5
2E

Universidad Nacional Autónoma de México

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
"ACATLAN"

UNA COMPARACION ESTADISTICA DE MODELOS ALTERNATIVOS PARA EL CALCULO DE LA PRIMA NETA EN EL SEGURO DE AUTOMOVILES

T E S I S

Que para obtener el título de:

A C T U A R I O

P r e s e n t a:

Héctor Abraham Campos Sarabia

Santa Cruz Acatlán, a 16 de Marzo 1993

TESIS CON
FALLA DE ORIGEN

VNIVERSIDAD NACIONAL AVTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN"

DIVISION DE MATEMATICAS E INGENIERIA

PROGRAMA DE ACTUARIA Y M.A.C.

SR. HECTOR ABRAHAM CAMPOS SARABIA
Alumno de la carrera de Actuaría
Presente.

De acuerdo a su solicitud presentada con fecha 23 de febrero de 1993, me complace notificarle que esta Jefatura tuvo a bien asignarle el siguiente tema de tesis: "UNA COMPARACION ESTADISTICA DE MODELOS ALTERNATIVOS PARA EL CALCULO DE LA PRIMA NETA EN EL SEGURO DE AUTOMOVILES", el cual se desarrollará como sigue:

Indice.

Introducción.

I.- Modelos que no consideran por separado a las componentes de la prima neta para el cálculo de la misma.

II.- Un modelo que si considera por separado a las componentes de la prima neta para el cálculo de la misma.

III.- Comparación estadística de los modelos enunciados.

Conclusiones.

Apéndices.

Anexos.

Bibliografía.

Asimismo fue designado como Asesor de Tesis el MTRO. LUCIO PEREZ RODRIGUEZ, profesor de esta Escuela.

Ruego a usted tomar nota que en cumplimiento de lo especificado en la Ley de Profesiones, deberá prestar servicio social durante un tiempo mínimo de seis meses como requisito básico para sustentar examen profesional, así como de la disposición de la Coordinación de la Administración Escolar en el sentido de que se imprima en lugar visible de los ejemplares de la tesis el título del trabajo realizado. Esta comunicación deberá imprimirse en el interior de la tesis.

E.N.E.P. ACATLAN

Atentamente
"POR MI RAZA HABLARA EL ESPIRITU"
Acatlán, Edo. de Méx., a 23 de junio de 1993.

ACT. LAURA MA. RIVERA BECERRA
Jefe del Programa de Actuaría y M.A.C.

JEFATURA DEL PROGRAMA DE ACTUARIA Y MATEMATICAS APLICADAS Y COMPUTACION

INDICE

página

página

INTRODUCCION

La función del seguro consiste en brindar protección contra posibles contingencias, y así ofrecer estabilidad económica a aquella persona o institución que lo adquiera.

Una vez ideado el tipo de seguro y la cobertura del mismo, el siguiente paso es obtener la cantidad monetaria (prima) que ha de requerirse al asegurado, para contratar dicho beneficio. La prima ha de obtenerse a partir de los datos del asegurado y/o de lo asegurado, basándose en la experiencia en siniestralidad que se tenga al respecto, y a partir de ésto tratar de predecir las futuras contingenicas.

Así, se busca idear un modelo que prediga de la forma más exacta posible la futura siniestralidad, para así cobrar al asegurado el equivalente a su pérdida potencial (prima neta) más una cantidad que constituirá los gastos inherentes al seguro y la ganancia de la compañía aseguradora.

El presente trabajo se enfoca al cálculo de la prima neta en el seguro de automóviles, comparando 5 distintos modelos que están divididos en 2 categorías: aquellos que no consideran por separado a las componentes de la prima neta (frecuencia y severidad del reclamo), y un modelo que sí hace dicha separación. Tratará de verificarse, por medio del análisis de regresión, que las componentes de la prima neta están influenciadas por factores no comumes, e incluso que factores comunes ejercen una distinta influencia sobre las mismas. De aquí que también se espere que este modelo ofrezca un mejor nivel de predicción.

Los datos de los cuales se tomaron los resultados provienen de la experiencia canadiense para los años de 1984 a 1987.

CAPITULO I

MODELOS QUE NO CONSIDERAN POR SEPARADO A LAS COMPONENTES DE LA PRIMA NETA PARA EL CALCULO DE LA MISMA

1.1 MODELO 1: MODELO DE LA CELDA MEDIA.

La Prima Neta es la cantidad monetaria que se le requiere a una persona que desee contratar un Seguro (en este caso, un Seguro para su automóvil), y que es igual a la pérdida esperada.

El propósito de un modelo para el cálculo de la prima neta es predecir, con la mayor exactitud posible, la futura siniestralidad. Esto tiene como objeto dos finalidades:

1. No cobrar una prima tan baja como para que no sea suficiente resarcir los daños futuros.

2. No cobrar una prima tan alta como para que los asegurados potenciales decidan asumir las consecuencias de un siniestro por cuenta propia.

En la medida en la que la predicción sea más acertada, se estará cobrando una prima más razonable.

1.1.1 MANEJO DE LA INFORMACION.

El uso de la pérdida promedio para el cobro de la prima neta sugiere una similitud en las características generales de las personas que desean contratar el seguro.

En una cartera se tienen asegurados con características distintas bastante marcadas. De aquí se divide a la cartera en una serie de categorías, en donde se asume que todos los individuos pertenecientes a una misma categoría o celda tienen las mismas características.

Una categoría estará definida de acuerdo a una compleja combinación de variables o características generales. El criterio para tomar en cuenta dichas variables, es suponer que ellas tengan un real impacto en el número y el monto de los reclamos. De esta forma, los puntos que se toman en cuenta para dicha división son:

- Características del automóvil: Modelo, marca, tipo, uso, etc.

- Características del principal conductor: Edad, sexo, estado civil, etc.

- Territorio: Urbano y no urbano, etc.

- Récord previo de manejo: 3 o más años sin reclamo, y otros, etc.

El uso de estos cuatro puntos para hacer una división en categorías es arbitrario, y dependerá del tipo de cobertura que se contrate y de la experiencia de la compañía aseguradora, así como de la información disponible.

Para este modelo y para los subsecuentes se utilizará la experiencia canadiense para los años de 1984 a 1987, en la cobertura de colisión para automóviles (se excluyen camiones y flotillas).

En el presente trabajo, se clasifican los datos, para la mencionada cobertura, en base a una tabla tridimensional, definida por los siguientes rubros:

i) Territorio: Urbano y No Urbano.

ii) Características del automóvil y del principal conductor: Tabla de Edad y Uso (Tabla I)

iii) Récord de Manejo: 3 o más años sin reclamo, y otros.

TABLA I

Tabla de Edad y Uso

Variables	Clases 1	2	3	4	5	6	7	8
Edad	1	1	1	0	1	0	0	0
Sexo	0	0	0	1	0	0	1	1
Estado Civil	0	0	0	0	0	0	1	0
Placer	1	0	0	0	0	0	0	0
Negocios	0	0	0	0	1	0	0	0
Distancia	0	0	1	0	0	0	0	0
Manejo Ocasional	0	0	0	1	0	0	0	0

Observaciones:

- Edad: 1 si el principal operador tiene 25 años o más;

  0 de otra forma.

- Sexo: 1 si es hombre;

  0 de otra forma.

- Estado Civil: 1 si es casado (a);

  0 de otra forma.

- Placer: 1 si maneja sólo por placer y una distancia anual de 16,000 km. ó menos;

  0 de otra forma.

- Negocios: 1 si el automóvil se usa para negocios;

  0 de otra forma.

- Distancia: 1 si la distancia al trabajo es de 16 km. ó más;

  0 de otra forma.

Esta clasificación arroja un total de 32 categorías, es decir, 32 distintas combinaciones.

Como puede observarse, habrá asegurados que, dadas sus características generales, no pertenecerán en un sentido estricto a alguna de las categorías aquí enunciadas. En este caso la compañía aseguradora decidirá en cuál categoría encaja dicho individuo. Un criterio para ello podría basarse en el comportamiento similar en la siniestralidad.

¿Por qué no mejor hacer un número mayor de categorías? Pareciera ser que entre más refinada sea la división de la cartera, más exacta será la predicción de la pérdida para cada celda.

Entre más se hagan subdivisiones en la cartera, menos elementos se tendrán en cada categoría. Así, habrá algunas celdas con un número de exposiciones bajo. Consecuentemente no es muy confiable calcular a la prima neta como la pérdida promedio, ya que se pueden presentar grandes desviaciones debido al número tan bajo de elementos. Este problema es conocido por los actuarios como el "Problema de la Credibilidad".

1.1.2 ESTIMACION.

Desde un punto de vista estadístico, se puede tomar a la prima neta (PN) como un parámetro, y estimarlo. Este parámetro, que es la pérdida promedio para una categoría particular y para un año dado, es:

$$\hat{PN}_g = \bar{y}_g = \sum_{h=1}^{n_g} y_{gh}/n_g \,, \qquad g=\overline{1,32} \qquad (1.1.2.1)$$

donde:

y_{gh} pérdida del h-ésimo individuo que pertenece a la g-ésima categoría

n_g número de exposiciones en la g-ésima categoría.

1.1.3 PROPIEDAD DEL ESTIMADOR.

Esta forma de cálculo de la prima neta nos brinda estimadores que son insesgados independientemente del modelo generatriz de la pérdida de cada individuo.

PRUEBA:

Tómese la g-ésima categoría. Puesto que se trabaja bajo el supuesto de que todos los individuos pertenecientes a una misma categoría tienen las mismas características, el comportamiento de la siniestralidad de cada uno se distribuye de forma idéntica. Sean $Y_{g1},\ldots,Y_{gn_g}$ variables aleatorias independientes idénticamente distribuidas con media $E[Y_g]$, que representan el comportamiento de la siniestralidad.

$$E[\hat{PN}] = E[\bar{Y}_g]$$

$$= E[\sum_{h=1}^{n_g} Y_{gh}/n_g]$$

$$= E[(1/n_g)\sum_{h=1}^{n_g} Y_{gh}]$$

$$= (1/n_a)E[\sum_{h=1}^{n_a} Y_{ah}]$$

$$= (1/n_a)\sum_{h=1}^{n_a} E[Y_{ah}]$$

$$= (1/n_a)\sum_{h=1}^{n_a} E[Y_a]$$

$$= (1/n_a)E[Y_a]\sum_{h=1}^{n_a} 1$$

$$= (1/n_a)E[Y_a](n_a)$$

$$= E[Y_a] \qquad \text{l.q.q.d.}$$

Esto quiere decir que, en promedio, el estimador propuesto será igual al valor verdadero que se tiene en consideración. En otras palabras, la estimación de la prima neta será igual, en promedio, a la pérdida esperada, lo que es el objeto de este modelo.

1.1.4 PRONOSTICO.

La Prima Neta obtenida como la pérdida promedio para una determinada categoría se basa en información pasada. De esta forma debe existir un ajuste en el cobro de la mencionada prima a los presentes asegurados. Dicho ajuste se hará contemplando procesos inflacionarios.

1.1.5 EJEMPLO ILUSTRATIVO.

Tómese la información del Estado de Massachusetts, E.E. U.U., para los años de 1974 y 1975 para una cobertura obligatoria combinada. Esta cobertura contiene lo siguiente: Responsabilidad Civil por daños a terceros en sus personas ($5,000 dólares / $10,000 dólares), gastos médicos a ocupantes, responsabilidad civil por daños a terceros en sus bienes ($5,000 dólares). (Nótese cómo de acuerdo al tipo de cobertura y la experiencia de la compañía y de una forma arbitraria varía la forma de división en categorías. Aquí una categoría está definida sólo por dos rubros: Clase de Conductor y Territorio).

Para el año de 1975, la Tabla II muestra la pérdida promedio obtenida para cada categoría. Dicha pérdida promedio es el cociente dado por la pérdida potencial total de cada categoría y su respectivo número de exposiciones. Este número de exposiciones se encuentra en la Tabla III.

Análogamente las Tablas IV y V para el año de 1976.

TABLA II

Monto Promedio de Reclamos Observado por Territorio y Clase de Conductor
Cobertura Obligatoria Combinada
Massachusetts, 1975
(Dólares)

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42
1	25.05	26.28	44.03	48.97	48.35	65.48	121.45
2	19.15	25.66	30.70	50.94	32.89	60.53	97.65
3	30.63	30.92	40.19	54.69	66.24	79.58	114.12
4	28.93	30.48	37.86	52.55	59.07	72.92	117.63
5	27.81	35.11	42.08	52.98	63.51	94.35	126.56
6	29.51	36.15	46.63	57.48	75.36	81.20	153.75
7	36.28	39.50	42.50	60.49	71.35	86.67	156.11
8	34.59	40.61	53.41	60.31	81.41	93.19	133.87
9	40.62	42.77	67.34	60.97	54.62	93.87	184.96
10	43.73	48.77	55.38	73.54	75.35	103.53	153.65
11	37.03	42.19	63.93	69.61	84.26	111.02	129.92
12	33.56	49.78	58.82	82.78	83.69	112.90	127.49
13	47.32	49.67	99.76	92.52	101.86	108.08	150.91
14	38.65	66.94	58.51	77.90	125.98	116.72	168.38
15	38.68	69.74	76.07	82.00	103.33	116.55	162.71

TABLA III

Distribución Conjunta de Exposiciones por Territorio y Clase de Conductor
Cobertura Obligatoria Combinada
Massachusetts, 1975

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42
1	6,967	44,738	3,309	5,781	1,101	1,638	4,107
2	6,103	50,974	3,082	7,329	1,457	2,235	4,797
3	17,745	192,369	13,626	28,218	6,504	8,531	19,738
4	14,978	157,357	12,539	22,038	5,326	7,495	14,448
5	19,552	217,426	19,293	31,470	6,721	10,075	20,600
6	27,850	195,661	16,490	28,207	6,227	9,165	17,883
7	17,485	201,263	16,706	30,501	6,561	9,578	19,427
8	22,417	333,416	21,719	35,338	7,993	12,335	22,018
9	5,204	49,283	3,692	7,818	2,087	2,315	4,493
10	13,375	110,071	9,136	15,954	4,282	5,040	9,895
11	2,733	27,629	1,821	5,341	909	1,632	2,077
12	2,036	24,837	1,716	2,791	659	969	2,149
13	1,323	16,718	666	2,318	615	764	1,715
14	1,350	16,091	830	2,290	562	903	1,394
15	8,209	91,947	5,305	9,995	2,656	3,457	6,734

TABLA IV

Monto Promedio de Reclamos Observado por Territorio y Clase de Conductor
Cobertura obligatoria combinada
Massachusetts, 1976
(Dolares)

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42
1	21.91	29.11	37.41	57.91	60.74	75.47	126.52
2	21.41	25.93	31.45	48.49	57.75	75.34	95.28
3	32.25	25.83	35.10	40.96	52.79	72.76	127.13
4	25.36	28.17	42.20	48.84	91.76	67.91	125.56
5	30.49	30.59	30.51	52.96	86.35	75.98	125.87
6	33.47	34.53	39.13	59.17	58.59	85.62	134.29
7	31.55	36.62	51.24	61.49	66.09	84.70	143.57
8	30.82	37.85	45.66	57.92	66.57	91.11	155.84
9	33.61	44.28	47.59	55.76	86.67	95.84	151.21
10	30.72	47.59	53.26	80.42	86.66	90.45	158.75
11	27.42	43.55	67.57	63.95	66.35	107.13	146.76
12	43.04	39.06	49.45	80.69	63.76	75.47	135.81
13	51.11	55.89	69.97	77.35	101.52	97.55	153.31
14	55.32	51.97	132.06	54.21	111.57	66.75	183.99
15	47.73	66.55	91.96	81.49	155.23	133.52	188.87

TABLA V

Distribución conjunta de exposiciones por Territorio y Clase de Conductor
Cobertura obligatoria combinada
Massachusetts, 1976

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42
1	8,345	45,534	3,197	5,926	1,085	1,599	4,880
2	7,613	51,591	3,471	7,351	1,201	2,173	4,151
3	31,043	195,103	12,644	28,442	6,347	8,363	19,213
4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8,347	[illegible]
5	[illegible]	219,222	18,235	32,151	6,227	10,106	20,376
6	[illegible]	[illegible]	[illegible]	[illegible]	5,731	[illegible]	[illegible]
7	[illegible]	201,310	15,406	30,576	5,923	[illegible]	[illegible]
8	[illegible]	235,487	20,393	39,527	7,399	12,977	22,231
9	6,510	[illegible]	[illegible]	7,636	[illegible]	[illegible]	[illegible]
10	[illegible]	[illegible]	[illegible]	13,426	[illegible]	[illegible]	[illegible]
11	[illegible]	[illegible]	[illegible]	[illegible]	857	1,586	[illegible]
12	[illegible]	[illegible]	[illegible]	[illegible]	357	[illegible]	[illegible]
13	[illegible]	[illegible]	806	[illegible]	[illegible]	698	[illegible]
14	[illegible]	[illegible]	[illegible]	[illegible]	401	[illegible]	[illegible]
15	9,646	66,359	4,415	6,602	2,112	3,391	5,377

donde:

Clase de Conductor	Descripción
15	Conductor con más de 65 años.
10&12	Conductor entre 25 y 65 años.
30&31	Uso de negocios.
24&26	Mujer soltera menor a 25 años.
50	Casado (a) menor a 25 años.
20&40	Hombre soltero menor a 25 años, no dueño del auto.
22&42	Hombre soltero menor a 25 años, dueño del auto.

