seguros o entidad reaseguradora.

⁴ La probabilidad de incumplimiento PD_i será la que corresponda a la Institución de Seguros o entidad reaseguradora del extranjero, en función de la calificación que tenga dicha entidad conforme a lo señalado en el Anexo 8.20.2 de la CUSF.

método Cape-Cod, considera los patrones de siniestralidad histórica y es sensible a los cambios en la composición de la cartera.

El modelo es alineado con la metodología para la estimación de las reservas de siniestros ocurridos no reportados descrita anteriormente, con la cual se completan los siniestros esperados conforme a la experiencia de la compañía y se evitan inconsistencias en la estimación de la siniestralidad.

Es decir una vez que hemos realizado el procedimiento para obtener la reserva SONR, ya tendremos nuestra matriz de siniestralidad Cape Cod completa.

$$SinCapeCod (CC)_i = C_{i,\min (I-i,J)} + \hat{\kappa}^{CC} \Pi_i^{Dev} \sum_{k=0}^{i-1} (\beta_{I-k} - \beta_{I-k-1}) v^{i-k} * 1_{i=1,2,\dots,I}$$

Donde v^i es el factor de descuento con la tasa libre de riesgo.

9,909	8,695	7,035	3,406	1,000
9,437	8,290	6,730	3,249	947
7,316	6,423	5,197	2,522	737
6,362	5,576	4,524	2,189	641
4,970	4,366	3,532	1,710	497
5,847	5,132	4,151	2,014	588
4,301	3,777	3,058	1,478	692
4,893	4,300	3,486	2,273	618
5,451	4,786	4,404	1,984	540
6,454	2,415	1,821	822	223

Figura 4.1 Matriz de siniestros completa.

De esta manera el índice de siniestralidad de cada período de inicio de vigencia κ_i , se calcula de la siguiente manera:

$$\hat{\kappa}_i = \frac{SinCapeCod (CC)_i}{\Pi_i^{Dev}}$$

Para toda $i = 0, \dots, I$.

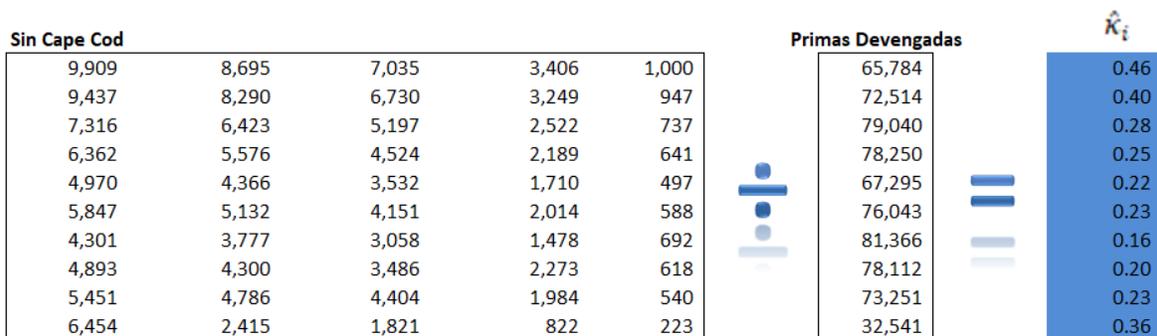


Figura 4.2 Se muestra la primera simulación de la matriz de siniestros entre la prima devengada para obtener los factores $\hat{\kappa}_i$

Con lo cual, el índice de siniestralidad se calcula como el promedio de $N=20,000$.

$$k = \frac{\sum_{n=1}^N \sum_{i=0}^I \hat{\kappa}_i}{N * I}$$

Donde:

$\hat{\kappa}_i$ = Índice de siniestralidad del período de inicio de vigencia i calculado con el método Cape-Cod.

N = número de simulaciones. Al menos se realizan 20,000 simulaciones, a partir de las cuales se verifica si al aumentar una simulación, el promedio del índice k no varía en más del 0.01%. En caso de que la variación sea superior, se adicionarán el número de simulaciones necesarias hasta que la variación sea menor o igual a 0.01%.

i = Período de inicio de vigencia de la póliza, donde denotaremos a I como el último período observado.

4.2.2 Mejor Estimación de la Reserva de Riesgos en Curso

Definición obtenida de la nota técnica (inciso 4 y 5 de la bibliografía).

La Mejor Estimación de la Reserva de Riesgos en Curso será igual a multiplicar la prima neta no devengada de cada endoso en vigor por el índice de siniestralidad última de cada ramo o subramo k_r , más el porcentaje de gastos de administración:

$$BEL_{RRC,r} = \sum_{i=1}^n \pi_i * FND_i * k_r + \sum_{i=1}^n \pi_i * FND_i * \alpha_r$$

$$BEL_{RRC,r} = BEL_{Riesgo,r} + BEL_{Gasto,r}$$

Fórmula extraída de la nota técnica de riesgos en curso (inciso 5 de la bibliografía).

