UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA

Universidad de El Salvador
Hacia la libertad por la cultura

"ESTADÍSTICA APLICADA AL ANÁLISIS ACTUARIAL"

TRABAJO DE GRADUACIÓN PRESENTADO POR:

ROSA MATILDE RIVERA ROSA

PARA OPTAR AL GRADO DE:

LICENCIADA EN ESTADÍSTICA

CIUDAD UNIVERSITARIA, AGOSTO DE 2011.

SAN SALVADOR EL SALVADOR CENTRO AMERICA

UNIVERSIDAD DE EL SALVADOR

RECTOR: **ING. RUFINO QUEZADA.**

SECRETARIO GENERAL: **LIC. DIOUGLAS VLADIMIR ALFARO.**

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO: **DR. RAFAEL ANTONIO GOMEZ ESCOTO.**

SECRETARIO: **LIC. MARÍA TRINIDAD TRIGUEROS DE CASTRO.**

ESCUELA DE MATEMÁTICA

DIRECTOR: **ING. CARLOS MAURICIO CANJURA LINARES.**

SECRETARIO: **LIC. ERNESTO AMÉRICO HIDALGO CASTELLANOS.**

CIUDAD UNIVERSITARIA, AGOSTO DE 2011.

UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA

TRABAJO DE GRADUACIÓN

"ESTADÍSTICA APLICADA AL ANÁLISIS ACTUARIAL"

PRESENTADO POR:

ROSA MATILDE RIVERA ROSA.

____

ASESOR:
Dr. JOSÉ NERYS FUNES TORRES.

CIUDAD UNIVERSITARIA, AGOSTO DE 2011.

Agradecimientos.

A mi buen Dios y Padre que esta en los cielos, por darme su amor, sabiduría, fortaleza y valor para seguir adelante en los momentos difíciles al final de la carrera y permitirme terminar lo que un día comencé.

A mi madre, por su gran amor, su comprensión y su incondicional apoyo en los momentos difíciles.

A mi hermano, hermana y a su querido esposo José Ulloa (Q.P.D.), por su amor, apoyo y comprensión.

Al Dr. José Nerys Funes Torres, por brindarme su valiosa y oportuna ayuda, por compartir sus conocimientos y experiencias y sobre todo por su comprensión y calidad humana.

A todos los docentes que fueron parte de mi proceso de enseñanza y aprendizaje, por compartir sus conocimientos y experiencias.

A todas las personas que en su momento me brindaron su cariño, amistad, ayuda y apoyo.

Dios bendiga grandemente la vida de aquellos que fueron pieza importante en este escalón más que hoy Dios me permite alcanzar.

ROSA MATILDE RIVERA ROSA.

INDICE

INTRODUCCIÓN

La *Estadística* o *Matemática Actuarial* es la ciencia que proporciona las herramientas necesarias para el estudio de ciertas actividades económicas que llevan a cabo las compañías de seguros. Estas actividades aparecen ligadas al término *riesgo*, habitual en los escenarios actuariales. Por *riesgo* se entenderá *cualquier suceso que pueda ocurrir y que suponga una pérdida, generalmente cuantificable en términos económicos.*

La clasificación más simple es la que distingue entre *seguros no vida y seguros vida.* Los primeros, denominados también seguros generales, cubren en su totalidad los seguros que habitualmente se contratan: seguro de automóviles, de accidentes, contra incendios, contra robos, hogar, etc. Por otro lado un contrato de vida se establece entre una empresa aseguradora y una persona, el asegurado, bajo el que la aseguradora se compromete a pagar a la muerte del asegurado una suma fija al o los beneficiaros designados por el mismo.

Se denomina *asegurado a la persona física o jurídica titular del bien o interés asegurado que está expuesto al riesgo.*

Un *siniestro es un suceso que produce un daño previsto en el contrato de seguro y que da lugar al cumplimiento de las obligaciones contraídas por la aseguradora mediante la reposición del bien o la indemnización al asegurado.* El *asegurador*, dependiendo del contrato del seguro, asume la obligación del pago de la indemnización, cantidad que debe pagar una aseguradora si se produce un siniestro, o de realizar determinadas prestaciones. La relación entre el asegurado y el asegurador se establece por medio del contrato de seguro, documento por el que el asegurador se encarga de cubrir riesgos ajenos mediante el cobro de un precio denominado *prima.* El contrato se formaliza a través de una póliza, en él debe constar por escrito el riesgo que se cubre, la suma asegurada, el importe de la prima, etc.

En el negocio del seguro aparecen incertidumbres para el asegurador que no suelen aparecer en otro tipo de negocio. Generalmente son de dos tipos: *el número de reclamaciones o siniestros y la cuantía de las mismas.*

Lo ideal, cuando se va a elaborar un procedimiento de cálculo de prima es trabajar con las variables aleatorias del número de reclamaciones y de la cuantía de las mismas para obtener lo que se denomina un *modelo compuesto*.

De acuerdo a lo anterior, este trabajo pretende realizar un estudio exhaustivo y comprensivo del análisis de la Estadística Actuarial, en donde se abordará el estudio de un modelo compuesto para el cálculo de *prima*, también se abordará el modelo de riesgo individual y el modelo no compuesto.

OBJETIVOS

OBJETIVO GENERAL

- Estudiar los fundamentos teóricos y las aplicaciones de la estadística en el análisis actuarial.

OBJETIVOS ESPECÍFICOS

- Estudiar las herramientas estadísticas básicas para el *análisis actuarial.*
- Abordar el problema del *cálculo de prima,* a partir del planteamiento de un modelo de variables aleatorias.
- Aplicar la teoría de la *estadística actuarial* en el estudio de casos.
- Inferir sobre los resultados obtenidos en el estudio de casos.
- Elaborar una propuesta metodológica de aplicación de la estadística en el análisis actuarial.

CAPÍTULO 1

INTRODUCCIÓN A LA ESTADÍSTICA ACTUARIAL

1.1. INTRODUCCIÓN.

En este capítulo se describe de una manera breve el objeto de estudio de la Estadística Actuarial y su aplicación en el negocio del seguro, también se estudiará la teoría básica que nos introducirá al estudio de la Estadística Actuarial.

Se incluirá el estudio de la *función de supervivencia*, la *distribución estacionaria de renovación* que juega un papel muy importante en la teoría de la ruina en tiempo continuo, *transformaciones de las variables aleatorias de aplicación en reaseguros* de gran utilidad en estadística actuarial y finalmente las *técnicas de inferencia bayesiana*.

1.2. OBJETO DE ESTUDIO DE LA ESTADÍSTICA ACTUARIAL.

La *Estadística* o *Matemática Actuarial* es la ciencia que proporciona las herramientas necesarias para el estudio de ciertas actividades económicas que llevan a cabo las compañías de seguros. Estas actividades aparecen ligadas al término *riesgo*, habitual en los escenarios actuariales. Por *riesgo* se entenderá *cualquier suceso que pueda ocurrir y que suponga una pérdida, generalmente cuantificable en términos económicos.*

La clasificación más simple es la que distingue entre *seguros no vida y seguros vida.* Los primeros, denominados también seguros generales, cubren en su totalidad los seguros que habitualmente se contratan: seguro de automóviles, de accidentes, contra incendios contra robos, hogar, etc. Por otro lado un contrato de vida se establece entre una empresa aseguradora y una persona, el asegurado, bajo el que la aseguradora se compromete a pagar a la muerte del asegurado una suma fija al o los beneficiaros designados por el mismo.

Se denomina *asegurado a la persona física o jurídica titular del bien o interés asegurado que está expuesto al riesgo.*

Un *siniestro es un suceso que produce un daño previsto en el contrato de seguro y que da lugar al cumplimiento de las obligaciones contraídas por la aseguradora mediante la reposición del bien o la indemnización al asegurado.* El *asegurador*, dependiendo del contrato del seguro, asume la obligación del pago de la indemnización, cantidad que debe pagar una aseguradora si se produce un siniestro, o de realizar determinadas prestaciones. La relación entre el asegurado y el asegurador se establece por medio del *contrato de seguro*, documento por el que el asegurador se encarga de cubrir riesgos ajenos mediante el cobro de un precio denominado *prima.* El contrato se formaliza a través de una *póliza*, documento o documentos en el que se recogen los acuerdos entre el asegurado y la compañía de seguros. En el mismo deben constar por escrito el riesgo que se cubre, la suma asegurada, el importe de la prima, etc.

Las herramientas estadísticas necesarias para abordar los seguros no vida son más sencillas que las necesarias para abordar los seguros vida. Aunque la muerte es un suceso seguro, es incierto el momento en que se producirá, dependiendo de elementos como el sexo, edad, factores genéticos o hereditarios, estado de salud, formas de vida, guerras, etc. Este tipo de seguros se fundamenta sobre las tablas de mortalidad, resumen de los riesgos de vida de un grupo representativo de individuos suficientemente grande.

Bajo una póliza de seguro no vida, los factores que influyen en el acontecimiento de un suceso contra el cual un bien puede asegurarse pueden ser múltiplos y diversos. Por ejemplo en el seguro de automóviles el acontecimiento de un siniestro puede obedecer a causas como la edad del conductor, antigüedad del vehiculo, cilindrada del mismo, etc. En cualquiera de los casos, el asegurado paga una cantidad de dinero, la prima del seguro al asegurador, la compañía de seguros, y desde ese instante comienza la *cobertura* sobre el bien asegurado.

En el negocio del seguro aparecen incertidumbres para el asegurador que no suelen aparecer en otro tipo de negocio. Generalmente son de dos tipos: *el número de reclamaciones o siniestros y la cuantía de las mismas.*

Lo ideal, cuando se va a elaborar un procedimiento de cálculo de prima es trabajar con las variables aleatorias del número de reclamaciones y de la cuantía de las mismas para obtener lo que se denomina un *modelo compuesto*.

De acuerdo a lo anterior, la Estadística Actuarial aborda el estudio de un modelo compuesto para el cálculo de *prima*, así como también el estudio del modelo de riesgo individual, el modelo no compuesto, la teoría de la *credibilidad* , así como diversas modalidades de *reaseguro*.

Un *reaseguro* no es más que el seguro de las compañías de seguro, de modo que si ven afectado su negocio por el volumen de primas contratadas, aseguran parte de las mismas en otra compañía de seguros (reaseguradora*)*

En el modelo IBNR del inglés: (*Incurred But Not Reported*) se considera un amplio espectro de situaciones de reservas que pueden ajustarse a dicho modelo: reservas para reclamaciones pendientes, reservas para reclamaciones conocidas por la compañía pero no completamente pagadas.

Controlar la solvencia de las empresas de seguros supone una de las principales ocupaciones de la ciencia actuarial. Obviamente para una compañía de seguros, y para cualquier empresa, es vital prevenirse de resultados negativos.

Una compañía es solvente si posee suficientes activos para hacer frente a sus pasivos. Aunque esto es cierto para cualquier empresa o negocio, en la compañía de seguros el asunto toma otro cariz. En el negocio del seguro los pasivos no son nunca conocidos y pueden extenderse por muchos años en el futuro. Suponiendo que la compañía aseguradora comienza con una cantidad de dinero positiva (reservas iniciales en el tiempo 0), recauda primas y paga las indemnizaciones correspondientes en la medida en que se producen. Si las reservas o capital disponible por la aseguradora llegan a ser negativas o nulas, se dice entonces que ocurre la ruina.

El objetivo fundamental de la teoría de la ruina es el cálculo de la probabilidad de ruina. Este es un indicador muy útil para la compañía de seguros ya que un valor alto de la misma indica inestabilidad, lo que supone que la compañía tome medidas como el reaseguro o cambie su política de primas.

En la teoría de la ruina el proceso de reservas de una compañía de seguros se modela de acuerdo a un proceso estocástico de riesgo.

1.3. FUNCIONES DE SUPERVIVENCIA Y DE AZAR.

Consideremos una variable aleatoria X continua no negativa con función de distribución $F_X(x)$ y una función de densidad $f_X(x)$.

Definición 1.1 *La función,*

$$S_X(x) = \Pr(X > x) = 1 - F_X(x) \tag{1.1}$$

Se denomina función de supervivencia

La función $S_X(x)$ recibe el nombre de función de supervivencia porque proporciona la probabilidad de que un individuo sobreviva más que x o que una variable de pérdida X exceda el valor x.

Definición 1.2 *Se denomina función de azar a la cantidad:*

$$r_X(x) = \frac{f_X(x)}{\Pr(X > x)} = \frac{f_X(x)}{1 - F_X(x)} \tag{1.2}$$

La función de azar recibe el nombre de tasa de fallo de fiabilidad y fuerza de mortalidad en seguros de vida. Aunque en estadística actuarial a menudo se trabaja con variables no negativas, la definición 1.2 es válida para este tipo de variable.

Supongamos que ahora X representa el tiempo de vida de un elemento. Entonces la función de azar se puede interpretar como la probabilidad de que el elemento sobreviva después del momento x. En efecto:

$$
\begin{aligned}
r_X(x) &= \lim_{\delta \to 0} \frac{\Pr(x < X \leq x + \delta / X > x)}{\delta} \\
&= \lim_{\delta \to 0} \frac{\dfrac{F_X(x+\delta) - F_X(x)}{\delta}}{1 - F_X(x)} = \frac{\dfrac{d}{d_x} F_X(x)}{1 - F_X(x)} \\
&= \frac{f_X(x)}{1 - F_X(x)}.
\end{aligned}
$$

Por tanto la función de azar se puede interpretar como la probabilidad instantánea de fallo dado que el elemento ha sobrevivido hasta el instante x.

A partir de la definición 1.1 se cumple que:

$$f_X(x) = r_X(x) S_X(x). \tag{1.3}$$

Es decir, la función de densidad es el producto de la función de azar por la función de supervivencia. Existe una correspondencia uno a uno entre función de distribución y función de supervivencia.

Teorema 1.1 *Sea X una variable aleatoria continua con función de azar $r_X(t)$, Entonces la función de distribución viene dada por,*

$$F_X(t) = 1 - \exp\left[- \int_{-\infty}^{x} r_X(t) dt \right]. \tag{1.4}$$

Demostración: para probar este resultado calculamos la derivada del logaritmo de la función de supervivencia:

$$\frac{d}{dx} \log[1 - F_X(x)] = \frac{-f_X(x)}{1 - F_x(x)} = -r_X(t),$$

e integrando desde $-\infty$ hasta x tenemos que:

$$\log[1 - F(x)] = -\int_{-\infty}^{x} r_X(t) dt,$$

A partir de la relación entre $F_X(x)$ y $r_X(x)$ dada en (1.4) se puede obtener unívocamente una a partir de la otra. Usando la expresión (1.1) y (1.4) se obtiene la relación entre función de supervivencia y función de azar:

$$S_X(x) = \exp\left[-\int_{-\infty}^{x} r_X(t)dt\right]. \qquad \blacksquare$$

1.4. VIDA RESIDUAL MEDIA

Sea X una variable aleatoria no negativa de tipo continuo de distribución $F(\cdot)$. Supongamos que la media $\mu_X = E(X)$ es finita y no nula. Se define a continuación la variable aleatoria vida residual.

Definición 1.3 *Se denomina vida residual de X y se denota por X_t, a la variable aleatoria que representa el tiempo de vida restante del individuo, sabiendo que ha sobre vivido hasta t, es decir,*

$$X_t = \{X - t / X > t\}, t > 0 \text{. Obviamente } X - t > 0$$

La función de distribución de la variable residual viene dada por:

$$F_{X_t}(x) = P_r(X_t \leq x) = P_r(X - t \leq x / X > t)$$

$$= \frac{P_r(t < X \leq x+t)}{\Pr(X > t)} = \frac{P_r(X > t) - P_r(X > x+t)}{P_r(X > t)}$$

$$= 1 - \frac{S(x+t)}{S(t)}, x \geq 0.$$

A continuación se define la vida residual media de X como la esperanza matemática de la vida residual.

Definición 1.4 Se denomina vida residual media de la variable X, y la denotaremos por $e_X(\cdot)$, a la esperanza matemática de la variable aleatoria vida residual.

$$e_X = E(X_t) = E(X - t / X > t) = \int_0^\infty [1 - F_{X_t}(x)]dx = \int_0^\infty \left[1 - \left(1 - \frac{S(x+t)}{S(x)}\right)\right]dx$$

$$= \int_0^\infty \frac{S(x+t)}{S(t)}dx = \frac{1}{S(t)}\int_0^\infty S(x+x)dx = \frac{1}{S(t)}\int_t^\infty S(x)dx, \qquad (1.5)$$

Para $t \geq 0$ siempre que $S(x) > 0$ y $e_X(t) = 0$ para los valores t tales que $S(t) = 0$.

La vida residual media permite estudiar el peso de la cola de la distribución. La función vida residual media viene determinada como prueba (1.5).

El siguiente resultado establece una relación entre la vida residual media y la función de azar.

Teorema 1. 2 *Sea X una variable aleatoria continua con vida residual media $e(x)$ y una función de supervivencia $S(x)$. Entonces, la función de azar viene dada por,*

$$r(x) = -\frac{S'(x)}{S(x)} = \frac{1+e'(x)}{e(x)} \tag{1.6}$$

Demostración: Se verifica que $e(x) = \frac{1}{S(x)}\int_x^\infty S(t)dt$

$$S(x).e(x) = \int_x^\infty S(t)dt\,,$$

Derivando se obtiene:

$$S(x)e'(x) + S'(x)e(x) = -S(x),$$

Dividiendo todo por $S(x)$,

$$\frac{S(x)e'(x)}{S(x)} + \frac{e(x)S'(x)}{S(x)} = \frac{-S(x)}{S(x)}$$

teniendo en cuenta que $r(x) = -\frac{S'(x)}{S(x)}$, obtenemos,

$$e'(x) - r(x)e(x) = -1$$

$$-r(x) = \frac{-1-e'(x)}{e(x)}$$

$$r(x) = \frac{1+e'(x)}{e(x)} \qquad \blacksquare$$

1.5 DISTRIBUCIÓN ESTACIONARIA DE RENOVACIÓN

La siguiente distribución de probabilidad juega un papel importante en la teoría de la ruina en tiempo continuo.

Definición 1.5 *Sea X una variable aleatoria continua no negativa. Se denomina variable aleatoria estacionaria de renovación X_e a la variable con función de densidad:*

$$f_{X_e}(x) = \frac{S_X(x)}{E(x)},\ \ x \geq 0 \tag{1.7}$$

La expresión (1.7) define una función de densidad genuina, puesto que si X es no negativa entonces:

$$E(X) = \int_0^\infty S_X(x)dx \Rightarrow 1 = \int_0^\infty \frac{S_X(x)}{E(X)}dx,$$

De donde se sigue el resultado. La función de supervivencia de X_e *es:*

$$S_{X_e}(x)=\int_x^{\infty} f_{X_e}(x)dx=\frac{\int_x^{\infty} S_X(t)dt}{E(X)},\ \ x\geq 0.$$

La función de azar de la distribución estacionaria es:

$$r_{X_e}(x)=\frac{f_{X_e}(x)}{S_{X_e}(x)}=\frac{S_X(x)}{\int_x^{\infty} S_X(t)dt}=\frac{1}{e_X(x)}.$$

Por lo tanto, la inversa de la función de azar de la distribución estacionaria es la vida residual media.

Por medio de estos resultados se puede probar el teorema de inversión de la vid residual media.

Teorema 1. 3 *Sea X una variable aleatoria no negativa con función de vida residual media $e_X(x)$. entonces, la función de distribución se puede obtener como:*

$$F_X(x)=1-\frac{e_X(0)}{e_X(x)}exp\left[-\int_0^x \frac{1}{e_X(t)}dt\right]$$

Demostración: Por el resultado (1.3) aplicado a la distribución estacionaria tenemos:

$$f_{X_e}(x)=r_{X_e}(x)S_{X_e}(x)=r_{X_e}(x)exp\left[-\int_0^x r_{X_e}(t)dt\right],$$

$$\frac{S_X(x)}{E(x)}=r_{X_e}\ exp\left[-\int_0^x r_{X_e}(t)dt\right]$$

$$S_X(x)=E(x).r_{X_e}\ exp\left[-\int_0^x r_{X_e}(t)dt\right]$$

lo que equivale:

$$S_X(x)=\frac{e_X(0)}{e_X(x)}exp\left[-\int_0^x \frac{1}{e_X(t)}dt\right],\quad \blacksquare$$

Puesto que $e_X(0)=E(X),$ de donde se obtiene el resultado.

1.6 TRANSFORMACIONES ÚTILES EN REASEGUROS

La variable X representa la cantidad reclamada, suponemos que X es no negativa y que tiene como función de distribución $F_X(x)$.

Transformación stop-loss

Definición 1.6 S*upongamos que la variable* X *representa una pérdida y por lo tanto es no negativa. Sea* d *un número real positivo que llamaremos retención. Se denomina transformación stop-loss de* X *a:*

$$(X-d)_+ = máx\{X-d,0\} = \begin{cases} X-d & si \;\; X>d, \\ 0 & si \;\; X \le d, \end{cases}$$

Se define a continuación la prima neta en un contrato de reaseguro de tipo stop-loss.