En este ejemplo se tiene un total de 105 categorías o celdas.

1.1.6 VENTAJA DEL MODELO.

Por medio de este modelo se obtienen estimadores insesgados de la prima neta para cada celda.

Si el número de exposiciones en cada celda fuera muy grande y el proceso de generación de reclamos fuera estable a través del tiempo, este estimador sería muy preciso.

1.1.7 DESVENTAJAS DEL MODELO.

Las desventajas que presenta el uso de este modelo son las siguientes:

- Este estimador no será tan preciso como sería deseable, ya que la experiencia en automóviles marca que la generación de reclamos no es tan estable año con año. Adicionalmente, es muy probable que se presenten casos de celdas con un número bajo de exposiciones.

Obsérvese en las Tablas III y V la categoría definida por la combinación Territorio 14 y Clase de Conductor 30&31. El número de exposiciones que se tiene es bajo. Para el año de 1975 se tiene un reclamo promedio de $58.51 dólares. Para 1976 $132.06 dólares. Esto es, hubo un incremento del 125.7% (más del doble) en el monto de reclamo promedio.

El hecho de usar así la pérdida promedio para el cálculo de la prima neta, trajo como consecuencia una importante pérdida monetaria en esta celda, dada la desviación desfavorable que se presentó.

- La división en categorías es arbitraria, a pesar de estar basada en la experiencia que se tenga del comportamiento de la siniestralidad en la cartera total.

No existe un marco teórico que lleve a concluir si una determinada combinación de características es la adecuada para definir una clase de riesgo o una categoría para la utilización de este modelo de la celda media. Este hecho puede además ser el que lleve a importantes desviaciones en cualquier celda.

- Se trabaja bajo el supuesto de que no existe interrelación alguna entre las distintas celdas; pero las celdas no están, de hecho, completamente aisladas, ya que son parte de un sistema de clasificación cruzada con muchas interrelaciones.

Al trabajar con una celda totalmente aislada se está desperdiciando información importante que nos pueden proporcionar las otras celdas. Por ejemplo, se desperdicia la información que nos puede proporcionar el comportamiento de otras clases en el mismo territorio en el que se encuentra la categoría en cuestión, aunque no así en la misma categoría.

- Las dos componentes de la prima neta, Frecuencia del Reclamo y Severidad del Reclamo, no están separadas. Bien podría ser que esta componentes estén influenciadas por características generales no comunes, e incluso que algunas de las características que sí les sean comunes a las dos componentes tengan un distinto impacto sobre cada una de ellas.

1.2 MODELO 2: MODELO MULTIPLICATIVO.

Una desventaja del modelo de la Celda Media es que considera, para el cálculo de la prima neta, un aislamiento total de las celdas, impidiendo así el hecho de considerar una interrelación entre las mismas dado que cuenten con características similares.

Esto lleva a perder valiosa información que puede ser proporcionada por las categorías complementarias. Obsérvese por ejemplo un punto que define a una categoría: el territorio. En el modelo de la celda media se está ignorando por completo el comportamiento en cuanto a siniestralidad de exposiciones que, aunque no se encuentran en la misma categoría, sí pueden encontrarse en el mismo territorio.

El modelo multiplicativo salva en cierta medida esta deficiencia del modelo de la celda media al considerar, para una categoría particular, la información que se tiene en las otras categorías que tienen uno o más puntos en común con aquella.

1.2.1 MANEJO DE LA INFORMACION.

La prima neta para este modelo se obtiene del producto de una prima neta global, obtenida por la pérdida promedio de la cartera total, y tres relatividades, dado que son tres puntos los que se consideran aquí para definir a una categoría. Estas relatividades representan el efecto de encontrarse en un determinado rubro que define a la categoría.

Sea P_{ijk} la prima neta indicada a cobrar a todo individuo que se encuentre en la categoría definida por la combinación i-ésimo territorio, j-ésima clase y k-ésimo récord de manejo. Se define:

$$P_{ijk} = p_{...}t_i c_j r_k \qquad (1.2.1.1)$$

$i = 0,1$

$j = \overline{1,8}$

$k = 0,1$

donde p... es la pérdida promedio general, es decir:

$$p_{...} = \sum_{ijk} p_{ijk} n_{ijk} / n_{...}$$

p_{ijk} es la pérdida promedio de la combinación (i,j,k), siendo n_{ijk} el número de exposiciones en tal categoría. n... el número total de exposiciones en la cartera:

$$n_{...} = \sum_{ijk} n_{ijk}$$

En palabras:

$$\text{pérdida promedio general} = \frac{\text{monto total de reclamos}}{\text{número total de exposiciones}}$$

Así, la prima neta para una cierta categoría será igual a una pérdida promedio general, que será multiplicada por factores que representan el efecto de pertenecer a los puntos que dan forma a la respectiva categoría. Si, por ejemplo, el territorio en el que se encuentra la exposición es de alto riesgo, es de esperarse que ésta pague una prima neta más alta con respecto al estándar, y de ahí que t_i tendría que tomar un valor mayor que 1, o menor que 1 si se encuentra en un territorio de menor riesgo. Así se ve también con la clase y el récord.

1.2.2 ESTIMACION.

Ahora se trata de buscar estimadores para las relatividades.

Una primera aproximación hacia este punto es el considerar los siguientes:

$$\hat{t}_i = \frac{p_{i..}}{p_{...}}$$

$$\hat{c}_j = \frac{p_{.j.}}{p_{...}}$$

$$\hat{r}_k = \frac{p_{..k}}{p_{...}}$$

donde:

$$p_{i..} = \sum_{jk} \frac{n_{ijk}\, p_{ijk}}{n_{i..}}$$

$$p_{.j.} = \sum_{ik} \frac{n_{ijk}\, p_{ijk}}{n_{.j.}}$$

$$p_{..k} = \sum_{ij} \frac{n_{ijk}\, p_{ijk}}{n_{..k}}$$

Traducido a palabras:

$$\hat{t}_i = \frac{\text{monto promedio de reclamos del territorio i}}{\text{monto promedio de reclamos de la cartera total}}$$

$$\hat{c}_i = \frac{\text{monto promedio de reclamos de la clase j}}{\text{monto promedio de reclamos de la cartera total}}$$

$$\hat{r}_k = \frac{\text{monto promedio de reclamos del récord k}}{\text{monto promedio de reclamos de la cartera total}}$$

$$p_{i..} = \frac{\text{monto total de reclamos del territorio i}}{\text{número total de exposiciones del territorio i}}$$

$$p_{.j.} = \frac{\text{monto total de reclamos de la clase j}}{\text{número total de exposiciones del territorio i}}$$

$$p_{..k} = \frac{\text{monto total de reclamos del récord k}}{\text{número total de exposiciones del territorio i}}$$

La pérdida promedio general es, en primera instancia, multiplicada por la relatividad territorial. Si el monto promedio de reclamos del territorio i_1 es mayor que el monto promedio de reclamos de la cartera total, el cociente que define a la relatividad territorial será mayor que 1, lo que trae como consecuencia un aumento en la prima neta a cobrar a la categoría aquí tratada con respecto a una prima neta general. Esto es bastante claro al tomar en cuenta que, como previamente se había mencionado, se cobra una prima neta más alta a una persona que circula regularmente por un territorio en donde el monto de cada siniestro es, en promedio, más alto que el estándar.

A la inversa: si el monto promedio de reclamos del territorio i_1 es menor que el monto promedio de reclamos de la cartera total, la relatividad territorial será menor que 1, lo que trae una disminución en la prima neta a cobrar a la presente persona con respecto a una prima neta global. La interpretación es análoga.

Lo mismo es para la relatividad de clase y para la relatividad de récord.

En resumen, la prima neta para una categoría particular es obtenida como el producto de una pérdida promedio general, y tres cocientes o relatividades menores o mayores que 1, que respectivamente disminuyen o aumentan (en forma relativa al estándar) la prima neta a cobrar, según sea el rubro tratado de menor o mayor riesgo que forma a la categoría.

1.2.3 PROPIEDADES DE LOS ESTIMADORES.

Una característica deseable en este tipo de estimadores es el que éstos sean "Balanceados".

Se dice que un estimador es balanceado cuando la Prima Neta obtenida por éste, perteneciente a un solo rubro completo, iguala a la pérdida promedio de ese mismo rubro. En otras palabras, la Prima Neta Indicada para un rubro debe ser igual a la pérdida promedio del mismo

Tomando los estimadores dados, se tiene que la prima neta indicada para el i-ésimo territorio es:

$$\hat{P}_{i..} = \sum_{jk} p_{...}\hat{t}_i\hat{c}_j\hat{r}_k n_{ijk}/n_{i..}$$

$$= \sum_{jk} p_{...}(p_{i..}/p_{...})(p_{.j.}/p_{...})(p_{..k}/p_{...})n_{ijk}/n_{i..}$$

$$= \sum_{jk} (p_{i..}p_{.j.}p_{..k}/p_{...}^{2})n_{ijk}/n_{i..}$$

$$= p_{i..}\left[\sum_{jk}(p_{.j.}p_{..k}/p_{...}^{2})n_{ijk}/n_{i..}\right]$$

Puesto que en general el factor entre corchetes es distinto de 1, la prima neta indicada para el j-ésimo territorio es distinta de la pérdida promedio observada para el mismo. Así, tenemos una "Falta de Balance".

Similarmente, para la clase y el récord:

$$\hat{P}_{.j.} = p_{.j.}\left[\sum_{jk} (p_{.i.}p_{..k}/p_{...}^{2})n_{ijk}/n_{.j.} \right]$$

$$\hat{P}_{..k} = p_{..k}\left[\sum_{jk} (p_{.i.}p_{.j.}/p_{...}^{2})n_{ijk}/n_{..k} \right]$$

De esta forma, se observa que los estimadores propuestos no son balanceados para cada rubro tratado.

AJUSTE DE LOS ESTIMADORES.

Obtener un estimador balanceado para t_i significa lograr que $\hat{P}_{i..} = p_{i..}$ Así, bastará con hacer que :

$$\hat{t}_i = \frac{p_{i..}}{p_{...}} \Big/ \left[\sum_{jk} (p_{.j.}p_{..k}/p_{...}^{2})n_{ijk}/n_{i..} \right]$$

$$= \frac{p_{i..}}{p_{...}} \Big/ \left[\sum_{jk} \hat{c}_j\hat{r}_k n_{ijk}/n_{i..} \right]$$

Análogamente:

$$\hat{c}_j = \frac{p_{.j.}}{p_{...}} \Big/ \left[\sum_{jk} (p_{i..}\,p_{..k}/p_{...}^{2})\,n_{ijk}/n_{.j.} \right]$$

$$= \frac{p_{.j.}}{p_{...}} \Big/ \left[\sum_{jk} \hat{t}_i \hat{r}_k n_{ijk}/n_{.j.} \right]$$

$$\hat{r}_k = \frac{p_{..k}}{p_{...}} \Big/ \left[\sum_{ij} (p_{i..}\,p_{.j.}/p_{...}^{2})\,n_{ijk}/n_{..k} \right]$$

$$= \frac{p_{..k}}{p_{...}} \Big/ \left[\sum_{jk} \hat{t}_i \hat{c}_j n_{ijk}/n_{..k} \right]$$

El problema que ahora se presenta es que no se da un balance simultáneo, ya que al balancear un estimador se desbalancea al otro puesto que son dependientes entre sí.

Podría tomarse el balance en un solo rubro y los otros dos estimadores dejarlos desbalanceados. Esto es, por ejemplo, para el territorio:

$$\hat{t}_i = \frac{p_{i..}}{p_{...}} \Big/ \left[\sum_{jk} \hat{c}_j \hat{r}_k n_{ijk}/n_{i..} \right]$$

$$\hat{c}_j = \frac{p_{.j.}}{p_{...}}$$

$$\hat{r}_k = \frac{p_{..k}}{p_{...}}$$

Como un asunto teórico sería deseable encontrar estimadores balanceados simultáneamente. Sin embargo, es sabido que en muchas situaciones prácticas el monto de esta falta de balance tiende a ser relativamente pequeño, y podría ser considerado tolerable.

1.2.4 PRONOSTICO.

Para efectos de pronóstico, la prima neta estimada se ajusta de la forma mencionada en el Modelo de la Celda Media.

1.2.5 EJEMPLO ILUSTRATIVO

A continuación se presenta un ejemplo para el Estado de Massachusetts, E.E. U.U. (1976), que ilustra el procedimiento que se enuncia para el modelo aquí tratado. Para este caso particular, la prima neta para una categoría específica se obtiene como el producto de una pérdida promedio general y dos relatividades, dado que las categorías se encuentran en un esquema de dos dimensiones:

$$\begin{bmatrix}\text{prima neta para}\\ \text{una categoría}\\ \text{particular}\end{bmatrix} = \begin{bmatrix}\text{pérdida}\\ \text{promedio}\\ \text{general}\end{bmatrix} \text{X} \begin{bmatrix}\text{relativi-}\\ \text{dad ter-}\\ \text{ritorial}\end{bmatrix} \text{X} \begin{bmatrix}\text{relatividad}\\ \text{de clase de}\\ \text{conductor}\end{bmatrix}$$

Tómense nuevamente las Tablas IV y V. Obsérvese que ahora se han agregado un renglón y una columna que definen, respectivamente, el reclamo promedio por clase de conductor, y el monto promedio por territorio.

TABLA IVa

Monto Promedio de Reclamos Observado por Territorio y Clase de Conductor
Cobertura obligatoria combinada
Massachusetts, 1976
(Dólares)

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42	promedio ponderado
1	21.91	29.11	37.41	57.91	60.74	75.47	126.52	$38.33
2	21.81	23.93	31.45	40.49	57.75	74.34	95.21	$33.98
3	22.25	29.83	35.10	48.96	52.79	72.76	122.13	$40.33
4	25.36	28.37	52.20	48.84	53.76	67.53	125.56	$38.86
5	30.49	30.59	38.51	52.96	66.35	75.90	126.87	$41.86
6	33.47	34.53	39.11	59.17	58.55	95.62	134.29	$45.36
7	31.55	36.62	51.24	61.49	66.05	81.78	143.67	$48.26
8	30.82	37.82	45.50	57.97	66.53	91.11	156.06	$49.51
9	33.51	[illegible]	47.59	55.46	66.82	95.86	151.83	$52.19
10	38.72	47.59	53.26	68.42	89.66	98.45	150.25	$57.07
11	27.42	43.55	67.57	63.95	65.36	107.15	146.46	$56.75
12	43.04	39.06	69.45	60.60	63.75	74.57	134.61	$40.25
13	51.11	55.09	69.97	73.29	106.52	97.55	163.37	$65.97
14	55.92	51.97	132.05	54.21	114.97	86.75	183.88	$62.59
15	47.73	66.55	91.96	81.59	155.23	133.52	188.97	$75.09
promedio ponderado	$32.21	$36.58	$45.63	$57.42	$72.60	$86.14	$137.20	$46.97

Usando los estimadores iniciales se tiene la siguiente tabla:

TABLA VI

Estimadores Iniciales del Modelo Multiplicativo

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42	promedio ponderado
1	26.29	29.05	37.24	46.86	59.24	70.29	111.96	$37.39
2	23.31	26.47	33.02	41.55	52.53	62.32	99.21	$33.72
3	27.59	31.33	39.00	49.18	62.10	73.78	117.51	$40.00
4	26.65	30.27	37.75	47.51	60.07	71.27	113.52	$38.83
5	28.16	31.98	39.89	50.20	63.47	75.31	119.96	$41.25
6	31.11	35.33	44.07	55.45	70.11	83.19	132.50	$45.35
7	33.10	37.59	46.89	59.00	74.60	88.51	140.98	$48.57
8	33.89	38.48	48.00	60.41	76.38	90.62	144.34	$49.37
9	36.29	41.11	51.28	64.56	81.59	96.81	154.20	$52.87
10	39.14	44.45	55.44	69.77	88.23	104.66	166.71	$56.43
11	37.55	42.64	53.19	66.94	84.63	100.41	159.94	$55.73
12	33.05	37.58	46.88	58.99	74.58	88.49	140.95	$46.57
13	44.52	50.55	63.06	79.35	100.32	119.03	189.60	$65.35
14	43.20	49.06	61.19	77.00	97.35	115.51	183.99	$62.33
15	51.30	58.49	72.96	91.81	116.07	137.72	215.37	$71.33
promedio ponderado	$32.28	$36.66	$45.15	$57.04	$73.03	$86.46	$135.73	$46.89

Obsérvese cómo efectivamente $p_{i.}$ es distinto de $\hat{P}_{i..}$, y $p_{.j}$ distinto de $\hat{P}_{.j}$ para $i=\overline{1,15}$, $j=\overline{1,7}$ (estimadores no balanceados).