Donde:

$$FND_i = 1 - \left(\frac{FecVal - FecInVig}{FecTeVig - FecInVig} \right)$$

$BEL_{RRC,r}$ = Mejor Estimación de la Reserva de Riesgos en Curso de cada póliza del ramo o subramo r .

π_i = Primas de tarifa descontando derechos de póliza e IVA, del endoso i .

k_r = Índice de Siniestralidad Última obtenido mediante el método de Cape Cod antes descrito del ramo o subramo r .

α_i = Porcentaje de Gastos de Administración que dé a conocer o utilice la CNSF para efectos del método estatutario.

$FecVal$ = Fecha de Valuación.

$FecInVig$ = Fecha de Inicio de Vigencia del endoso i .

$FecTeVig$ = Fecha de Término de Vigencia del endoso i .

n = Número de endosos en vigor de cada póliza.

4.2.3 Importes Recuperables de Reaseguro

Una vez que calculamos nuestro índice de siniestralidad última, los importes recuperables de reaseguro son la estimación de la cantidad que se recuperará de los contratos de reaseguro proporcionales.

Se calculan de acuerdo a lo establecido la fracción I de disposición 5.3.5, e intuitivamente no es más que el BEL calculado anteriormente multiplicado por el porcentaje de cesión a reaseguro, por 1 menos la probabilidad de default (nos indica que tan probable es que la reaseguradora incumpla en el pago de sus obligaciones).

$$IRR = BEL_{Riesgo,r} * RC_i * FCR_i$$

Donde:

$BEL_{Riesgo,r} = \Pi_i * FND_i * k_r$ Mejor Estimación de la Reserva de Riesgos en Curso sin considerar el gasto de administración del endoso i .

RC_i = Porcentaje de prima cedida en contratos de reaseguro que implican una transferencia cierta de riesgo del endoso i .

FCR_i = Factor de calidad de reaseguro y es igual a $(1-PD_i)^5$ de la entidad reaseguradora al momento de la valuación.

5. Mejora planteada al modelo mediante una “Cópula Gaussiana”

Una de las posibles mejoras al modelo es el poder modelar la correlación entre cada año de retraso de las reclamaciones, por ejemplo: resulta lógico pensar que cuando se tiene una gran siniestralidad para en $X_{i,0}$ (es decir se ha tenido una gran cantidad de siniestros reportados en el año i con 0 años de retraso) ésta se mantendrá alta para las reclamaciones con j años de retraso ($X_{i,1}, X_{i,2}, X_{i,3}, \dots, X_{i,j}$), y de igual manera para los años en que se ha tenido una siniestralidad baja ésta se mantendrá baja para las reclamaciones con j años de retraso.

Por lo que para poder medir la correlación entre cada año de retraso se puede usar una Copula Gaussiana.

A continuación describiré breve e intuitivamente como nos ayudaran las cópulas a lograr nuestro objetivo con un ejemplo extraído de “Simon Barthelme, *Copulas made easy*” (inciso 8 de la bibliografía).

Supongamos que queremos generar muestras de un distribución con función de densidad $f(x)$.

Esto lo podemos hacer a partir de la generación de variables aleatorias uniformes.

Pues si u es una variable aleatoria con distribución uniforme sobre $[0,1]$, y $F(x)$ es la función de distribución correspondiente a la función de densidad $f(x)$ entonces $F^{-1}(u)$ se distribuye como la función de densidad f .

⁵ La probabilidad de incumplimiento PD_i será la que corresponda a la Institución de Seguros o entidad reaseguradora del extranjero, en función de la calificación que tenga dicha entidad conforme a lo señalado en el Anexo 8.20.2 de la CUSF.

Esto mismo se cumple en la otra dirección, si tomamos x y la evaluamos en $F(x)$ tendremos $u \sim U[0,1]$.

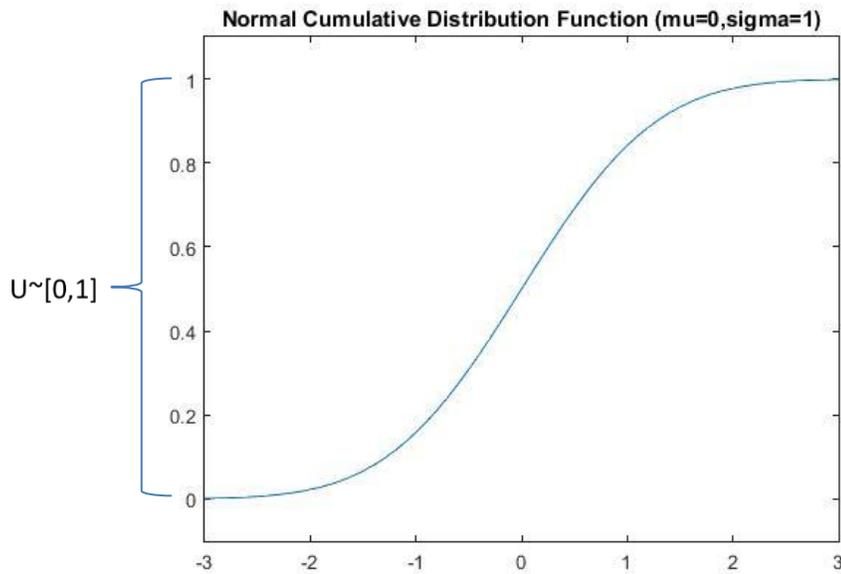


Figura 5.1 Grafica de la función distribución Normal $\sim [0,1]$, mostrando en el eje “y” la variable aleatoria uniforme $\sim [0,1]$.