Definición 1.7 *Se define la cantidad* $\pi_X(d)$ *como la esperanza matemática de la variable stop-loss* $(X-d)_+$,

$$\pi_X(d) = E[(X-d)_+], \quad d \ge 0.$$

A $\pi_X(d)$ se le conocerá posteriormente prima neta del reaseguro. La existencia de π_X presupone que $\mu_X < \infty$. Obviamente si $d=0$, $\pi_X(0)=\mu_X$,

Si la variable aleatoria X es de tipo discreto con función de probabilidad

$p_X = Pr(X = x_k)$, $k = 0,1,2,...$ entonces:

$$\pi_X(d) = E[(X-d)_+] = \sum_{\{x_k > d\}} (x_k - d) p_k.$$

Si ahora X es continua con función de densidad $f_X(x)$, entonces:

$$\pi_X(d) = E[(X-d)_+] = \int_d^\infty (x-d) f_X(x) dx.$$

Si se quiere calcular los momentos de orden superior a uno de la variable $(X-d)_+$ se aplica las fórmulas usuales. Por ejemplo, si X es de tipo continuo,

$$E\{(X-d)_+^k\} = \int_d^\infty (x-d)^k f_X(x) dx.$$

Para el cálculo de $\pi_X(d)$ conviene tener en cuenta las siguientes fórmulas:

- Caso discreto:

$$\pi_X(d) = E(X-d)_+ = \mu_X - d - \sum_{\{x_K \le d\}} x_k p_k + d \sum_{\{x_k \le d\}} p_k.$$

- Caso continuo.

$$\pi_X(d) = E(X-d)_+ = \mu_X - d - \int_0^d x f_X(x) dx + d \int_0^d f_X(x) dx.$$

Una fórmula alternativa para el cálculo de $\pi_X(d)$ es:

$$\pi_X(d) = E[(X-d)_+] = \int_d^{\infty} [1-F_X(x)]dx, \;\; d \geq 0. \tag{1.8}$$

Finalmente, se puede observar que al derivar se obtiene:

$$\pi'_X(d) = F_X(d) - 1.$$

Por lo que en ocasiones puede ayudar al cálculo de prima neta.

1.7 VARIABLE LIMITE DE PÉRDIDA

Sea X la pérdida en que ha incurrido un asegurado, y supongamos que la póliza de su seguro sólo cubre hasta un límite u. Por tanto, la cantidad que paga la compañía al asegurado es,

$mín\{X, u\}$

$$mín\{X, u\}$$

donde $mín\{X, u\}$significa el menor valor de entre X y u.

Definición 1.8 *Se denomina variable límite de pérdida* X *a:*

$$mín\{X, u\} = \begin{cases} X & si \;\; X < u \\ u & si \;\; X \geq u \end{cases}$$

Se denomina valor límite de pérdida esperado a la esperanza matemática de la variable límite de pérdida, es decir $E(mín\{X, u\})$.

La variable límite de pérdida es en realidad una variable censurada por la derecha. Si se observa que,

$$X = mín\{X, u\} + (X - d_+),$$

entonces

$$E(X) = E(mín\{X, u\}) + E(X - d_+),$$

Esto significa que comprando una póliza con un límite de d (un seguro) y otra con un deducible de d, es equivalente a comprar toda la cobertura de riesgo. Para el cálculo del valor límite de pérdida esperado y en general para los momentos de $X \wedge u$ se usan las siguientes fórmulas,

- Caso discreto,

$$E[(mín\{X, u\})^k] = \sum_{x_i \leq u} x_i^k \, pr(X = x_i) + u^k [1 - F_X(u)].$$

- Caso continuo:

$$E\left[(mín\{X,u\})^k\right] = \int_{-\infty}^{u} x^k f_X(x)dx + u^k\left[1 - F_X(u)\right].$$

En particular, en el caso continuo:

$$E\left[(mín\{X,u\})\right] = \int_{-\infty}^{u} x f_X(x)dx + u\left[1 - F_X(u)\right]. \tag{1.9}$$

A partir de (1.9) se puede deducir una fórmula alternativa en términos de las funciones de distribución y supervivencia, que viene dada por:

$$E\left[(mín\{X,u\})\right] = \int_{-\infty}^{u} F_X(x)dx + \int_{0}^{u} S_X(x)dx. \tag{1.10}$$

1.8 ESTADÍSTICA BAYESIANA

La perspectiva clásica o frecuentista y la bayesiana están fundamentadas en diferentes nociones de probabilidad. De acuerdo con la perspectiva frecuentista, sólo los sucesos susceptibles de ser repetidos tienen probabilidad. En la perspectiva bayesiana, la probabilidad describe incertidumbre, en un sentido amplio del término.

Un suceso puede ser incierto por el hecho de ser intrínsecamente impredecible, es decir, por estar sujeto a variabilidad aleatoria, como por ejemplo la pérdida asociada a un siniestro. También puede ser incierto por el hecho de que tengamos un conocimiento imperfecto sobre el mismo, como por ejemplo, la respuesta media a un medicamento entre los pacientes de la población. La perspectiva frecuentista únicamente el primer tipo de incertidumbre, mientras que la aproximación bayesiana considera ambos tipos de incertidumbre.

Por lo tanto, desde una perspectiva frecuentista, la probabilidad se define como una frecuencia relativa de un suceso que se repite de un número elevado de veces. Por tanto, este método está basado firmemente en una definición de la probabilidad basada en conteo, razón por la cual se denomina método "frecuentista". La estadística bayesiana en cambio, está basada en la interpretación de la probabilidad como el grado personal de creencia.

Los métodos estadísticos están generalmente formulados con el objetivo de realizar inferencia sobre parámetros desconocidos. Por tanto, los parámetros representan valores desconocidos y generalmente corresponden con propiedades de la población en estudio. Cualquier cuestión de interés en el análisis puede ser expresada como una cuestión sobre el valor de estos parámetros. Los parámetros son específicos para cada problema, y generalmente no están sujetos a variabilidad aleatoria. Se tratará por tanto de valores fijos aunque desconocidos, por lo que no se puede asignar probabilidades a sus valores. En cambio, la estadística bayesiana sí permite asignar probabilidades a los parámetros por el simple hecho de que son desconocidos.

1.9 TEOREMA DE BAYES PARA EL CASO DISCRETO Y CONTINUO

El teorema de Bayes es uno de los resultados básicos de la teoría de la probabilidad. Se estudiará su interpretación como mecanismos de aprendizaje sobre estas cantidades desconocidas. En su versión más elemental se puede escribir como:

$$P_r(A/B) = \frac{P_r(A).P_r(B/A)}{P_r(B)},$$

Para dos sucesos A y B con $P_r(B) > 0$. El teorema de Bayes es interpretado como un mecanismo de aprendizaje sobre las cantidades de interés, puede ser aplicado reiteradas veces. Así las probabilidades a posteriori obtenidas en una fase pueden ser utilizadas como probabilidades a priori en la siguiente fase, y así sucesivamente.

El teorema de Bayes permite una interpretación nueva de las probabilidades: el hecho de que éstas son revisables cuando se combinan probabilidades iniciales con la información muestral proporcionada por los datos.

En general, el análisis estadístico de unos datos observados $\underline{x}$ suele comenzar con una evaluación descriptiva mediante la cual puede surgir algún modelo probabilístico $\{f(x/\theta); \theta \in \Theta\}$ que represente, para algún valor (desconocido) de θ, el mecanismo probabilístico que ha generado los datos $\underline{x}$ observados. El paradigma bayesiano establece que es necesario asignar una distribución a priori $\pi(\theta)$ sobre el espacio paramétrico Θ que describa el conocimiento disponible sobre el valor θ antes de haber observado los datos. Se sigue entonces que por la teoría de la probabilidad que, si el modelo de probabilidad es correcto, toda la información disponible sobre el valor de θ después de observar $\underline{x}$ estará contenida en la densidad a posteriori $\pi(\theta/\underline{x})$ obteniendo mediante:

$$P_r(\theta_i/\underline{x}) = \frac{f(\underline{x}/\theta_i)P_r(\theta_i)}{\sum_{j=1}^{n} f(\underline{x}/\theta_j)P_r(\theta_j)} \;\alpha\; f(\underline{x}/\theta_i)P_r(\theta_i), \quad i = 1,\ldots,n, \qquad (1.11)$$

O bien que $P_r(\theta_i/\underline{x})$ es proporcional a $f(\underline{x}/\theta_i)P_r(\theta_i)$ y lo representaremos por

$$P_r(\theta_i/\underline{x}) \;\alpha\; f(\underline{x}/\theta_i)P_r(\theta_i)$$

En el caso continuo se tiene:

$$\pi(\theta_i / \underline{x}) = \frac{f(\underline{x}/\theta)\pi(\theta)}{\int\limits_{\Theta} f(\underline{x}/\theta)\pi(\theta)d\theta} \;\; \alpha \;\; f(\underline{x}/\theta)\pi(\theta). \tag{1.12}$$

Las expresiones (1.11) y (1.12) constituyen el teorema de Bayes para el caso discreto y continuo, respectivamente.

Observados unos datos $\underline{x}$, la cantidad que aparece en el denominador (denominada distribución *predictiva*, ya sea en el caso continuo o discreto) es constante y por eso es común escribir el teorema de Bayes en su interpretación como actualización mediante:

Conocimiento a posteriori α conocimiento a priori $\times$ información muestral.

Al contrario que en la estadística clásica, los parámetros no son cantidades fijas, si no que se consideran variables aleatorias. Esto no es una descripción de su variabilidad (los parámetros pueden suponerse fijos pero desconocidos) sino que más bien es una descripción de la incertidumbre sobre su verdadero valor.

El paradigma bayesiano se basa en el aprendizaje. Así, la misión de los datos es añadir información a nuestros conocimientos y, de esta forma, actualizar nuestras creencias sobre los parámetros de interés y la hipótesis relevantes. Se debe por tanto especificar nuestras creencias anteriores al análisis de los datos, lo que se denomina como información a priori.

La información a priori expresa lo que se conoce, acerca de los parámetros de interés antes de observar los datos. Esta información es entonces combinada con los datos para producir la distribución a posteriori, que se expresa lo que se conoce de los parámetros de interés tras el análisis de los datos.

La forma más sencilla de expresar el teorema de Bayes es el siguiente:

"la distribución a posteriori es proporcional al producto de la distribución a priori por la verosimilitud"

Es decir, los nuevos juicios sobre el sistema se forman combinando los juicios iniciales (o a priori) con la información muestral (o verosimiltud).

En efecto vemos como nuestros nuevos juicios contienen la información de la distribución a priori y los datos. La estimación a posteriori (posterior) es, por tanto, un punto medio entre las creencias a priori y los datos. De esta forma obtiene estimaciones más precisas que cada información por separado. Esta es una de las ventajas del análisis bayesiano, la posibilidad de hacer uso de más información y obtener a sí resultados más precisos.

1.10 USO SECUENCIAL DEL TEOREMA DE BAYES

Uno de los aspectos sobresalientes del método bayesiano es la posibilidad de utilizar secuencialmente el teorema de Bayes, es decir, que *"la distribución a posteriori de hoy es la distribución a priori de mañana"*. En efecto, si se tiene una muestra inicial $\underline{x}$, se conoce que:

$$\pi(\theta / \underline{x}) \;\alpha\; \pi(\theta).L(\underline{x} / \theta).$$

Donde L denota la función de verosimilitud de los datos:

$$L(\underline{x} / \theta) = \prod_{i=1}^{n} f(x_i / \theta),$$

Siendo n el tamaño muestral. Esta notación se utilizará en todo lo que sigue.

Supongamos que se desea incorporar un conjunto de observaciones z independientes e igualmente distribuidas que las primeras observaciones.

Entonces:

$$\pi(\theta / \underline{x}, z) \;\alpha\; \pi(\theta), L(\underline{x}, z / \theta).$$

La independencia implica que:

$$f(y, z / \theta) = f(y / \theta).f(z / \theta)$$

Y por tanto,

$$L(\underline{x}, z / \theta) \;\alpha\; L(\underline{x} / \theta).L(z / \theta).$$

En consecuencia,

$$\pi(\theta / \underline{x}, z) \;\alpha\; \pi(\theta).L(\underline{x} / \theta).L(z / \theta) \;\alpha\; \pi(\theta / \underline{x}).L(z / \theta).$$

La distribución a posteriori de θ dados $\underline{x}$ y z se obtiene considerando la distribución a posteriori de θ dado $\underline{x}$ como la a priori para el proceso cuando se incorporan los datos z.

1.11 LA DISTRIBUCIÓN PREDICTIVA

El factor de proporcionalidad que convierte en igualdad el ajuste del juicio a posteriori mediante la verosimilitud y la a priori es la distribución marginal:

$$p(\underline{x}) = \int_{\Theta} L(\underline{x} / \theta)\pi(\theta)d\theta,$$

En su versión continua. En el caso discreto, basta sustituir el operador integral por medio del sumatoria.

Definición 1.9 *La distribución predictiva a priori es la distribución de los datos* $\underline{x}$ *para el modelo de verosimilitud dado por* $L(\underline{x}/\theta)$ *y la densidad a priori* $\pi(\theta)$, *definida por*

$$p(\underline{x}) = \int_{\Theta} L(\underline{x}/\theta)\pi(\theta)d\theta,$$

Análogamente, para el caso de un nuevo conjunto de observaciones z *independientes de las anteriores, se definirá la distribución predictiva (a posteriori) por*

$$p(z/\underline{x}) = \int_{\Theta} L(z/\theta)\pi(\theta/\underline{x})d\theta.$$

1.12 ESTIMACIÓN BAYESIANA PUNTUAL

En el análisis bayesiano son tres los elementos fundamentales. Por un lado el modelo generador de los datos que, en común con la estadística clásica viene dado por la expresión de la función de verosimilitud $L(\underline{x}/\theta)$. *Por otro lado, los juicios iniciales del investigador deben ser expresados en términos de una densidad a priori* $\pi(\theta)$ *que mejor refleje el conocimiento sobre el parámetro de interés. El teorema de Bayes permite obtener la densidad a posteriori,* $\pi(\theta/\underline{x})$, *siendo esta la función que mejor representa los datos observados.*

1.12.1 Estimación puntual

Habitualmente se tiene la necesidad de resumir el conocimiento que se tiene sobre el parámetro de interés (expresado a través de la distribución a posteriori $\pi(\theta/\underline{x})$) en un único valor, $\tilde{\theta}$. En tal caso se utilizan las medidas clásicas de localización: media, mediana, moda. En muchas ocasiones, la distribución a posteriori puede ser unimodal y simétrica y en consecuencia, cualquiera de las tres medidas resulta igualmente idónea como estimador puntual puesto que coinciden en valor. Sin embargo, en otras ocasiones la distribución a posteriori presenta cierta asimetría, lo que hace que la mediana sea el estimador puntual bayesiano preferido.

Si se admite que la distribución a posteriori del parámetro refleja todo el conocimiento disponible sobre él una vez que se ha observado los datos y se necesita un valor que concentre todo este conocimiento, parece ser entonces que se debe utilizar la moda a posteriori como dicho estimador bayesiano, ya que la moda es el valor donde se hace máxima dicha distribución a posteriori y en cierto sentido donde existe mayor plausibilidad de ocurrencia.

La moda también es conocida como *estimador bayesiano de máxima verosimilitud.*

1.13 INTERVALOS BAYESIANOS DE CREDIBILIDAD

Junto con las medidas de localización es habitual en inferencia estadística aportar en torno a los intervalos que contienen al parámetro con cierta probabilidad. Se construirá ahora los intervalos bayesianos de credibilidad como alternativa a los intervalos de confianza clásicos. En ocasiones se llamará también intervalos bayesianos a dichos intervalos.

Definición 1.10 *Los puntos* a_1 *y* a_2 *definen un intervalo bayesiano de credibilidad con probabilidad* $100(1-\alpha)\%$, *si se verifica que*

$$\Pr(a_1 \leq \theta \leq a_2 / x) = \int_{a_1}^{a_2} \pi(\theta / x) d\theta \geq 1-\alpha$$

Se ha definido el intervalo de credibilidad utilizando la distribución a posteriori. Es posible definir un intervalo (a priori) de credibilidad sin más que sustituir la densidad a posteriori por la a priori.

Una de las principales ventajas de los intervalos bayesianos de credibilidad es que se puede interpretar en términos de probabilidad. Así por ejemplo, se encuentra que la $\Pr(a_1 \leq \theta \leq a_2 / x) = 0.95$ se puede decir que (una vez observados los datos $\underline{x}$) con probabilidad 0.95 el intervalo $[a_1, a_2]$ contiene al verdadero valor del parámetro.

Esta forma de definir los intervalos de credibilidad no asegura su unicidad y básicamente se pueden presentar varias situaciones:

1. Intervalos de una cola. En esta situación se estaría hablando de los cuantiles de la distribución a priori. Por ejemplo, para el caso continuo uniparamétrico con $\Theta = (-\infty, \infty)$, un intervalo bayesiano de credibilidad al $100(1-\alpha)\%$ (de una cola) nos lo proporciona cualquier intervalo (θ^*, ∞) que cumpla

$$\Pr(\theta > \theta^* / \underline{x}) = \int_{\theta^*}^{\infty} \pi(\theta / \underline{x}) d\theta = 1-\alpha.$$

2. Intervalos de dos colas con igual área. Este es el caso habitual en el que se considera cualquier intervalo (θ_*, θ^*) que cumple que $\Pr(\theta_* \leq \theta \leq \theta^* / \underline{x}) = 1-\alpha$ y cada extremo del intervalo verifica que $\Pr(\theta_* < \theta / \underline{x}) = \frac{\alpha}{2}$ y $\Pr(\theta^* < \theta / \underline{x}) = \frac{\alpha}{2}$.

3. Intervalos de alta densidad a posteriori (en notación, HPD o también HDI). Cuando la distribución es asimétrica puede ocurrir que, los intervalos con extremos que tienen igual área no sea el intervalo de menor longitud.

Definición 1.11 *Un conjunto* C *se dice que es un conjunto de credibilidad de alta densidad con probabilidad* $100(1-\alpha)\%$, *en notación,* $100(1-\alpha)\%$, *HPD si verifica:*

1. $\Pr(\theta \in C/\underline{x}) = \int_C \pi(\theta/\underline{x})d\theta = 1-\alpha.$
2. *Dado cualquier otro conjunto* C_2 *que cumpla la condición anterior, para cualquier* $\theta_1 \in C$ *y cualquier* $\theta_2 \in C_2$ *se cumple que* $\pi(\theta_1/\underline{x}) \geq \pi(\theta_2/\underline{x})$.

Un intervalo HPD es aquel intervalo que alcanzando una probabilidad posteriori dada tiene menor longitud. En ocasiones, un conjunto creíble HPD puede estar formado por la unión de dos intervalos. El siguiente resultado señala condiciones para asegurar la coincidencia entre un intervalo de dos colas y el HPD.

Proposición 1.1 *Si la densidad a posteriori* $\pi(\theta/\underline{x})$ *es unimodal y continua. Entonces el intervalo bayesiano de credibilidad con probabilidad* $1-\alpha$ *con menor longitud* $l = b - a$ es la única solución de:

$$\int_a^b \pi(\theta/\underline{x})d\theta = 1-\alpha,$$

$$\pi(a/\underline{x}) \quad = \quad \pi(b/\underline{x}).$$

Este intervalo además coincide con el PHD.

1.14 TEST DE HIPÓTESIS BAYESIANOS

La teoría estadística de los contrastes de hipótesis también admite una formulación bayesiana.

La forma más simple de plantear un test de hipótesis para un parámetro desconocido $\theta \in \Theta$ puede ser la siguiente.

Supongamos que el espacio paramétrico Θ está participando en dos conjuntos Θ_0 y Θ_1 tales que: $\Theta_0 \cup \Theta_1 = \Theta$ y $\Theta_0 \cap \Theta_1 = \emptyset$.

Normalmente el interés se centra en conocer si se puede considerar que: $H_0 : \theta \in \Theta_0$ (denominada *hipótesis nula*) es cierta (en probabilidad) o por el contrario lo es $H_1 : \theta \in \Theta_1$ (*hipótesis alternativa*). Para ello se usará un conjunto de observaciones muestreadas todas ellas de manera independiente y de la misma distribución $f(x/\theta)$.

Para el contraste bayesiano se necesita disponer de las probabilidades a posteriori de cada una de las hipótesis:

$p_0 : P_r(H_0 \ cierta / x) = P_r(\theta \in \Theta_0 / x), \qquad p_1 : P_r(H_1 \ cierta / x) = P_r(\theta \in \Theta_1 / x)$

Y decidir entre $H_0 \ y \ H_1$.

Para ello se necesita disponer de las probabilidades a priori de cada una de las hipótesis:

$\pi_0 = P_r(H_0 \ cierta) = P_r(\theta \in \Theta_0), \qquad \pi_1 = P_r(H_1 \ cierta) = P_r(\theta \in \Theta_1)$

Definición 1.12 *Se define el "odds" a priori de* H_0 *frente a* H_1 *como el cociente de* $\frac{\pi_0}{\pi_1}$.

Análogamente se define el "odds" a posteriori de H_0 *frente a* H_1 *como el cociente de* $\frac{p_0}{p_1}$.

El "odds" a priori necesita ser asignado por el investigador mientras que el "odds" a posteriori se obtiene mediante el mecanismo de revisión de juicios que es el teorema de Bayes.
Valores del "odds" a priori (a posteriori) próximos a 1 indicarán que H_0 es igualmente probable que H_1 a priori (a posteriori) y valores marcadamente mayores que 1 indicarán evidencia a priori (a posteriori) a favor de H_0 frente a H_1. Inversamente para valores menores que 1.

Definición 1.13 *Se denominará factor Bayes a favor de* H_0 *frente a* H_1 *al cociente*

$$B_{01} = \frac{p_0 / p_1}{\pi_0 / \pi_1} = \frac{p_0 \pi_1}{p_1 \pi_0}.$$

Luego, se deduce que:

$$\frac{p_0}{p_1} = B_{01} \frac{\pi_0}{\pi_1},$$

Por lo tanto, el factor Bayes es la B_{01} es la cantidad que multiplicada por el "odds" a priori proporciona el "odds" a posteriori.

HIPÓTESIS NULA SIMPLE FRENTE A ALTERNATIVA SIMPLE.