Enseguida se tiene la tabla con la relatividad balanceada para el territorio:

TABLA VII

Modelo Multiplicativo con Ajuste de Estimadores para el Territorio

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42	promedio ponderado
1	26.94	30.60	30.16	48.03	60.72	72.06	114.75	$38.33
2	23.49	26.67	33.27	41.07	52.93	62.03	100.06	$33.90
3	37.20	30.90	38.65	48.53	61.49	72.36	116.21	$40.33
4	26.67	30.20	37.77	47.53	60.05	71.30	113.57	$38.06
5	28.06	31.44	39.71	49.97	63.10	76.56	119.51	$41.06
6	31.12	35.34	44.00	55.47	70.13	83.21	132.54	$45.36
7	33.02	37.27	46.19	58.51	73.97	87.76	139.00	$48.26
8	33.56	38.05	47.51	59.78	75.39	89.68	142.05	$49.41
9	36.29	41.27	51.10	64.60	81.78	97.03	156.56	$52.79
10	39.50	44.35	56.07	70.56	89.21	105.65	160.53	$57.07
11	36.00	41.40	52.26	65.74	83.12	98.67	157.09	$54.75
12	34.32	38.90	50.62	61.10	77.35	91.78	146.39	$48.25
13	45.35	50.36	62.15	79.05	99.96	110.58	188.80	$65.31
14	43.65	49.57	61.83	77.03	98.38	116.72	185.92	$62.93
15	54.22	61.50	76.81	96.65	122.20	144.99	236.95	$75.03
promedio ponderado	$32.37	$36.74	$45.18	$57.07	$73.12	$86.56	$135.70	$46.97

Ahora véase que $p_{i.}$ es idéntico a $\hat{P}_{i..}$, y el valor de $\hat{P}_{.j}$ no se aleja tanto de $p_{.j}$.

1.2.6 VENTAJA DEL MODELO.

Este modelo, como se mencionó, toma en cuenta la información adyacente a una cierta exposición enclavada en una determinada categoría. Esto es muy importante, ya que una categoría no puede estar realmente aislada respecto de las otras celdas. Así, este modelo da la pauta para considerar los efectos que tiene la información que le compete, pero que no necesariamente está dentro de su celda.

1.2.7 DESVENTAJAS DEL MODELO

- El modelo multiplicativo produce estimadores con las desviaciones grandes más frecuentes de los valores observados para aquellas exposiciones que se encuentran en territorios de alto riesgo, y al mismo tiempo en clases y récords también de alto riesgo. En otras palabras, la prima neta a cobrar en este tipo de categorías estará muy por arriba de la pérdida promedio que se espera tener para ellas.

Para ilustrar lo dicho, obsérvese la siguiente Tabla:

TABLA VIII

Tabla de Diferencias (Modelo Multiplicativo)

Clase de Conductor

Territorio	15	10&12	30&31	24&26	50	20&40	22&42	promedio ponderado
1	-5.03	-1.49	-0.75	9.00	0.02	3.43	11.77	0.00
2	-1.60	-0.75	-1.02	6.62	4.02	11.33	-5.76	0.00
3	4.97	-1.15	-3.47	0.53	-0.79	-7.70	10.92	0.00
4	-1.31	-2.11	4.43	1.31	31.67	-3.39	11.53	0.00
5	2.45	-1.25	-1.20	2.97	3.37	0.34	-5.58	0.00
6	2.35	-0.61	-4.37	3.70	-11.54	12.47	3.75	0.00
7	-1.27	-0.55	4.75	2.90	-7.70	-2.38	3.67	0.00
8	-3.72	-0.27	-1.65	-1.08	-9.06	1.43	13.31	0.00
9	-2.60	3.07	-3.81	-9.22	5.05	-1.15	-13.35	0.00
10	-0.81	2.66	-2.01	-2.14	-2.55	-7.38	-18.76	0.00
11	-9.56	1.67	15.33	-1.79	-16.70	0.53	-18.87	0.00
12	8.72	0.90	9.47	1.42	-13.33	-17.31	-31.38	0.00
13	6.76	4.73	7.35	-5.76	-6.58	-21.03	-45.35	0.00
14	12.27	2.60	70.23	-23.60	16.53	-35.37	-22.95	0.00
15	-6.49	4.97	15.15	-15.76	33.03	-31.47	-61.38	0.00
promedio ponderado	-0.16	-0.16	0.45	0.35	-0.52	-0.43	1.42	0.00

Esta tabla de diferencias se obtuvo de la resta de las tablas IVa y VII. Se resaltan aquellas desviaciones que en valor absoluto son mayores a 10. Puede también notarse la falta de balance de los estimadores para la clase de conductores.

Este defecto del presente modelo era de esperarse dada la forma de su fórmula matemática. Así, dicho defecto se traduce en un error sistemático, lo que es totalmente indeseable dado que el método de cálculo de la prima neta basado en la división en categorías debe funcionar bien para todas ellas.

- Sólo se tendrá un estimador de la relatividad balanceado, y aunque la desviación de los otros dos estimadores es en la práctica tolerable, teóricamente sería lo deseable el que todos los estimadores fueran balanceados a un mismo tiempo.

- No se considera en forma separada a las componentes de la prima neta.

1.3 MODELO 3: MODELO ADITIVO.

Para evitar un crecimiento de tipo exponencial en el cobro de una prima neta a las categorías definidas por renglones de alto riesgo, y así evitar el comentado error sistemático, la propuesta inmediata se encuentra en sumar, en lugar de multiplicar, los efectos de los rubros en la prima neta tratados en el modelo anterior. Así, en vez de relatividades se hablará de diferenciales.

1.3.1 MANEJO DE LA INFORMACION.

A partir de la experiencia en la siniestralidad pasada, agrupada en categorías o celdas en el mismo esquema de tres dimensiones, el monto promedio de reclamos para una categoría particular, será igual a la suma de los siguientes cuatro términos:

(1) Monto promedio de reclamos de la cartera total.

(2) Un diferencial del territorio.

(3) Un diferencial de la clase.

(4) Un diferencial del récord.

Así, se define la prima neta indicada de la siguiente forma:

$$P_{ijk} = \mu + t_i + c_j + r_k \qquad (1.3.1.1)$$

donde cada sumando es, respectivamente, cada uno de los cuatro puntos enunciados arriba.

De igual forma, la prima neta indicada para un individuo que se encuentre en la categoría definida por la combinación (i_1, j_1, k_1) será igual a una prima neta general, más el efecto o impacto que ocasiona el encontrarse en el i_1-ésimo territorio, más el efecto o impacto que ocasiona el estar en la j_1-ésima clase, más el efecto dado por pertenecer al k_1-ésimo récord de manejo.

Nuevamente el punto está en encontrar los estimadores para los diferenciales.

La condición que ahora deben cumplir los estimadores es, por ejemplo tomando al territorio, $t_i > 0$ si el territorio que define a la categoría en cuestión es de más alto riesgo con respecto al estándar, y $t_i < 0$ si el territorio que define a la categoría en cuestión es de más bajo riesgo con respecto al estándar. El mismo tratamiento se da para los estimadores de los diferenciales de la clase y del récord previo de manejo.

Finalmente, se tiene la prima neta para una categoría particular como la suma de un monto promedio de reclamos y cantidades positivas o negativas que respectivamente incrementan o disminuyen la cantidad a cobrar, según se trate de un rubro de alto o bajo riesgo con respecto al estándar.

1.3.2 ESTIMACION.

La estimación de los parámetros de este modelo se hará por medio del método de Mínimos Cuadrados Ordinario.

El modelo lineal aditivo se define con la siguiente expresión:

$$P_{ijkh} = \mu + t_i + c_j + r_k + \varepsilon_{ijkh} \qquad h=\overline{1,n_{ijk}}$$

donde:

P_{ijkh} = Prima neta indicada para el h-ésimo duo que se encuentra en la categoría definida por la combinación (t_i, c_j, r_k).

t_i, c_j, r_k definidos como antes.

ε_{ijkh} = Término perturbación.

Se tiene por supuesto que los términos perturbación son independientes entre sí, cada uno con un valor esperado de cero.

Ahora hay que estimar P_{ijkh}, es decir:

$$\hat{P}_{ijkh} = \hat{\mu} + \hat{t}_i + \hat{c}_j + \hat{r}_k$$

Los residuales están dados por:

$$e_{ijkh} = P_{ijkh} - \hat{P}_{ijkh}$$

Por el método de Mínimos Cuadrados Ordinario, el problema a resolver es el de minimizar la suma de los cuadrados de los residuales, para de esta forma tener una recta, plano o hiperplano (según sea la dimensión en la que se esté trabajando) más "ajustado" a los datos. Gráficamente, en el caso de dos dimensiones:

$$\min_{\mu, t_i, c_j, r_k} \{Q = \sum_{ijkh} e_{ijkh}^2\}$$

Condiciones de primer orden:

$$\frac{\delta Q}{\delta \hat{\mu}} = -2 \sum_{ijkh} (P_{ijkh} - \hat{\mu} - \hat{t}_i - \hat{c}_j - \hat{r}_k) = 0$$

$$\frac{\delta Q}{\delta \hat{t}_i} = -2 \sum_{jkh} (P_{ijkh} - \hat{\mu} - \hat{t}_i - \hat{c}_j - \hat{r}_k) = 0$$

$$\frac{\delta Q}{\delta \hat{c}_j} = -2 \sum_{ikh} (P_{ijkh} - \hat{\mu} - \hat{t}_i - \hat{c}_j - \hat{r}_k) = 0$$

$$\frac{\delta Q}{\delta \hat{r}_k} = -2 \sum_{ijh} (P_{ijkh} - \hat{\mu} - \hat{t}_i - \hat{c}_j - \hat{r}_k) = 0$$

Condiciones de segundo orden:

$$\frac{\delta^2 Q}{\delta \hat{\mu}^2} = 2n... > 0$$

$$\frac{\delta^2 Q}{\delta \hat{t}_i^2} = 2n... > 0$$

$$\frac{\delta^2 Q}{\delta \hat{c}_j^2} = 2n... > 0$$

$$\frac{\delta^2 Q}{\delta \hat{r}_k^{2}} = 2n_{...} > 0$$

De las condiciones de primer orden se obtienen las ecuaciones normales:

$$(1) \qquad n_{...}\hat{\mu} + \sum_i n_{i..}\hat{t}_i + \sum_j n_{.j.}\hat{c}_j + \sum_k n_{..k}\hat{r}_k = \sum_{ijkh} P_{ijkh}$$

$$(2) \qquad n_{i..}\hat{\mu} + \sum_i n_{i..}\hat{t}_i + \sum_j n_{ij.}\hat{c}_j + \sum_k n_{j.k}\hat{r}_k = \sum_{jkh} P_{ijkh}$$

$$(3) \qquad n_{.j.}\hat{\mu} + \sum_i n_{ij.}\hat{t}_i + \sum_j n_{.j.}\hat{c}_j + \sum_k n_{.jk}\hat{r}_k = \sum_{ikh} P_{ijkh}$$

$$(4) \qquad n_{..k}\hat{\mu} + \sum_i n_{i.k}\hat{t}_i + \sum_j n_{.jk}\hat{c}_j + \sum_k n_{..k}\hat{r}_k = \sum_{ijh} P_{ijkh}$$

Suponiendo estimadores balanceados:

$$\sum_{ijkh} P_{ijkh} = n_{...}p_{...}$$

$$\sum_{jkh} P_{ijkh} = n_{i..}p_{i..}$$

$$\sum_{ikh} P_{ijkh} = n_{.j.}p_{.j.}$$

$$\sum_{ijh} P_{ijkh} = n_{..k}p_{..k}$$

donde $p_{i..}$ es el monto promedio de reclamos del i-ésimo territorio, $p_{.j.}$ es el monto promedio de la j-ésima clase, y $p_{..k}$ es el monto promedio del k-ésimo récord.

Así, se tiene el sistema de ecuaciones en forma matricial:

$$
\begin{bmatrix}
n_{...} & n_{1..} & n_{2..} & \cdots & n_{I..} & n_{.1.} & n_{.2.} & \cdots & n_{.J.} & n_{..1} & n_{..2} & \cdots & n_{..K} \\
n_{1..} & n_{1..} & 0 & \cdots & 0 & n_{11.} & n_{12.} & \cdots & n_{1J.} & n_{1.1} & n_{1.2} & \cdots & n_{1.K} \\
n_{2..} & 0 & n_{2..} & \cdots & 0 & n_{21.} & n_{22.} & \cdots & n_{2J.} & n_{2.1} & n_{2.2} & \cdots & n_{2.K} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{I..} & 0 & 0 & \cdots & n_{I..} & n_{I1.} & n_{I2.} & \cdots & n_{IJ.} & n_{I.1} & n_{I.2} & \cdots & n_{I.K} \\
n_{.1.} & n_{11.} & n_{21.} & \cdots & n_{I1.} & n_{.1.} & 0 & \cdots & 0 & n_{.11} & n_{.12} & \cdots & n_{.1K} \\
n_{.2.} & n_{12.} & n_{22.} & \cdots & n_{I2.} & 0 & n_{.2.} & \cdots & 0 & n_{.21} & n_{.22} & \cdots & n_{.2K} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{.J.} & n_{1J.} & n_{2J.} & \cdots & n_{IJ.} & 0 & 0 & \cdots & n_{.J.} & n_{.J1} & n_{.J2} & \cdots & n_{.JK} \\
n_{..1} & n_{1.1} & n_{2.1} & \cdots & n_{I.1} & n_{.11} & n_{.21} & \cdots & n_{.J1} & n_{..1} & 0 & \cdots & 0 \\
n_{..2} & n_{1.2} & n_{2.2} & \cdots & n_{I.2} & n_{.12} & n_{.22} & \cdots & n_{.J2} & 0 & n_{..2} & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{..K} & n_{1.K} & n_{2.K} & \cdots & n_{I.K} & n_{.1K} & n_{.2K} & \cdots & n_{.JK} & 0 & 0 & \cdots & n_{..K}
\end{bmatrix}
\begin{bmatrix}
\hat{\mu} \\ \hat{t}_1 \\ \hat{t}_2 \\ \vdots \\ \hat{t}_I \\ \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_J \\ \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_K
\end{bmatrix}
=
\begin{bmatrix}
n_{...}p_{...} \\ n_{1..}p_{1..} \\ n_{2..}p_{2..} \\ \vdots \\ n_{I..}p_{I..} \\ n_{.1.}p_{.1.} \\ n_{.2.}p_{.2.} \\ \vdots \\ n_{.J.}p_{.J.} \\ n_{..1}p_{..1} \\ n_{..2}p_{..2} \\ \vdots \\ n_{..K}p_{..K}
\end{bmatrix}
$$

Nótese que el rango de esta matriz no será I+J+K+1, dado que, por ejemplo, el primer renglón puede ser obtenido como una combinación lineal de los otros renglones. Esto indica que el sistema tiene múltiples soluciones, dado que el primer elemento del vector de soluciones también se obtiene con dicha combinación lineal de elementos de su vector.

Analícense las ecuaciones normales. De (1) :

$$\hat{\mu} = p_{...} - \sum_i n_{i..}\hat{t}_i/n_{...} - \sum_j n_{.j.}\hat{c}_j/n_{...} - \sum_k n_{..k}\hat{r}_k/n_{...}$$

Por definición, μ representa al monto promedio de reclamos totales, y por lo tanto se sugiere como un estimador de μ a $p_{...}$ Esto se cumple si:

$$\sum_i n_{i..}\hat{t}_i/n_{...} = 0$$

$$\sum_{j} n_{.j.}\hat{c}_j/n_{...} = 0$$

$$\sum_{k} n_{..k}\hat{r}_k/n_{...} = 0$$

Estas tres condiciones pueden ayudar en la búsqueda de una solución única para el sistema de ecuaciones.

Una vez determinado el estimador para como $p_{...}$, el sistema de ecuaciones inicial se reduce a un sistema de I+J+K ecuaciones con I+J+K incógnitas, es decir:

$$\begin{bmatrix}
n_{1..} & 0 & \cdots & 0 & n_{11.} & n_{12.} & \cdots & n_{1J.} & n_{1.1} & n_{1.2} & \cdots & n_{1.K} \\
0 & n_{2..} & \cdots & 0 & n_{21.} & n_{22.} & \cdots & n_{2J.} & n_{2.1} & n_{2.2} & \cdots & n_{2.K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & n_{I..} & n_{I1.} & n_{I2.} & \cdots & n_{IJ.} & n_{I.1} & n_{I.2} & \cdots & n_{I.K} \\
n_{11.} & n_{21.} & \cdots & n_{I1.} & n_{.1.} & 0 & \cdots & 0 & n_{.11} & n_{.12} & \cdots & n_{.1K} \\
n_{12.} & n_{22.} & \cdots & n_{I2.} & 0 & n_{.2.} & \cdots & 0 & n_{.21} & n_{.22} & \cdots & n_{.2K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1J.} & n_{2J.} & \cdots & n_{IJ.} & 0 & 0 & \cdots & n_{.J.} & n_{.J1} & n_{.J2} & \cdots & n_{.JK} \\
n_{1.1} & n_{2.1} & \cdots & n_{I.1} & n_{.11} & n_{.21} & \cdots & n_{.J1} & n_{..1} & 0 & \cdots & 0 \\
n_{1.2} & n_{2.2} & \cdots & n_{I.2} & n_{.12} & n_{.22} & \cdots & n_{.J2} & 0 & n_{..2} & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1.K} & n_{2.K} & \cdots & n_{I.K} & n_{.1K} & n_{.2K} & \cdots & n_{.JK} & 0 & 0 & \cdots & n_{..K}
\end{bmatrix}
\begin{bmatrix}
\hat{t}_1 \\ \hat{t}_2 \\ \vdots \\ \hat{t}_I \\ \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_J \\ \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_K
\end{bmatrix}
=
\begin{bmatrix}
n_{1..}(p_{1..}-p_{...}) \\ n_{2..}(p_{2..}-p_{...}) \\ \vdots \\ n_{I..}(p_{I..}-p_{...}) \\ n_{.1.}(p_{.1.}-p_{...}) \\ n_{.2.}(p_{.2.}-p_{...}) \\ \vdots \\ n_{.J.}(p_{.J.}-p_{...}) \\ n_{..1}(p_{..1}-p_{...}) \\ n_{..2}(p_{..2}-p_{...}) \\ \vdots \\ n_{..K}(p_{..K}-p_{...})
\end{bmatrix}$$

Ahora obsérvese que el determinante de la matriz de coeficientes también será cero puesto que, por ejemplo, el primer renglón puede ser obtenido como una combinación lineal de los demás renglones, y el sistema tiene soluciones múltiples puesto que el primer elemento del vector solución puede ser obtenido con la misma combinación lineal de sus elementos.