Ahora queremos generar 2 variables aleatorias correlacionadas x y y con distribuciones marginales $f(x)$ y $g(y)$.

Podemos seguir esta serie de pasos:

- Generar a, b de una distribución normal multivariada, con la correlación que deseemos que tengan nuestras variables aleatorias x y y .
- De a y b podemos obtener $u, v \sim U[0,1]$ con el proceso descrito anteriormente como $u=\Phi(a)$ y $v=\Phi(b)$.
- Finalmente, obtener x y y correlacionadas como : $x=F^{-1}(u)$, $y=G^{-1}(v)$

Aplicado a nuestro problema, lo que queremos es incluir en el modelo la correlación que se tiene de siniestralidad entre los años de retraso, es decir, poder simular factores de desarrollo f_i 's correlacionados.

Por lo que el primer paso será obtener la correlación se tiene entre columnas de nuestra matriz de siniestralidad, es decir, la correlación que se tiene entre años de retraso.

Para poder medir la correlación entre las columnas, obtendremos la función de distribución de la log normal evaluada en cada factor de desarrollo f_i con los parámetros obtenidos mediante (4) y (5).

Fac. Desarrollo	0	1	2	3
0	1.77000	1.33932	1.08107	1.01050
1	1.94000	1.40701	1.12149	1.03142
2	1.90000	1.24158	1.13231	1.01519
3	1.98000	1.46525	1.19686	1.08224
4	1.92000	1.46479	1.10471	1.04739
5	1.98000	1.47020	1.11833	1.01905
6	1.74000	1.33598	1.17604	
7	1.77000	1.30452		
8	1.90000			

Figura 5.2 Matriz de factores de desarrollo.

$$\hat{\xi}_j = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \log\left(\frac{C_{i,j+1}}{C_{i,j}}\right) \text{ para } j = 0, 1, \dots, J-1 \quad (4)$$

$$\hat{\sigma}_j^2 = \frac{1}{I-j} \sum_{i=0}^{I-j-1} \left(\log\left(\frac{C_{i,j+1}}{C_{i,j}}\right) - \hat{\xi}_j \right)^2 \text{ para } j = 0, 1, \dots, J-1 \quad (5)$$

Con lo que obtendremos una matriz de variables uniformes [0,1].

$$p = F(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^x \frac{e^{-\frac{(\ln(t)-\mu)^2}{2\sigma^2}}}{t} dt$$

```
for i = (1:xC)
    xMatrizUniformesLogN(:,i) = logncdf(xMatrizfi(:,i+1),mu,sigma);
end
```

A esta matriz le calcularemos la correlación la cual está definida como:

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2)(\sum x)^2} \sqrt{n(\sum y^2)(\sum y)^2}}$$

El comando de Matlab “corr” calcula el coeficiente lineal de correlación entre cada par de columnas de la matriz xMatrizUniformesLogN.

```
xMatrizCorrelacion = corr(xMatrizUniformesLogN);
```

Matriz Correlación	0	1	2	3
0	1.00000	0.57230	0.67821	0.55698
1	0.57230	1.00000	0.24545	0.59352
2	0.67821	0.24545	1.00000	0.79712
3	0.55698	0.59352	0.79712	1.00000

Figura 5.3 Matriz de correlación.

Por lo que podemos ver que todas nuestras columnas se encuentran positivamente correlacionadas, por ejemplo los factores de desarrollo del año 0, se encuentran positivamente relacionados 57.230% con los factores de desarrollo del año 1.

Con esto ya podemos generar la cópula para poder obtener los factores de desarrollo simulados correlacionados.

$$C_R^{Gauss}(u) = \varphi_R(\varphi^{-1}(u_1), \dots, \varphi^{-1}(u_d))$$

```
xCopulaUniforme = copularnd('Gaussian', xMatrizCorrelacion, 20000);
```

Realmente lo que hace la función de matlab “copularnd” es el paso a) y b) de la explicación anterior, por lo que con la función “copularnd” obtendremos $m= 20,000$ variables aleatorias uniformes $u_0^{sim_m}, u_2^{sim_m}, \dots, u_3^{sim_m}$, con las correlaciones especificadas.

Graficaré las variables aleatorias $u_2^{sim_m}$ y $u_3^{sim_m}$ para poder ver su correlación la cual es del 79.712% .

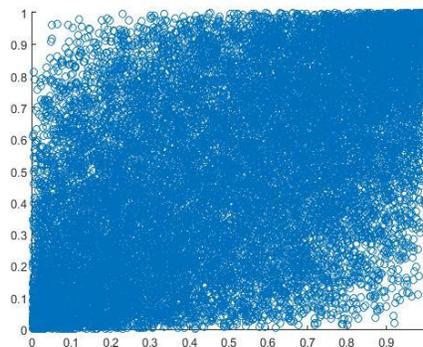


Figura 5.3 Gráfica de las variables aleatorias $u_2^{sim_m}$ y $u_3^{sim_m}$.