El problema es el siguiente:

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_1 : \theta = \theta_1.$$

En tal caso, $\pi_i = P_r(\theta = \theta_i), \ i = 0,1,$ deben ser asignadas por el investigador y para un modelo de verosimilitud dado, $f(x / \theta_i), \ i = 0,1.$

De esta forma se obtiene que el “odds” a posteriori se deduce de

$$\frac{p_0}{p_1} = \frac{f(x/\theta_0)}{f(x/\theta_1)} \cdot \frac{\pi_0}{\pi_1},$$

Es decir, para este caso el factor Bayes a favor de H_0 frente a H_1 es:

$$B_{01} = \frac{f(x/\theta_0)}{f(x/\theta_1)}.$$

El factor Bayes coincide con el cociente de verosimilitudes de H_0 frente a H_1 y por tanto puede ser interpretado como la evidencia que única y exclusivamente aportan los datos $\mathbf{X}$ a favor de H_0 frente a H_1.

El factor Bayes se interpreta como la evidencia que sólo los datos aportan a la hipótesis nula frente a la alternativa.

Se observa que para este caso (y para cualquier otro en el que $\pi_0 + \pi_1 = p_0 + p_1 = 1$), se tiene que:

$$p_0 = p_1 B_{01} \frac{\pi_0}{\pi_1} \Leftrightarrow p_0 = (1 - p_0) B_{01} \frac{\pi_0}{1 - \pi_0},$$

Es decir,

$$p_0 = \frac{B_{01}}{B_{01} + \frac{1 - \pi_0}{\pi_0}} \Leftrightarrow p_0 = \frac{1}{1 + \frac{1 - \pi_0}{\pi_0} \frac{1}{B_{01}}},$$

Si se tiene el factor de Bayes “controlado” es posible obtener la probabilidad a posteriori de la hipótesis nula sin más que conocer su probabilidad a priori.

HIPÓTESIS NULA COMPUESTA FRENTE A ALTERNATIVA COMPUESTA.

Las situaciones a estudiar son aquellas que puedan formularse de la siguiente forma:

$$H_0 : \theta \in \theta_0 \quad \text{vs} \quad H_1 : \theta \in \theta_1.$$

Esta situación contempla tanto los test unilaterales: $H_0 : \theta < \theta_0$ vs $H_1 : \theta \geq \theta_0$ (en cualquiera de sus posibles formulaciones) como los contrastes del tipo: $H_0 : \theta \in (\theta_0, \theta_1)$ vs $H_0 : \theta \notin (\theta_0, \theta_1)$, pudiendo ser $\theta_0 = -\infty$ y $\theta_1 = +\infty$.

Asignada una densidad a priori para θ sobre el espacio paramétrico Θ, cada una de las hipótesis tendrá unas probabilidades a priori de ser ciertas obtenidas mediante

$$\pi_0 = \int_{\Theta_0} \pi(\theta) d\theta, \qquad \pi_1 = \int_{\Theta_1} \pi(\theta) d\theta,$$

Y observados los datos $\mathbf{X}$ con verosimilitud $f(x/\theta)$, cada una de las probabilidades a priori se trasformarán en a posteriori mediante el teorema de Bayes $(\pi(\theta/x) \;\alpha\; f(x/\theta)\pi(\theta))$:

$$p_0 = P_r(H_0 \; cierta/x) = P_r(\theta \in \Theta_0 / x) = \int_{\Theta_0} \pi(\theta/x) d\theta,$$

$$p_1 = P_r(H_1 \; cierta/x) = P_r(\theta \in \Theta_1 / x) = \int_{\Theta_1} \pi(\theta/x) d\theta,$$

HIPÓTESIS NULA SIMPLE FRENTE ALTERNATIVA COMPUESTA.

Se construirá el procedimiento para el contraste de hipótesis del tipo:

$$H_0 : \theta = \theta_0 \text{ vs } H_1 : \theta \neq \theta_0.$$

En este caso la densidad a priori sobre las hipótesis estará definida como una mixtura de distribuciones discreta y continua, respectivamente. Es decir:

$$\pi(\theta) = \begin{cases} \pi_0, & si \;\; \theta = \theta_0 \\ (1-\pi_0)\pi_1(\theta), & si \;\; \theta \neq \theta_0 \end{cases}$$

Siendo $\pi_1(\theta)$ una densidad sobre el espacio $\Theta_1 = \Theta - \{\theta_0\}$. La distribución predictiva necesaria para el cálculo de las probabilidades a posteriori se obtiene por:

$$p(x) = \pi_0 f(x/\theta_0) + (1-\pi_0) p_1(x),$$

Siendo $p_1(x) = \int_{\{\theta \neq \theta_0\}} \pi_1(\theta) f(x/\theta_0) d\theta.$

El cálculo de las probabilidades a posteriori es inmediato:

$$p_0 = P_r(H_0 \; cierta/x) = \frac{\pi_0 f(x/\theta_0)}{p(x)},$$

$$p_1 = P_r(H_1 \; cierta/x) = \frac{(1-\pi_0) p_1(x)}{p(x)}$$

De donde se deduce que el factor Bayes es:

$$B_{01} = \frac{p_0 / p_1}{\pi_0 / \pi_1} = \frac{f(x/\theta_0)}{p_1(x)}.$$

1.15 INFERENCIA BAYESIANA Y TEORÍA DE LA DECISIÓN

El problema de inferencia bayesiana puede ser visto de forma general como un problema de decisión. La función de pérdidas forma parte de este problema.

Definición 1.14 *Una función de pérdidas* $L(\widetilde{\theta},\theta)$ *describe la pérdida en la que incurre el investigador cuando utiliza* $\widetilde{\theta}$ *como estimador de* θ, *el verdadero valor del parámetro.*

Definición 1.15 *El estimador bayesiano de un parámetro para una función de pérdidas dada, es aquel que minimiza la pérdida esperada a posteriori.*

El estimador bayesiano de cada situación dependerá de la función de pérdidas elegida.
Un problema estadístico desde el punto de vista de la teoría de la decisión está compuesta por:

1. Un espacio de estados de la naturaleza o espacio paramétrico Θ.
2. Un conjunto de acciones posibles llamadas decisiones, y que se denotan por D.
3. Una función de pérdida:
$$L:\Theta\times D\rightarrow\Re^{+}.$$
4. Un modelo estadístico $(\mathcal{X},\mathcal{A},\mathcal{P})$ donde se puede observar una variable aleatoria cuya distribución depende de los estados de la naturaleza.

Valores "grandes" de $L(\theta,d)$ indican que d es más incorrecto, mientras que valores "pequeños" indican que d es más correcto. En una función de utilidad se habla de ganancias" y no de pérdidas, por lo que la utilidad puede interpretarse simplemente por $-L(\theta,d)$.

Algunas de las propiedades de las funciones de pérdidas son:

1. $L(\theta,d)\geq 0.$
2. $L(\theta,\theta)=0$
3. $L(\theta,d)\leq L(\theta,d^{'}),\ \ si\ \ \left|\theta-d\right|\leq\left|\theta-d^{'}\right|.$

En general en la teoría de la estimación, el espacio de decisiones suele ser un sub espacio de $\Re^{+}$. Una función de decisión no aleatorizada es una aplicación
$$d:\mathcal{X}\rightarrow D,$$
Que a cada observación se le asocia una decisión, y la función de riesgo es:
$$R(\theta,d)=E\left[L(\theta,d(X))\right].$$

El problema central consiste en encontrar funciones de decisión adecuadas. Se dice que d_1 es preferible a d_2 si:

$$R(\theta, d_1) \leq R(\theta, d_2), \forall \theta \in \Theta.$$

En el caso de que se haga variar ambas variables (θ, d) entonces la desigualdad anterior es sólo pre-orden parcial, en el sentido de que puede haber decisiones no comparables, es decir,

$$R(\theta_1, d_1) < R(\theta_1, d_2) \quad \text{y} \quad R(\theta_2, d_1) < R(\theta_2, d_2).$$

Por tanto, no hay decisiones óptimas.

Elegida una función de pérdidas, el estimador Bayes (llamada también regla de Bayes) para una distribución a priori π, es la cantidad que minimiza la expresión.

$$r(x, d) = \int_{\Theta} L(\theta, d)\pi(\theta / x)d\theta,$$

Para $d \in D.$

1.16 FUNCIONES DE PÉRDIDAS EN ESTADÍSTICA ACTUARIAL

Definición 1.16 (Pérdidas absolutas) *La función de pérdida absoluta está definida por la expresión:*

$$L(\theta, d) = |\theta - d|.$$

Definición 1.17 (Pérdidas 0 - 1) *La función de pérdida* 0 – 1 *está definida por la expresión:*

$$L(\theta, d) = \begin{cases} 0, & si \;\; \theta = d, \\ 1, & si \;\; \theta \neq d \end{cases}$$

Definición 1.18 (Pérdidas bilineales) *para a y b dos constantes positivas, se define la función de pérdidas bilineal como:*

$$L(\theta, d) = \begin{cases} a(\theta - d) & si \;\; d \leq \theta, \\ b(d - \theta) & si \;\; d \geq \theta. \end{cases}$$

Definición 1.19 (Pérdidas cuadráticas ponderadas) *una función de pérdidas cuadráticas ponderadas viene definida por la expresión* $L(d, \theta) = w(\theta)(\theta - d)^2$, *siendo* $w(\theta)$ *una función positiva de* θ.

Un caso particular de las pérdidas cuadráticas ponderadas lo constituye el caso de pérdidas cuadráticas sin más que considerar $w(\theta)=1$. La siguiente función de pérdidas utilizada en problemas de estimación:

$$L(\theta,d)=\begin{cases}(\theta-d)^2 & si\ \ d\le\theta,\\ K(d-\theta)^2 & si\ \ d\ge\theta.\end{cases}$$

Con $K>1$, penaliza la sobre-estimación, más que la sub-estimación. En ocasiones también se puede considerar la función

$$L(\theta,d)=\frac{(d-\theta)^2}{|\theta|+1},$$

Donde se penalizan los errores en la estimación más cuando θ está cerca de cero, que cuando $|\theta|$ es grande.

Definición 1.20 (Pérdidas LINEX) *La función LINEX* (LINear EXponential), *de carácter asimétrico propuesta por Varian (1975) y usada por Zellner (1986), viene definida por:*

$$L(\theta,d)=\exp\{c(d-\theta)\}-c(d-\theta)-1,$$

Definición 1.21 (Pérdidas exponenciales) *se hablará de pérdidas exponenciales cuando:*

$$L(\theta,d)=[\exp(c\theta)-\exp(cd)]^2,$$

1.17. ESPECIFICACIÓN DE DENSIDADES A PRIORI.

El establecimiento de algunas hipótesis es el paso previo en el proceso de especificación de la distribución a priori. Se puede explorar el rango de posibles especificaciones a priori que interpreten de forma razonable la evidencia a priori, permitiendo así cierta imprecisión. Si la decisión a posteriori es insensible a las diferentes distribuciones a priori se puede afirmar que el grado de imprecisión admitida no es preocupante. Esta práctica, también conocida como análisis de sensibilidad con respecto a la distribución a priori, es muy común en la práctica de la metodología bayesiana.

La precisión necesaria en la especificación depende de la evidencia de los nuevos datos. Así, si se cuenta una nueva base de datos de gran tamaño, el efecto de la distribución a priori sobre las conclusiones finales será relativamente pequeño por lo que se puede permitir un menor grado de imprecisión. Si los nuevos datos no tienen un peso tan relevante, la información a priori tendrá un efecto apreciable, por lo que el análisis de sensibilidad es esencial.

Se puede distinguir diferentes tipos de distribuciones a priori:

- *Informativa:* representa el conocimiento a priori acerca de los parámetros de interés.

El proceso mediante el cual se asigna los conocimientos a priori a distribuciones de probabilidad se denomina: *elícitación o asignación.*
Para el caso informativo, se debe distinguir claramente el espacio parámetrico discreto del continuo.
El caso discreto, $\Theta = \{\theta_1, \theta_2, ..., \theta_n\}$ se resuelve de manera inmediata puesto que el investigador debe asignar, en función de su conocimiento, las probabilidades a priori $\pi_i = \Pr(\theta = \theta_i),\ i = 1,...,n.$ Para el caso continuo, $\Theta \subset \Re^k\ (k \geq 1)$ el problema de construir una densidad a priori $\pi(\theta)$ es más complicado.
Algunas aproximaciones consisten en discretizar el problema, como es el caso de la *aproximación del histograma.* Cuando Θ es un intervalo de la recta real, esta aproximación consiste en dividir el espacio parámetrico en intervalos, de tal manera que el investigador determina la probabilidad subjetiva de cada intervalo y entonces se construye el histograma de probabilidad. Con este histograma se debe entonces "aproximar" una densidad $\pi(\theta)$. Esta aproximación presenta algunos inconvenientes: ¿Cuántos intervalos se debe considerar? ¿De que tamaño debe ser cada intervalo?, no se conoce el comportamiento de las colas d las densidades así construidas, etc.
El método más utilizado consiste en suponer que la densidad a priori $\pi(\theta)$tiene una forma funcional dada (normal, beta,...) y se trata de elegir aquella densidad que con esta forma funcional mejor represente a los juicios del experto.
Por ejemplo si se supone que el parámetro $\theta \in \Re$ y se supone que la densidad a priori de θ tiene forma funcional $\mathcal{N}(\mu, \sigma^2)$ entonces se necesitará asignar valores de la media y la varianza a priori que mejor representen los juicios del experto. De a misma manera, para $\theta \in (0,1)$ con una forma funcional $\mathcal{Be}(\alpha, \beta)$uno puede asignar la media y la varianza a priori $\mu\ y\ \sigma^2$ y entonces obtener los valores a priori de α y β mediante las relaciones:

$$\mu = \frac{\alpha}{\alpha + \beta}, \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$$

Es decir, suponer una forma funcional bastará con asignar un conjunto de momentos de la densidad a priori para tenerla totalmente determinada. Este método presenta la dificultad de que las colas de una densidad pueden tener un efecto drástico en sus momentos. Por ejemplo, si se elige una forma funcional para $(\theta \in (0,+\infty))$ del tipo $c\theta^{-2}$, esta densidad no tiene momentos.

Una alternativa a este método que produce mejores resultados y además resulta más intuitivo de asignar a los expertos consiste en estimar subjetivamente varios cuantiles de la distribución a priori y entonces elegir los parámetros de una forma funcional dada que tenga esos cuantiles.

1.17.1 FAMILIAS CONJUGADAS

En ocasiones la distribución a priori está determinada por el propio estudio. En ese caso es necesario conocer el valor a priori para el experto de unos pocos parámetros. Aunque la elección de la distribución a priori de esta forma es arbitraria y discutible, lo cierto es que cualquier otra distribución que recoja de forma fiel los conocimientos del experto es probable que sea muy similar la fijada de antemano, por lo que daría lugar a estimaciones parecidas. Resulta por tanto útil realizar la elección de la distribución a priori siguiendo los criterios de simplicidad y conveniencia.

Matemáticamente, en algunos problemas estadísticos sencillos existe una clase de distribuciones a priori conocidas como *distribuciones conjugadas* que son particularmente recomendables.

En primer lugar, *porque si la distribución a priori es de la familia conjugada, la distribución a posteriori será también de la misma familia.*

En segundo lugar, *Las distribuciones conjugadas son lo suficientemente sencillas como para permitir realizar gran cantidad de inferencia sin necesidad de métodos computacionales complejos Su asignación correspondería con la aproximación de una forma funcional dad pero especialmente bien comportada en el cálculo de la distribución a posteriori.*

Si para un determinado problema existe una familia a priori conjugada que ajuste de forma correcta la distribución a priori, entonces debemos utilizar el análisis conjugado.

Definición 1.21 (Familia conjugada) *Supongamos que el método que genera los datos* x *viene determinado por la distribución* $f(x/\theta)$. *Una familia de densidades a priori* $\mathcal{F}$ *para el parámetro* θ, *se dice que la conjugada para el muestreo por* $f(x/\theta)$. *si para cualquier densidad a priori*

$\pi(\theta) \in \mathcal{F}$ *se verifica que la densidad a posteriori* $\pi(\theta / \mathrm{x}) \ \alpha \ L(\mathrm{x}/\theta)\pi(\theta)$ *es también una densidad de la familia* $\mathcal{F}$.

Verosimilitud Distribución a priori	Distribución a posteriori
$X \sim \mathcal{P}(\theta)$ $\theta \sim \mathcal{G}\left(a, \frac{1}{b}\right)$	$\mathcal{G}\left(a + n\bar{x}, \frac{1}{(b+n)}\right)$
$X \sim \mathcal{P}(\theta)$ $\theta \sim \mathcal{IG}(\mu, \beta)$	$\mathcal{GIG}\left(n\bar{x} - \frac{1}{2}, \mu\sqrt{\frac{1}{2\beta n+1}}, \frac{\beta}{2\beta n+1}\right)$
$X \sim \mathcal{BN}(r, \theta)$ $\theta \sim \mathcal{Be}(a, b)$	$\mathcal{B}(a + nr, b + n\bar{x})$
$X \sim \mathcal{B}(m, \theta)$ $\theta \sim \mathcal{Be}(a, b)$	$\mathcal{B}(a + n\bar{x}, b + mn - n\bar{x})$
$X \sim \mathcal{G}\left(\alpha, \frac{1}{\theta}\right)$ $\theta \sim \mathcal{G}\left(a, \frac{1}{b}\right)$	$\mathcal{G}\left(n\alpha + a, \frac{1}{(n\bar{x}+b)}\right)$
$X \sim \mathcal{N}(\theta, \sigma^2)$ $\theta \sim \mathcal{N}(a, r^2)$	$\mathcal{N}\left(\frac{a\sigma^2 + n\bar{x}r^2}{\sigma^2 + tr^2}, \frac{\sigma^2 r^2}{\sigma^2 + tr^2}\right)$

TABLA 1. *Distribuciones a priori conjugadas*

La tabla 1 recoge, a modo de resumen algunas distribuciones a priori conjugadas respecto a una verosimilitud dada. Se puede observar que todas las densidades a posteriori pueden ser expresadas en función del estadístico suficiente (si existe) en el muestreo. En efecto esta es una propiedad general de la densidad a posteriori sea o no conjugada, ya que si **t** es un estadístico suficiente para una población $f(x/\theta)$, *por el teorema de factorización ocurre que la función de de verosimilitud cumple que:*

$$L(\mathrm{x}/\theta) \ \alpha \ L(t/\theta),$$

Y por lo tanto la densidad a posteriori de θ dado x es la misma que dado **t** ya que

$$L(\theta/\mathrm{x}) \ \alpha \ L(\mathrm{x}/\theta)\pi(\theta) \ \alpha \ L(t/\theta)\pi(\theta) \ \alpha \ \pi(\theta/t).$$

1.18 ANÁLISIS BAYESIANO PARA DATOS NORMALES

Otra de las situaciones más frecuentes en la práctica actuarial es aquella en la que los datos procedentes de una población normal.

1.18.1 CASO DE MEDIA DESCONOCIDA Y VARIANZA CONOCIDA

Consideremos ahora una primera situación en la que la varianza σ^2 es conocida y que por tanto, el único parámetro desconocido será la media μ, sobre la que se desea hacer inferencia. La función de verosimilitud viene dada por:

$$L(\mathrm{x}/\mu) \ \alpha \ \exp\left\{-\frac{1}{2}\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}\right\}, \qquad (1.13)$$

Donde se ha prescindido en términos de proporcionalidad de la parte conocida, El único parámetro a estimar en este modelo será la media de distribución normal.

Consideremos que para este caso, una densidad a priori para μ del tipo $\mathcal{N}(\mu_o,\sigma_0^2)$:

$$\mu \sim \mathcal{N}(\mu_o,\sigma_0^2) \qquad (1.14)$$

Con μ_o y σ_0^2 constantes conocidas.

Teorema 1.4. *Para el caso de verosimilitud una normal con varianza σ^2 conocida y con función de densidad a priori $\pi(\mu)$ de tipo $\mathcal{N}(\mu_o,\sigma_0^2)$,* se verifica que la densidad a posteriori es de nuevo normal con parámetros a posteriori dados por las fórmulas:

$$E(\mu/\mathrm{x}) = \mu_0 . \frac{(\sigma_0^2)^{-1}}{(\sigma_0^2)^{-1}+(\sigma^2/n)^{-1}} + \bar{x}.\frac{(\sigma_0^2)^{-1}}{(\sigma_0^2)^{-1}+(\sigma^2/n)^{-1}}, \qquad (1.15)$$

$$Var(\mu/\mathrm{x}) = \frac{1}{(\sigma_0^2)^{-1}+n(\sigma^2)^{-1}}. \qquad (1.16)$$

Además la distribución predictiva de una observación futura es de nuevo de tipo normal.

Demostración: Multiplicando la función de verosimilitud (1.13) por la función de densidad (1.14), se obtiene:

$$\pi(\mu/\mathrm{x}) \;\alpha\; \exp\left\{-\frac{1}{2}\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{\sigma^2}\right\}.\exp\left\{-\frac{1}{2}\frac{(\mu-\mu_0)^2}{\sigma_0^2}\right\}$$

$$\alpha\; \exp\left\{-\frac{1}{2}\left(\frac{\sum_{i=1}^{n}x_i^2-2\mu\sum_{i=1}^{n}x_i+n\mu^2}{\sigma^2}+\frac{\mu^2-2\mu\mu_0+\mu_0^2}{\sigma_0^2}\right)\right\}$$

$$\alpha\; \exp\left\{-\frac{1}{2}\left[\mu^2\left(\frac{n}{\sigma^2}+\frac{1}{\sigma_0^2}\right)+\mu\left(\frac{n\overline{x}}{\sigma}+\frac{\mu_0}{\sigma_0}\right)\right]\right\}$$

$$\alpha\; \exp\left\{-\frac{1}{2}\frac{(\mu-\mu_1)^2}{\sigma_1^2}\right\},$$

Donde μ_1 y σ_1^2 vienen definidos por (1.15) y (1.16), respectivamente. Se puede observar que $\overline{X} \sim \mathcal{N}\left(\mu,\sigma^2/n\right)$ y por tanto para obtener la distribución predictiva de una nueva observación y sólo se tiene que considerar que $\mathrm{y}=(\mathrm{y}-\mu)+\mu$ y teniendo en cuenta que ambos sumandos son independientes y que

$$\mathrm{y}-\mu \sim \mathcal{N}\left(0,\sigma^2\right),$$
$$\mu \sim \mathcal{N}\left(\mu_1,\sigma_1^2\right)$$

Se tiene que $\mathrm{y} \sim \mathcal{N}\left(\mu_1,\sigma^2+\sigma_1^2\right)$. ■

Si se escribe la relación (1.16) en términos de la precisión (inversa de la varianza) se tiene que la precisión a posteriori verifica la relación:

$$\tau_1=\tau_0+n\tau,$$

Y por tanto la precisión a posteriori es la suma de las precisiones a priori más n veces la precisión de los datos, que se supone conocida.