Puesto que se está buscando una base para así tener una solución única, se puede cambiar este primer renglón por la primera condición que se impuso para obtener un estimador de sesgo nulo para μ:

$$
\begin{bmatrix}
n_{1..} & n_{2..} & \cdots & n_{I..} & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & n_{2..} & \cdots & 0 & n_{21.} & n_{22.} & \cdots & n_{2J.} & n_{2.1} & n_{2.2} & \cdots & n_{2.K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & n_{I..} & n_{I1.} & n_{I2.} & \cdots & n_{IJ.} & n_{I.1} & n_{I.2} & \cdots & n_{I.K} \\
n_{11.} & n_{21.} & \cdots & n_{I1.} & n_{.1.} & 0 & \cdots & 0 & n_{.11} & n_{.12} & \cdots & n_{.1K} \\
n_{12.} & n_{22.} & \cdots & n_{I2.} & 0 & n_{.2.} & \cdots & 0 & n_{.21} & n_{.22} & \cdots & n_{.2K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1J.} & n_{2J.} & \cdots & n_{IJ.} & 0 & 0 & \cdots & n_{.J.} & n_{.J1} & n_{.J2} & \cdots & n_{.JK} \\
n_{1.1} & n_{2.1} & \cdots & n_{I.1} & n_{.11} & n_{.21} & \cdots & n_{.J1} & n_{..1} & 0 & \cdots & 0 \\
n_{1.2} & n_{2.2} & \cdots & n_{I.2} & n_{.12} & n_{.22} & \cdots & n_{.J2} & 0 & n_{..2} & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1.K} & n_{2.K} & \cdots & n_{I.K} & n_{.1K} & n_{.2K} & \cdots & n_{.JK} & 0 & 0 & \cdots & n_{..K}
\end{bmatrix}
\begin{bmatrix}
\hat{t}_1 \\ \hat{t}_2 \\ \vdots \\ \hat{t}_I \\ \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_J \\ \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_K
\end{bmatrix}
=
\begin{bmatrix}
0 \\
n_{2..}(\bar{y}_{2..}-\bar{y}_{...}) \\
\vdots \\
n_{I..}(\bar{y}_{I..}-\bar{y}_{...}) \\
n_{.1.}(\bar{y}_{.1.}-\bar{y}_{...}) \\
n_{.2.}(\bar{y}_{.2.}-\bar{y}_{...}) \\
\vdots \\
n_{.J.}(\bar{y}_{.J.}-\bar{y}_{...}) \\
n_{..1}(\bar{y}_{..1}-\bar{y}_{...}) \\
n_{..2}(\bar{y}_{..2}-\bar{y}_{...}) \\
\vdots \\
n_{..K}(\bar{y}_{..K}-\bar{y}_{...})
\end{bmatrix}
$$

Análogamente, es fácil ver que el (I+1)-ésimo renglón y el (I+J+1)-ésimo renglón pueden ser obtenidos por una combinación de los renglones restantes, y que el sistema tiene soluciones múltiples, y así introducimos las dos restantes condiciones en estos renglones, para de esta manera ver si finalmente se ha obtenido una base:

$$
\begin{bmatrix}
n_{1..} & n_{2..} & \cdots & n_{I..} & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & n_{2..} & \cdots & 0 & n_{21.} & n_{22.} & \cdots & n_{2J.} & n_{2.1} & n_{2.2} & \cdots & n_{2.K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & n_{I..} & n_{I1.} & n_{I2.} & \cdots & n_{IJ.} & n_{I.1} & n_{I.2} & \cdots & n_{I.K} \\
0 & 0 & \cdots & 0 & n_{.1.} & n_{.2.} & \cdots & n_{.J.} & 0 & 0 & \cdots & 0 \\
n_{12.} & n_{22.} & \cdots & n_{I2.} & 0 & n_{.2.} & \cdots & 0 & n_{.21} & n_{.22} & \cdots & n_{.2K} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1J.} & n_{2J.} & \cdots & n_{IJ.} & 0 & 0 & \cdots & n_{.J.} & n_{.J1} & n_{.J2} & \cdots & n_{.JK} \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & n_{..1} & n_{..2} & \cdots & n_{..K} \\
n_{1.2} & n_{2.2} & \cdots & n_{I.2} & n_{.12} & n_{.22} & \cdots & n_{.J2} & 0 & n_{..2} & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
n_{1.K} & n_{2.K} & \cdots & n_{I.K} & n_{.1K} & n_{.2K} & \cdots & n_{.JK} & 0 & 0 & \cdots & n_{..K}
\end{bmatrix}
\begin{bmatrix}
\hat{t}_1 \\ \hat{t}_2 \\ \vdots \\ \hat{t}_I \\ \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_J \\ \hat{r}_1 \\ \hat{r}_2 \\ \vdots \\ \hat{r}_K
\end{bmatrix}
=
\begin{bmatrix}
0 \\
n_{2..}(\bar{y}_{2..}-\bar{y}_{...}) \\
\vdots \\
n_{I..}(\bar{y}_{I..}-\bar{y}_{...}) \\
0 \\
n_{.2.}(\bar{y}_{.2.}-\bar{y}_{...}) \\
\vdots \\
n_{.J.}(\bar{y}_{.J.}-\bar{y}_{...}) \\
0 \\
n_{..2}(\bar{y}_{..2}-\bar{y}_{...}) \\
\vdots \\
n_{..K}(\bar{y}_{..K}-\bar{y}_{...})
\end{bmatrix}
$$

$$(M) \qquad (\hat{\beta}) = (Y)$$

Ahora obsérvese que efectivamente ningún renglón puede obtenerse como una combinación lineal de los restantes, con lo que ya se tiene una base, y por lo tanto este sistema de ecuaciones tiene solución única.

Finalmente se obtienen los estimadores de los diferenciales de territorio, clase y récord como el producto de la inversa de la última matriz de coeficientes (N), y el último vector columna solución:

$$\hat{\beta} = N^{-1}Y$$

1.3.3 PROPIEDADES DE LOS ESTIMADORES.

- Estimadores de Sesgo Nulo.

De las ecuaciones normales se obtiene lo siguiente:

$$\hat{\mu} = \sum_{ijkh} P_{ijkh}/n_{...} - \sum_{i} n_{i..}\hat{t}_i/n_{...} - \sum_{j} n_{.j.}\hat{c}_j/n_{...} - \sum_{k} n_{..k}\hat{r}_k/n_{...}$$

$$\hat{t}_i = \sum_{jkh} P_{ijkh}/n_{i..} - \hat{\mu} - \sum_{j} n_{ij.}\hat{c}_j/n_{i..} - \sum_{k} n_{i.k}\hat{r}_k/n_{i..}$$

$$\hat{c}_j = \sum_{ikh} P_{ijkh}/n_{.j.} - \hat{\mu} - \sum_{i} n_{ij.}\hat{t}_i/n_{.j.} - \sum_{k} n_{.jk}\hat{r}_k/n_{.j.}$$

$$\hat{r}_k = \sum_{ijh} P_{ijkh}/n_{..k} - \hat{\mu} - \sum_{i} n_{i.k}\hat{t}_i/n_{..k} - \sum_{j} n_{.jk}\hat{c}_j /n_{..k}$$

Al considerar los supuestos de independencia entre los términos perturbación y valor esperado de cero de cada uno de ellos, los estimadores obtenidos por el método de mínimos cuadrados ordinario tienen sesgo nulo, siempre y cuando todos ellos tengan sesgo nulo a un mismo tiempo.

PROPOSICION:

$\hat{\mu}$ es un estimador de sesgo nulo de μ si $E\ \varepsilon_{ijkh} = 0$

para toda i,j,k,h, y $\hat{t}_i$, $\hat{c}_j$, $\hat{r}_k$ son estimadores de sesgo nulo de t_i, c_j, r_k respectivamente

DEM.:

$$E[\hat{\mu}] = E[\sum_{ijkh} P_{ijkh}/n_{\cdots} - \sum_i n_{i\cdot\cdot}\hat{t}_i/n_{\cdots} - \sum_j n_{\cdot j\cdot}\hat{c}_j/n_{\cdots} - \sum_k n_{\cdot\cdot k}\hat{r}_k/n_{\cdots}]$$

$$= \sum_{ijkh} E[\ \mu + t_i + c_j + r_k + \varepsilon_{ijkh}\ ]/n_{\cdots} - \sum_i n_{i\cdot\cdot} E[\hat{t}_i]/n_{\cdots}$$

$$- \sum_j n_{\cdot j\cdot} E[\hat{c}_j]/n_{\cdots} - \sum_k n_{\cdot\cdot k} E[\hat{r}_k]/n_{\cdots}$$

$$= \mu + \sum_i n_{i\cdot\cdot} t_i/n_{\cdots} + \sum_j n_{\cdot j\cdot} c_j/n_{\cdots} + \sum_k n_{\cdot\cdot k} r_k/n_{\cdots} + \sum_{ijkh} E[\varepsilon_{ijkh}]$$

$$- \sum_i n_{i\cdot\cdot} t_i/n_{\cdots} + \sum_j n_{\cdot j\cdot} c_j/n_{\cdots} + \sum_k n_{\cdot\cdot k} r_k/n_{\cdots}$$

$$= \mu$$

De esta proposición se desprende el siguiente corolario:

COROLARIO:

Cada estimador es de sesgo nulo si los restantes estimadores también son de sesgo nulo.

- Estimadores Balanceados.

Otra característica de estos estimadores es que son balanceados, es decir: $\hat{P}_{i..} = p_{i..}$, $\hat{P}_{.j.} = p_{.j.}$, $\hat{P}_{..k} = p_{..k}$:

$$\hat{P}_{i..} = \sum_{jk} \hat{P}_{ijk} n_{ijk}/n_{i..}$$

$$= \sum_{jk} (\hat{\mu} + \hat{t}_i + \hat{c}_j + \hat{r}_k) n_{ijk}/n_{i..}$$

$$= \hat{\mu} + \hat{t}_i + \sum_j \hat{c}_j n_{ij}/n_{i..} + \sum_k \hat{r}_k n_{i.k}/n_{i..}$$

$$= p_{...} + (p_{i..} - p_{...} - \sum_j \hat{c}_j n_{ij.}/n_{i..} - \sum_k \hat{r}_k n_{i.k}/n_{i..})$$

$$+ \sum_j \hat{c}_j n_{ij.}/n_{i..} - \sum_k \hat{r}_k n_{i.k}$$

$$= p_{i..}$$

De manera análoga es inmediato ver que los demás estimadores cumplen con esta propiedad.

Otra característica interesante de estos estimadores se presenta al considerar, que las distribuciones de exposiciones para territorios, clases y récordes son independientes:

Territorio: $$\frac{n_{i..}}{n_{...}} = \frac{n_{ij.}}{n_{.j.}} = \frac{n_{i.k}}{n_{..k}}$$

Clase: $$\frac{n_{.j.}}{n_{...}} = \frac{n_{ij.}}{n_{i..}} = \frac{n_{.jk}}{n_{..k}}$$

Récord: $$\frac{n_{..k}}{n_{...}} = \frac{n_{i.k}}{n_{i..}} = \frac{n_{.jk}}{n_{.j.}}$$

Si ésto se cumple, entonces dado que se consideró que:

$$\sum_i n_{i..}\hat{t}_i/n_{...} = 0$$

$$\sum_j n_{.j.}\hat{c}_j/n_{...} = 0$$

$$\sum_k n_{..k}\hat{r}_k/n_{...} = 0$$

los estimadores quedarán de la siguiente manera:

$$\hat{t}_i = p_{i..} - p_{...}$$

$$\hat{c}_j = p_{.j.} - p_{...}$$

$$\hat{r}_k = p_{..k} - p_{...}$$

es decir, los estimadores no dependen uno de otro.

1.3.4 PRONOSTICO.

Con el fin de predecir la siniestralidad, los estimadores aquí obtenidos tendrán igualmente un ajuste por inflación y por otras variables que influyan en la variación de las sumas aseguradas.

1.3.5 EJEMPLO ILUSTRATIVO.

Retomando los datos utilizados en el Modelo Multiplicativo, ahora tenemos que:

$$\begin{bmatrix}\text{prima neta para}\\ \text{una categoría}\\ \text{particular}\end{bmatrix} = \begin{bmatrix}\text{pérdida}\\ \text{promedio}\\ \text{general}\end{bmatrix} + \begin{bmatrix}\text{diferen-}\\ \text{cial ter-}\\ \text{ritorial}\end{bmatrix} + \begin{bmatrix}\text{diferencial}\\ \text{de}\\ \text{clase}\end{bmatrix}$$

La siguiente Tabla ilustra el valor de los estimadores para este conjunto de datos:

TABLA IX

Estimadores del Modelo Aditivo (1976)

Territorio	15	10&12	30&31	24&26	50	20&40	22&42	promedio ponderado
1	24.50	28.98	38.60	50.29	64.86	78.50	130.27	$38.33
2	19.43	23.85	33.55	45.36	59.72	73.37	125.14	$33.30
3	24.80	29.20	38.96	50.51	65.06	78.72	130.49	$40.23
4	23.30	28.30	38.00	49.69	64.26	77.90	129.67	$38.66
5	25.93	30.35	40.04	51.66	66.22	79.86	131.63	$41.06
6	30.47	34.87	44.57	56.18	70.75	84.39	136.15	$45.28
7	32.96	37.36	47.06	58.67	73.26	86.88	138.65	$48.26
8	34.02	38.42	48.11	59.73	74.29	87.93	139.70	$49.41
9	38.00	42.40	52.10	63.71	78.28	91.92	143.69	$52.79
10	42.70	47.10	56.80	68.41	82.38	96.52	148.39	$57.07
11	39.00	43.40	53.10	64.71	79.28	92.92	144.69	$54.25
12	35.05	39.45	49.15	60.76	75.33	88.97	140.74	$50.25
13	49.85	54.24	63.95	75.55	90.12	103.76	155.53	$64.91
14	48.59	52.99	62.69	74.30	88.87	102.51	154.28	$62.39
15	62.50	66.98	76.66	88.29	102.86	116.50	168.27	$75.03
promedio ponderado	$32.21	$36.58	$45.63	$57.42	$72.60	$86.14	$137.20	$46.97

Obsérvese cómo los estimadores obtenidos por este método son balanceados tanto para el Territorio como para la Clase de Conductor (compárese con la Tabla IVa).

1.3.6 VENTAJAS DEL MODELO.

- El modelo aditivo, como puede verse en la Tabla X, no presenta un patrón de diferencias altas definido para un conjunto de categorías en especial, como lo presentó el modelo multiplicativo. Es decir, no se está sobre-estimando o bajo-estimando en forma sistemática a algún conjunto de celdas con ciertas características semejantes en cuanto a nivel de riesgo.

TABLA X

Tabla de Diferencias (Modelo Aditivo)

Territorio	15	10&12	30&31	24&26	50	20&49	22&42	promedio ponderado
1	-2.67	0.13	-1.27	7.62	-4.12	-3.83	-7.75	0.00
2	7.36	2.08	-3.18	3.33	-1.97	0.37	-23.06	0.00
3	7.45	0.63	-3.73	-1.55	-12.39	-5.45	-3.36	0.00
4	1.38	-0.21	4.12	-0.05	21.50	-9.99	-4.11	0.00
5	4.56	0.24	-1.53	1.28	0.13	-3.96	-6.76	0.00
6	3.00	-0.34	-5.46	2.99	-12.16	11.23	-1.07	0.00
7	-1.41	-0.74	5.10	2.02	-7.35	-2.10	4.02	0.00
8	-3.28	-0.60	-2.65	-1.01	-7.76	3.18	16.36	0.00
9	-4.55	1.88	-4.51	-0.25	0.55	3.97	-2.60	0.00
10	-3.93	4.49	-3.54	0.91	3.66	1.85	1.05	0.00
11	-11.50	0.15	14.47	-0.76	-12.95	16.23	1.77	0.00
12	7.89	-0.35	0.38	7.86	-11.57	-15.50	-6.13	0.00
13	1.27	0.65	6.93	-2.26	71.40	-8.73	-12.29	0.00
14	7.33	-3.82	65.37	-20.05	26.10	-35.76	0.99	0.00
15	-14.85	-0.43	15.28	-6.80	52.37	-17.02	0.70	0.00
promedio ponderado	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Análogamente esta Tabla se obtuvo de la resta de las tablas IVa y IX. Otra vez se resaltan aquellas diferencias mayores, en valor absoluto, a 10. No se encuentra un patrón de diferencias altas para algún conjunto de categorías con similitud en el grado de su riesgo, en comparación al Modelo multiplicativo.