Por lo que para obtener los factores f_i 's simulados correlacionados, es decir el paso c) de nuestra explicación, bastara con aplicar $f_j^{sim} = F^{-1}(u_j^{sim_m})$ con los parámetros μ y σ calculados con 4) y 5)

`fiSim = logninv(xCopulaUniforme(:,j),mu(j),sigma(j))`

Por lo que ahora tendremos una matriz de factores de desarrollo simulados en R^3 de [9,4,20000].

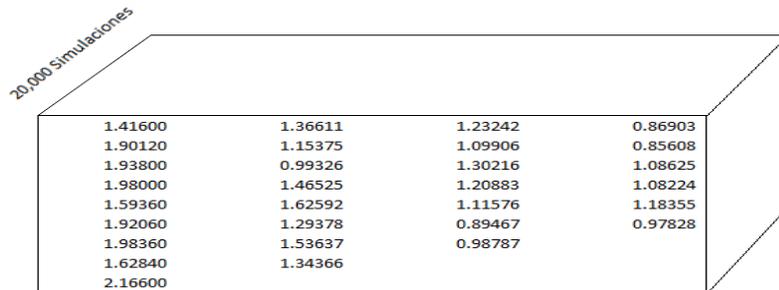


Figura 5.4 Matriz de factores de desarrollo.

Una vez que tenemos nuestros factores de desarrollo simulados, obtendremos la matriz acumulada simulada, A partir de nuestra ecuación (1).

$$E[C_{i,j}|C_{i,0}, \dots, C_{i,j-1}] = E[C_{i,j}|C_{i,j-1}] = \hat{C}_{i,j}^{CL} = f_{j-1}C_{i,j-1} \quad (1)$$

Por lo que con esto podremos obtener nuestros 20,000 triángulos simulados

Simulado/Acumulado	0	1	2	3	4
0	9,909	18,604	25,639	29,045	30,045
1	9,437	17,727	24,457	27,705	28,652
2	7,316	13,739	18,935	21,458	22,195
3	6,362	11,938	16,462	18,651	19,292
4	4,970	9,336	12,867	14,578	15,075
5	5,847	10,979	15,130	17,144	17,732
6	4,301	8,078	11,136	12,614	
7	4,893	9,193	12,678		
8	5,451	10,237			
9	6,454				

Figura 5.5 Matriz mostrando la primera simulación de la matriz de triángulos superiores simulados.

Para después calcular el Cape-Cod descrito anteriormente sobre cada uno de los triángulos simulados y obtener la estimación SONR y el índice de siniestralidad para calcular la reserva de riesgos en curso.

Por lo que de esta manera podríamos obtener la reserva SONR y la reserva RRC, tomando en cuenta la correlación que se puede tener entre los años de retraso.

6. Resultados Obtenidos

6.1 Prueba de BackTesting

De acuerdo al modelo descrito anteriormente, para poder evaluar su idoneidad, precisión y validez se realiza la prueba de Back-Testing.

Lo que busca el modelo es realizar las estimaciones del valor esperado de las obligaciones futuras de la compañía, para poder analizar el comportamiento de dichas estimaciones, la prueba de back-testing se basa en la comparación de las estimaciones realizadas con el modelo respecto de los valores reales observados.

Por ejemplo, si hoy es 31/12/2018 comparamos la siniestralidad real que se tiene actualmente, y comparamos con lo que pronosticamos un año anterior el 31/12/2017.

Para determinar la validez del modelo se utilizan períodos de 24 meses. En cada uno de los meses que conforman la ventana de observación se contrasta las estimaciones de siniestralidad en período t respecto de los valores reales de ese mismo período.

El modelo consiste en analizar el número de veces que las obligaciones reales observadas superan en magnitud el límite de tolerancia definido. El nivel de confianza y la amplitud del límite varían conjuntamente, de forma que un nivel de confianza más amplio tendrá más probabilidad de acierto, mientras que para un nivel más pequeño, determina una estimación más precisa.

A partir de las simulaciones de los índices de siniestralidad media, como se describe en las secciones anteriores, se obtiene el percentil 99.5% el cual se establece como límite superior. Es decir, que los valores reales observados se contrastan con este límite a fin de determinar el número de fallos o excepciones (número de veces en que el valor real supera la estimación del modelo) a partir de los cuales se considera si el modelo es adecuado, no es completamente confiable o no es aceptable.