Para la media a posteriori también podemos deducir que:

$$E(\mu/\overline{x})=\mu_1=\mu_0.\frac{\tau_0}{\tau_0+n\tau}+\overline{x}.\frac{\tau_1}{\tau_0+n\tau},$$

Es decir, la esperanza a posteriori se puede expresar como una media ponderada de la media a priori y la media muestral. La familia de distribuciones a priori normal bajo muestreo también normal (en el caso de la varianza conocida) es una familia conjugada.

1.18.2. CASO DE MEDIA CONOCIDA Y VARIANZA DESCONOCIDA.

Analizamos ahora el caso en que la media $\mu = \mu_0$ es conocida y σ^2 es el parámetro desconocida de esta situación sobre el que se necesita hacer inferencia.

La función de verosimilitud viene dada por:

$$L\left(\underline{x}/\sigma^2\right) \ \alpha \ \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2} \frac{\sum_{i=1}^{n}\left(x_i - \mu_0\right)^2}{\sigma^2} \right) = \left(\sigma^2\right)^{-n/2} \exp\left(-\frac{1}{2} \frac{S_{\mu_0}}{\sigma^2} \right),$$

Donde $S_{\mu_0} = \sum_{i=1}^{n}\left(x_i - \mu_0\right)^2$. Como distribución a priori del parámetro σ^2 se elige una distribución chi-cuadrado inversa de parámetros S_0 y v_0 (grados de libertad) cuya función de densidad es la siguiente:

$$\pi\left(\sigma^2\right) \ \alpha \ \left(\sigma^2\right)^{-v_0/2 - 1} \exp\left(-\frac{S_0}{2\sigma^2} \right). \qquad (1.17)$$

La media, la varianza y la moda de (1.17) vienen dados respectivamente por las fórmulas:

$$E\left(\sigma^2\right) = \frac{S_0}{v_0 - 2}, \ \ si \ \ v_0 > 2,$$

$$Var\left(\sigma^2\right) = \frac{2S_0^2}{\left(v_0 - 2\right)^2\left(v_0 - 4\right)}, \ \ si \ \ v_0 > 4,$$

$$Moda\left(\sigma^2\right) = \frac{S_0}{v_0 + 2}.$$

Se denota $\sigma^2 \sim$ ✠$^{-2}\left(S_0, v_0\right)$ y es una distribución a priori conjugada. La distribución a para σ^2 viene dada por:

$$\sigma^2/\underline{x} \sim ✠^{-2}\left(S_0 + S_{\mu_0}, v_0 + n\right).$$

1.18.3. CASO DE MEDIA Y VARIANZA DESCONOCIDA

Este caso es donde la media y la varianza son parámetros desconocidos. La función de verosimilitud es:

$$L\left(\mathrm{x}/\mu, \sigma^2\right) \ \alpha \ \left(\sigma^2\right)^{-n/2} \exp\left\{ -\frac{1}{2} \frac{\left[S + n\left(\overline{x} - \mu\right)^2\right]}{\sigma^2} \right\},$$

Siendo $S = \sum_{i=1}^{n} (x_i - \bar{x})^2$. la especificación de las distribuciones a priori es la siguiente:

$$\mu / \sigma^2 \sim \mathcal{N}\left(\mu_0, \sigma^2 / n_0\right),$$
$$\sigma^2 \sim X^{-2}\left(s_0, v_0\right)$$

Los casos anteriores pueden considerarse casos particulares de esta situación puesto que coinciden con ella bajo el supuesto de que ambos parámetros son independientes y cada uno de ellos conocido, en cada caso. La distribución a prior chi-cuadrado inversa.

$$\pi(\mu, \sigma^2) \ \alpha \ (\sigma^2)^{-\frac{(v_0+1)}{2}-1} \exp\left\{-\frac{1}{2}\frac{\left[S_0 + n_0(x\mu - \mu_0)^2\right]}{r}\right\}$$

$$\alpha \ (\sigma^2)^{-\frac{(v_0+1)}{2}-1} \exp\left\{-\frac{1}{2} Q_0(\mu) \Big/ \sigma^2\right\},$$

Donde $Q_0(\mu)$ es la forma cuadrática

$$Q_0(\mu) = n_0\mu^2 - 2(n_0\mu_0)\mu + \left(n_0\mu_0^2 + S_0\right).$$

La distribución a posteriori conjunta se obtiene como combinación de la distribución a priori y la verosimilitud normal

$$\pi(\mu, \sigma^2 / x) \ \alpha \ \pi(\mu, \sigma^2).\ell(x / \mu, \sigma^2) \alpha (\sigma^2)^{-(v_1+1)/2-1} \exp\left(-\frac{1}{2} Q_1(\mu) / \sigma^2\right), \qquad (1.18)$$

Donde $v_1 = v_0 + n$ y la expresión cuadrática $Q_1(\mu)$ es:

$$Q_1(\mu) = S_1 + n_1(\mu - \mu_1)^2 = n_1\mu^2 - 2(n_1\mu_1)\mu + \left(n_1\mu_1^{\ 2} + S_1\right)$$

CAPITULO 2

ESTADISTICA ACTUARIAL

2.1. INTRODUCCIÓN.

En este capítulo se introducen los diferentes sistemas de tarificación más utilizados en la literatura actuarial. Se estudiarán sus propiedades, así como el sistema de tarificación comúnmente empleado en Europa en el Sector Automovilístico, el llamado sistema *bonus-malus*, así como también la *teoría de la credibilidad, la metodología bayesiana y el reaseguro,* un instrumento que permite a la compañía aseguradora acomodar su estructura de riesgos a su capacidad financiera.

Lo ideal en estadística actuarial es trabajar con las distribuciones del número de reclamaciones y de la cantidad reclamada. Dicha distribución constituye uno de los tópicos más importante en estadística actuarial, así, se abordará el problema bajo el modelo colectivo e individual de la teoría del riesgo.

2.2. TARIFICACIÓN

La cobertura de un riesgo por parte de una compañía aseguradora se establece con la garantía de un contrato, la póliza exige al asegurado a pagar un precio, *la prima.*

2.2.1. Principios de cálculo de prima

La prima es el precio para el seguro (o reaseguro) vendido por la compañía aseguradora. Puesto que lo que la compañía vende es la cobertura de un riesgo, podemos especificar más la definición anterior como sigue.

Definición 2.1: *La prima es el pago que un asegurado hace a un asegurador por la cobertura total o parcial contra un riesgo.*

De forma reducida, una prima mínima técnica está compuesta de los siguientes elementos:

- Prima pura de riesgo.
- Sobre prima de seguridad.
- Costo adicional para el beneficio.

Por tanto, la prima o precio del servicio es el costo que para una empresa suponen los siniestros más el margen de beneficio. En este contexto nos centraremos en las dos primeras componentes y no haremos mención a la tercera, que en un principio parece no tener un componente estocástico.

El precio correcto, que es llamado *rating,* es vital, pues si es demasiado bajo representa una pérdida para la compañía aseguradora y si es demasiado alto se pierde competitividad frente a otras. Por tanto, una de las labores del actuario consiste en encontrar métodos de cálculo de primas, generalmente llamados en la literatura actuarial *principios de cálculo de prima.*

Si se denota por la X la variable aleatoria número de reclamaciones o cantidad reclamada (o una combinación de ambas), un principio de cálculo de prima se define como sigue:

Definición 2.2: *un principio de cálculo de prima es una función* $H(X)$ *que asigna un riesgo* X*, un número real, que es la prima.*

2.2.2. Propiedades

Una prima debe de satisfacer una serie de propiedades ideales o axiomas. Sin embargo, no existe en la literatura actuarial un sistema axiomático comúnmente aceptado de propiedades que un principio de cálculo de prima debería satisfacer.

Gerber (1979) sostiene que las cinco propiedades que un principio de cálculo de prima $H(X)$ debería satisfacer son:

1. ***Sobreprima de seguridad no negativa.***

   $$\mathrm{H}(x) \geq \mathrm{E}(x),$$

   Esto significa que para evitar la ruina técnica la ganancia esperada $\mathrm{H}(x) - \mathrm{E}(x)$ será no negativa.

2. ***No estafa.*** La prima no excederá a la reclamación máxima posible, rx

   $$\mathrm{H}(x) \leq rx$$

3. ***Consistencia.*** Para cada riesgo X y cada constante c,

   $$\mathrm{H}(x+c) = \mathrm{H}(x) + c$$

   Esto significa que si el beneficio se incrementa en una constante esta constante tiene que ser añadida a la prima.

4. ***Aditividad.*** Si X_1 y X_2 son riesgos independientes, entonces:

   $$\mathrm{H}(X_1 + X_2) = \mathrm{H}(X_1) + \mathrm{H}(X_2).$$

Esto quiere decir que la incorporación de riesgos independientes no afectan a la prima total.

5. ***Iteratividad.*** Si X y Θ son dos riesgos arbitrarios dependientes, entonces

$$\mathrm{H}(X) = \mathrm{H}[\mathrm{H}(X/\Theta)].$$

Esto significa que la prima para X puede calcularse en dos pasos. Primero calcular la prima condicionada para X, $\mathrm{H}(X/\Theta)$ aplicando H a la distribución condicional de X. Esta prima condicional es una función de Θ y por lo tanto, una variable aleatoria en sí misma. Entonces se aplica H a la distribución de $\mathrm{H}(X/\Theta)$ para obtener $\mathrm{H}[\mathrm{H}(X/\Theta)]$

Heilman (1989) solo presta atención a la primera de estas propiedades, mientras que Hurliman (1994) no considera la quinta y, sin embargo, añade otras:

6. $H(c) = c$, para toda constante $c \geq 0$.

Esto significa que para un riesgo no aleatorio $X = c$, con $P_r(X = c) = 1$, la prima a cobrar será c.

7. ***Homogenidad positiva***. $H(cX) = cH(X)$, para todo $c \geq 0$.

Que resulta conveniente para corregir efectos inflaccionarios.

2.2.3. Prima de riesgo, colectiva y Bayes

Una vez establecido un principio de cálculo de prima a aplicar a un riesgo X el siguiente paso consistirá en calcular la prima asociada a X conforme a una determinada distribución de probabilidad asociada de riesgo. En este sentido es conveniente comentar que en algunos casos las variables aleatorias que intervienen en el proceso de riesgo degeneran en variables deterministas. Por ejemplo en muchas formas de los seguros de vida la cantidad reclamada es fija. En otras ocasiones, tanto los costes como el número de siniestros o reclamaciones son variables aleatorias, como ocurre en seguros de accidentes, especialmente en seguros de automóviles.

La forma de recogida de datos por parte de las compañías aseguradoras determinará la metodología de trabajo. Se señala los siguientes:

1. En algunas ocasiones, las compañías recogen datos solamente de la cantidad total reclamada en unidades monetarias generada por cada póliza y año. En este caso la única vía para trabajar es utilizar esta cantidad y todos los modelos y/o estimadores se referirán a la distribución de la cantidad total reclamada.

2. Los datos se recogen separadamente para el número de reclamaciones y el coste de cada uno de ellos. En este caso, el modo en que se trabaja es componer los modelos, del número de reclamaciones y de la cuantía de las mismas, para obtener la distribución de la cantidad total reclamada, esto es, la distribución de la variable aleatoria compuesta, $X = \sum_{i=1}^{N} X_i$, donde N es variable aleatoria asociada al número de reclamaciones y X_i la variable aleatoria asociada a la cuantía del *i*-ésimo siniestro.

3. Por último en algunas ocasiones se cree que una vez ocurrido un siniestro la cuantía del mismo está fuera de control del asegurador, de modo que el número de reclamaciones o siniestros es la única componente a considerar. En este caso, el modo de trabajar es por medio del número de reclamaciones y no de la cantidad total reclamada, este es el caso habitual en el ramo de los seguros de automóviles, en el que para conductores precavidos, mediante bonificaciones en la prima, se puede estimar al mismo o estimar una conducta más prudente.

Un riesgo X representará indistintamente el número, la cuantía o la cantidad total o agregada.
La siguiente metodología de cálculo de prima se basa en Heilmann (1989), donde se construyen diversos principios de cálculo de prima mediante el uso de funciones de pérdidas en el escenario de la teoría de la decisión, equivalente en algunas ocasiones a utilizar funciones de utilidad.
Una función de pérdida $\mathcal{L}(\overline{\theta}, \theta)$ describe la pérdida soportada por un decidor que elige el estimador θ en vez del verdadero valor del parámetro θ.
Consideremos ahora una función de pérdida $\mathcal{L} : \Re^2 \rightarrow \Re$ que atribuya a algún $(x, P) \in \Re^2$
La pérdida soportada por un decidor que toma la acción P y se encuentra con el resultado X de algún experimento aleatorio. La prima de riesgo se define de la siguiente manera.

Definición 2.3: Dados un riesgo X con función de densidad $f(x)$ y una función de pérdida $\mathcal{L} : \Re^2 \rightarrow \Re$, la prima de riesgo es el valor de P que minimiza la pérdida esperada.

$$\mathcal{L}(x,P)\, f(x)dx = E_f\ \mathcal{L}[(x,P)], \qquad (2.1)$$

Donde X es el resultado del experimento aleatorio X y P la prima cobrada por tomar X
Obviamente, si X es discreta P deberá minimizar la pérdida esperada

$$\sum_{x=0}^{\infty} \mathcal{L}(x,P) f(x)$$

Donde ahora $f(x) = Pr(X = x)$ es la función de la densidad discreta. Para obtener las diversas primas de riesgo consideramos funciones de pérdidas de la forma

$$\mathcal{L}(x, P) = g(x)\left[h(x) - h(P)\right]^2 \qquad (2.2)$$

Donde $g(x)$ y $h(x)$ son funciones apropiadas cuyas esperanzas bajo $f(.)$ existen.

Ahora utilizando (2.1) y el siguiente resultado podemos obtener la prima de riesgo.

Teorema 2.1: Si $h(x)$ es estrictamente creciente y diferenciable (y por tanto invertible) y $g(x)$ es no negativa, entonces para todo riesgo X con $E_f\left[g(X)h(X)\right] < \infty$ Se tiene

$$H(X) = h^{-1}\left(\frac{E_f\left[g(X)h(X)\right]}{\mathrm{E}_f\left[g(X)\right]}\right)$$

Demostración: Derivando con respecto a P el funcional

$$\int L(x,P)f(x)dx = \int g(x)\left[h(x) - h(P)\right]^2 f(x)dx$$

resulta

$$-2\int g(x)\left[h(x) - h(P)\right]h'(P)f(x)dx = 0$$

de donde

$$h'(P)\int g(x)h(x)f(x)dx = h(P)h'(P)\int g(x)f(x)dx = 0$$

Ahora puesto que $h'(P) > 0$ resulta:

$$h(P) = \frac{\int g(x)h(x)f(x)dx}{\int g(x)f(x)dx}$$

Y por tanto

$$P = H(X) = h^{-1}\left(\frac{\int g(x)h(x)f(x)dx}{\int g(x)f(x)dx}\right) = h^{-1}\left(\frac{Ef\left[g(X)h(X)\right]}{Ef\left[g(X)\right]}\right) \qquad \blacksquare$$

PRIMAS DE RIESGOS PARA FUNCIONES PARTICULARES $f(x)$ **Y** $g(x)$.

Proposición 2.1 Si consideramos la función de pérdida cuadrática, dada por

$$\mathcal{L}(x,P) = (x - P)^2$$

Resulta

$$P = E_f(X) \tag{2.3}$$

denominado principio de prima neta o de equivalencia

Demostración:

Derivando $\mathcal{L}(x,P) = (x - P)^2$ con respecto a P, se tiene:

$$\int \mathcal{L}(x,P) f(x)dx = \int (x - P)^2 f(x)dx = -2\int (x - P) f(x)dx = 0$$

$$\int (x - P) f(x)dx = 0$$

$$\int x f(x)dx - P\int f(x)dx = 0$$

$$P\int f(x)dx = \int xf(x)dx$$

$$P = \int xf(x)dx, \quad \text{puesto que} \quad \int f(x)dx = 1$$

Y por tanto:

$$P = E_f(X) \qquad \blacksquare$$

Proposición 2.2 Si consideramos la función de pérdida exponencial, dada por

$$\mathcal{L}(x,P) = \frac{1}{\alpha}\left(e^{\alpha x} - e^{\alpha P}\right)^2$$

Con $\alpha > 0$ resulta

$$P = \frac{1}{\alpha}\log E_f\left(e^{\alpha X}\right) \tag{2.4}$$

denominado principio de utilidad exponencial.

Demostración:

Derivando $\mathcal{L}(x,P) = \frac{1}{\alpha}\left(e^{\alpha x} - e^{\alpha P}\right)^2$ con respecto a P, se tiene:

$$\int \mathcal{L}(x,P) f(x)dx = \int \frac{1}{\alpha}\left(e^{\alpha x} - e^{\alpha P}\right)^2 f(x)dx = \frac{1}{\cancel{\alpha}}(-2\cancel{\alpha})\int \left(e^{\alpha x} - e^{\alpha P}\right)\left(e^{\alpha P}\right) f(x)dx = 0$$

$$= -2e^{\alpha P}\int \left(e^{\alpha x} - e^{\alpha P}\right) f(x)dx = \int \left(e^{\alpha x} - e^{\alpha P}\right) f(x)dx = 0$$

$$\int e^{\alpha x} f(x)dx - \int e^{\alpha P} f(x)dx = 0$$

$$\int e^{\alpha P} f(x)dx = \int e^{\alpha x} f(x)dx$$

$$e^{\alpha P} \int f(x)dx = \int e^{\alpha x} f(x)dx$$

$$e^{\alpha P} = \int e^{\alpha x} f(x)dx, \quad \text{puesto que } \int f(x)dx = 1$$

$$log\, e^{\alpha P} = log \int e^{\alpha x} f(x)dx$$

$$\alpha P = log \int e^{\alpha x} f(x)dx$$

$$P = \frac{1}{\alpha} log \int e^{\alpha x} f(x)dx$$

Y por tanto:

$$P = \frac{1}{\alpha} \log E_f\left(e^{\alpha X}\right) \quad \blacksquare$$

Este principio de cálculo de prima viene dado en términos del logaritmo de la función generatriz de momentos de la variable aleatoria X . En teoría de la decisión a ∞ se le denomina *constante de aversión al riesgo* (también llamada medida de Arroz-Pratt) asociada al decidor que toma la función de pérdida de $\mathcal{L}(x,P)$. Cuanto mayor ∞ es más adverso al riesgo será el decisor (en nuestro caso la compañía aseguradora).

Proposición 2.3 Si consideramos la función de pérdida cuadrática ponderada con peso $g(x) = e^{\alpha x}$, dada por

$$\mathcal{L}(x,P) = e^{\alpha x}(x-P)^2$$

Con $\alpha > 0$ entonces

$$P = \frac{E_f\left(Xe^{\alpha X}\right)}{E_f\left(e^{\alpha X}\right)} \tag{2.5}$$

denominado principio Esscher.

Demostración:

Derivando $\mathcal{L}(x,P) = e^{\alpha x}(x-P)^2$ con respecto a P , se tiene:

$$\int \mathcal{L}(x,P) f(x)dx = \int e^{\alpha x}(x-P)^2 f(x)dx = -2\int e^{\alpha x}(x-P) f(x)dx = 0$$

$$\int e^{\alpha x}(x-P) f(x)dx = 0$$

$$\int xe^{\alpha x} f(x)dx - P\int e^{\alpha x} f(x)dx = 0$$

$$P\int e^{\alpha x} f(x)dx = \int xe^{\alpha x} f(x)dx$$

$$P = \frac{\int xe^{\alpha x} f(x)dx}{\int e^{\alpha x} f(x)dx}$$

Y por tanto:

$$P = \frac{E_f\left(Xe^{\alpha X}\right)}{E_f\left(e^{\alpha X}\right)} \qquad \blacksquare$$

El parámetro α de (2.5) tiene la misma interpretación que en el caso anterior.

Proposición 2.4 Si consideramos la función de perdida cuadrática ponderada con peso g(x)=x, dada por

$$\mathcal{L}(x,P) = x(x-P)^2$$

Entonces

$$P = \frac{E_f\left(X^2\right)}{E_f\left(X\right)} = E_f(X) + \frac{Varf\left(X\right)}{Ef\left(X\right)} \tag{2.6}$$

denominado principio de varianza.