La forma aditiva, así, está evitando que se penalice más de lo necesario a las personas cuya categoría está conformada por rubros de alto riesgo dada la forma multiplicativa del modelo anterior.

- Este modelo también usa toda la información disponible competente para el cálculo de la prima neta de una determinada categoría, con lo que no se está aislando a cada celda del resto de la información.

- Los estimadores arrojados por el modelo aditivo son balanceados para los tres puntos que definen a una categoría. La pasada tabla lo muestra para el caso que se ha venido usando como ejemplo de dos dimensiones. Compárese con la tabla VIII.

- Los estimadores fueron obtenidos ya no de una forma eurística, sino por medio de un método (método de mínimos cuadrados ordinario) que arroja estimadores con ciertas características deseables, bajo ciertos supuestos establecidos previamente.

1.3.7 DESVENTAJA DEL MODELO.

- No se considera por separado a las componentes de la prima neta.

1.4 MODELO 4: MODELO DE REGRESION DE UNA SOLA ECUACION.

Una técnica más sofisticada desde el punto de vista estadístico para el cálculo de la pérdida potencial nos la proporciona el Análisis de Regresión.

1.4.1 MANEJO DE LA INFORMACION.

Se asume que la pérdida potencial de el h-ésimo individuo (y_h) en un período dado de tiempo está caracterizada por la siguiente expresión:

$$y_h = \beta_1 + \beta_2 x_{2h} + \ldots + \beta_m x_{mh} + \epsilon_h \qquad h=\overline{1,n} \qquad (1.4.1.1)$$

En forma matricial:

$$y = X\beta + \epsilon \qquad (1.4.1.2)$$

donde y es un vector de nx1 variables dependientes, X una matriz de nxm variables explicativas, β un vector de mx1 constantes, y ε un vector de nx1 términos perturbación.

EL TERMINO PERTURBACION.

La justificación para añadir un término perturbación en esta ecuación descansa fundamentalmente en tres razones:

- Pueden existir factores que sean omitidos dentro del conjunto de variables explicativas que tengan efecto sobre la variable dependiente; sin embargo, muchos de estos factores solo tienen efectos muy ligeros sobre la variable dependiente, y no por ésto no se desea incluirlos explicitamente en la ecuación. Por eso se considera el efecto neto conjunto, el cual se resume en ϵ_n.

- La gente no es totalmente predecible. La variable dependiente puede interpretarse como la combinación de una componente deterministica con una componente no sistemática o no determinística, que representa el elemento impredecible. Sirve para integrar a la parte sistemática con el mundo real.

- Puede haber errores de medida por parte del investigador, tanto por su lado como por los instrumentos de que se auxilia.

Bajo los supuestos del modelo de regresión lineal múltiple clásico:

i) $E[\epsilon] = 0$

Esto significa que, en promedio, se espera que el error sea igual a cero.

ii) $E[\epsilon\epsilon'] = \sigma^2 I_n$

La varianza del término perturbación es constante (varianza homocedástica).

iii) X es no estocástica.

iv) $rho(X) = m < n$

Esto es: el orden de la matriz más grande cuyo determinante es distinto de cero, es igual a m.

Se asumen los anteriores supuestos a razón de que éstos facilitan el manejo del modelo, y arrojan estimadores con ciertas características deseables, en caso de ser éstos ciertos.

Adicionalmente, se manejará que el término perturbación sigue una ley de distribución normal.

LA TRANSFORMACION LOGARITMICA.

Teniendo el conocimiento de la siniestralidad, bien puede ser que ésta no se comporte de una manera lineal (como lo propone la ecuación (1.4.1.1)), sino más bien exponencial. Así, se sugiere la transformación logarítmica de la pérdida potencial:

$$y_h = \beta_1 \, \beta_2^{x_{2h}} \, \beta_3^{x_{3h}} \ldots \beta_m^{x_{mh}} \exp(\epsilon_h) \qquad (1.4.1.3)$$

Tomando logaritmos de ambos lados de la igualdad:

$$\log y_h = \log\beta_1 + x_{2h}\log\beta_2 + x_{3h}\log\beta_3 + \ldots + x_{mh}\log\beta_m + \epsilon_h$$

Redefiniendo:

$$z_h = \log y_h$$

$$\phi_1 = \log \beta_1$$

$$\vdots$$

$$\phi_m = \log \beta_m$$

Así:

$$z_h = \phi_1 + \phi_2 x_{2h} + \phi_3 x_{3h} + \ldots + \phi_m x_{mh} + \epsilon_h \qquad (1.4.1.4)$$

De esta forma se observa que el método de estimación que se use para la ecuación (1.4.1.1) puede ser el mismo que para (1.4.1.4).

Para efectos prácticos, pueden proponerse formas funcionales lineales o no lineales, y posteriormente tomar aquella que ofrezca un mejor ajuste y/o un mejor nivel de predicción.

SELECCION DE VARIABLES.

Dado el tipo de información con el que se trabaja, las variables serán de tipo cualitativo y cuantitativo.

Las de tipo cualitativo son aquellas que, como su nombre lo dice, toman un valor según las cualidades o características de la persona: edad, sexo, etc.

Se manejará una variable cuantitativa (variable de tendencia), la cual se describe adelante.

La selección de variables siempre estará sujeta al tipo de información de que se disponga, y por consiguiente no se tiene una total libertad de selección.

El presente punto puede estudiarse tan ampliamente como se desee, en el sentido de proponer diversos conjuntos de variables explicativas, y finalmente tomar aquél que se ajuste más a nuestras necesidades.

El conjunto final de variables explicativas se escogerá de acuerdo a los siguientes criterios:

1. Coeficientes estimados estadísticamente significativos.

2. Coeficientes estimados consistentes con los supuestos "a priori" (signos correctos).

3. Nivel de predicción.

El punto número uno habla de que exista regresión a un nivel porcentual dado, es decir, que efectivamente una variable escogida tenga un impacto real en la función (en este caso, en el reclamo por siniestro) con un cierto nivel de confianza o cierto nivel de significación.

El segundo punto se refiere a que no haya contradicción con el valor de un coeficiente de regresión con la estadística de que se dispone, y a la forma de planteamiento del problema. Si por ejemplo, se tiene la siguiente función de regresión de reclamo monetario:

$$y = \alpha + \beta x + \epsilon$$

con:

$$x = \begin{cases} 1 & \text{si es hombre} \\ 0 & \text{si es mujer} \end{cases}$$

Si estadísticamente es claro que los hombres tienen reclamos más costosos que las mujeres, entonces nuestros supuestos "a priori" serán:

$$\hat{\beta} > 0$$

$$\hat{\alpha} > 0 \qquad \text{(no hay reclamos negativos)}$$

El tercer punto nos habla de que nuestras predicciones sean lo más atinadas posibles, para así cobrar primas más justas y permitir un buen margen de ganancia.

Para el presente trabajo se proponen 2 conjuntos de variables explicativas, basados en las características que han dado forma a los anteriores modelos.

X_1 = edad, sexo, estado civil, placer, negocios, distancia, manejo ocasional, urbano, récord previo de manejo, variable de tendencia.

X_2 = clases 1-8, urbano, récord previo de manejo, variable de tendencia.

De esta forma, se tiene a y_h como:

X_1 :

$$y_h = \beta_0 + \beta_1 x_{1h} + \beta_2 x_{2h} + \ldots + \beta_{10} x_{10h} + \epsilon_h$$

EDAD: $$x_{1h} = \begin{cases} 1 & \text{si h tiene 25 o más años} \\ 0 & \text{de otra forma} \end{cases}$$

SEXO: $$x_{2h} = \begin{cases} 1 & \text{si h es hombre} \\ 0 & \text{de otra forma} \end{cases}$$

ESTADO CIVIL: $$x_{3h} = \begin{cases} 1 & \text{si h es casado (a)} \\ 0 & \text{de otra furma} \end{cases}$$

PLACER: $$x_{4h} = \begin{cases} 1 & \text{si h maneja por placer y una distancia anual menor a 16 km.} \\ 0 & \text{de otra forma} \end{cases}$$

NEGOCIOS: $$x_{5h} = \begin{cases} 1 & \text{si el uso del auto es para negocios} \\ 0 & \text{de otra forma} \end{cases}$$

DISTANCIA: $$x_{6h} = \begin{cases} 1 & \text{si la distancia al trabajo es de 16 km. o más} \\ 0 & \text{de otra forma} \end{cases}$$

MANEJO OCASIONAL: $$x_{7h} = \begin{cases} 1 & \text{si el manejo es ocasional} \\ 0 & \text{de otra forma} \end{cases}$$

URBANO: $x_{8h} = \begin{cases} 1 & \text{si h maneja en un ambiente considerado urbano} \\ 0 & \text{de otra forma} \end{cases}$

RECORD PREVIO DE MANEJO: $x_{9h} = \begin{cases} 1 & \text{si h tiene 5 o más años sin accidente} \\ 0 & \text{de otra forma} \end{cases}$

x_{10h} variable de tendencia

ε_h término perturbación

X_2:

$$y_h = \beta_0 + \beta_1 x_{1h} + \beta_2 x_{2h} + \ldots + \beta_{11} x_{11h} + \varepsilon_h$$

CLASE: $x_{jh} = \begin{cases} 1 & \text{si h pertenece a la clase j } (j=1,\ldots,8) \\ 0 & \text{de otra forma} \end{cases}$

URBANO: $x_{9h} = \begin{cases} 1 & \text{si h maneja en un ambiente considerado urbano} \\ 0 & \text{de otra forma} \end{cases}$

RECORD PREVIO DE MANEJO: $x_{10h} = \begin{cases} 1 & \text{si h tiene 5 o más años sin accidente} \\ 0 & \text{de otra forma} \end{cases}$

x_{11h} variable de tendencia

ε_h término perturbación

Las variables que sólo pueden tomar uno de dos posibles valores son conocidas como "variables dicótomas" o "variables dummy". En este caso son las variables cualitativas.

La variable de tendencia puede tomar valores enteros positivos: 0,1,2,3,... En la mayoría de los análisis de regresión múltiple que incluye, como aquí, datos de series de tiempo se acostumbra introducir esta variable por las siguientes razones:

- La variable de tendencia puede ser un "reemplazo" de alguna variable básica que afecta a la variable dependiente. Sin embargo, esta variable puede ser no observable o no tener datos sobre ella, y así la suponemos como función del tiempo (e.g., la tecnología) por conveniencia.

- A veces puede creerse que una variable básica que afecta a la variable dependiente está estrechamente relacionada con el tiempo, y es preferible, al menos en términos de costos, introducir la variable de tendencia que la básica.

Esta variable de tendencia también es conocida como "variable de tiempo". Si pasa un año contado a partir de este momento, esta variable toma el valor de 1; si pasan 2, 2; etc. Se toma anual por así manejar nuestra información.

1.4.2 ESTIMACION.

Para la estimación de los coeficientes de regresión del presente modelo se usará el método de máxima verosimilitud (ML).

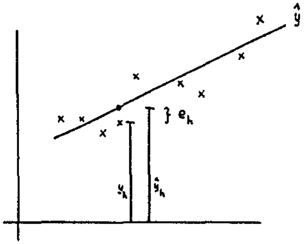
Acorde a la ecuación (1.4.1.2), se cuenta con el modelo:

$$y = X\beta + \epsilon$$

De aquí se busca llegar a:

$$\hat{y} = X\hat{\beta}$$

A manera de ilustración, gráficamente y para un modelo en R^2 (i.e., $y = \alpha + \beta x + \epsilon$), se tiene lo siguiente:



Los puntos x son los valores observados; los puntos sobre la recta $\hat{y}$ son los valores ajustados; e_h es el residual o estimador de ϵ_h. Se usa e_h cuando se trata de una muestra; ϵ_h cuando se trata de la población.

Como puede verse en la gráfica:

$$e_h = y_h - \hat{y}_h = y_h - (\hat{\alpha} + \hat{\beta}x_h)$$

ML consiste en maximizar la Función de Verosimilitud (L), ésto es, la función de densidad de probabilidad conjunta de la muestra aleatoria.

Bajo los supuestos dados:

$$\varepsilon \sim NI(\,0,\sigma^2_\varepsilon I_n\,)$$

$$\Rightarrow f(\varepsilon) = (2\pi)^{-n/2}|\sigma^2_\varepsilon I_n|^{-1/2}\exp\{-1/2(\varepsilon-E\varepsilon)'(\sigma^2_\varepsilon I_n)^{-1}(\varepsilon-E\varepsilon)\}$$

$$\Rightarrow f(\varepsilon) = (2\pi)^{-n/2}\sigma^{-n}_\varepsilon \exp\{-1/(2\sigma^2_\varepsilon)\ \varepsilon'\varepsilon\ \}$$

De aquí, L para la muestra es:

$$f(e) = (2\pi)^{-n/2}\sigma^{-n}_\varepsilon \exp\{-1/(2\sigma^2_\varepsilon)\ e'e\}$$

Es evidente que L es una función que depende de dos parámetros: β y σ^2_ε:

$$L(\hat{\beta},\hat{\sigma}^2_\varepsilon) = f(e) = (2\pi)^{-n/2}\sigma^{-n}_\varepsilon\exp\{-1/\sigma^2_\varepsilon(y-X\hat{\beta})'(y-X\hat{\beta})\}$$

El problema de optimización matemática es:

$$\max_{\hat{\beta},\hat{\sigma}^2_\varepsilon} L(\hat{\beta},\hat{\sigma}^2_\varepsilon)$$

Puesto que el obtener el máximo de una función es equivalente a obtener el máximo del logaritmo de la misma, el problema se puede expresar como:

$$\max_{\hat{\beta},\hat{\sigma}^2_\varepsilon} \operatorname{Log} L(\hat{\beta},\hat{\sigma}^2_\varepsilon)$$

Así:

$$\operatorname{Log} L(\hat{\beta},\hat{\sigma}^2_\varepsilon) = -(n/2)\operatorname{Log}(2\pi)-n\operatorname{Log}\hat{\sigma}_\varepsilon-(1/\hat{\sigma}^2_\varepsilon)(y - X\hat{\beta})'(y - X\hat{\beta})$$

$$= -(n/2)\operatorname{Log}(2\pi)-n\operatorname{Log}\hat{\sigma}_\varepsilon-(1/\hat{\sigma}^2_\varepsilon)[y'y-y'X\hat{\beta}-\hat{\beta}'X'y+\hat{\beta}'X'X\hat{\beta}]$$

Para encontrar el máximo:

$$\frac{\delta[\log L(\hat{\beta},\hat{\sigma}^2_e)]}{\delta\hat{\beta}} = 0$$

$$\Rightarrow \qquad \frac{\delta L}{\delta \beta} = -X'y + X'X\hat{\beta} = 0$$

$$\hat{\beta}^{ML} = (X'X)^{-1}X'y$$

donde es posible obtener la inversa de X'X, dado que se trabaja bajo el supuesto de que rho(X) = m.

Similarmente, para obtener un estimador para la varianza:

$$\frac{\delta L}{\delta\hat{\sigma}^2_e} = -\frac{n}{\hat{\sigma}^2_e} + \frac{(y-X\hat{\beta})'(y-X\hat{\beta})}{2(\hat{\sigma}^2_e)^2} = 0$$

$$\therefore \qquad \hat{\sigma}^{2\,ML}_e = S^2_e = \frac{(y-X\hat{\beta})'(y-X\hat{\beta})}{n}$$

1.4.3 PROPIEDADES DE LOS ESTIMADORES.

A continuación se procede a enunciar y demostrar las propiedades de los estimadores ML de β y σ^2_e:

1) PROPIEDADES DE ESTIMACION DE $\hat{\beta}^{ML}$:

TEOREMA: (Teorema de Gauss-Markov)

Bajo los supuestos del modelo de regresión lineal múltiple clásico, $\hat{\beta}^{ML}$ es lineal e insesgado, y es el mejor de todos los estimadores lineales insesgados, i.e., $\hat{\beta}^{ML}$ tiene varianza mínima, es el más eficiente dentro de la clase de los estimadores lineales insesgados de β. Así, se dice que $\hat{\beta}^{ML}$ es BLUE (Best Linear Unbiased Estimator).

DEM.:

- $\hat{\beta}^{ML}$ ES LINEAL.

$$\hat{\beta}^{ML} = (X'X)^{-1}X'y$$

$$= Wy$$

donde $W = (X'X)^{-1}X'$

- $\hat{\beta}^{ML}$ ES INSESGADO.

$$\hat{\beta}^{ML} = (X'X)^{-1}\ X'y$$

$$= (X'X)^{-1}\ X'(\beta X + \epsilon)$$

$$= (X'X)^{-1}\ X'X\beta + (X'X)^{-1}X'\epsilon$$

$$= \beta + (X'X)^{-1}X'\epsilon$$

$$\Rightarrow \qquad E[\hat{\beta}^{ML}] = E[\beta + (X'X)^{-1}X'\epsilon]$$

$$= \beta + (X'X)^{-1}X'E\epsilon$$

$$\therefore \qquad E\,\hat{\beta}^{ML} = \beta$$

i.e., $\hat{\beta}^{ML}$ es un estimador insesgado de β.