Para el caso de SONR el percentil 99.5% se obtiene de los 20,000 triángulos simulados y la "siniestralidad real" se obtendrá a partir de los siniestros con retraso que se reportaron en este año, es decir los siniestros que se encuentran en la diagonal:

Año	Período de Inicio de	Retraso				
		0	1	2	3	4
2009	0	9,909	7,630	5,951	1,904	267
2010	1	9,437	8,871	7,452	3,130	908
2011	2	7,316	6,584	3,358	2,283	297
2012	3	6,362	6,234	5,860	3,633	1,817
2013	4	4,970	4,572	4,435	1,464	732
2014	5	5,847	5,730	5,444	2,014	363
2015	6	4,301	3,183	2,514	1,760	
2016	7	4,893	3,768	2,637		
2017	8	5,451	4,906			
2018	9	6,454				

Figura 6.1 En este caso para ver la idoneidad del modelo comparamos la diagonal marcada en azul (siniestros con retraso que se reportaron durante 2018), con la estimación que se realizó en 2017.

Para el caso de los la Reserva de Riesgos en Curso, se obtiene el percentil 99.5 sobre las simulaciones de los índices de siniestralidad y una vez que obtenemos este índice el $BEL_{RRC99.5\%}$ se obtendrá de manera análoga al BEL, para poder comparar el percentil con la siniestralidad real.

$$BEL_{RRC99.5\%} = \sum_{i=1}^n \pi_i * FND_i * k_{99.5\%}$$

Año	Período de Inicio de	Retraso				
		0	1	2	3	4
2009	0	9,909	7,630	5,951	1,904	267
2010	1	9,437	8,871	7,452	3,130	908
2011	2	7,316	6,584	3,358	2,283	297
2012	3	6,362	6,234	5,860	3,633	1,817
2013	4	4,970	4,572	4,435	1,464	732
2014	5	5,847	5,730	5,444	2,014	363
2015	6	4,301	3,183	2,514	1,760	
2016	7	4,893	3,768	2,637		
2017	8	5,451	4,906			
2018	9	6,454				

Figura 6.2 En este caso para ver la idoneidad del modelo comparamos la celda marcada en azul (siniestros de pólizas suscritas durante 2017 cuyos siniestros se reportaron ese mismo año), con la estimación de Reserva de Riesgos en Curso que se realizó en 2017 (la cuál es la estimación de la siniestralidad que tendrían las pólizas suscritas durante 2017).

Esta prueba compara el nivel de significancia con la proporción de excepciones del modelo, de tal manera que cada una de las observaciones se considera como un ensayo Bernoulli, con una probabilidad α de caer fuera del límite.

La siguiente tabla muestra las probabilidades exactas de obtener un número particular de excepciones en una muestra de 24 observaciones (2 años de observaciones mensuales).

Excepciones	$1 - P(X = x)$
0	11.33%
1	89.31%
2	99.38%
3	99.98%
4	100%

Lo que nos dice está tabla es que la probabilidad que se tenga 1 excepción (fracaso) en 24 ensayos Bernoulli cada uno con 99.5% de probabilidad de éxito (en un modelo adecuado), es del $100\% - 89.31\% = 10.7\%$, y la probabilidad de que se tengan 3 excepciones es del $100\% - 99.98\% = 0.02\%$ y así sucesivamente, hasta que se llega a que la probabilidad de se presenten 4 excepciones es casi nula, es decir si en la prueba de back-testing tenemos 4 o más excepciones el modelo no es adecuado.

Las tres zonas por color son definidos a manera de balancear dos tipos de errores estadísticos: el error Tipo I⁶ y el error Tipo II⁷, Con base en el número de excepciones, calculando la probabilidad del número de excepciones con una distribución binomial.

Los resultados de la tabla no consideran ningún umbral de excepciones que considere tanto una probabilidad baja de rechazar erróneamente un modelo preciso ni una probabilidad baja de aceptar erróneamente un modelo impreciso. El método propuesto por el Comité de Basilea clasifica los resultados en tres categorías, que involucran la definición de tres umbrales:

- Zona verde, los resultados de la prueba son consistentes con un modelo preciso y la probabilidad de aceptar erróneamente un modelo impreciso es baja.
- Zona amarilla donde los resultados del backtesting pueden ser consistentes con modelos precisos o imprecisos.

⁶ Tipo I La probabilidad de rechazar la hipótesis “el modelo es preciso” cuando en realidad lo es.

⁷ Tipo II La probabilidad de no rechazar la hipótesis “el modelo es preciso” cuando en realidad no lo es.

- Zona roja, los resultados de la prueba son sumamente improbables de haber sido generados por un modelo exacto, y la probabilidad de rechazar erróneamente un modelo preciso es casi nula o nula.

Por lo que en caso de que los resultados de la prueba se encuentren en el rango de la zona roja, lleva a la conclusión de que existe un problema en la confiabilidad o en la precisión del modelo.