Demostración:

Derivando $\mathcal{L}(x,P) = x(x-P)^2$ con respecto a P, se tiene:

$$\int \mathcal{L}(x,P) f(x)dx = \int x(x-P)^2 f(x)dx = \int (x^3 - 2x^2 P + xP^2) f(x)dx$$

$$\int (-2x^2 + 2xP) f(x)dx = 0$$

$$2\left(-\int x^2 f(x)dx + P\int xf(x)dx\right) = 0$$

$$P\int xf(x)dx = \int x^2 f(x)dx$$

$$P = \frac{\int x^2 f(x)dx}{\int xf(x)dx}$$

$$P = \frac{E_f(X^2)}{E_f(X)}$$

$$P = \frac{E_f^2(X) + var(X)}{E_f(X)}$$

Y por tanto:

$$P = E_f(X) + \frac{var(X)}{E_f(X)} \qquad \blacksquare$$

La ventaja de este principio es que no solo estima la siniestralidad media del riesgo, sino que proporciona además el recargo de seguridad que debe llevar la prima para atender a las desviaciones aleatorias de la siniestralidad. En muchos textos la expresión de P se presenta como

$$P = E(X) \; + \; \delta \, var(x)$$

Siendo $\alpha > 0$ un parámetro y se dice entonces que la sobreprima de seguridad es proporcional a la varianza.

El siguiente resultado se muestra útil para probar la propiedad de sobreprima de seguridad no negativa de los principios de cálculo de prima antes presentados.

Teorema 2.2 Si $h(x) \; = \; x$, es decir, $\mathcal{L}(x,P) = g(x)(x-P)^2$ y $g : \Re^+ \to \Re^+$ entonces $g(x)$ es creciente (decreciente) si y solo si para todo riesgo X con $Pr(X > 0) = 1$ se verifica:

$$H(X) = \frac{Ef\left[g(X)X\right]}{Ef\left[g(X)\right]} \geq (\leq) Ef(X) \tag{2.7}$$

Existen otra serie de principios, estos son:

- **Principio del valor esperado.** La prima de riesgo viene dada mediante:

$$P = (1+\lambda)E(X), \quad \lambda > 0.$$

- **Principio de desviación típica.** La prima de riesgo, en este caso, viene dada por:

$$P = E(X) + \lambda\sqrt{Var(X)} \qquad \lambda > 0$$

- **Principio de prima suizo,** la prima de riesgo P es la solución de la ecuación

$$E\left[v(X - zP)\right] = v((1-z)P)$$

  En la que $v(.)$ es una función peso verificando $v'(z) > 0, v''(z) \geq 0, si \;\; z > 0$

- **Principio de Orlicz.** La prima de riesgo es la solución de la ecuación:

$$\mathrm{E}\left[\phi\left(\frac{X}{P}\right) = 1\right]$$

  donde Φ es una función peso con $\Phi' > 0$, $\Phi'' \geq 0$

- **Principio de Wang.** La prima de riesgo viene dada por:

$$P = \int_0^\infty g\left[Sx(t)\right] \, dt,$$

Donde $S_X(t) = Pr(X > t)$ es la función de supervivencia de X y g es una función no decreciente con dominio en $[0,1]$

- **Principio holandés.** La prima de riesgo es:

$$P = E(X) + \lambda E[(X - \alpha E(X))]_+, \quad \alpha \geq 1, 0 < \lambda \leq 1$$

donde: $(X - \infty E(X))_+ = máx\{(X - \alpha E(X)), 0\}$

Por otro lado, cabe también la posibilidad de generar principios de cálculo de primas considerando otras funciones de pérdidas.
De este modo, bajo la función de pérdida:

$$\mathcal{L}(x,P) = \frac{(x-P)^2}{x}, \; x > 0, Pr(X > 0) = 1$$

Se obtiene la prima de riesgo,

$$P = \frac{1}{E\left(\frac{1}{X}\right)}$$

Si se considera la función de pérdida

$$\mathcal{L}(x,P) = \frac{(x-P)^2}{x(x+1)}, \; x > 0, Pr(X > 0) = 1$$

Se obtiene la prima de riesgo

$$P = \frac{E\left(\frac{1}{x+1}\right)}{E\left(\frac{1}{x(x+1)}\right)}$$

Finalmente, bajo la función de pérdida

$$\mathcal{L}(x,P) = (log\, x - log\, P)^2, x > 0, P > 0, Pr(X > 0) = 1$$

La prima de riesgo resultante es

$$P = exp\{E[log(X)]\}.$$

Los principios de cálculo de primas mostrados pueden estudiarse siempre que la función de densidad $f(x)$ sea conocida. En la Literatura actuarial es habitual considerar que todos o algunos de los parámetros de los que depende esta densidad de probabilidad son desconocidos. Así se supondrá ahora que la densidad $f(x)$ depende de un parámetro desconocido

$\theta, \theta \in \Theta$, entonces la densidad de probabilidad será $f(x;\theta)$ o $f(x/\theta)$ Dependiendo que el parámetro sea fijo o aleatorio. Podemos ahora suponer que dicho parámetro se distribuye entre toda la cartera de seguros de acuerdo a cierta función de densidad $\pi(\theta)$. Desde un punto de vista Bayesiano esta no es más que la distribución a priori, denominada en el escenario actuarial función estructura. Ahora la prima de riesgo P depende del parámetro desconocido y por tanto será notada como $P(\theta)$. En principio, la mejor estimación que puede obtenerse de la misma es la prima colectiva que se define de la siguiente forma.

Definición 2.4 *Dados un riesgo X con función de densidad $f(x/\theta)$, siendo θ un parámetro desconocido con función de densidad a priori $\pi(\theta)$ y una función de pérdida, $\mathcal{L}: \Re^2 \rightarrow \Re$ la prima colectiva es el valor P' que minimiza la pérdida esperada*

$$\int_\Theta \mathcal{L}(P(\theta), P')\pi((\theta)d\theta,$$

Siendo $P(\theta)$ la prima de riesgo definida en (2.1).

La prima colectiva tal y como está definida arriba representa la mejor decisión que estima la prima de riesgo (obviamente desconocida). Observemos que para calcularla se necesitará que el actuario defina una distribución de probabilidad (distribución a priori) para el valor del parámetro desconocido θ. Para ello será fundamental la experiencia de lo acontecido en los periodos precedentes o en otros contratos similares.

Ahora se supondrá que la distribución de X esta especificada mediante un parámetro desconocido y además se incorpora experiencia de siniestralidad individual.

En este caso el análisis bayesiano nos permitirá combinar la información inicial o a priori que se tiene sobre el parámetro θ. Con la información muestral se obtendrá la distribución a posteriori del parámetro. Si $\pi(\theta)$ es la densidad a priori (que refleja las creencias sobre θ antes de obtener la información muestral), y $x = (x_1 \ldots\ldots\ldots x_t)$ es un vector de datos observados que recoge la información muestral, la verosimilitud del dato observado la denotaremos por $L(x/\theta)$ y el teorema de Bayes nos permitirá obtener la distribución a posteriori $\pi(x/\theta)$, de la siguiente manera:

$$\pi(\theta/x) = \frac{L(x/\theta)\pi(\theta)}{\int_\Theta L(x/\theta)\pi(\theta)d\theta} \alpha \ L(x/\theta)\pi(\theta),$$

Es decir, $\pi(x/\theta)$ es el cociente entre la distribución conjunta $\mathcal{L}(x/\theta)\pi(\theta)$ y la distribución predictiva $p(x) = \int_\Theta L(x/\theta)\pi(\theta)d\theta$ como ya hemos visto. Esto permite construir la prima de Bayes o a posteriori de la siguiente manera:

Definición 2.5 *Dados un riesgo* X *con función de densidad de probabilidad dada por* $f(x/\theta)$ *siendo* θ *un parámetro desconocido con distribución a priori* $\pi(\theta)$ *una función de pérdida* $\mathcal{L}:\Re^2 \rightarrow \Re$ *y un vector de datos observados* x*, la prima Bayes es el valor* $P(x)$ *que minimiza:*

$$\int \mathcal{L}[P(\theta),P(x)]\pi(\theta/x)d\theta$$

Siendo $\pi(\theta/x)$ *la distribución a posteriori de* θ *dada la muestra y* $P(\theta)$ *la prima de riesgo definida anteriormente.*

Se expone a continuación el siguiente resultado que muestra que la media de la distribución a posteriori de θ es igual a la media de la distribución predictiva X_{t+1}, siempre que $E(X_i/\theta)=\theta,\ i=1,2,...,t.$

Teorema 2.3. Sean $X_1,X_2,...,X_t$ variables aleatorias $i.i.d.$ Tales que $E(X_i/\theta)=\theta,\ i=1,2,...,t.$ Entonces:

$$E(X_{t+1}/x_1,x_2,...,x_t)=E(\theta/x_1,x_2,...,x_t).$$

Demostración:

$$\begin{aligned}
E(X_{t+1}/x_1,x_2,...,x_t) &= \int_x xf(x_{t+1}\mid x_1,x_2,...,x_t)dx \\
&= \int_x x\left(\int_\Theta f(x_{t+1}\mid\Theta)\pi(\Theta\mid x_1,x_2,...,x_t)d\Theta\right)dx \\
&= \int_\Theta \left(\int_x xf(x_{t+1}\mid\Theta)dx\right)\pi(\Theta\mid x_1,x_2,...,x_t)d\Theta \\
&= E(\Theta\mid x_1,x_2,...,x_t). \qquad \blacksquare
\end{aligned}$$

Por tanto, la prima neta de Bayes coincide con la media de la distribución predictiva.

2.2.4. La teoría de la credibilidad.

El problema básico sobre la teoría de la credibilidad es el siguiente:

Supongamos que disponemos para un asegurado o contrato de la experiencia de siniestralidad $X_1...X_t$ de modo que $E(X_j)=\varepsilon$ y $Var(X_j)=\sigma^2$, para todo $j=1,....,t$. El objetivo de la aseguradora es decidir qué prima cargar a esa póliza o asegurado. Existen las siguientes alternativas:

1. Ignorar la experiencia de siniestralidad y cargar lo que en literatura actuarial se conoce como prima manual o de libro, M. Esta prima está basada en la experiencia de otros contratos similares.

2. Cobrarle $\overline{X} = \frac{1}{t}\sum_{j=1}^{t} X_j$, es decir dar credibilidad total a la experiencia del asegurado.

3. Cobrar una prima que venga dada como un punto medio o combinación lineal convexa entre la experiencia individual y M. Estos es:

$$Prima = Z(t)\overline{X} + [1 - Z(t)]M, \tag{2.8}$$

Donde el factor, $Z \in [0,1]$, recibe el nombre de *factor de credibilidad* y la fórmula (2.8) *de credibilidad.*

2.2.4.1 credibilidad total

Parece lógico que los asegurados con una experiencia de reclamación que les sea favorable quieran que la prima que tengan que pagar esté basada únicamente en su propia experiencia de siniestralidad, es decir que la aseguradora le asigna a esta un 100% de credibilidad. Sin embargo, desde el punto de vista de la aseguradora, esto solo será posible si la experiencia de reclamación es estable.

Una manera de resolver este problema es suponer que $\overline{X}$ es estable si existe una probabilidad alta de que la diferencia entre $\overline{X}$ y ε sea pequeña. En términos matemáticos esto supondría suponer credibilidad total si

$$Pr\left(\left|\overline{\mathrm{X}} - \varepsilon\right| \le c\varepsilon\right) = Pr\left((1-c)\varepsilon \le \overline{\mathrm{X}} \le (1+c)\varepsilon\right) \ge p, \tag{2.9}$$

Siendo $0 < p < 1 \ y \ c > 0.$ En la práctica lo razonable es elegir un valor de p cercano a 1 y un valor de c cercano a 0. Normalmente suelen considerarse 0.9 y 0.05 para p y c, respectivamente.

Se reescribe (2.9) en la forma:

$$\Pr\left(\left|\frac{\overline{\mathrm{X}} - \varepsilon}{\sigma / \sqrt{t}}\right| \le x_p\right) = p.$$

y se define ahora x_p, como

$$X_P = \underset{x}{\text{ínf}} \left\{ Pr\left(\left| \frac{\overline{X} - \varepsilon}{\sigma / \sqrt{t}} \right| \le x \right) \ge p \right\}.$$

Suponiendo que $\overline{X}$ sigue una distribución de tipo continua, esta última expresión es equivalente a:

$$Pr\left(\left| \frac{\overline{X} - \varepsilon}{\sigma / \sqrt{t}} \right| \le x_p \right) = p. \qquad (2.10)$$

Por tanto, la condición que ha de verificarse para suponer credibilidad total es

$$\frac{c\varepsilon\sqrt{t}}{\sigma} \ge x_p \, ,$$

O de forma equivalente:

$$\frac{\sigma}{\varepsilon} \le \frac{c}{x_p}\sqrt{t} = \sqrt{\frac{t}{\lambda_0}}, \qquad (2.11)$$

Siendo $\lambda_0 = \left({x_p}/{c} \right)^2$. La expresión (2.11) puede interpretarse en el sentido siguiente: se supone credibilidad total si el coeficiente de variación es menor o igual que $\sqrt{t/\lambda_0}$. También, a partir de (2.11) se observa que se supone credibilidad total si:

$$var(\overline{X}) = \frac{o^2}{t} \le \frac{\varepsilon^2}{\lambda_0},$$

Y, por otro lado, el valor que ha de tomar t para suponer credibilidad total ha de cumplir:

$$t \ge \lambda_0 \left(\frac{\sigma}{\varepsilon} \right)^2$$

En la práctica, si la experiencia del asegurado es suficientemente grande, de acuerdo al teorema central del límite, entonces $(\overline{x} - x)/\left(\sigma\sqrt{t}\right)$ sigue aproximadamente una distribución normal con media cero y desviación típica 1. Así (2.10) puede entonces escribirse como $p = 2\Phi(x_p) - 1,$ en donde $\Phi(x)$ es la función de la distribución normal tipificada.

2.2.4.2. Credibilidad parcial

Para muchos asegurados la experiencia de siniestralidad es suficiente para suponer credibilidad total, es decir que el factor $Z(t)$ sea igual a 1. Ahora se supone que la prima a cargar sea una combinación lineal entre la experiencia del asegurado y la experiencia del colectivo o prima manual, de modo que:

$$P = Z(t)\overline{X} + [1 - Z(t)]M$$

Y habrá una, por tanto, que determinar el valor $Z(t)$ para obtener la prima. Dado que:

$$Var(P) = Var\left[Z(t)\overline{X} + [1 - Z(t)]M\right] = Z(t)^2 Var(\overline{X}) = Z(t)^2 \frac{\sigma^2}{t},$$

Igualando este último término a $\varepsilon^2 / \lambda_0$ resulta,

$$Z(t) = (\varepsilon / \sigma)\sqrt{t / \lambda_0},$$

De modo que se elige $Z(t)$ de acuerdo a la expresión:

$$Z(t) = min\left\{\frac{\varepsilon}{\sigma}\sqrt{\frac{t}{\lambda_0}}, 1\right\}. \qquad (2.12)$$

2.2.4.3. Credibilidad e inferencia bayesiana

El uso de distribuciones a priori, con un marcado carácter subjetivo, resulta útil en el mercado de seguros, sobre todo si se tiene en cuenta que cuando se quiere tarificar un riesgo nuevo no se cuenta con datos disponibles, algunas primas obtenidas mediante la metodología bayesiana pueden escribirse como fórmulas de credibilidad.

El problema de la teoría de la credibilidad consiste en determinar las ponderaciones que afectan a la experiencia de siniestralidad de una póliza respecto a la experiencia de un colectivo al que pertenece dicha póliza.

Consideremos la cartera de seguros que aparece en la tabla 2.1, que consta de k pólizas o asegurados y t periodos de las mismas.

Tabla 2.1: Cartera de seguros

	1	2	$\cdots$	j	$\cdots$	k
1	x_{11}	x_{21}	$\cdots$	x_{j1}	$\cdots$	x_{k1}
2	x_{12}	x_{22}	$\cdots$	x_{j2}	$\cdots$	x_{k1}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$
t	x_{1t}	x_{2t}	$\ddots$	x_{jt}	$\ddots$	x_{kt}

La cuestión básica de la teoría de la credibilidad es determinar una prima establecida como una combinación lineal o convexa entre la experiencia particular de un asegurado y la experiencia del colectivo, esto es de toda la cartera. Una expresión válida sería:

$$P_j = [1 - Z(t)]P_0 + Z(t)\tilde{P}_j,$$

Donde:

- P_j : Prima a aplicar a los asegurados al riesgo j.

- P_0 : Prima A aplicar a un colectivo al que pertenece el asegurado j.

- $Z(t)$: Factor de credibilidad que verifica $lim_{t\to\infty} Z(t) = 1$, siendo t el número de expuestos al riesgo j o el periodo de observación de la póliza j. Por tanto, si $Z(t) = 1$ la experiencia del asegurado es creíble al 100%, mientras que si $Z(t) = 0, P_j = P_0$ y la prima del asegurado j coincide con la del colectivo al que pertenece dicha póliza.

La fórmula de credibilidad puede por tanto interpretarse también de la siguiente manera: puede considerarse a P_0 como la información a priori; $\tilde{P}_j$ la nueva información obtenida mediante la observación de la siniestralidad del riesgo j y P_j el resultado de combinar la información a priori con la información adquirida. Por tanto,

Prima (a posterior) = $[1 - Z(t)]$ Prima a priori $+ Z(t)$ Experiencia observada.

La teoría de la credibilidad sigue un esquema bayesiano, donde se da entrada a la información a priori con la información muestral, para obtener finalmente un estimador revisado de la prima.

Definición 2.6 *La teoría de la credibilidad es el mecanismo que permite el ajuste sistemático de las primas de seguros a medida que se obtiene la experiencia de siniestralidad.*

Una de las principales aplicaciones de la teoría de la credibilidad se presenta en el seguro de automóviles, en el que la prima inicial se va modelando sucesivamente a medida que se incorpora la información de la siniestralidad. Son los denominados **sistemas de tarificación bonus-malus.** Se parte de un nivel x neutro, de modo que para niveles superiores a x el asegurado entra en la escala *malus* y para niveles inferiores a x el asegurado entra en la escala *bonus.*

2.2.4.4 Modelo de Bühlmann de distribución libre

Dentro de los modelos de credibilidad clásicos más destacados figuran el modelo de Bühlmann de distribución libre y el modelo de Bühlmann-straub. Ambos modelos constituyen el punto de partida de la moderna teoría de la credibilidad.

El objetivo de ambos modelos es estimar la prima correspondiente a un asegurado o grupo de asegurados que conforman una póliza en una cartera de seguros, restringiéndose a las primas lineales y utilizando el método de los mínimos cuadrados. La diferencia fundamental entre ambos modelos radica en que el segundo admite observaciones ponderadas. Lo relevante del modelo propuesto es la no necesidad de establecer hipótesis alguna, ni sobre la distribución que gobierna los riesgos individuales, ni sobre la distribución a priori de los parámetros de riesgo, de ahí el nombre de modelo de distribución libre.

Para su descripción se considera de nuevo la cartera de seguros que aparece en la tabla 2.1 en que θ_j representa el parámetro de riesgo por la póliza $j-ésima$. Se trata de una variable estructural que describe las características de riesgo del contrato $j-ésimo.$ En la ciencia actuarial es costumbre considerar a dicho parámetro desconocido aleatorio.

La variable X_{ij} representa la experiencia de reclamaciones para la póliza $j-ésima$ en el periodo $i-ésimo$. Se trata de una variable aleatoria con realizaciones observables.

Se denotará mediante:

$$\mu(\theta_j) = E(X_{ij} / \theta_j),$$

A la prima de riesgo para la póliza j,

$$m = E\lfloor P(\theta_j)\rfloor,$$

Al valor esperado de todas las primas de riesgo, es decir la prima colectiva. Finalmente,

$$a = Var\lfloor \mu(\theta_j)\rfloor,$$

Es la varianza de las primas de riesgo, que es un indicador de la heterogeneidad de la cartera.

El objetivo del modelo de Bühlmann consiste en calcular la mejor prima lineal

$$\mathcal{H}\left(\mu(\theta_j) / X_{j1}, X_{j2}, ..., X_{jt}\right),$$

Dependiente de los datos observados, mediante el método de los mínimos cuadrados. Para ello establecemos la siguiente notación previa, en la que prescindiremos del subíndice j:

- $\mu(\theta) = E(X / \theta)$: Prima de riesgo individual

- $m = E_{total}(X) = E[\mu(\theta)]$: Prima de riesgo colectiva. Valor esperado de todas las primas de riesgo individuales.
- $a = Var[E(X/\theta)] = Var[\mu(\theta)]$: Varianza de las primas de riesgos individuales, indicador de la heterogeneidad de la cartera.
- $s^2 = E[Var(X/\theta)]$: Medida global de la dispersión de la siniestralidad individual.

Se supone también que $X_1/\theta, X_2/\theta, ..., X_t/\theta$ están idénticamente distribuidas con media y varianza comunes $\mu(\theta)$ y $\sigma^2(\theta)$, respectivamente. Antes de entrar en el modelo de Bühlmann necesitamos el siguiente resultado.