- $\hat{\beta}^{ML}$ TIENE VARIANZA MINIMA.

$$\text{Var-Cov } \hat{\beta}^{ML} = E\{(\hat{\beta}^{ML} - E\hat{\beta}^{ML})(\hat{\beta}^{ML} - E\hat{\beta}^{ML})'\}$$

$$= E(\hat{\beta}^{ML} - \beta)(\hat{\beta}^{ML} - \beta)'$$

Usando la regla para matrices $(ABC)' = C'B'A'$ y notando que $(X'X)^{-1}$ es una matriz simétrica, $(X'X)$ es una matriz simétrica.

$$\Rightarrow \qquad \text{Var-Cov } \hat{\beta}^{ML} = E[\beta + (X'X)^{-1}X\text{-}\ \beta][\beta + (X'X)^{-1}X'\epsilon - \beta]'$$

$$= E[(X'X)^{-1}X'\epsilon][\epsilon'X(X'X)^{-1}]$$

$$= (X'X)^{-1}X'\sigma^2_{\epsilon}IX)X'X)$$

$$= \sigma^2_{\epsilon}\ (X'X)^{-1}X'X(X'X)^{-1}$$

$$= \sigma^2_{\epsilon}\ (X'X)^{-1}$$

Sea $\tilde{\beta} = [(X'X)^{-1}X' \ + P]$ y , donde P es una matriz de perturbaciones no estocástica. Este estimador de β define:

- Un conjunto de todos los estimadores de β que están determinados con una P dada.

- Un conjunto de todos los estimadores de β que son lineales en y.

$$\bar{\beta} = \beta + (X'X)^{-1}X'\varepsilon + PX\beta + P\varepsilon$$

$$E[\bar{\beta}] = \beta + PX\beta$$

$\therefore$ $\bar{\beta}$ es insesgado si $PX = 0$

Como P puede ser cualquier matriz de perturbaciones sujeta sólo a estas condiciones, la clase de estimadores definida por $\bar{\beta}$ contiene a todos los estimadores lineales insesgados de β.

- Por demostrar que la matriz de varianzas y covarianzas de β^{ML} es la menor dentro de esta clase de estimadores:

$$\text{Var-Cov } \bar{\beta} = E(\bar{\beta}-\beta)(\bar{\beta}-\beta)'$$

$$= E[(X'X)^{-1} X'\varepsilon+P\varepsilon][\varepsilon'X(X'X)^{-1}+\varepsilon'P']$$

$$= \sigma^2_{\varepsilon} [(X'X)^{-1} + PP']$$

$$\Rightarrow \quad \text{Var-Cov } \hat{\beta}^{ML} \leq \text{Var-Cov } \bar{\beta}$$

$\therefore$ $\hat{\beta}^{ML}$ es el estimador más eficiente en el conjunto de estimadores lineales insesgados, i.e., el mejor de todos los estimadores lineales insesgados.

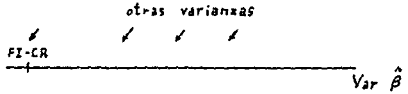
$\therefore$ $\hat{\beta}^{ML}$ es BLUE.

TEOREMA:

Bajo los supuestos del modelo de regresión lineal múltiple clásico normal, $\hat{\beta}^{ML}$ es Eficiente.

DEM.:

Por demostrar que Var-Cov $\hat{\beta}^{ML}$ alcanza la FI-CR (Frontera Inferior de Cramer y Rao):



$$\text{FI-CR} = -E\left[\frac{\delta^2 \text{LogL}}{\delta\hat{\beta}\hat{\beta}'}\right]^{-1}$$

$$\text{Log L} = -\frac{n}{2}\log 2\pi - n\log\sigma_u - \frac{1}{2\sigma^2_u}(y-X\hat{\beta})'(y-X\hat{\beta})$$

$$= -\frac{1}{2\sigma^2_u}(-2X'y + 2X'X\hat{\beta})$$

$$= \frac{1}{\sigma^2_u}(X'y - X'X\hat{\beta})$$

$$\frac{\delta^2 \text{LogL}}{\delta\hat{\beta}\hat{\beta}'} = \frac{X'X}{\sigma^2_u}$$

$\therefore$ FI-CR para estimadores regulares insesgados de β:

$$\text{FI-CR} = -E\left[\frac{-X'X}{\sigma^2_\varepsilon}\right]^{-1}$$

$$= \sigma^2_\varepsilon (X'X)^{-1}$$

$$= \text{Var-Cov } \hat{\beta}^{ML}$$

$\therefore$ $\hat{\beta}^{ML}$ es eficiente.

2) PROPIEDADES DE ESTIMACION DE S^2_ε :

Analicemos si este estimador para σ^2_ε es insesgado:

$$E[S^2_\varepsilon] = \frac{E e'e}{n} \stackrel{?}{=} \sigma^2_\varepsilon$$

Resultado Preliminar:

$$y = X\beta + \varepsilon$$

$$\hat{\beta} = (X'X)^{-1}X'y$$

$$\Rightarrow \quad y - X\hat{\beta} = y - X(X'X)^{-1} X'y$$

$$= X\beta + \varepsilon - X(X'X)^{-1}X'(X\beta + \varepsilon)$$

$$= X\beta + \varepsilon - X(X'X)\ X'X\beta - X(X'X)^{-1}X'\varepsilon$$

$$= X\beta + \varepsilon - X\beta - X(X'X)^{-1}X'\varepsilon$$

$$= \varepsilon - X(X'X)^{-1}X'\varepsilon$$

De aquí:

$$e'e = (y-X\hat{\beta})'(y-X\hat{\beta})$$

$$= [\,\epsilon - X(X'X)^{-1}X'\epsilon]'[\epsilon - X(X'X)^{-1}X'\epsilon]$$

$$= [\epsilon' - \epsilon'X(X'X)^{-1}X'][\epsilon - X(X'X)^{-1}X'\epsilon]$$

$$= \epsilon'\epsilon - \epsilon'X(X'X)^{-1}X'\epsilon$$

$$\Rightarrow \quad E[e'e] = E\epsilon'\epsilon - E[\epsilon'X(X'X)^{-1}X']$$

Sea $A = X(X'X)^{-1}X'$

$$\Rightarrow \quad E[e'e] = E[\epsilon'\epsilon] - E[\epsilon'A\epsilon]$$

$$= E\sum_{h=1}^{n}\epsilon^2_h - E\sum_{h=1}^{n}\sum_{h_1=1}^{n}\epsilon_h\,\epsilon_{h_1}a_{hh_1}$$

Se observa en el segundo término que todos los términos tales que h distinto de h_1 serán igual a 0 dados los supuestos del modelo.

$$\Rightarrow \quad E[e'e] = \sum_{h=1}^{n}E[\epsilon_h^2] - \sum_{h=1}^{n}E[\epsilon^2_h]\,a_{hh}$$

$$= n\sigma_\epsilon^2 - \sigma_\epsilon^2\sum_{h=1}^{n}a_{hh}$$

$$= n\sigma_\epsilon^2 - (\sigma_\epsilon^2)(\mathrm{tr}\ A)$$

Puesto que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$:

$$\mathrm{tr}(A) = \mathrm{tr}[X(X'X)^{-1}X']$$

$$= \mathrm{tr}[X'X(X'X)^{-1}]$$

$$= tr[I_m] = m$$

$$\therefore \quad E[e'e] = \sigma^2_e \, (n-m)$$

$$\therefore \quad E[S^2_e] = \frac{E(e'e)}{n} = \frac{\sigma^2_e(n-m)}{n} = \sigma^2_e - \frac{m}{n}\,\sigma^2_e$$

$\therefore$ S^2_e es sesgado.

Así, se propone un segundo estimador del que es inmediato probar que es de sesgo nulo:

$$s^2_e = \hat{\sigma}^2_e = \frac{e'e}{n-m}$$

Por lo que ahora se usará s^2_e en lugar de S^2_e como estimador para σ^2_e.

1.4.4 PRONOSTICO.

La inflación y otras variables (como la tecnología) provocan un aumento en el monto de los siniestros, y por ende en la Prima Neta a cobrar.

Para tomar en cuenta este efecto que está relacionado con el tiempo, se implementó la Variable de Tendencia.

"A priori", esta Variable de Tendencia debe tener un valor positivo, para así ir incrementando (ajustando) el valor de la Prima Neta conforme transcurre el tiempo.

1.4.5 EJEMPLO ILUSTRATIVO.

Supóngase que el individuo h_i cuenta con las siguientes características:

(1)	Edad:	29 años.
(2)	Sexo:	Mujer.
(3)	Estado Civil:	Soltero.
(4)	Manejo por Placer:	No.
(5)	Negocios:	Sí.
(6)	Distancia al Trabajo:	13 km.
(7)	Manejo Ocasional:	No.
(8)	Urbano:	Sí.
(9)	Récord Previo:	6 años sin accidente.
(10)	Año de Valuación:	Tercero.

- MODELO LINEAL.

a) Conjunto de variables X_i.

Para el presente caso, se tiene que la pérdida potencial para el h_i-ésimo individuo está dada por:

$$y_{hi} = \beta_o + \sum_{t=1}^{10} \beta_t x_{thi} + \varepsilon_{hi}$$

Una vez obtenidos los estimadores para los parámetros por el método de máxima verosimilitud:

$$\hat{y}_{hi} = \hat{\beta}_o + \sum_{t=1}^{10} \hat{\beta}_t x_{thi}$$

Así, la pérdida potencial que este individuo tendrá acorde al modelo estará dada por:

$$\hat{y}_{hi} = \hat{\beta}_0+\hat{\beta}_1(1)+\hat{\beta}_2(0)+\hat{\beta}_3(0)+\hat{\beta}_4(0)+\hat{\beta}_5(1)+\hat{\beta}_6(0)+\hat{\beta}_7(0)+$$
$$\hat{\beta}_8(1)+\hat{\beta}_9(1)+\hat{\beta}_{10}(3)$$

$$\hat{y}_{hi} = \hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_5 + \hat{\beta}_8 + \hat{\beta}_9 + 3\hat{\beta}_{10}$$

Lo cual constituye la prima neta a cobrar al presente individuo.

b) Conjunto de variables X_2 .

$$y_{hi} = \beta_0 + \sum_{t=1}^{11} \beta_t x_{thi} + \epsilon_{hi}$$

$$\hat{y}_{hi} = \hat{\beta}_0 + \sum_{t=1}^{11} \hat{\beta}_t \, x_{thi}$$

Dado el conjunto de caracteristicas, la persona aquí tratada pertenece claramente a la clase número 5. De esta manera:

$$\hat{y}_{hi} = \hat{\beta}_0+\hat{\beta}_1(0)+\hat{\beta}_2(0)+\hat{\beta}_3(0)+\hat{\beta}_4(0)+\hat{\beta}_5(1)+\hat{\beta}_6(0)+\hat{\beta}_7(0)+$$
$$\hat{\beta}_8(0)+\hat{\beta}_9(1)+\hat{\beta}_{10}(1)+\hat{\beta}_{11}(3)$$

$$\hat{y}_{hi} = \hat{\beta}_0 + \hat{\beta}_5 + \hat{\beta}_9 + \hat{\beta}_{10} + 3\hat{\beta}_{11}$$

- TRANSFORMACION LOGARITMICA.

a) Conjunto de variables X_1 .

Como se hizo mención anteriormente, este modelo con Transformación Logarítmica estará representado por:

$$y_{hi} = \beta_0 \left(\prod_{t=1}^{10} \beta_t^{x_{thi}} \right) \exp(\varepsilon_{hi})$$

Tomando logaritmos en ambos lados:

$$\log(y_{hi}) = \log\beta_0 + \sum_{t=1}^{10} x_{thi} \log\beta_t + \varepsilon_{hi}$$

Redefiniendo términos:

$$z_{hi} = \phi_0 + \sum_{t=1}^{10} x_{thi}\phi_t + \varepsilon_{hi}$$

donde:

$$z_{hi} = \log y_{hi}$$

$$\phi_t = \log \beta_t \qquad t=\overline{0,10}$$

y una vez obtenidos los estimadores:

$$\hat{z}_{hi} = \hat{\phi}_0 + \sum_{t=1}^{10} x_{thi}\hat{\phi}_t$$

La prima neta o pérdida potencial para el h_i-ésimo individuo, entonces estará dada por:

$$\hat{y}_{hi} = \exp(\hat{z}_{hi}) = \exp(\hat{\phi}_0 + \sum_{t=1}^{10} x_{thi}\hat{\phi}_t)$$

A partir de datos obtenidos de la experiencia canadiense para los años 1984-1986, se obtuvieron los siguientes resultados para los estimadores :

Constante:	$\hat{\phi}_0$	$= -0.7373$
Edad:	$\hat{\phi}_1$	$= -0.8148$
Sexo:	$\hat{\phi}_2$	$= 0.7368$
Estado Civil:	$\hat{\phi}_3$	$= -0.4737$
Placer:	$\hat{\phi}_4$	$= -0.2904$
Negocios:	$\hat{\phi}_5$	$= 0.5696$
Distancia:	$\hat{\phi}_6$	$= 0.2573$
Manejo Ocasional:	$\hat{\phi}_7$	$= -2.1795$
Urbano:	$\hat{\phi}_8$	$= 0.5229$
Récord:	$\hat{\phi}_9$	$= -0.8207$
Var. de Tendencia:	$\hat{\phi}_{10}$	$= 0.0903$

Con estos valores, la pérdida potencial para el individuo ejemplificado h_1 será:

$$\hat{y}_{h1} = \exp(\hat{\phi}_0 + \hat{\phi}_1 + \hat{\phi}_5 + \hat{\phi}_8 + \hat{\phi}_9 + 3\hat{\phi}_{10})$$

$$= \exp(-1.0094)$$

$$= 0.36$$

Por lo tanto, este individuo deberá pagar una prima neta equivalente a 36 centavos de dólar.

b) Conjunto de variables X_2.

De forma análoga, se parte de:

$$y_{hi} = \beta_o \left(\prod_{t=1}^{11} \beta_t^{x_{thi}} \right) \exp(\epsilon_{hi})$$

y se llega a que la pérdida potencial es:

$$\hat{y}_{hi} = \exp(\hat{\phi}_o + \sum_{t=1}^{11} x_{thi} \hat{\phi}_t)$$

Usando los mismos datos de la experiencia canadiense, se obtuvo lo siguiente para este conjunto de variables:

Constante:	$\hat{\phi}_o$	=	3.9330
Clase 1:	---		
Clase 2:	$\hat{\phi}_2$	=	0.1007
Clase 3:	$\hat{\phi}_3$	=	0.2651
Clase 4:	$\hat{\phi}_4$	=	-0.2148
Clase 5:	$\hat{\phi}_5$	=	0.4027
Clase 6:	$\hat{\phi}_6$	=	0.4609
Clase 7:	$\hat{\phi}_7$	=	0.5662
Clase 8:	$\hat{\phi}_8$	=	0.9274
Urbano:	$\hat{\phi}_9$	=	0.0339
Récord:	$\hat{\phi}_{10}$	=	-0.2902
Var. de Tendencia:	$\hat{\phi}_{11}$	=	0.0981

La pérdida potencial para el individuo h_1 será:

$$\hat{y}_h = \exp(\hat{\beta}_0 + \hat{\beta}_8 + \hat{\beta}_7 + \hat{\beta}_{10} + 3\hat{\beta}_{11})$$

$$= \exp(4.3737)$$

$$= 79.34$$

Con este tratamiento el asegurado tendrá que pagar 79 dólares con 34 centavos por concepto de prima neta.

Las estimaciones de los parámetros aquí enunciadas serán retomadas en el Capítulo III. Cabe mencionar que los resultados de la estimación de las desviaciones estándar de los estimadores también se tratará en aquel capítulo.

1.4.6 VENTAJAS DEL MODELO.

- El modelo de regresión de una sola ecuación está basado en ideas ya no tan intuitivas, sino en ideas con firmes bases teóricas y técnicas estadísticas más sofisticadas.

- Este tipo de modelo nos permite saber (como se verá en el Capítulo III) si una variable tiene real influencia o impacto en la variable dependiente, estableciendo a la vez un nivel de confianza con el cual se da dicho efecto.

- El estudio del modelo así dado nos permite no sólo obtener un estimador puntual, sino un intervalo de confianza en el cual se puede encontrar dicho estimador, lo que constituye un arma estadística más poderosa.

1.4.7 DESVENTAJAS DEL MODELO.

- No se considera por separado a las componentes de la prima neta.

- Se ha trabajado bajo ciertos supuestos para el Término Perturbación, los que de no cumplirse traerán como consecuencia serios sesgos en la estimación de los parámetros y pérdida de propiedades deseables en los estimadores.

CAPITULO II

UN MODELO QUE SI CONSIDERA POR SEPARADO A LAS COMPONENTES DE LA PRIMA NETA PARA EL CALCULO DE LA MISMA

2.1 MODELO 5: MODELO DE REGRESION DE DOS PARTES.

Los modelos hasta ahora tratados no consideran la separabilidad de las componentes de la prima neta: Ocurrencia y Severidad del Reclamo.

Como se ha mencionado, bien puede ser que dichas componentes estén influenciadas por distintos factores, e incluso que un factor común a las componentes ejerza una distinta influencia.

El presente es un modelo basado en el Análisis de Regresión que sí considerará, para el cálculo de la prima neta, la separabilidad de sus componentes.