A continuación se muestran los resultados de la prueba de back-testing:

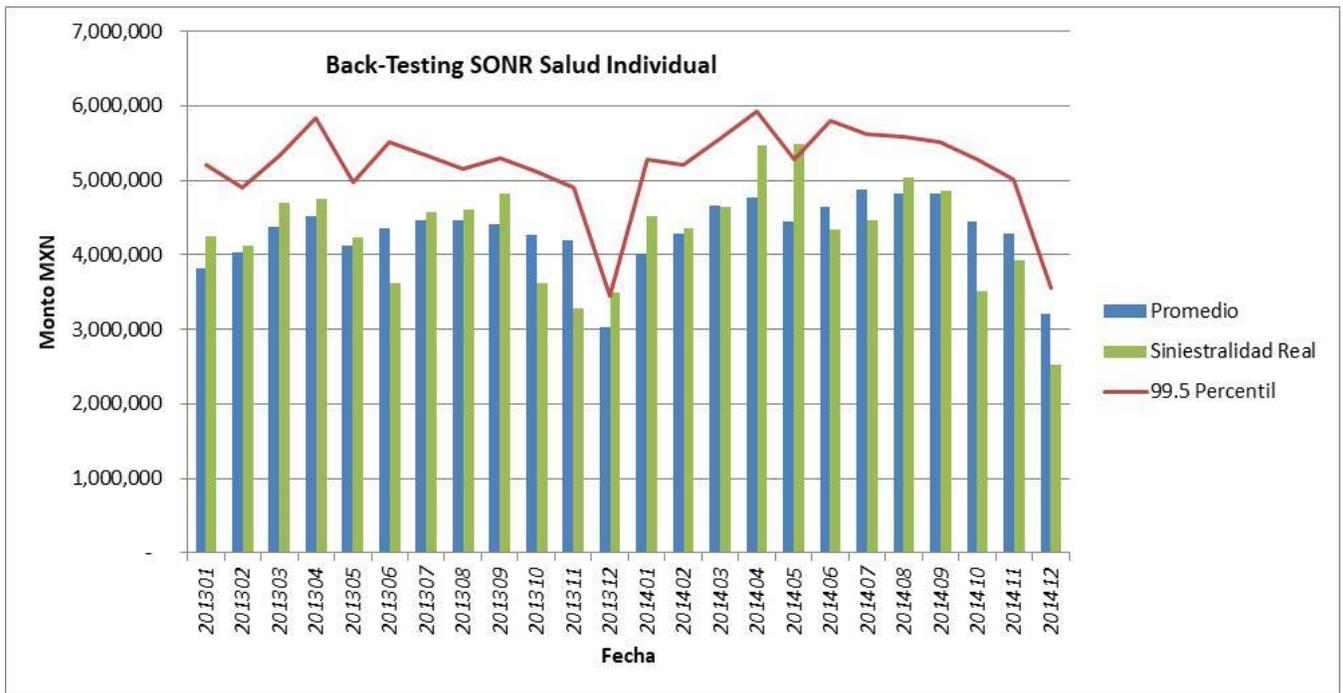


Figura 6.3 Se puede observar que sólo se tiene un exceso al percentil 99.5, por lo que se concluyó que el modelo pronostica de manera correcta las reclamaciones de SONR para el ramo de salud individual.