Proposición 2.5 si X y Y son variables aleatorias con distribución conjunta dependiente de la variable aleatoria Θ se verifica:

$$E(X) = E_\Theta[E_x(X/\Theta)],$$

$$Cov(X,Y) = E[Cov(X,Y/\Theta)] + Cov[E(X/\Theta), E(Y/\Theta)] \qquad (2.13)$$

Demostración:

Respecto a la primera relación, se tiene:

$$\begin{aligned} E_\Theta[E(X/\Theta)] &= \int E(X/\Theta=\theta) f_\Theta(\theta)d\theta \\ &= \int\int x f_{X/\Theta}(x/\theta)dx f_\Theta(\theta)d\theta \\ &= \int x \int f_{X/\Theta}(x/\theta) f_\Theta(\theta)d\theta\, dx \\ &= \int x \int f_{X,\Theta}(x,\theta)d\theta\, dx \\ &= \int x f_X(x)dx \\ &= E(X). \end{aligned}$$

Y por tanto:

$$E(X) = E_\Theta[E_x(X/\Theta)],$$

En cuanto a la segunda tenemos:

$$\begin{aligned} Cov(X,Y) &= E\{[X - E(X)]\ [Y - E(Y)]\} \\ &= E\{[X - E(X/\Theta) + E(X/\Theta) - E(X)] \times [Y - E(Y/\Theta) + E(Y/\Theta) - E(Y)]\} \\ &= E_\Theta E_{X,Y/\Theta}\{[X - E(X/\Theta)]\ [Y - E(Y/\Theta)]\} \end{aligned}$$

$$+E_\Theta E_{X,Y/\Theta}\{[X-E(X/\Theta)]\ [E(Y/\Theta)-E(Y)]\}$$
$$+E_\Theta E_{X,Y/\Theta}\{[E(X/\Theta)-E(X)]\ [Y-E(y/\Theta)]\}$$
$$+E_\Theta E_{X,Y/\Theta}\{[E(X/\Theta)-E(X)]\ [E(Y/\Theta)-E(Y)]\}$$
$$=E[Cov(x,y/\Theta)]+0+0+Cov[E(X/\Theta),E(Y/\Theta)].$$

Obsérvese que de (2.13) se deduce:

$$Var(X)=E[Var(X/\theta)]+Var[E(X/\theta)], \qquad \blacksquare \qquad (2.14)$$

Haciendo $X=Y$

Teorema 2.4 La mejor aproximación lineal a $\mathcal{H}\left(\mu(\theta)/X_1...,X_t\right)$ es:

$$a+b\overline{X}=a+b\frac{1}{t}\sum_{i=1}^{t}X_i,$$

Donde

$$a=(1-b)m,$$
$$b=\frac{t}{t+k},$$
$$k=\frac{E[\sigma^2(\theta)]}{Var[\mu(\theta)]}.$$

Demostración:

Se quiere encontrar la mejor estimación de la prima neta de riesgo que depende linealmente de los datos observados, esto es:

$$\mathcal{H}\left(\mu(\theta)/X_1...,X_t\right)=c_0+\sum_{s=1}^{t}c_sX_s.$$

Para ello se hace mínima la esperanza del cuadrado de la desviación de la prima de riesgo individual respecto a $\mathcal{H}\left(\mu(\theta)/X_1...,X_t\right)$, esto es:

$$\min_{c_i} E\left[\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)^2\right].$$

Calculando las derivadas $\frac{\partial H}{\partial Co}$ y $\frac{\partial H}{\partial Cs}$, se tiene:

$$\frac{\partial H}{\partial Co}=-2\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)=0$$

$$\mu(\theta)-Co-\sum_{s=1}^{t}CsXs=0$$

$$\frac{\partial H}{\partial Cs}=-2\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)X_r=0$$

$$\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)X_r=0$$

Luego, se obtiene el sistema de ecuaciones:

$$\begin{cases} E\left[\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right]=0, & (1)\\ E\left[\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)X_r\right]=0, \quad r=1,2,\ldots,t. & (2)\end{cases}$$

Multiplicando la ecuación (1) por $E(X_r)$, se tiene:

$$E(X_r)E\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)=0$$

$$E(X_r)\left(E(\mu(\theta))-E(Co)-E\left(\sum_{s=1}^{t}CsXs\right)\right)=0$$

$$\left(E(X_r)E(\mu(\theta))-E(Co)E(X_r)-E(X_r)E\left(\sum_{s=1}^{t}CsXs\right)\right)=0 \qquad (3)$$

$$E\left[\left(\mu(\theta)-Co-\sum_{s=1}^{t}CsXs\right)X_r\right]=0$$

$$E\left[\left(X_r\mu(\theta)-X_rCo-X_r\sum_{s=1}^{t}CsXs\right)\right]=0$$

$$E(X_r\mu(\theta))-E(X_rCo)-E\left(X_r\sum_{s=1}^{t}CsXs\right)=0 \qquad (4)$$

Restando la ecuación (3) de la ecuación (4), se obtiene:

$$E(X_r\mu(\theta))-E(X_r)E(\mu(\theta))-E(X_rCo)+E(Co)E(X_r)-E\left(X_r\sum_{s=1}^{t}CsXs\right)+E(X_r)E\left(\sum_{s=1}^{t}CsXs\right)=0$$

$$E(X_r\mu(\theta))-E(X_r)E(\mu(\theta))-CoE(X_r)+CoE(X_r)-E\left(X_r\sum_{s=1}^{t}CsXs\right)+E(X_r)E\left(\sum_{s=1}^{t}CsXs\right)=0$$

$$E(X_r\mu(\theta))-E(X_r)E(\mu(\theta))=E\left(X_r\sum_{s=1}^{t}CsXs\right)-E(X_r)E\left(\sum_{s=1}^{t}CsXs\right)$$

$$Cov[X_r,\mu(\theta)]=\sum_{s=1}^{t}CsCov(X_r,Xs), \quad r=1,2,\ldots,t \qquad (2.15)$$

Teniendo en cuenta ahora:

$$Cov(X_r, X_s) = E[Cov(X_r, X_s/\theta)] + Cov[E(X_r/\theta), E(X_s/\theta)] = s^2 + a, \;\; r \neq s$$

$$Cov(X_r, X_s) = Cov[\mu(\theta), \mu(\theta)] = \mathrm{var}[\mu(\theta)] = a, \quad r = s,$$

El sistema (2.15) puede reescribirse como:

$$\begin{cases} s^2 c_r + \sum_{s=1}^{t} c_s a = a \\ c_0 = m - m\sum_{s=1}^{t} c_s, \end{cases}$$

Donde se ha tenido en cuenta que, $E\left[\mu(\theta) - Co - \sum_{s=1}^{t} CsXs\right] = 0,$ y por tanto:

$$Co = E[\mu(\theta)] - \sum_{s=1}^{t} CsE(Xs) = m - m\sum_{s=1}^{t} Cs.$$

Debido a la simetría del sistema resulta que $c_1 = ... = c_t$, luego:

$$\begin{cases} s^2 c + atc = a, \\ c_0 + mtc = m. \end{cases}$$

De donde se deduce:

$$\begin{cases} c = \dfrac{a}{s^2 + at} \\ c_0 = m(1 - tc) = m\left(1 - \dfrac{ta}{s^2 + at}\right) = m\dfrac{s^2}{s^2 + at} \end{cases}$$

Por tanto,

$$\mathcal{H}(\mu(\theta)/X_1..., X_t) = c_0 + \sum_{s=1}^{t} c_s X_s = m\frac{s^2}{s^2 + at} + ct\bar{x}$$

$$= m\frac{s^2}{s^2 + at} = \frac{at}{s^2 + at}\bar{x} = [1 - Z(t)]m + Z(t)\bar{x}$$

Con:

$$Z(t) = \frac{at}{at + s^2} = \frac{tVar[\mu(\theta)]}{tVar[\mu(\theta)] + E[\sigma^2(\theta)]}$$

Dividiendo ambos miembros de la fracción por $Var[\mu(\theta)]$, se tiene:

$$Z(t) = \frac{t}{t + k} \qquad \blacksquare$$

Obsérvese que el resultado no depende de la distribución de probabilidad de X ni de la distribución de probabilidad de θ, de ahí el término de distribución libre.

Las cantidades $a = Var[\mu(\theta)]$, $s^2 = E[\sigma^2(\theta)]$ y $m = \overline{X}$ suelen llamarse parámetros estructurales del modelo y pueden estimarse a partir de:

$$\widehat{m} = \frac{1}{k}\sum_{j=1}^{k}\overline{x}_j = \frac{1}{k}\sum_{j=1}^{k}\sum_{s=1}^{t}\frac{x_{js}}{t}, \tag{2.16}$$

$$\widehat{s}^2 = \frac{1}{k}\sum_{j=1}^{k}\widehat{s}_j^2,\ \widehat{s}_j^2 = \frac{1}{t-1}\sum_{s=1}^{t}(x_{js} - \overline{x}_j)^2, \tag{2.17}$$

$$\widehat{a} = \frac{1}{k-1}\sum_{j=1}^{k}(\overline{x}_j - \overline{x})^2 - \frac{1}{t}\widehat{s}^2 \tag{2.18}$$

Finalmente, puede probarse que estos estimadores son insesgados y consistentes.

Esto es:

$$E(\hat{m}) = m,$$
$$E(\hat{s})^2 = s^2,$$
$$E(\hat{\alpha}) = \alpha,$$
$$(\hat{m}, \hat{s}, \hat{\alpha}) \to (m, s^2.\alpha),\ cuando\ t \to \infty$$

Propiedades del factor de credibilidad.

Se verifica que:

- $Z(t)$ es una función creciente en t, de modo que $Z(t) \to 1$ cuando $t \to \infty$, mientras que $Z(t) \to 0$ cuando $t \to 0$. Por tanto $t = 0$ supone que no se dispone de experiencia para el asegurado (se trata de un contrato nuevo), y la prima a cobrar en este caso es simplemente la prima colectiva. En la medida en que aumenta t, y por tanto se dispone de más datos, pesa más la información individual.
- $Z(t)$ es también una función creciente de la varianza de las medidas teóricas, $\alpha = Var[E(X/\Theta)]$, con límite 1 cuando aquella tiende a infinito y cero cuando tiende a cero. Esto es lógico, pues si la cartera no es heterogénea, $\alpha = 0$, entonces la prima colectiva es el mejor estimador de la prima individual, mientras que una mayor heterogeneidad de la cartera debe suponer dar mayor peso a la información individual.

- $Z(t)$ es una función decreciente respecto al valor esperado de la varianza teórica, $s^2 = E[Var(X/\theta)]$, de modo que cuanto mayor sea la varianza del individuo menor peso se da a su experiencia individual y mayor a la del colectivo.

El siguiente resultado muestra que el estimador de Bühlmann de la prima neta y el estimador bayesiano coinciden cuando ambas distribuciones pertenecen a la familia exponencial.

Teorema 2.5 Dados un riesgo X con función de densidad $f(x/\theta)$, y la distribución a priori del parámetro conjugada para esa verosimilitud, entonces el estimador de Bühlmann de la prima neta y el estimador bayesiano (La prima neta Bayes) coinciden cuando ambas distribuciones pertenecen a la familia exponencial.

Demostración:

La demostración se llevará a cabo considerando la familia exponencial continua. De manera análoga se demuestra para el caso discreto. Así, dada la familia exponencial de la forma:

$$f(x/\theta) = \frac{a(x)e^{-\theta x}}{c(\theta)}, \theta \in \Theta,$$

En la que $c(\theta)$ es la constante de normalización. La distribución a priori conjugada natural para esta verosimilitud es:

$$\pi(\theta) = \frac{[c(\theta)]^{-n_0} e^{-\theta x_0}}{d(n_0, x_0)} \tag{2.19}$$

Donde $d(n_0, x_0)$ es de nuevo una constante de normalización y n_0 y x_0 dos parámetros de la que depende. La distribución a posteriori es de nuevo del tipo (2.19), pero con los parámetros actualizados:

$$n_0 \to n_0 + t,$$

$$x_0 \to x_0 + \sum_{i=1}^{t} x_i$$

La prima neta de riesgo y la varianza de X vienen dadas por:

$$P(\theta) = \mu(\theta) = \frac{c'(\theta)}{c(\theta)},$$

$$Var(X/\theta) = \frac{c''(\theta)c(\theta) - c'(\theta)^2}{c(\theta)^2} = -\frac{d}{d\theta}[P(\theta)].$$

Derivando (2.19) con respecto a θ se obtiene:

$$\pi'(\theta)=\frac{1}{d(n_0,x_0)}\left\{-n_0[c(\theta)]^{-n_0-1}c'(\theta)e^{-\theta x_0}-x_0e^{-\theta x_0}[c(\theta)]^{-n_0}\right\}$$

$$\pi'(\theta)=\frac{e^{-\theta x_0}[c(\theta)]^{-n_0}}{d(n_0,x_0)}\left\{\frac{-n_0c'(\theta)}{c(\theta)}-x_0\right\}$$

$$\pi'(\theta)=\pi(\theta)\left[\frac{-n_0c'(\theta)}{c(\theta)}-x_0\right]$$

$$\pi'(\theta)=\pi(\theta)[n_0\mu(\theta)-x_0] \tag{2.20}$$

Integrando ahora (2.20) sobre Θ tenemos:

$$\pi(\theta)\Big|_{\Theta}=n_0\int_{\Theta}\mu(\theta)\pi(\theta)d\theta-x_0$$

Y suponiendo que $\pi(\theta)$ se anula en los extremos de Θ resulta:

$$\int_{\Theta}\mu(\theta)\pi(\theta)d\theta=\frac{x_0}{n_0}=m.$$

Entonces:

$$\int_{\Theta}\mu(\theta)\pi(\theta/x_1,...,x_t)d\theta=\frac{x_0+\sum_{i=1}^{t}x_i}{n_0+t}=[1-Z(t)]m+Z(t)\bar{x},$$

Con $Z(t)=\dfrac{t}{n_0+t}$, derivando (2.20) con respecto a Θ queda:

$$\pi''(\theta)=\pi'(\theta)[n_0\mu(\theta)-x_0]+\pi(\theta)n_0\frac{-c''(\theta)c(\theta)+c'(\theta)^2}{c(\theta)^2}$$

Sustituyendo $\pi'(\theta)=\pi(\theta)[n_0\mu(\theta)-x_0]$ y $Var(X/\theta)=\dfrac{-c''(\theta)c(\theta)+c'(\theta)^2}{c(\theta)^2}$, se obtiene:

$$\pi'(\theta)=\pi(\theta)[n_0\mu(\theta)-x_0]^2-\pi(\theta)n_0Var(X/\theta).$$

Finalmente, integrando la expresión anterior con respecto a Θ resulta:

$$\int_{\Theta}\pi''(\theta)d\theta=\int_{\Theta}[n_0(\mu(\theta)-m)]^2\pi(\theta)-\pi(\theta)n_0E[Var(x/\theta)]d\theta$$

$$\int_{\Theta}\pi''(\theta)d\theta=\int_{\Theta}[n_0(\mu(\theta)-m)]^2\pi(\theta)d\theta-\int_{\Theta}n_0E[Var(x/\theta)]\pi(\theta)d\theta$$

$$\int_{\Theta}\pi''(\theta)d\theta=\int_{\Theta}n_0^2(\mu(\theta)-m)^2\pi(\theta)d\theta-n_0\int_{\Theta}\pi(\theta)E[Var(x/\theta)]d\theta$$

$$\int_{\Theta}\pi''(\theta)d\theta=\int_{\Theta}n_0^2\pi(\theta)(\mu(\theta)-m)^2d\theta-n_0\int_{\Theta}\pi(\theta)E[Var(x/\theta)]d\theta$$

$$\int_{\Theta}\pi''(\theta)d\theta = n_0^2 Var[\mu(\theta)] - n_0 E[Var(X/\theta)]$$

$$\int_{\Theta}\pi''(\theta)d\theta = n_0\{n_0 Var[\mu(\theta)] - E[Var(X/\theta)]\}$$

Teniendo en cuenta que si $\pi(\theta)$ se anula en los extremos de Θ, también lo hará su derivada, concluimos que:

$$n_0 = \frac{E[Var(X/\theta)]}{Var[E(X/\theta)]} = \frac{s^2}{a}$$

$$n_0 = \frac{s^2}{a}, \qquad \blacksquare$$

Y, por tanto, el factor de credibilidad $Z(t)$ es igual al de Bühlmann y es igual al estimador bayesiano de la prima en un gran numero de casos. Esto ocurre, por ejemplo, si la distribución a priori es conjugada y la verosimilitud es un miembro de la familia exponencial, entonces el estimador de credibilidad de Bühlmann de la prima neta coincide con el estimador de credibilidad bayesiano.

2.2.5 Sistemas bonus-malus

El sistema bonus-malus es un sistema de tarificación en el que la prima inicial se va modelando a medida que se incorpora la experiencia de siniestralidad. Para ello se parte de un nivel x neutro; para niveles inferiores a x el asegurado entra en la escala bonus, mientras que para niveles superiores a x, el asegurado se incorpora a la escala malus.

Este tipo de sistemas está generalmente basado en el número de reclamaciones, y no en la cuantía, expresando la prima como una función del número medio de reclamaciones $\overline{\mathrm{x}},\ t\overline{\mathrm{x}} = \sum_{t=1}^{t} x_i$, y del periodo de tiempo t, representaremos la prima bonus-malus mediante $P_{BM}(\overline{\mathrm{x}},t)$.

Lo relevante de este sistema de tarificación es que un asegurado que en el periodo actual no presente reclamación se verá bonificado en el siguiente periodo mediante un descuento en la prima a pagar. Por el contrario, si experimenta reclamación se vera penalizado con un incremento de la prima. Luego, tendrán que verificarse las siguientes reglas de transición:

$$\frac{\partial P_{BM}(\overline{\mathrm{x}},t)}{\partial\overline{\mathrm{x}}} > 0, \qquad\qquad \frac{\partial P_{BM}(\overline{\mathrm{x}},t)}{\partial\overline{\mathrm{x}}} < 0.$$

2.2.5.1. Cálculo de primas bonus-malus. Método bayesiano

Una forma de obtener primas que cumplan las reglas de transición, utilizando la metodología bayesiana, consiste en dividir las prima Bayes entre la prima colectiva para los principios de cálculo de prima estudiados. Denotando ahora mediante $P(\mathrm{x},t)$ a la prima Bayes, la prima bonus-malus puede obtenerse como:

$$P_{BM}(\overline{\mathrm{x}},t) = \frac{P(\overline{\mathrm{x}},t)}{P'} \tag{2.21}$$

Donde P' representa la prima colectiva.

La mayoría de estos sistemas penalizan injustamente a determinados asegurados, haciéndoles pagar más de lo que realmente les corresponde. Por otro lado, en muchas ocasiones las bonificaciones establecidas son pequeñas, lo que puede acarrear serios problemas de competitividad y, en consecuencia, de equilibrio financiero a la compañía aseguradora. Además, es acostumbrado en el mercado de seguros de automóviles, que el asegurado cambie de compañía aseguradora buscando precios más competitivos, más si tenemos en cuenta que cuando un asegurado cambia de compañía se lleva consigo su historial de siniestralidad, pudiendo por tanto mantener su bonificación. Esto obviamente representa un problema para la compañía aseguradora, pues los asegurados situados en la clase bonus $(t\overline{\mathrm{x}} = 0)$ pueden optar por abandonar dicha compañía, el perder los ingresos que le reporta estos clientes supone no poder compensar el balance de la empresa, pues aunque los asegurados situados en la clase malus $(t\overline{\mathrm{x}} > 0)$ paguen más, es generalmente una población menos numerosa la que figura en estas clases.

Por otro lado, también pueden argumentarse en este otro sentido. La experiencia demuestra que los asegurados situados en la clase bonus prefieren pagar una prima ligeramente mayor ahora de modo que si incurren en una reclamación al año siguiente su prima no se incrementa demasiado.

Las ideas sobre la penalización fueron inicialmente propuestas por Lemaire (1979), donde el autor expone un procedimiento para reducir un porcentaje de incremento de la prima que permite mantener el ajuste presupuestario de la compañía aseguradora. Para ello se utilizara el principio exponencial, minimizando la diferencia entre la prima $P(\mathrm{x},t)$ y el valor del parámetro θ , sujeto a la restricción presupuestaria. Así, denominando m al número máximo de clases en la cartera, N_{x} al número de asegurados en dichas clases y $N = \sum_{x=0}^{m} N_x$, formalmente el problema a resolver tendrá la siguiente formulación:

$$\begin{cases} Min. \dfrac{1}{N}\displaystyle\sum_{x=0}^{m} N_{\overline{\mathbf{X}}} \int_{\Theta} \dfrac{1}{\alpha}\left[e^{\alpha\theta - e^{\alpha P(\overline{\mathbf{X}},t)}}\right]^{2} \pi(\theta / \overline{\mathbf{X}}) d\theta, \\ s.a \;\; \dfrac{1}{N}\displaystyle\sum_{x=0}^{m} N_{\overline{\mathbf{X}}} P(\overline{\mathbf{X}},t) = E(\theta). \end{cases} \qquad (2.22)$$

Mediante la formulación de este problema el asegurado minimiza la pérdida esperada de la operación de aseguramiento sujeta a la restricción que permite el equilibrio financiero.
La solución de este problema de optimización restringida se presenta en la siguiente proposición.