2.1.1 OCURRENCIA DEL RECLAMO.

El primer paso para el cálculo de la prima neta (pérdida esperada) consiste en obtener la probabilidad de que una persona con determinadas características efectúe uno o más reclamos en un período determinado de tiempo.

Para este caso frecuentemente se usa un proceso Poisson con el fin de aproximar la distribución de el número de reclamos asociados con un sólo individuo en un período dado de tiempo; sin embargo, de acuerdo a la experiencia de las compañías aseguradoras, la mayoría de los asegurados no tienen reclamos, y solo muy pocos tiene más de un reclamo en un año. De aquí es más plausible considerar que un asegurado, o no tiene reclamo o sólo efectúa uno, y así estimar la probabilidad de ocurrencia de dicho reclamo.

Puesto que se trabajará bajo dos únicos posibles eventos, se manejará una variable dicótoma, es decir, una variable que sólo tome uno de dos posibles valores: 1 si hay ocurrencia de reclamo; 0 si no hay ocurrencia de reclamo. Esto se basará en las características de la persona. De esta forma, se plantea el modelo con la siguiente variable dicótoma:

$$w_h = \begin{cases} 1 & \text{si } y_{ih} > 0 \\ 0 & \text{si } y_{ih} \leq 0 \end{cases}$$

donde:

$$y_{ih} = x'_{ih}\, \delta_i + \varepsilon_{ih} \qquad h = \overline{1,n}$$

x_{ih} y δ_i denotan vectores $m_i x1$ de variables explicativas y parámetros respectivamente, y ε_i es el término perturbación.

Como se mencionó, la ocurrencia del reclamo se dará o no con cierta probabilidad de acuerdo a las características del asegurado en cuestión, lo que queda expresado en las condicionantes que tienen los posibles valores de la variable dicótoma.

Nuevamente se asumen los supuestos clásicos del modelo de regresión lineal múltiple normal:

i) $\varepsilon_i \sim NI(0,\sigma^2_i)$.

ii) X_i es no estocástica.

iii) $rho(X_i) = m_i < n$

2.1.2 SEVERIDAD DEL RECLAMO.

Una vez establecida la probabilidad de la ocurrencia del reclamo, nos enfocamos en la severidad del mismo.

La ocurrencia del reclamo condiciona a la severidad del reclamo, es decir, se hablará de la severidad siempre y cuando se tenga previamente ocurrencia del reclamo. De aquí que la muestra considerada para estimar la severidad del reclamo consistirá exclusivamente de aquellos asegurados que hayan tenido reclamo.

Siguiendo con el modelo de regresión, se establece la variable Severidad del Reclamo (y_{2h}) la que puede ser lineal o una transformación logaritmica, y que está condicionada sobre la ocurrencia del reclamo, i.e., $w_h = 1$:

$$y_{2h} = x'_{2h}\delta_2 + \varepsilon_{2h}$$

donde x_{2h} y δ_2 son vectores de m_2x1 variables explicativas y parámetros respectivamente, y ε_{2h} es el término perturbación, el cual se supone independiente de ε_{1h} y x_{1h}.

El que ε_{2h} sea independiente de ε_{1h} significa que la severidad del reclamo es independiente de la ocurrencia del mismo.

Siguiendo con los supuestos clásicos del modelo de regresión lineal múltiple normal:

i) $\varepsilon_2 \sim NI(0,\sigma^2_2)$.

ii) X_2 es no estocástica.

iii) $rho(X_2) = m_2 < n$

Finalmente, la prima neta será obtenida como el producto de la probabilidad de ocurrencia del reclamo y su severidad esperada, condicionada sobre la ocurrencia del mismo.

$$PN_h = \{Pr(w_h=1)\} \{E[y_{2h} \mid w_h=1]\}$$

2.1.3 MANEJO DE LA INFORMACION.

Como se discutió en el modelo anterior, las variables serán de tipo cualitativo dado el tipo de información, y sólo una será de tipo cuantitativo (la variable de tendencia). Así también se establecieron en el modelo anterior algunos criterios para seleccionar uno entre diversos conjuntos de variables explicativas.

Las variables a escoger para este modelo serán las mismas que se han venido manejando. Se considerarán conjuntos de variables para la ocurrencia y la severidad del reclamo, las cuales diferirán en algunos puntos para de esta forma permitir la posibilidad de que distintas variables afecten a las componentes de la prima neta.

- OCURRENCIA DEL RECLAMO.

Para la Variable de Ocurrencia del Reclamo (y_{1n}), se proponen los siguientes dos conjuntos de variables:

X_1 = edad, sexo, estado civil, placer, negocios, distancia, manejo ocasional, urbano, récord previo de manejo, variable de tendencia.

X_2 = clases 1-8, urbano, récord previo de manejo, variable de tendencia.

De esta forma, se tiene a y_{1n} como:

X_1 :

$$y_{1n} = \delta_{10} + \delta_{11}x_{11n} + \delta_{12}x_{12n} + \ldots + \delta_{110}x_{110n} + \varepsilon_{1n}$$

X_2 :

$$y_{1n} = \delta_{10} + \delta_{11}x_{11n} + \delta_{12}x_{12n} + \ldots + \delta_{111}x_{111n} + \varepsilon_{1n}$$

donde las variables explicativas se definen de forma análoga a la descrita en el modelo anterior.

- SEVERIDAD DEL RECLAMO.

Para la Variable Severidad del Reclamo (y_{2h}) se tienen los siguientes dos conjuntos:

X_1 = edad, sexo, estado civil, placer, negocios, urbano, récord previo de manejo, variable de tendencia.

X_2 = clases 1-8, urbano, récord previo de manejo, variable de tendencia.

De esta forma, se tiene a y_{2h} como:

X_1 :

$$y_{2h} = \delta_{20} + \delta_{21}x_{21h} + \delta_{22}x_{22h} + \ldots + \delta_{28}x_{28h} + \varepsilon_{2h}$$

X_2 :

$$y_{2h} = \delta_{20} + \delta_{21}x_{21h} + \delta_{22}x_{22h} + \ldots + \delta_{211}x_{211h} + \varepsilon_{2h}$$

donde las variables explicativas se definen de forma análoga a la descrita en el modelo anterior.

2.1.4 ESTIMACION.

- OCURRENCIA DEL RECLAMO.

Para la estimación de la ocurrencia del reclamo se propone el modelo de resultado binario "Probit". Se prefiere este modelo en lugar de los modelos de resultado binario "Logit" y "Probabilidad Lineal", dado que se conoce que da ligeramente mejores predicciones.

Se definen las siguientes probabilidades:

$$P_h = Pr(w_h = 1)$$

$$1-P_h = Pr(w_h = 0)$$

$$\Rightarrow \qquad f(w_h) = P_h^{w_h}(1-P_h)^{w_h} \qquad\qquad w_h = 0,1$$

y:

$$w_h = \begin{cases} 1 & \text{si } y_{1h} > 0 \\ 0 & \text{si } y_{1h} \leq 0 \end{cases}$$

La probabilidad de ocurrencia del reclamo será:

$$P_h = Pr(w_h = 1)$$

$$= Pr[y_{1h} > 0]$$

$$= Pr[\Phi_{1h} > -x'_{1h}\delta_1]$$

Para el modelo Probit se considera que el término perturbación sigue una ley de distribución normal estándar, es decir:

$$\epsilon_h \sim N(0,1).$$

$$\Rightarrow \qquad P_h = Pr[\epsilon_h > -x'_{1h}\delta_1]$$

$$= 1-Pr[\epsilon_h \leq -x'_{1h}\delta_1]$$

$$= 1-F(-x'_{1h}\delta_1)$$

$$= F(x'_{1h}\delta_1)$$

Al ser $P_h = F(x'_{1h}\delta_1)$ se ve que la probabilidad de ocurrencia del reclamo siempre estará en el intervalo [0,1].

La estimación de la probabilidad de ocurrencia del reclamo se hará por medio del método de Máxima Verosimilitud.

$$L = \prod_{h=1}^{n} f(w_h)$$

$$= \prod_{h=1}^{n} P_h^{w_h}(1-P_h)^{1-w_h}$$

$$= \prod_{h=1}^{n} [F(x'_{1h}\delta_1)]^{w_h}[1-F(x'_{1h}\delta_1)]^{1-w_h}$$

Obteniendo el logaritmo de la Función de Verosimilitud:

$$\text{Log } L = \sum_{h=1}^{n} y_h \text{Log}[F(x'_{1h}\delta_1)] + \sum_{h=1}^{n} (1-y_h)\text{Log}[1-F(x'_{1h}\delta_1)]$$

Derivando el logaritmo de la función de verosimilitud con respecto a δ_1 e igualando a cero para obtener el valor crítico:

$$\frac{\delta \text{Log} L}{\delta \delta_1} = \sum_{h=1}^{n} y_h \frac{f}{F} x'_{1h} - \sum_{h=1}^{n}(1-y_h)\frac{f}{1-F} x'_{1h} = 0 \qquad (2.1.4.1)$$

donde $f = f(x'_{1h}\delta_1)$, $F = F(x'_{1h}\delta_1)$. f y F no son funciones lineales de $(x'_{1h}\delta_1)$, y por lo tanto (2.1.4.1.) no puede ser resuelta directamente.

Para obtener la solución a esta última ecuación recurrimos a un método iterativo para ecuaciones no lineales. Se propone el Método de Newton-Rapson:

$$\hat{\delta}_{1(t+1)} = \hat{\delta}_{1t} - \left[\frac{\delta^2 \text{LogL}}{\delta\delta_1 \delta\delta_1'}\bigg|_{\hat{\delta}_{1t}}\right]^{-1} \left[\frac{\delta \text{LogL}}{\delta\delta_1}\bigg|_{\hat{\delta}_{1t}}\right]$$

y:

$$\frac{\delta^2 \text{LogL}}{\delta\delta_1 \delta\delta_1'} = -\sum_{h=1}^{n} f\left[y_h \frac{f+(x_h'\delta_1)F}{F^2} + (1-y_h) f - \frac{(x_h'\delta_1)(1-F)}{(1-F)^2}\right] x_h x_h'$$

Un estimador de la Matriz de Varianzas y Covarianzas Asintótica que puede ser usada como una base para pruebas de hipótesis o intervalos de confianza es:

$$\hat{\sigma}^2_1 = -\left[\frac{\delta^2 \text{LogL}}{\delta\delta_1 \delta\delta_1'}\right]^{-1}$$

evaluada en el conjunto final de estimadores paramétricos δ_1 (Véase Judge, et al. pag. 792)

- SEVERIDAD DEL RECLAMO.

La estimación de la Severidad del Reclamo se hará por medio del Método de Máxima Verosimilitud.

La estimación de la matriz de coeficientes y la matriz de varianzas y covarianzas es la siguiente:

$$\hat{\delta}_2 = (X'X)^{-1}X'y$$

$$\hat{\sigma}_2 = \vartheta_2 = \frac{e'e}{n_2 - m_2}$$

que fué obtenida de forma totalmente análoga a la presentada en el modelo anterior. n_2 es el numero de individuos que sufrieron el accidente.

2.1.5 PROPIEDADES DE ESTIMACION

- OCURRENCIA DEL RECLAMO.

Es inmediato verificar que la derivada del Logaritmo de la Función de Verosimilitud es negativa, y por lo tanto se tendrá un máximo.

Dado que el estimador de este renglón se obtuvo por un método iterativo, se busca que dicho método converja a un solo valor, el cual será el máximo absoluto. Esta convergencia se dará si:

$$\hat{\delta}_{1(n+1)} = \hat{\delta}_{1(n)} \qquad (n \to \infty)$$

En cuanto al estimador de la Matriz de Varianzas y Covarianzas, se sabe que es Consistente (Ver Judge, et al. pag. 792).

- SEVERIDAD DEL RECLAMO.

Las propiedades de los estimadores para esta parte del modelo son las mismas que las enunciadas en el modelo anterior.

2.1.6 PRONOSTICO.

Con el fin de hacer pronósticos se incluyó la Variable de Tendencia. Este tópico se trató en el capitulo anterior.

2.1.7 EJEMPLO ILUSTRATIVO.

Para este ejemplo se usará la Regresión Lineal para la ecuación Ocurrencia del Reclamo, y para la Severidad del Reclamo se usará la Transformación Logaritmica. El conjunto utilizado será el de variables X_2.

Ocurrencia del Reclamo.

Tomando al individuo ejemplificado en el anterior modelo, se tiene lo siguiente:

$$\hat{y}_{1hi} = \hat{\delta}_{10} + \hat{\delta}_{18} + \hat{\delta}_{19} + \hat{\delta}_{110} + 3\hat{\delta}_{111}$$

En base a los mismos datos de la Experiencia Canadiense, se obtuvieron las siguientes estimaciones:

Constante:	$\hat{\delta}_{10} = 1.4370$
Clase 1:	---
Clase 2:	$\hat{\delta}_{12} = 0.0651$
Clase 3:	$\hat{\delta}_{13} = 0.1093$
Clase 4:	$\hat{\delta}_{14} = -0.4013$
Clase 5:	$\hat{\delta}_{15} = 0.1778$
Clase 6:	$\hat{\delta}_{16} = 0.2512$
Clase 7:	$\hat{\delta}_{17} = 0.3140$
Clase 8:	$\hat{\delta}_{18} = 0.4312$
Urbano:	$\hat{\delta}_{19} = 0.1880$
Récord:	$\hat{\delta}_{110} = -0.2094$
Var. de Tend.	---

- **Severidad del Reclamo.**

En este caso:

$$\hat{y}_{2h_i} = \exp\{\hat{\Phi}_{20} + \hat{\Phi}_{25} + \hat{\Phi}_{29} + \hat{\Phi}_{210} + \hat{\Phi}_{211}\}$$

Constante:	$\hat{\Phi}_{20} = 7.3070$
Clase 1:	$\hat{\Phi}_{21} = -0.0074$
Clase 2:	---
Clase 3:	---
Clase 4:	---
Clase 5:	$\hat{\Phi}_{25} = 0.1303$
Clase 6:	$\hat{\Phi}_{26} = 0.0798$
Clase 7:	$\hat{\Phi}_{27} = 0.0928$
Clase 8:	$\hat{\Phi}_{28} = 0.2850$
Urbano:	$\hat{\Phi}_{29} = -0.2512$
Récord:	$\hat{\Phi}_{210} = 0.0108$
Var. de Tend.	$\hat{\Phi}_{211} = 0.0942$

- **Cálculo de la Prima Neta.**

Como se mencionó, la prima neta a cobrar al h_i-ésimo individuo, se obtendrá como el producto de la probabilidad de ocurrencia del reclamo y la esperanza del monto estimado del mismo.

Evaluando $x'_{1hi}\hat{\delta}_1$:

$$x'_{1h}\hat{\delta}_1 = \hat{\delta}_{10} + \hat{\delta}_{18} + \hat{\delta}_{19} + \hat{\delta}_{110}$$

$$= -1.4370 + 0.1778 + 0.1880 - 0.2094$$

$$= -1.2806$$

y se obtuvo que $Pr(w_{hi} = 1) = F(x'_{1h}\hat{\delta}_1)$:

$$Pr(w_{hi} = 1) = F(x'_{1hi}\hat{\delta}_1)$$

$$= F(-1.2806)$$

$$= 0.09695$$

Es decir, la probabilidad de que una persona con las características dadas efectúe un reclamo para el tercer año de valuación es de poco más del 0.09.

Procediendo como en el anterior modelo, el valor de la Severidad del Reclamo es:

$$\hat{y}_{2hi} = \exp(7.3070 + 0.1303 + -0.2512 + 0.0108 + 3*0.0942)$$

$$= 1771.35$$

es decir, el monto del reclamo que efectuaría una persona con las características dadas, dado que efectivamente efectuó reclamo es de 1771 dólares con 35 centavos.

En resumen, la pérdida esperada (prima neta) por parte del individuo ejemplificado será de:

$$\hat{PN}_{hi} = (0.09695)(1771.35)$$

$$= 171.73$$

El individuo tendrá que pagar una prima neta de 171 dólares con 73 centavos.

2.1.8 VENTAJAS DEL MODELO.

- Este modelo considera por separado a las componentes de la prima neta.

- Los métodos de estimación tienen una firme base teórica, arrojando estimadores con ciertas caracteristicas muy deseables.

- Este tipo de modelo permite saber si una variable tiene real influencia o impacto en la variable dependiente, estableciendo a la vez un nivel de confianza con el cual se da dicho efecto.

- El estudio del modelo así dado nos permite no sólo obtener un estimador puntual, sino un intervalo de confianza en el cual se puede encontrar dicho estimador, lo que constituye un arma estadística más poderosa.

2.1.9 DESVENTAJAS DEL MODELO

- De no cumplirse los supuestos bajo los que se trabajó para el Término Perturbación, se tendrán serios sesgos y pérdida de propiedades deseables en un estimador.

- Se supuso que los términos perturbación de la ocurrencia del reclamo y la severidad del mismo son independientes. De no cumplirse ésto, también traerá como consecuencia serios sesgos en las estimaciones.

CAPITULO III

COMPARACION ESTADISTICA DE LOS MODELOS ENUNCIADOS

3.1 COMPARACION DE LOS MODELOS DE REGRESION POR SIGNIFICACION ESTADISTICA DE LOS REGRESORES.

En esta sección se hará una comparación de los modelos 4 y 5 en cuanto al nivel de significación estadística de los coeficientes de regresión estimados.