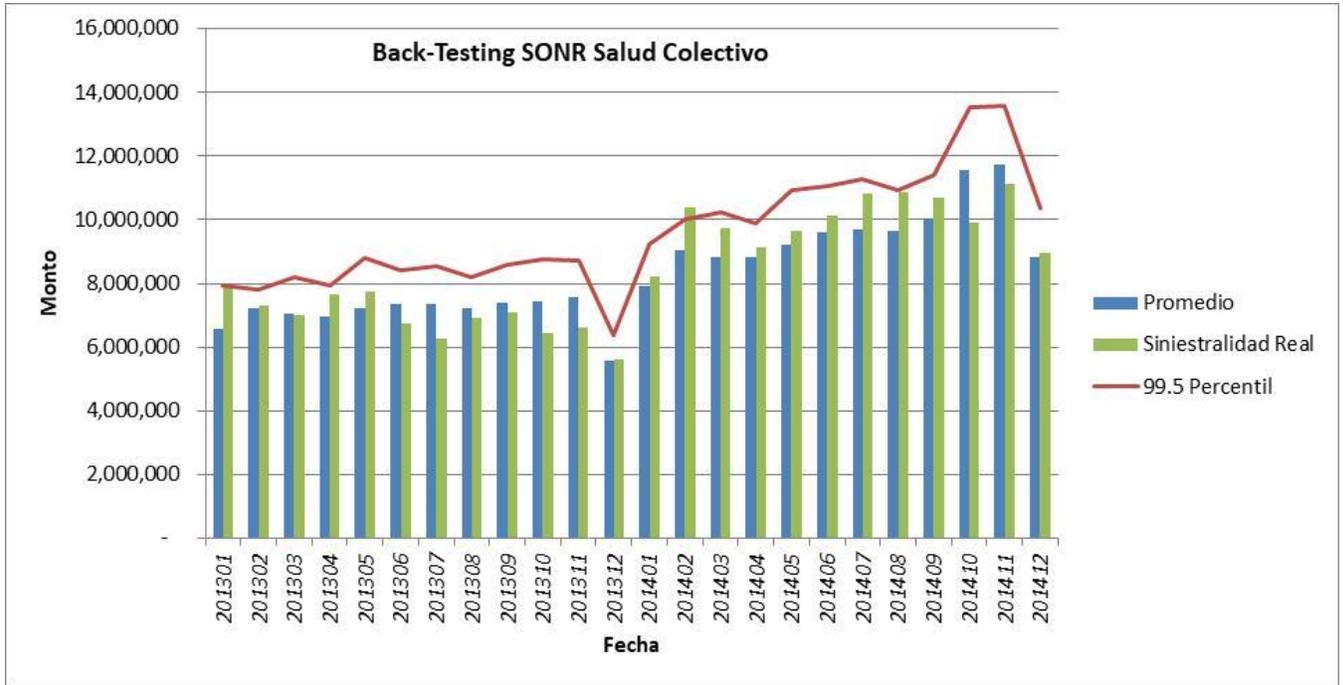


Figura 6.4 Se puede observar que sólo se tiene un exceso al percentil 99.5 , por lo que se concluyó que el modelo pronostica de manera correcta las reclamaciones de SONR para el ramo de salud colectivo.

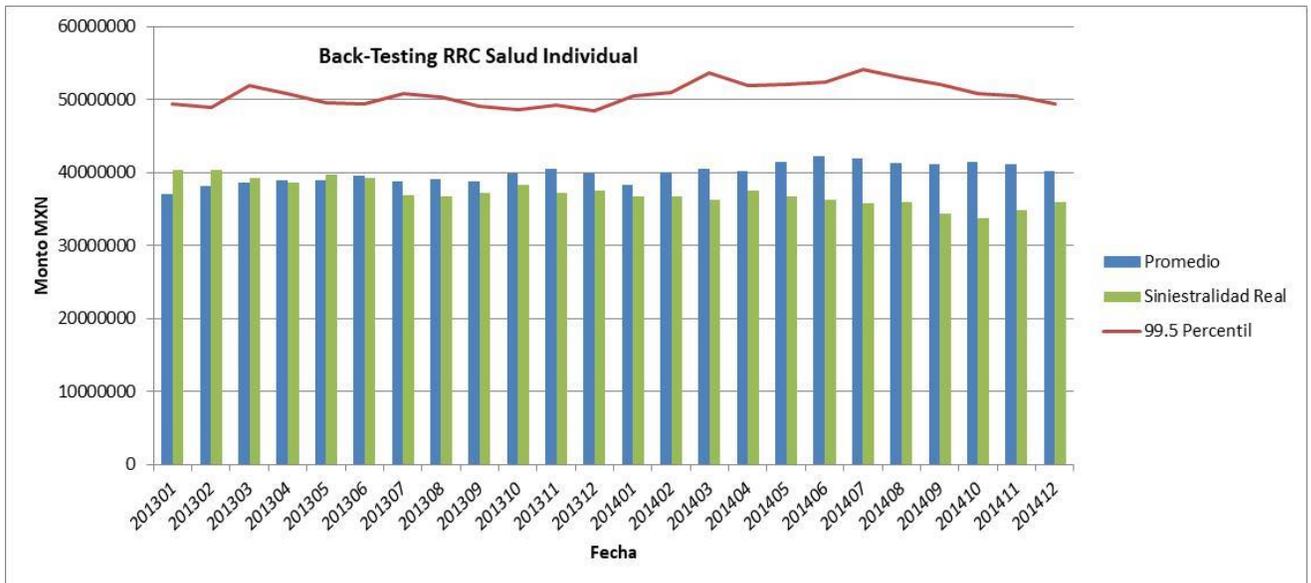


Figura 6.5 Se puede observar que no se tiene ningún exceso, por lo que se concluyó que el modelo pronostica de manera correcta las reclamaciones de riesgos en curso para el ramo de Salud Individual.