Proposición 2.6 La solución del problema de optimización (2.22) viene dada por:

$$P(\mathrm{x},t) = E(\theta) + \frac{1}{\alpha}\left[\frac{1}{N}\sum_{x=0}^{m} N_0 \log\int_{\Theta} e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta - \log\int_{\Theta} e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta\right]. \quad (2.23)$$

Demostración: El problema planteado equivale al problema de maximización, en donde se determinará la prima máxima que el asegurado deberá pagar en un tiempo determinado de modo que si incurre en una reclamación al año siguiente su prima no se incremente demasiado.

$$Máx. \frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} \int_{\Theta} \frac{1}{\alpha}\left\{1 - e^{\alpha[\theta - P(\overline{\mathbf{X}},t)]}\right\}\pi(\theta/\mathrm{x})d\theta,$$

$$s.a \;\; \frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} P(\mathrm{x},t) = E(\theta). \qquad (2.24)$$

La función lagrangiana es:

$$\psi = \frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} \int_{\Theta} \frac{1}{\alpha}\left\{1 - e^{\alpha[\theta - P(\overline{\mathbf{X}},t)]}\right\}\pi(\theta/\mathrm{x})d\theta - \beta\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} P(\overline{\mathbf{X}},t) - E(\theta)\right),$$

En la que β es el correspondiente multiplicador de Lagrange. Ahora,

$$\frac{\partial\psi}{\partial\beta} = 0 - \left[0 + 1\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} P(\overline{\mathbf{X}},t) - E(\theta)\right)\right],$$

$$\frac{\partial\psi}{\partial\beta} = 0 \Rightarrow E(\theta) = \frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m} N_{\overline{\mathbf{X}}} P(\overline{\mathbf{X}},t),$$

$$\psi=\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\left[\int_{\Theta}\frac{1}{\alpha}\pi(\theta/\mathrm{x})d\theta-\int_{\Theta}\frac{1}{\alpha}e^{\alpha[\theta-P(\overline{\mathbf{X}},t)]}\pi(\theta/\mathrm{x})d\theta\right]-\beta\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}P(\mathrm{x},t)-E(\theta)\right),$$

$$\frac{\partial\psi}{\partial P(\mathrm{x},t)}=\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\left[-\frac{1}{\cancel{\alpha}}\int_{\Theta}\pi(\theta/\mathrm{x})e^{\alpha[\theta-P(\overline{\mathbf{X}},t)]}(-\cancel{\alpha})d\theta\right]-\beta\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\right),$$

$$\frac{\partial\psi}{\partial P(\overline{\mathbf{X}},t)}=0\Rightarrow 0=\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\left[\int_{\Theta}\pi(\theta/\mathrm{x})e^{\alpha[\theta-P(\overline{\mathbf{X}},t)]}d\theta\right]-\beta\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\right)$$

$$\beta\left(\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\right)=\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}\left[\int_{\Theta}\pi(\theta/\mathrm{x})e^{\alpha[\theta-P(\overline{\mathbf{X}},t)]}d\theta\right]$$

$$\beta=\int_{\Theta}\pi(\theta/\mathrm{x})e^{\alpha[\theta-P(\overline{\mathbf{X}},t)]}d\theta$$

$$\beta=\int_{\Theta}e^{\alpha\theta}e^{-\alpha P(\overline{\mathbf{X}},t)}\pi(\theta/\mathrm{x})d\theta,\quad \overline{\mathbf{X}}=0,\dots,m.$$

Tomando logaritmos en ambos lados en la última igualdad y despejando P(x, t) se obtiene:

$$P(\mathrm{x},t)=\frac{1}{\alpha}\log\beta-\frac{1}{\alpha}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta. \tag{2.25}$$

Dividiendo ahora por N , multiplicando por $N_{\overline{\mathbf{X}}}$ y sumando para toda $\overline{\mathbf{X}}$ resulta:

$$\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}P(\mathrm{x},t)=\frac{1}{\alpha}\log\beta-\frac{1}{N\alpha}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta.$$

Finalmente, despejando de esta expresión $\frac{1}{\alpha}\log\beta,$:

$$\left[\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}P(\overline{\mathbf{X}},t)+\frac{1}{N\alpha}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta\right]=\frac{1}{\alpha}\log\beta$$

sustituyendo en (2.25)

$$P(\mathrm{x},t)=\left[\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}P(\mathrm{x},t)+\frac{1}{N\alpha}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta\right]-\frac{1}{\alpha}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta.$$

y teniendo en cuenta que $E(\theta)=\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{m}N_{\overline{\mathbf{X}}}P(\mathrm{x},t),$ se obtiene:

$$P(\mathrm{x},t)=\left[E(\theta)+\frac{1}{N\alpha}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta\right]-\frac{1}{\alpha}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta.$$

$$P(\mathrm{x},t)=E(\theta)+\frac{1}{\alpha}\left[\frac{1}{N}\sum_{\overline{\mathbf{X}}=0}^{\mathbf{m}}N_{\overline{\mathbf{X}}}\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta-\log\int_{\Theta}e^{-\alpha\theta}\pi(\theta/\mathrm{x})d\theta\right].\qquad\blacksquare$$

Ahora se puede construir una prima bonus-malus como el cociente:

$$P_{BM}(\overline{\mathbf{X}},t)=\frac{P(\overline{\mathbf{X}},t)}{P(0,0)}. \tag{2.26}$$

2.2.5.3 Método markoviano

Un sistema bonus-malus puede verse también como un proceso de markov, en el sentido de que el asegurado se mueve de un estado a otro en el tiempo. Veamos un ejemplo sencillo para ilustrar esta idea. Supongamos que una compañía de seguros divide su póliza en 3 clases $C_1, C_2\ y\ C_3$, en la que la prima para la clase C_i es menor que la prima de la clase C_j para *i* < *j.* El asegurado comienza en la clase C_1, si reclama por un siniestro por el que la compañía tenga que pagar una indemnización a un tercero pasará de la clase C_1 en la que actualmente está situado a la clase C_2 al principio del año siguiente. Si el asegurado no reclama permanecerá en la clase en la que estaba, esto es C_1. Ahora, si un asegurado esta en la clase C_2 pasará a la clase C_3 si reclama y volverá a la clase C_1 si no reclama finalmente, si el asegurado esta en la clase C_3 y experimenta reclamación permanecerá en esta clase el año siguiente, mientras que pasará a la clase C_1 si no reclama. De esta forma, el asegurado puede ver cada año bonificada o penalizada su prima si no experimenta o experimenta reclamación, respectivamente, a excepción si esta en la clase C_3 en la que permanecerá igual que estaba o se le bonificará.

Vamos a suponer además que la probabilidad de experimentar reclamación en un año es P y que no depende de cómo el asegurado llego a la clase en la que esta situado actualmente considerando el problema desde un punto de vista estocástico podemos decir que cada una de las *clases* corresponde a un estado de sistema, cuya matriz de transición de probabilidades es $\mathbf{P}=(P_{ij}),\ i,j=1,2,3,$ donde P_{ij} representa la probabilidad de transición del estado i al estado j. La matriz de transición para esta situación viene dada por :

$$P=\begin{bmatrix} q & p & 0 \\ q & 0 & p \\ q & 0 & p \end{bmatrix}$$

En la que $q=1-p$ es la probabilidad de no experimentar reclamación. Ahora P_{ij} es la probabilidad de transición de la clase C_i a la clase C_j. La matriz P es una matriz estocástica en la que cada fila representa una distribución de probabilidad de cada uno de los estados existentes. Luego $\sum_{j=1}^{3} p_{ij} = 1$, para todo i. El objetivo ahora es hallar la probabilidad de que los asegurados se encuentren en una clase determinada después de *t* años para ello podemos proceder de la siguiente forma, si suponemos que para $t = 0$, la probabilidad de que cada asegurado se sitúe en el estado $j = 1,2,3$ es e_j, y obviamente $\sum_{j=1}^{3} p_{ij} = 1$, entonces,

$$m(1) = (e_1, e_2, e_3)P = (q, e_1 p, (e_2 + e_3)p),$$

$$m(2) = m(1)P = (q, pq, p^2),$$

$$m(3) = m(2)P = m(1)P^2 = (q, pq, p^2),$$

$$m(4) = m(5) = ... = m(\infty) = (q, pq, p^2),$$

2.3 Modelos de Riesgo Colectivo e Individual

Lo ideal en estadística actuarial es trabajar con las distribuciones del número de reclamaciones y de la cantidad reclamada para obtener la función de distribución de la cantidad total reclamada.
En el *modelo colectivo* se considera que la cartera consta de un número desconocido de pólizas. En este modelo el número total de reclamaciones ocurridas en un período de tiempo es aleatorio, y se supone que las cuantías asociadas a las reclamaciones son independientes entre sí, idénticamente distribuidas e independiente del número de reclamaciones. Se supone también que cada riesgo o póliza puede dar lugar solamente a un siniestro.
El *modelo individual* considera un número n de pólizas, donde n no es aleatorio, se supone independencia entre las pólizas.

2.3.1 Modelo de riesgo colectivo

Para el modelo de riesgo colectivo se calculará la distribución de la cantidad total reclamada en un cierto periodo. Ahora vemos la cartera de clientes como un colectivo cuyo número es aleatorio. Representamos la cantidad total reclamada como una suma S de un número aleatorio de reclamaciones N, donde las cantidades reclamadas son $\{X_1, ..., X_N\}$. De este modo:

$$S = X_1 + X_2 + ... + X_N, \quad N = 0,1,2,... \qquad (2.30)$$

Donde por convenio $X = 0 \; si \; N = 0$. Con objeto de hacer el modelo tratable se establecen dos hipótesis:

1. Las variables aleatorias $X_1, X_2, ..., X_N$ son independientes e igualmente distribuidas con función de distribución $F_X(.)$
2. Las variables aleatorias $N, X_1, X_2, ..., X_N$ son mutuamente independientes.

Definición 2.11 Un modelo de riesgo colectivo representa la cantidad total reclamada como la suma S de un número aleatorio N de cantidades $X_1, ..., X_N$, *i.e.* $S = \sum_{i=1}^{N} X_i$, donde se supone que las X_i son variables aleatorias $i.i.d.$ e independientes de N.

Proposición 2.7 La variable aleatoria $S = \sum_{i=1}^{N} X_i$ tiene como función de distribución

$$F_S(x) = \sum_{n=0}^{\infty} p_n F^{*n}(x) \; ,$$

Donde $F(x) = \Pr(X \le x)$ es la función de distribución de las $X_j, p_n = \Pr(N = n)$ y $F^{*n}(x)$ es la convolución $n-ésima$ de la función de distribución de X.

Demostración:

Se obtiene a partir de:

$$F(x) = \Pr(X \le x) = E\left[\Pr(X \le x / N)\right]$$
$$= \sum_{n=0}^{\infty} p_n \Pr(X \le x / N = n) = \sum_{n=0}^{\infty} p_n F^{*n}(x)$$

Donde se ha hecho uso primeramente de la hipótesis de independencia entre X_i y N, luego de la hipótesis de $i.i.d.$ de X_i.

Se verifica que la función de densidad de probabilidad de S viene dada por:

$$f_S(x) = \sum_{n=0}^{\infty} P_n f^{*n}(x), \tag{2.31}$$

Donde $f(x)$ es la función de densidad de X. ■

Definición 2.12 Dado el modelo colectivo compuesto definido en (2.30), se denominan distribuciones primaria y secundaria a las distribuciones del número de reclamaciones y de la cantidad reclamada, respectivamente.

2.3.1.1 Resultados generales

Teorema 2.6 La función generatriz de momentos y la función generatriz de probabilidad de S definida en (2.30) vienen dadas por:

$$M_S(t) = M_N\left[\log M_X(t)\right], \tag{2.32}$$

$$P_S(z) = P_N\left[P_X(z)\right], \tag{2.33}$$

Respectivamente, siendo $M_N(t)$ y $P_N(z)$ la función generatriz de momentos y de probabilidades de la distribución primaria y $M_X(t)$ y $P_X(z)$ la función generatriz de momentos y de probabilidades de la distribución secundaria.

Demostración:

- $M_S(t) = M_N\left[\log M_X(t)\right]$

Sea $S = X_1 + X_2 + ... + X_N$, $N = 0,1,2,...$ donde $X_1, X_2, ..., X_n$ son variables independientes e idénticamente distribuidas, se tiene que:

$$M_s(t) = \sum_{n=0}^{\infty} \left(e^{t(x_1+x_2+...+x_N)} / N = n\right) p(N = n)$$

$$M_s(t) = \sum_{n=0}^{\infty} \left(e^{t(x_1+x_2+...+x_N)}\right) p(N = n)$$

$$M_s(t) = \sum_{n=0}^{\infty} \left(M_x(t)\right)^n p(N = n)$$

$$M_s(t) = E\left(\left(M_x(t)\right)^N\right)$$

$$M_s(t) = E\left(e^{N\log(M_X(t))}\right)$$

$$M_s(t) = M_N \log\left(M_X(t)\right)$$

- $P_S(z) = P_N\left[P_X(z)\right]$

Por la propiedad de la esperanza condicionada se tiene que:

$$P_s(Z) = E\left(Z^s\right) = E\left[E\left(Z^s / N\right)\right]$$

Condicionando $N = n$:

$$E\left(Z^s / N = n\right) = E\left(Z^{X_1+X_2+...+X_n}\right)$$

Por la independencia de las variables, se tiene:

$$E\left(Z^s / N = n\right) = \left[P_X(z)\right]^n$$

Por tanto:

$$E\left(Z^{s} / N\right)=\left[P_X(z)\right]^N$$

De donde:

$$E\left(t^{s}\right)=E\left\{\left[P_X(z)\right]^N\right\}$$

$$E\left(t^{s}\right)=\sum_{n=0}^{\infty}\left[P_X(z)\right]^n P_r(N=n)$$

$$E\left(t^{s}\right)=\sum_{n=0}^{\infty}\left[P_X(z)\right]^n P_r(N=n)$$

$$E\left(t^{s}\right)=P_N\left[P_X(z)\right] \qquad \blacksquare$$

Corolario 2.1 La esperanza y la varianza de S vienen dadas por:

$$E(S)=E(N)E(X), \tag{2.34}$$

$$Var(S)=E(N)Var(X)+E^2(X)Var(N), \tag{2.35}$$

respectivamente.

Demostración:

Puesto que:

$M_S^{'}(t)=M_N^{'}\left(\log M_X(t)\right)\dfrac{M_X^{'}(t)}{M_X(t)}$, tenemos que:

$M_S^{'}(0)=M_N^{'}\left(\log M_X(0)\right)\dfrac{M_X^{'}(0)}{M_X(0)}=M_N^{'}(0)M_X^{'}(0)=E(N)E(X)$, donde se ha tenido en cuenta que $M_X(0)=1$.

Por otro lado tenemos:

$M_S^{''}(t)=M_N^{''}\left(\log M_X(t)\right)\left[\dfrac{M_X^{'}(t)}{M_X(t)}\right]^2+M_N^{'}\left(\log M_X(t)\right)\dfrac{M_X^{''}(t)M_X(t)-M_X^{'}(t)^2}{M_X(t)^2}$, y por lo tanto:

$M_S^{''}(0)=M_N^{''}(0)M_X^{'}(0)^2+M_N^{'}(0)\left[M_X^{''}(0)-M_X^{'}(0)^2\right]$.

Luego $E\left(S^2\right)=E\left(N^2\right)E^2(X)+E(N)Var(N)$, y en definitiva:

$$\begin{aligned} Var(S)&=E\left(S^2\right)-E^2(S)=E\left(N^2\right)E^2(X)+E(N)Var(X)-E^2(N)E^2(X)\\ &=E(N)Var(X)+E^2(X)Var(N) \end{aligned} \qquad \blacksquare$$

Entre las distribuciones que se utilizan para obtener el modelo de riesgo colectivo e individual están la distribución de Poisson compuesto, exponencial compuesto, geométrica compuesto, binomial negativa compuesto y logarítmico compuesto.

Modelo de Poisson compuesto.

Cuando la distribución primaria es Poisson con parámetro $\theta > 0$ y la distribución secundaria tiene por función de distribución F, se dice que S tiene la distribución de Poisson compuesta con parámetros θ y F , entonces se verifica que:

$$E(S) = \theta E(X),$$
$$Var(S) = \theta E(X^2).$$

Proposición 2.8 La distribución de Poisson compuesta tiene coeficiente de asimetría positivo con valor

$$\gamma_1(S) = \frac{E(X^3)}{\sqrt{\theta\left[E(X^2)\right]^3}} \tag{2.36}$$

Demostración:

De (2.32) se deduce que $M_S(t) = e^{\theta\left[e^{\log M_X(t)-1}\right]} = e^{\theta\left[M_X(t)-1\right]}$

Se verifica que:

$$M_S'''(0) = E(S^3) = \theta E(X^3) + 3E(S)E(S^2) - 2E^3(S), \tag{2.37}$$

y por tanto

$$\begin{aligned} E\left[S - E(S)\right]^3 &= E\left[S^3 - 3S^2E(S) + 3SE^2(S) - E^3(S)\right] \\ &= E(S^3) - 3E(S)E(S^2) + 3E^2(S)E(S) - E^3(S) \end{aligned} \tag{2.38}$$

Sustituyendo ahora (2.37) en (2.38) se obtiene:

$$E\left[S - E(S)\right]^3 = \theta E(X^3) \tag{2.39}$$

Finalmente,

$$\gamma_1(S) = \frac{E\left[S - E(S)\right]^3}{\sqrt{Var(S)^3}} = \frac{E(X^3)}{\sqrt{\theta\left[E(X^2)\right]^3}} \succ 0 \qquad \blacksquare$$

Proposición 2.9 Si $\{S_i\}_{i=1}^{n}$ son variables aleatorias independientes con distribución de Poisson compuesta con parámetros (θ_i, F_i), entonces $S=\sum_{i=1}^{n} S_i$ tiene la distribución de Poisson compuesta con parámetros (Θ, F), con $\Theta=\sum_{i=1}^{n}\theta_i$ y $F=\frac{1}{\Theta}\sum_{i=1}^{n}\theta_i F_i(x)$

Demostración:

El resultado se prueba usando la función generatriz de momentos de S.

$$M_S(t)=E\left(e^{tS}\right)=\prod_{i=1}^{n} E\left(e^{tS_i}\right)=\prod_{i=1}^{n} e^{\theta_i\left[M_i(t)-1\right]}$$

$$=e^{\sum_{i=1}^{n}\theta_i\left[M_i(t)-1\right]}=e^{\Theta\left[\sum_{i=1}^{n}\frac{\theta_i M_i(t)}{\Theta}-1\right]}$$

■

Teorema 2.7 Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas con distribución exponencial de parámetro $\lambda>0$, entonces se verifica:

$$f^{*n}(x)=\frac{\lambda^n}{(n-1)!}x^{n-1}e^{-\lambda x},\quad n=1,2,\ldots \tag{2.40}$$

que es la función de densidad de una variable aleatoria gamma $(n, 1/\lambda)$.

Demostración:

Haciendo uso de convoluciones de funciones de densidad y de

$$f_Z(z)=\int_0^z f_X(x)f_Y(z-x)dx.$$

Se tiene:

$$f*f(x)=\int_0^x f(x-x_1)f(x_1)dx_1=\int_0^{x_1}\lambda e^{-\lambda(x-x_1)}\lambda e^{-\lambda x_1}dx_1$$
$$=\lambda^2 x e^{-\lambda x}$$

$$f*(f*f)(x)=\int_0^x \lambda e^{-\lambda(x-x_1)}\lambda^2 e^{-\lambda x}dx_1=\frac{\lambda^3}{2}x^2e^{-\lambda x}$$

Por inducción matemática, supongamos cierto que $f^{*n}(x)=\frac{\lambda^n}{(n-1)!}x^{n-1}e^{-\lambda x}$, entonces:

$$f^{*(n+1)}(x)=f*f^{*n}(x)=\int_0^x \lambda e^{-\lambda(x-x_1)}\lambda^n e^{-\lambda x_1}\frac{x_1^{n-1}}{(n-1)!}dx_1$$

$$=\frac{\lambda^{n+1}}{n!}x^n e^{-\lambda x}$$

■

Lo que prueba el resultado.

Modelo general exponencial.

Teorema 2.8 Si la distribución primaria tiene por función de probabilidad p_n y la distribución secundaria es exponencial de parámetro $\lambda \succ 0$, entonces la función de densidad, así como la función de distribución de $S = \sum_{i=1}^{N} X_i$ vienen dadas por:

$$f_S(x) = \frac{e^{-\lambda x}}{x} \sum_{n=1}^{\infty} p_n \frac{(\lambda \alpha)^n}{(n-1)!}, \quad x \succ 0, \tag{2.41}$$

$$F_S(x) = 1 - e^{-\lambda x} \sum_{j=0}^{\infty} \overline{P}_j \frac{(\lambda x)^j}{j!}, \quad x \succ 0$$

donde $\overline{P}_j = \sum_{n=j+1}^{\infty} p_n, \quad j=0,1,...$

Demostración:

La función de densidad se obtiene utilizando las fórmulas (2.31) y (2.40), se tiene:

$$f_s(x) = \sum_{n=1}^{\infty} P_n f^{*n}(x)$$

$$f_s(x) = \frac{e^{-\lambda x}}{x} \sum_{n=1}^{\infty} P_n \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, \quad n = 1,2,...$$

$$f_s(x) = \frac{e^{-\lambda x}}{x} \sum_{n=1}^{\infty} P_n \frac{(\lambda x)^n}{(n-1)!}$$

Para obtener la función de distribución hacemos:

$$F_S(x) = \int_0^x f_S(t)\,dt = \int_0^x \frac{e^{-\lambda t}}{t} \sum_{n=1}^{\infty} p_n \frac{(\lambda t)^n}{(n-1)!} dt$$

$$= \sum_{n=1}^{\infty} \frac{p_n}{(n-1)!} \lambda^n \int_0^x t^{n-1} e^{-\lambda t} dt$$

Ahora utilizando el resultado:

$$\int_0^x t^{n-1} e^{-\lambda t} dt = 1 - \sum_{j=0}^{n-1} \frac{(\lambda t)^j e^{-\lambda t}}{j!},$$

se deduce que

$$F_S(x) = 1 - e^{-\lambda x} \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} \sum_{n=j+1}^{\infty} p_n = 1 - e^{-\lambda x} \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} \qquad \blacksquare$$

Proposición 2.10 Si la distribución primaria es de tipo Poisson con parámetro $\theta > 0$ y la distribución secundaria es exponencial con parámetro $\lambda > 0$, entonces la función de densidad de probabilidad se S viene dada por:

$$f_S(x) = \frac{e^{-(\theta+\lambda x)}}{x} \sum_{n=1}^{\infty} \frac{(\lambda \theta x)^n}{n!(n-1)!}, \; x \succ 0 \qquad (2.42)$$

con $f_S(0) = e^{-\theta}$.

Nótese que en este caso S es una variable aleatoria mixta.