El que un estimador paramétrico sea significativo estadísticamente hablando quiere decir que éste tenga un real impacto en la variable dependiente (reclamo potencial), estableciendo un cierto nivel de confianza con el cual se dará dicho impacto. En otras palabras, un estimador paramétrico causará un cierto impacto en el reclamo potencial si su valor es distinto de cero, a un nivel de confianza dado. De esta forma, se hace el planteamiento de una Prueba de Hipótesis.

3.1.1 METODO GENERAL DE PRUEBAS DE HIPOTESIS.

Siguiendo el Método General de Pruebas de Hipótesis, tenemos lo siguiente:

1o. Proponer una Hipótesis.

. Una hipótesis siempre es propuesta buscando rechazarla. En este caso buscamos rechazar que no existe regresión, es decir, nuestra Hipótesis Nula será:

$$Ho: \quad \beta_j = 0$$

versus toda alternativa ($\beta_j \neq 0$).

2o. Buscar una Estadística de Prueba.

Usando una Prueba de Razón de Verosimilitud (PRV), se tiene:

$$T_{\alpha} = \frac{\hat{\beta} - \beta \mid H_0}{\sqrt{[\sigma^2_{\omega}(X'X)^{-1}]}} = \frac{\hat{\beta}}{\sqrt{[\sigma^2_{\omega}(X'X)^{-1}]}}$$

3o. Encontrar la distribución de la prueba.

Bajo el supuesto de que el Término Perturbación sigue una ley de distribución Normal, se tiene que:

$$\hat{\beta} \sim N[\beta, \sigma^2_{\omega}(X'X)^{-1}]$$

Estandarizando:

$$\frac{\hat{\beta} - \beta}{\sqrt{[\sigma^2_{\omega}(X'X)^{-1}]}} \sim N(0,1)$$

Puesto que casi nunca se conoce la verdadera varianza, usamos el estimador de la varianza. Esto nos lleva a una Distribución t de Student:

$$\frac{\hat{\beta} - \beta}{\sqrt{[\hat{\sigma}^2_{\omega}(X'X)^{-1}]}} \sim t_{n-1}$$

$$\Rightarrow \qquad T_{\alpha} \sim t^{*}_{n-1,\alpha}$$

4o. Escoger un nivel de significación α o un nivel de confianza $1-\alpha$.

Seleccionemos $1-\alpha = 0.95$.

5o. Comparar niveles significativos.

Rechazamos Ho si:

$$T_c > t^*_{n-1,0.975}$$

ó

$$T_c < -t^*_{n-1,0.975}$$

No rechazamos Ho si:

$$-t^*_{n-1,0.975} \leq T_c \leq t^*_{n-1,0.975}$$

Consideremos que el tamaño de la muestra es muy grande ($n \to \infty$). Se toma una muestra así dadas las desviaciones que son características en el Seguro de Automóviles. Este tipo de desviación se ilustró en el Modelo I cuando una celda tiene un número bajo de elementos.

En Tablas de t de Student, se tiene que:

$$t^*_{\infty,0.975} = 1.96$$

3.1.2 ESTIMACION POR CLASES.

La estimación de los regresores se basó en la Experiencia Canadiense para los años de 1984 a 1987.

La siguiente Tabla se obtuvo como resultado de la estimación de la ecuación "Ocurrencia del Reclamo" por medio del Modelo Probit.

TABLA XI

Modelo Probit para la Ocurrencia del Reclamo

Constante	Clase 2	Clase 3	Clase 4	Clase 5	Clase 6	Clase 7	Clase 8	Urbano	Récord Previo de Manejo
-1.4370	0.0651	0.1093	-0.4013	0.1778	0.2512	0.3140	0.4312	0.1880	-0.2094
(0.0026)	(0.0013)	(0.0022)	(0.0040)	(0.0028)	(0.0027)	(0.0046)	(0.0029)	(0.0012)	(0.0024)
(-552.69)	(50.07)	(49.68)	(-100.33)	(63.50)	(93.05)	(68.26)	(148.69)	(156.67)	(-87.25)

Los números que no se encuentran entre paréntesis son los valores estimados de los regresores, mientras que los números entre paréntesis son las desviaciones estándar de los regresores y los cocientes de los estimadores entre sus desviaciones estándar, respectivamente.

A continuación se presentan los resultados de la estimación de la ecuación "Severidad del Reclamo". Esta estimación se basó en la Transformación Logarítmica.

TABLA XII

El Logaritmo de los Reclamos Potenciales por Clasificación Múltiple

	Modelo IV	Modelo V
Constante	3.9330 (2.5620) (1.53)	7.3070 (0.0110) (664.27)
Clase 1		-0.0074 (0.0010) (-7.40)
Clase 2	0.1007 (0.0750) (1.34)	. . .
Clase 3	0.2651 (0.1250) (2.12)	. . .
Clase 4	-0.2148 (0.4700) (-0.46)	. . .
Clase 5	0.4027 (0.2030) (1.98)	0.1303 (0.0070) (18.61)
Clase 6	0.4609 (0.2850) (1.62)	0.0798 (0.0070) (11.40)
Clase 7	0.5662 (0.3560) (1.59)	0.0928 (0.0210) (4.42)
Clase 8	0.9274 (0.4850) (1.91)	0.2150 (0.0070) (30.71)
Urbano	0.0339 (0.2170) (0.16)	-0.2512 (0.0020) (-125.60)
Récord Previo	-0.2902 (0.2370) (-1.22)	0.0180 (0.0180) (1.00)
Variable de Tendencia	0.0981 (0.0030) (32.70)	0.0942 (0.0010) (94.20)

- SIGNIFICACION ESTADISTICA.

En la ecuación "Severidad del Reclamo" para el Modelo IV puede observarse que solamente los coeficientes de la Clase 3, 5, y de la Variable de Tendencia son estadisticamente significativos a un nivel de confianza del 95%, mientras que en el Modelo V todos lo son, a excepción del Récord Previo de Manejo.

- SIGNOS.

Para efectos del monto de reclamación que se espera haga una persona con determinadas características, se postulan, basándonos en la experiencia, los siguientes supuestos "a priori":

COEFICIENTES	VALOR	JUSTIFICACION
Constante	> 0	Se debe cobrar una prima base positiva.
Clases	cualquiera	Dependerá de la combinación manejada.
Urbano	> 0	Condiciones de manejo menos favorables.
Récord previo	cualquiera	No se puede generalizar.
Var. de Tend.	> 0	Procesos inflacionarios, tecnologias, etc.

Estos puntos pueden ser ampliamente discutidos y rebatidos; sin embargo, se toman aqui como un marco de referencia al ver los resultados de la estimación.

Obsérvese cómo en la Tabla XII la estimación de el Término Constante y de la Variable de Tendencia es positiva para ambos modelos.

Ahora centrémonos en la estimación del coeficiente Urbano. Para el Modelo IV es positivo (como se estableció "a priori"), mientras que es negativo para el Modelo V. Sin embargo, vemos que para la ecuación Ocurrencia del Reclamo es positivo.

En este punto se observa que un parámetro común a la Ocurrencia y Severidad del Reclamo tiene una distinta influencia.

La explicación para este hecho es que, al manejar en un ambiente urbano, se incrementa la probabilidad de tener un accidente dado que hay un mayor número de vehiculos; pero este hecho ocasiona embotellamientos, lo cual hace suponer que gran número de estos siniestros se den a una baja velocidad, por lo que el accidente no será tan cuantioso en términos monetarios.

Un cambio análogo se presenta en el Récord Previo de Manejo.

Una persona con 5 o más años sin accidente tendrá una menor probabilidad de sufrir un siniestro, pero en el momento en que se siniestre su impacto será mayor. Aunque cabe aclarar que ésto no se da a un nivel de confianza del 95%.

3.1.3 ESTIMACION POR ATRIBUTOS PERSONALES Y CONDICIONES DE MANEJO.

La Tabla abajo enunciada da la estimación de la ecuación Severidad del Reclamo. También se obtuvo por el planteamiento de la Transformación Logaritmica.

TABLA XIII

El Logaritmo de los Reclamos Potenciales por Clasificación Múltiple

	Modelo IV	Modelo V
Constante	-0.7373 (2.1880) (-0.34)	7.3830 (0.0120) (615.25)
Edad	-0.8160 (0.1970) (-4.14)	-0.0780 (0.0070) (-11.14)
Sexo	0.7368 (0.1450) (5.08)	0.2066 (0.0100) (20.66)
Estado Civil	-0.4737 (0.0790) (-5.99)	-0.1935 (0.0220) (-8.79)
Placer	-0.2906 (0.0780) (-3.72)	-0.0073 (0.0010) (-7.30)
Negocios	0.5696 (0.1170) (4.87)	0.1766 (0.0090) (19.63)
Distancia	0.2573 (0.0430) (5.98)	.
Manejo Ocasional	-2.1795 (0.6730) (-3.24)	.
Urbano	0.5229 (0.2090) (2.50)	-0.2511 (0.0020) (-125.55)
Récord Previo	-0.8207 (0.2280) (-3.59)	0.0149 (0.0100) (1.49)
Variable de Tendencia	0.0903 (0.0010) (90.30)	0.0942 (0.0010) (94.20)

- SIGNIFICACION ESTADISTICA.

Exceptuando al cociente de la Constante en el Modelo IV, y al Récord Previo de Manejo en el Modelo V, todos los parámetros estimados son estadisticamente significativos (distintos de cero) a un nivel de confianza del 95%.

Nótese cómo, a excepción del Récord Previo de Manejo, todos los cocientes del Modelo V son más grandes que los cocientes del Modelo IV.

Al tener cocientes cada vez más grandes, se abre la posibilidad de ir proponiendo niveles de confianza cada vez mayores, con lo que se tendrá una mejor base para el planteamiento del impacto de un factor en una variable dependiente.

- SIGNOS.

Los supuestos "a priori" para los posibles valores de los coeficientes estimados son los siguientes:

COEFICIENTES	VALOR	JUSTIFICACION
Constante	ya visto.	
Edad	< 0	Una persona más joven en general es menos conciente que una persona de mayor edad.
Sexo	> 0	Un hombre es más audaz una mujer.
Estado Civil	< 0	Una persona casada es más cuidadosa que cualquier otra.
Placer	< 0	Se conduce una distancia anual corta.
Negocios	> 0	Se recorren grandes distancias.
Distancia	> 0	Distarcia más larga al trabajo.
Manejo Ocasional	< 0	Se maneja poco.
Urbano	ya visto.	
Récord Previo	ya visto.	
Var. de Tend.	ya visto.	

Todos los supuestos concuerdan en los dos modelos, a excepción de la Constante, Urbano y Récord Previo de Manejo.

La Constante en el Modelo IV es negativa, lo que es totelmente contradictorio. Además de no ser significativa a un nivel de confianza del 95%.

La discusión para los parámetros Urbano y Récord Previo para el Modelo V es como se discutió anteriormente.

3.2 COMPARACION DE TODOS LOS MODELOS POR NIVEL DE PREDICCION.

Se dice que un modelo tiene un mejor funcionamiento a medida que éste tiene una predicción más acertada. Así, un modelo será más deseable de implementar con respecto de otros cuando sus predicciones tengan una menor desviación con respecto de los valores reales.

Para hacer una comparación de los modelos en cuanto a su nivel de predicción se echará mano de un Error Total. Este Error Total combina el error de la variable y su sesgo, el cual es conocido como la "Raiz del Error Cuadrado Medio", y suele reemplazar al Error Estándar.

Se define el Error Total :

$$\text{Error Total} = \sqrt{[\text{Var} + (\text{Sesgo})^2]}$$

donde:

$$\text{Var} = E(\bar{y}_e - E\bar{y}_e)^2$$

$$\text{Sesgo} = E(\bar{y}_e) - \bar{Y}_{verdadero}$$

La siguiente Tabla presenta la Raíz de los Errores de Predicción Cuadrados Medios:

TABLA XIV

Raíz de los Errores de Predicción Cuadrados Medios

Modelos	1984	1985	1986	1987
Modelo I	7 563	7.033	16.141	18.299
Modelo II	37.678	44.540	52.681	56.032
Modelo III	11.092	8.211	12.378	16.085
Modelo IV(a)	197.683	220.792	247.634	145.457
(b)	197.683	220.792	247.635	28028.6
Modelo V(a)	10.800	9.554	11.057	13.476
(b)	9.888	8.701	10.240	12.671

(a) La Ecuación Severidad del Reclamo se basó en la Tabla XII.
(b) La Ecuación Severidad del Reclamo se basó en la Tabla XIII.

Los resultados de 1984 y 1985 se basaron en estimaciones usando sólo datos de 1984 y 1985. Los resultados de 1986 son errores de predicción de los parámetros estimados basados en los datos de 1984 y 1985. Los resultados de 1987 están basados en los parámetros estimados usando los datos de 1984 a 1986.

Se observa que el Modelo IV tiene las peores predicciones en todos los años. Para los años de 1984 y 1985 el mejor funcionamiento es el del Modelo de la Celda Media, mientras que para 1986 y 1987 el mejor es el Modelo de Regresión de Dos Partes.

De aquí podemos decir que el Modelo de la Celda Media provee un mejor "ajuste" de los datos, mientras que las mejores predicciones son las dadas por el Modelo que considera la separación de las componentes de la Prima Neta.

Adicionalmente se ve que el basar la ecuación de la Severidad del Reclamo tanto en Clases como en Características Personales provee similares resultados.

CONCLUSIONES

Basándose en la Experiencia Canadiense para los años de 1984 a 1987, se llegó a la conclusión de que las componentes de la Prima Neta están influenciadas por componentes comunes de una forma distinta. Véase por ejemplo la influencia de la Variable Explicativa Urbano.

Así mismo se vio que este modelo que plantea la separabilidad de las componentes de la Prima Neta nos lleva a predicciones más exactas con respecto a los otros modelos que no plantean dicha separabilidad.

Sin embargo, ésto no quiere decir que el Modelo de Regresión de Dos Partes dará siempre mejores predicciones.

Dada la naturaleza altamente variable de los reclamos en el Seguro de Automóviles, bien puede ser que, para otros años u otra información, otros modelos funcionen mejor.

LA EXPERIENCIA MEXICANA

El cálculo de la prima neta en el seguro de automóviles en México se basa en la pérdida promedio, y considerando solamente el modelo y tipo de vehículo.

Al tomar en cuenta exclusivamente estas características del vehículo, se está desperdiciando valiosa información acerca del asegurado y su contorno. Como se vio, estas características personales sí influyen en la frecuencia y monto de la siniestralidad. Además, este tipo de información sí es colectada por las Compañías Aseguradoras.

El Modelo de la Celda Media es bueno; sin embargo, debería tratar de implementarse modelos estadísticos más sofisticados para el cálculo de la prima neta.

Así, bien se ve que en el campo del Seguro de Daños, y más específicamente en el Seguro de Automóviles, aún hay mucho por hacer.

BIBLIOGRAFIA

BURDEN, F., 1988
"Numerical Analysis"
Boston, McGraw-Hill.

CHANG, L. y Fairley, W.B., 1979
"Pricing Automobile Insurance under Multivariate Classification of Risks: Additive versus Multiplicative"
The Journal of Risk & Insurance 49, pp. 539-563.

CHIANG, A.C., 1974
"Fundamental Methods of Mathematical Economics"
Nueva York, McGraw-Hill.

GUJARATI, D., 1981
"Econometría Básica"
México, McGraw-Hill.

HADLEY, G., 1961
"Linear Algebra"
Mass., Addison-Wesley.

HARRIS, B.
"Theory of Probability"
Addison Wesley.

HOGG, R. & CRAIG, A., 1978
"Introduction to the Mathematics of Statistics"
New york, Macmillan.

HSIAO, C., Kim, C. y Taylor, G., 1990
"A Statistical Perspective on Insurance Rate-Making"
Journal of Econometrics 44, pp. 5-24.

HUANG, D.S., 1970
"Regression & Econometric Methods"
E.E.U.U., John Wiley & Sons

INTRILIGATOR, M.D., 1990
"Modelos Econométricos, Técnicas y Aplicaciones"
México, F.C.E.

JOHNSTON, J., 1972
"Econometric Methods"
Nueva York, McGraw-Hill.

JUDGE, G.G. et al, 1988
"Introduction to the Theory & Practice of Econometrics"
República de Singapur, John Wiley & Sons.

KENDALL, M.G. y Stuart, A., 1963
"The Advanced Theory of Statistics", V. I
Londres, Charles Griffin & Co.

KISH, L., 1975
"Muestreo de Encuestas"
México, Trillas.

MODD, A.M., y Graybill, F.A., 1974
"Introduction to the Theory of Statistics"
Nueva York, McGraw-Hill.

PARZEN, E., 1987
"Teoría Moderna de Probabilidades y sus Aplicaciones"
México, Limusa.

SANT, D.T., 1980
"Estimating Expected Losses in Auto Insurance"
The Journal of Risk & Insurance 47, pp. 133-151.

STWEART, M.B., y Wallis, K.F., 1984
"Introducción a la Econometría"
Madrid, Alianza Editorial.

TAYLOR, L.D., 1974
"Probability & Mathematical Statistics"
Nueva York, Harper & Row Publishers Incorporated.

WEISBERG, H.I. y Tomberlin, T.J., 1982
"A Statistical Perspective on Actuarial Methods for Estimating Pure Premiums from Cross-Classified Data"
The Journal of Risk & Insurance 49, pp. 539-563