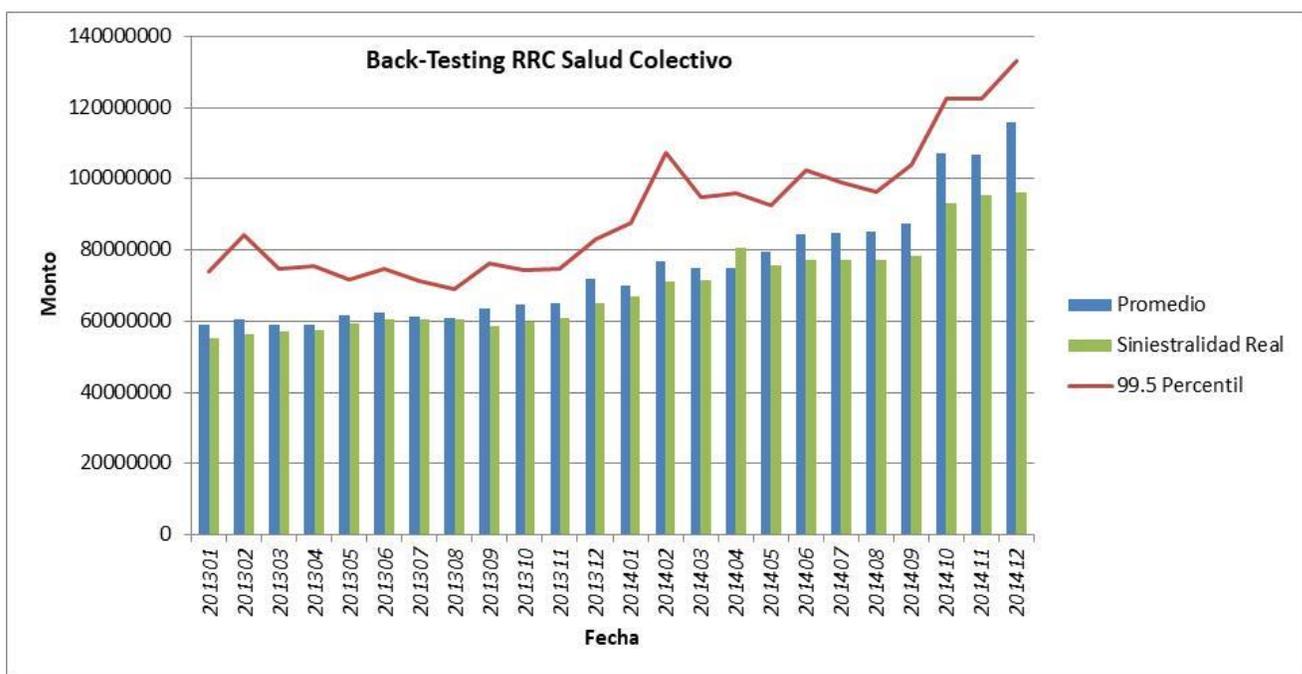


Figura 6.6 Se puede observar que no se tiene ningún exceso, por lo que se concluyó que el modelo pronostica de manera correcta las reclamaciones de riesgos en curso para el ramo de Salud Colectivo.

7. Conclusiones:

Mediante el uso de comandos vectoriales en Matlab se logró que el tiempo de cálculo de ambas reservas (SONR y RRC) sea menor a 3 minutos y el uso de la memoria RAM de las computadoras en las que se efectúa el cálculo sea óptimo al agrupar la información mediante las consultas en SQL.

El programa se implementó de tal forma que las 3 fases del proceso de reservas son completamente automáticas:

- 1) Extracción de información de las bases de datos directo del servidor, mostrando inconsistencias (fechas inconsistentes).
- 2) Proceso de la información.
- 3) Impresión de la resultados en distintos formatos:
 - a) Archivo de Excel para poder analizar los resultados.
 - b) Archivo en formato "csv" con el formato adecuado (cuentas y subcuentas contables) para su carga automática a contabilidad.
 - c) Archivos de Excel y en formato "txt", para entrega del Reporte Regulatorio 3 a la Comisión Nacional de Seguros y Fianzas.

Al automatizar completamente el cálculo de las reservas se disminuyó de manera considerable el riesgo operacional que conllevaba su cálculo de manera semiautomática, además de la reducción en el tiempo no solo en el cálculo mismo sino en la elaboración de los reportes regulatorios que podía llevar hasta 3-4 horas y ahora es menor a 5 minutos.

Con la implementación mediante del modelo mediante el método de bootstrap paramétrico se logró la correcta medición de las obligaciones futuras de la compañía como se pudo observar en el back-testing.

Se logró que las estimaciones sean consistentes y no tengan “picos” o comportamientos extremos.

8. Bibliografía

- 1- Wüthrich, M.V., Merz M. (2008), *Stochastic Claims Reserving Methods in Insurance*. John Wiley & Sons, Ltd
- 2 - Hertig, J. (1985). *A statistical approach to the IBNR-reserves in marine insurance*. Astin Bulletin 15/2, 171–183.
- 3- M. Wüthrich, ETH Zürich & M. Merz, Uni Tübingen (2006), *Stochastic Claims Reserving Methods in Non-Life Insurance*.
- 4 - *Ley de Instituciones de Seguros y Fianzas, Circular Única de Seguros y Fianzas*, Diario Oficial de la Federación del 4 de abril de 2013, Estados Unidos Mexicanos.
- 5- *Nota Técnica SONR*, Ramo Salud, Empresa de seguros especializada en el ramo de salud.
- 6- *Nota Técnica RRC*, Ramo Salud, Empresa de seguros especializada en el ramo de salud.
- 7- Rogelio Salinas-Gutiérrez, Arturo Hernández-Aguirre, Mariano J. J. Rivera-Meraz, and Enrique R. Villa-Diharce, *Using Gaussian Copulas in Supervised Probabilistic Classification*, Springer Berlin / Heidelberg, 355-372, Vol 318, 2011.
- 8- Wendy, 8 de enero de 2010, *Reservas IBNR*, recuperado de: <http://reservasibnr.blogspot.mx/>
- 9- Javier Rodríguez Coca, 11 de noviembre 2011, *Método para calcular Reserva para IBNR*, Piensa Ama, Recuperado de: https://piensaama.files.wordpress.com/2011/11/01_el-reseguro-np_ibnr_ama_20111111_f.pdf

10- Simon Barthelme, *Copulas made easy*, 28 de Octubre de 2011, dahtah, recuperado de:
<https://dahtah.wordpress.com/2011/10/28/hello-world/>