Demostración:

$$f_s(x) = \sum_{n=1}^{\infty} P_n f^{*n}(x)$$

$$f_s(x) = \sum_{n=1}^{\infty} \frac{e^{-\theta}\theta^n}{n!} \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$$

$$f_s(x) = \sum_{n=1}^{\infty} e^{-(\theta+\lambda x)} \frac{\theta^n \lambda^n x^{n-1}}{n!(n-1)!}$$

$$f_s(x) = \frac{e^{-(\theta+\lambda x)}}{x} \sum_{n=1}^{\infty} \frac{(\theta\lambda x)^n}{n!(n-1)!} \qquad x > 0$$

Modelo geométrica compuesto.

Proposición 2.11 Si la distribución primaria es geométrica con parámetro $\theta \in (0,1)$ y la distribución secundaria es exponencial con parámetro $\lambda \succ 0$, entonces la función de densidad de probabilidad de S viene dada por:

$$f_s(x) = \theta(1-\theta)\lambda e^{-\theta\lambda x}, \; x \succ 0 \qquad (2.43)$$

Siendo $f_s(0) = 0$

Demostración: puede obtenerse a partir de (2.41), o bien utilizando (2.31), (2.40) y el hecho de que $\Pr(N = n) = \theta(1-\theta)^n$, de donde resulta:

$$
\begin{aligned}
f_s(x) &= \sum_{n=1}^{\infty} \theta\left(1-\theta\right)^n \frac{\lambda^n e^{-\lambda x} x^{n-1}}{\left(n-1\right)!} \\
&= \frac{\theta}{x} e^{-\lambda x} \sum_{n=1}^{\infty} \frac{\left[\left(1-\theta\right)\lambda x\right]^n}{\left(n-1\right)!} \\
&= \frac{\theta}{x} e^{-\lambda x} \left(1-\theta\right)\lambda x \sum_{n=1}^{\infty} \frac{\left[\left(1-\theta\right)\lambda x\right]^{n-1}}{\left(n-1\right)!} \\
&= \theta\left(1-\theta\right)\lambda e^{-\theta\lambda x}
\end{aligned}
$$

2.4 Fórmula de recursión de Panjer

En la mayoría de las ocasiones resulta difícil obtener una expresión analítica para la distribución de S, puesto que desde el punto de vista actuarial resulta interesante calcular $P_r(S \geq s)$, es decir la probabilidad de que la cantidad total reclamada exceda a un cierto nivel s. De ahí que se hayan desarrollado métodos basados en algoritmos recursivos de cálculo numérico que permitan calcular o bien obtener un valor aproximado de dicha probabilidad.

Esta fórmula evita usar convoluciones o algún método de cálculo tedioso.

Una distribución discreta $\{P_n\}_{n=0}^{\infty}$ pertenece a la clase $(a,b,0)$ si

$$
\frac{P_n}{P_{n-1}} = a + \frac{b}{n}, \qquad n = 1,2,... \tag{2.45}
$$

Donde a y b son constantes y la probabilidad en 0, $p_{0.}$

Las distribuciones binomial, Poisson, geométrica y binomial negativa son las únicas distribuciones de esta clase.

La fórmula de recurrencia de Panjer permite evaluar la distribución de la cantidad total reclamada, cuando el número de reclamaciones pertenece a esta clase y la cantidad reclamada es una variable aleatoria que puede ser tanto de tipo discreto como continuo.

Proposición 2.14 Sea $X = \sum_{i=1}^{n} X_i$, con X_i variables aleatorias $i.i.d.$, entonces para todo $j,k \in N, \quad n = 1,2,...$ se verifica:

$$E\left(X_1 \middle| \sum_{i=1}^{n} X_i = j \right) = \frac{j}{n}, \tag{2.46}$$

$$\Pr\left(X_1 = k \middle| \sum_{i=1}^{n} X_i = j \right) = \frac{f(k)\, f^{*(n-1)}(j-k)}{f^{*n}(j)} \tag{2.47}$$

$$f^{*n}(x) = \frac{n}{ks} \sum_{j=1}^{s} j f^{*k}(j)\, f^{*(n-k)}(s-j) \tag{2.48}$$

Donde $f(x) = \Pr(X = k), \quad k = 0,1,\ldots$

Demostración: Puesto que X_i están idénticamente distribuidas, para $X = \sum_{i=1}^{n} X_i$ tenemos:

$$nE\left(X_1 \middle| X = j\right) = \sum_{k=1}^{n} E\left(X_k \middle| X = j\right) = E\left(\sum_{k=1}^{n} X_k \middle| X = j \right)$$
$$= E\left(X \middle| X = j\right) = j$$

Por otro lado, puesto que $X_1, X_2, \ldots$ son independientes:

$$\Pr\left(X_1 = k \middle| \sum_{i=1}^{n} X_i = j \right) = \frac{\Pr\left(X_1 = k, \sum_{i=2}^{n} X_i = j-k\right)}{\Pr\left(\sum_{i=1}^{n} X_i = j\right)}$$

$$= \frac{\Pr(X_1 = k)\Pr\left(\sum_{i=2}^{n} X_i = j-k\right)}{\Pr\left(\sum_{i=1}^{n} X_i = j\right)}$$

$$= \frac{f(k)\, f^{*(n-1)}(j-k)}{f^{*n}(j)}$$

Finalmente,

$$\Pr\left(X_1 + \ldots + X_k = j \middle| X_1 + \ldots + X_n = i\right) = \frac{\Pr\left(X_1 + \ldots + X_k = j, X_1 + \ldots + X_n = i\right)}{\Pr\left(X_1 + \ldots + X_n = i\right)}$$

$$= \frac{\Pr\left(X_1 + \ldots + X_k = j, X_{k+1} + \ldots + X_n = i - j\right)}{\Pr\left(X_1 + \ldots + X_n = i\right)}$$

$$= \frac{f^{*k}(j)\, f^{*(n-k)}(i-j)}{f^{*n(i)}},$$

De donde,

$$\frac{ki}{n} = E\left(X_1 + \ldots + X_k = j \middle| X_1 + \ldots + X_n = i\right)$$

$$= \frac{1}{f^{*n}(i)} \sum_{j=1}^{i} j f^{*k}(j)\, f^{*(n-k)}(i-j)$$

Teorema 2.9 (Fórmula de recursión de Panjer: versión discreta) La función de densidad de probabilidad de $S=\sum_{i=1}^{N} X_i$, siendo X_i variables aleatorias $i.i.d.$ e independientes de *N*, satisface la siguiente fórmula recursiva:

$$f_s(x)=\begin{cases} P_N\left(f(0)\right) & x=0 \\ \dfrac{1}{1-af(0)}\displaystyle\sum_{k=1}^{s}\left(a+\frac{bk}{x}\right)f(k)f_s(x-k), & x=1,2,\ldots \end{cases} \tag{2.49}$$

Siendo $f(x)$ la función de probabilidad de la variable aleatoria X discreta y positiva asociada a la cantidad de reclamación, y *N* pertenece a la clase $(a,b,0)$.

Demostración: Por el teorema de las probabilidades totales,

$$f_S(0)=\sum_{n=0}^{\infty}P_r(X_1+\ldots+X_n=0)P_r(N=n)$$

$$=\sum_{n=0}^{\infty}f(0)^n P_r(N=n)=P_N\left(f(0)\right)$$

Por otro lado, utilizando (2.45) y (2.46) tenemos

$$f_S(x)=\sum_{n=0}^{\infty}P_n f^{*n}(x)=p_o f^{*0}(0)+\sum_{n=1}^{\infty}P_n f^{*n}(x)$$

$$=\sum_{n=1}^{\infty}\left(a+\frac{b}{n}\right)p_{n-1}f^{*n}(x)$$

$$=\sum_{n=1}^{\infty}\left[a+bE\left(\frac{X_1}{j}\sum_{i=1}^{n}X_j=j\right)\right]p_{n-1}f^{*n}(x)$$

$$=\sum_{n=1}^{\infty}\left[a+b\sum_{k=0}^{j}\frac{k}{j}P_r\left(X_1=k\sum_{i=1}^{n}X_j=j\right)\right]p_{n-1}f^{*n}(x).$$

Usando ahora (2.47) resulta:

$$f_s(x)=\sum_{n=1}^{\infty}\left[a+b\sum_{k=0}^{x}\frac{k}{x}\frac{f(k)f^{*(n-1)}(x-k)}{f^{*n}(x)}\right]p_{n-1}f^{*n}(x)$$

$$=\sum_{n=1}^{\infty}\sum_{k=0}^{x}\left(a+\frac{bk}{x}\right)f(k)f^{*(n-1)}(x-k)p_{n-1}$$

$$=\sum_{k=0}^{x}\left(a+\frac{bk}{x}\right)f(k)f^{*(n-1)}(x-k)p_{n-1}$$

$$=af(0)g(x)+\sum_{k=1}^{x}\left(a+\frac{bk}{x}\right)f(k)f_s(x-k)$$

Por tanto,

$[1-af(0)]f_s(x)=\sum_{k=1}^{x}\left(a+\frac{bk}{x}\right)f(k)f_s(x-k)$, de donde despejando $fs(x)$ se obtiene el resultado.

Teorema 2.10 Los momentos de orden $r, r=1,2,...$ de la variable aleatoria $S=\sum_{i=1}^{N}X_i$ pueden obtenerse de manera recursiva como:

$$E(S^r)=\frac{1}{1-a}\sum_{i=0}^{r-1}\left[a\binom{r}{i}+b\binom{r-1}{i}\right]E(S^i)E(X_1^{r-i})$$

Demostración:

En efecto:

$$\begin{aligned}E(S^r)&=\sum_{s=0}^{\infty}S^r f_s(s)\\&=\frac{1}{1-af(0)}\sum_{k=1}^{\infty}s^r\sum_{k=1}^{s}\left(a+\frac{bk}{s}\right)f(k)f_s(s-k)\\&=\frac{1}{1-af(0)}\sum_{k=1}^{\infty}\sum_{s=k}^{\infty}(as^r+bk^{r-1})f(k)f_s(s-k)\end{aligned}$$

Haciendo $s=k+t$ y utilizando el desarrollo del binomio de Newton resulta:

$$\begin{aligned}E(S^r)&=\frac{1}{1-af_s(0)}\sum_{k=1}^{\infty}\left[a(t+k)^r+bk(t+k)^{r-1}\right]f_s(t)\\&=\frac{1}{1-af_s(0)}\sum_{k=1}^{\infty}f(k)\sum_{t=0}^{\infty}\left[a\sum_{i=0}^{r}\binom{r}{i}t^ik^{r-i}f_s(t)+bk\sum_{i=0}^{r-1}\binom{r-1}{i}t^ik^{r-1-i}(t+k)^r f_s(t)\right]\\&=\frac{1}{1-af_s(0)}\left[a\sum_{i=0}^{r}\binom{r}{i}\left(\sum_{t=0}^{\infty}t^if_s(t)\right)\sum_{k=1}^{\infty}k^{r-i}f(k)+b\sum_{i=0}^{r-1}\binom{r-1}{i}\left(\sum_{t=0}^{\infty}t^if_s(t)\right)\sum_{k=1}^{\infty}k^{r-i}f(k)\right]\\&=\frac{1}{1-af_s(0)}\left\{\sum_{i=0}^{r-1}\left[a\binom{r}{i}+b\binom{r-1}{i}\right]E(S^i)E(X^{r-i})+aE(S^r)\sum_{k=1}^{\infty}f(k)\right\}\end{aligned}$$

Teniendo ahora en cuenta que $\sum_{k=1}^{\infty}f(k)=1-f(0)$ y después de algunos cálculos adicionales se obtiene el resultado.

Teorema 2.11 Se verifica que,

$$\int_0^x yf(y)\, f^{*n}(x-y)\,dy = \frac{x}{n+1} f^{*(n+1)}(x), \quad n=1,2,\ldots \quad (2.50)$$

$$\int_0^x f(y)\, f^{*n}(x-y)\,dy = f^{*(n+1)}(x), \quad n=1,2,\ldots \quad (2.51)$$

Teorema 2.12 ***(Fórmula de recursión de Panjer: versión continua)*** La función de densidad de $S=\sum_{i=1}^{N} X_i$, siendo X_i variables aleatorias $i.i.d.$ e independientes de *N*, satisface la siguiente fórmula recursiva:

$$f_s(x)=\begin{cases} P_N(f_s(0)), & x=0 \\ p_1 f(x)+\int_0^x \left(a+\frac{by}{x}\right) f(y) f_s(x-y)dy, & x>0 \end{cases} \quad (2.52)$$

Siendo $f(x)$ la función de densidad de la variable aleatoria X positiva asociada a la cantidad de reclamación y N pertenece a la clase $(a,b,0)$.

Demostración:

En efecto, se tiene:

$$f_s(x)=\sum_{n=1}^{\infty} P_n f^{*n}(x) = p_1 f(x)+\sum_{n=2}^{\infty} p_n f^{*n}(x)$$

$$= p_1 f(x)+\sum_{n=1}^{\infty} P_{n+1} f^{*(n+1)}(x)$$

Utilizando ahora (2.45) tenemos:

$$g_s(x)= p_1 f(x)+\sum_{n=1}^{\infty} p_n \left(a+\frac{b}{n+1}\right) f^{*(n+1)}(x)$$

$$= p_1 f(x)+\sum_{n=1}^{\infty} p_n \left(af^{*(n+1)}(x)+\frac{b}{x} f^{*(n+1)}(x)\frac{x}{n+1}\right)$$

Utilizando ahora (2.50) tenemos:

$$f_s(x)= p_1 f(x)+\sum_{n=1}^{\infty} p_n \left[af^{*(n+1)}(x)+\frac{b}{x}\int_0^x yf(y) f^{*n}(x-y)dy\right],$$

y utilizando (2.51) se obtiene:

$$f_s(x) = p_1 f(x) + \sum_{n=1}^{\infty} p_n \int_0^{\infty} \left(a + \frac{by}{x} \right) f(y) f^{*n}(x-y) dy$$

$$= p_1 f(x) + \int_0^{\infty} \left(a + \frac{by}{x} \right) f(y) \sum_{n=1}^{\infty} p_n f^{*n}(x-y) dy$$

$$= p_1 f(x) + \int_0^{x} \left(a + \frac{by}{x} \right) f(y) f_s(x-y) dy$$

Teorema 2.13 La función de densidad de $S = \sum_{i=1}^{N} X_i$, siendo X_i variables aleatorias $i.i.d.,$ e independientes de N , satisface la siguiente fórmula recursiva:

$$f_s(x) = \frac{1}{1 - af(0)} \left[\sum_{k=1}^{x} \left(a + \frac{bk}{x} \right) f(k) f_s(x-k) + \left(p_1 - (a+b) p_0 f(x) \right) \right] \qquad (2.53)$$

Donde la variable aleatoria N pertenece a la clase $(a,b,1)$, siendo a y b constantes y la probabilidad en 1, p_1.

Definición 2.13 Una distribución discreta se dice que pertenece a la clase $R2(a,b,c)$ si su función de probabilidad puede calculare recursivamente a partir de la expresión:

$$P_n = \left(a + \frac{b}{n} \right) P_{n-1} + \frac{c}{n} P_{n-2}, \quad n = 1,2,... \qquad (2.54)$$

Teorema 2.14 Si N pertenece a la clase $R2(a,b,c)$, entonces la función de densidad de $S = \sum_{i=1}^{N} X_i$, siendo X_i variables aleatorias $i.i.d.,$ e independientes de N , satisface la siguiente fórmula recursiva:

$$f_s(x) = \frac{1}{1 - af(0)} \sum_{k=1}^{x} \left[\left(a + \frac{bk}{x} \right) f(k) + \frac{ck}{2x} f(k)^{*2} \right] f_s(x-k)$$

Demostración: Utilizando (2.46), (2.48) y (2.54) se obtiene:

$$f_s(x) = \sum_{n=1}^{\infty} \left[\left(a + \frac{b}{n} \right) p_{n-1} + \frac{c}{n} p_{n-2} \right] f^{*n}(x)$$

$$= af(0)f_s(x)+\sum_{k=1}^{x}\left(a+\frac{bk}{x}\right)f(k)f_s(x-k)+\sum_{n=1}^{\infty}\frac{c}{n}p_{n-2}f^{*n}(x)$$

$$= af(0)f_s(x)+\sum_{k=1}^{x}\left(a+\frac{bk}{x}\right)f(k)f_s(x-k)+\sum_{n=1}^{\infty}\frac{c}{n}p_{n-2}\left[\frac{n}{2x}\sum_{k=1}^{x}kf^{*2}(s)f^{*(n-2)}(x-k)\right]$$

$$= af(0)f_s(x)+\sum_{k=1}^{x}\left[\left(a+\frac{bk}{x}\right)f(k)+\frac{c}{2x}kf^{*2}(x)\right]f_s(x-k)$$

Ahora, teniendo en cuenta que $p_{-1}=0$ y despejando $fs(x)$ se obtiene el resultado.

Definición 2.14 Dada la variable aleatoria S son distribución $F_s(x)$ y función generatriz de momentos $M(t)$, para cualquier número real h la función de distribución $F_s(x)$ dada por

$$dF_s(x)=\frac{1}{M(h)}e^{hx}dF(x) \qquad (2.55)$$

Proposición 2.15 La función generatriz de momentos de $F_s(x)$ viene dada por

$$M(t)=\frac{M(t+h)}{M(h)} \qquad (2.56)$$

Definición 2.15 Un modelo de riesgo individual representa la cantidad total reclamada como la suma de las cantidades pagadas en cada una de las pólizas o contratos que componen la cartera de riesgos. Así, $S=\sum_{j=1}^{n}X_j, donde\ X_j$ es la cantidad asegurada (pagada) en la póliza $j-ésima$ y en el que se supone que $X_1,...,X_n$ son variables aleatorias independientes.

Teorema 2.15 (Fórmula de recursión de De Pril) Sea S la cantidad total reclamada con función de probabilidad $F_s(x)$. Sea también:

$$a(j,k)=\left(\frac{q_j}{p_j}\right)^k, \qquad (2.57)$$

$$A(i,k)=(-1)^{k+1}i\sum_{j=1}^{b}c_{ij}a(j,k) \qquad (2.58)$$

Entonces se verifica la siguiente recursión:

$$f_s(0)=\prod_{j=1}^{b}p_j^{cj}, \qquad (2.59)$$

$$f_s(x)=\frac{1}{x}\sum_{i=1}^{mín\{a,x\}} A(i,k)\sum_{k=1}^{[x/i]} f_s(x-ki), \quad x=1,2,\ldots,m, \qquad (2.60)$$

Siendo

$[x/i]=máx\{n\in N : n\leq x/i\}.$

Demostración: obsérvese en primer lugar que la función generatriz de probabilidad de la variable aleatoria cantidad reclamada por un asegurado con tasa de mortalidad q_j y suma asegurada i es $P_{ij}(u)=1-q_j+q_{ij}u^i$.

Ahora, la función generatriz de probabilidad de la variable aleatoria S es $p_s(u)=\sum_{s=0}^{m} f_s(s)s^u = f_s(0)+\sum_{s=1}^{m} f_s(s)u^s$.

Puesto que $X_1, X_2,\ldots, X_n$ son independientes, tenemos:

$$p_s(u)=P_{X_1,X_2,\ldots,X_n}(u)=\prod_{i=1}^{n} P_{X_i}(u)$$

$$p_s(u)=\prod_{i=1}^{a}\prod_{j=1}^{b}\left[P_{X_{ij}}(u)\right]^{c_{ij}} = \prod_{i=1}^{a}\prod_{j=1}^{b}\left(p_j+q_ju^i\right)^{c_{ij}}$$

Haciendo ahora $u=0$ resulta $P_s(0)=\prod_{j=1}^{b} p_j^{cj} = f_s(0)$. Por otro lado tenemos:

$\frac{d}{du}\log P_s(u)=\frac{P_s'(u)}{P_s(u)}=\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij}iu^{i-1}\frac{q_j}{p_j+q_ju^i}$, de modo que:

$$
\begin{aligned}
uP_s'(u) &= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij} i u^i \frac{q_j}{p_j + q_j u^i} \\
&= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij} i u^i \frac{q_j}{p_j}\left(1+\frac{q_j}{p_j}u^i\right)^{-1} \\
&= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij} i u^i \frac{q_j}{p_j}\sum_{k=0}^{\infty}(-1)^k\left(\frac{q_j}{p_j}u^i\right)^k \\
&= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij} i u^i \frac{q_j}{p_j}\sum_{k=1}^{\infty}(-1)^{k-1}\left(\frac{q_j}{p_j}u^i\right)^{k-1} \\
&= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} c_{ij} i \sum_{k=1}^{\infty}(-1)^{k-1}\left(\frac{q_j}{p_j}\right)^k u^{ik} \\
&= P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} u^{ki}(-1)^{k-1} i \sum_{k=1}^{\infty} c_{ij} a(j,k)
\end{aligned}
$$

Definiendo:

$$
A(i,k) = \begin{cases} (-1)^{k-1} i \sum_{j=1}^{b} c_{ij} a(j,k), & i = 1, 2, \ldots, a, \\ 0, & en\ otro\ caso, \end{cases}
$$

tenemos:

$$
uP_s'(u) = P_s(u)\sum_{i=1}^{a}\sum_{j=1}^{b} A(i,k) u^{ki} .
$$

Puesto que $P_s(u) = \sum_{x=0}^{\infty} u^x f_s(x)$, se tiene que $uP_s'(u) = \sum_{x=0}^{\infty} x u^x f_s(x)$, luego:

$$
\sum_{x=0}^{\infty} x u^x f_s(x) = \sum_{x=0}^{\infty} u^x f_s(x) \sum_{i=1}^{a}\sum_{k=1}^{\infty} A(i,k) u^{ki} .
$$

Ahora el coeficiente de u^x en el lado izquierdo es $xf_s(x)$, mientras que en el lado derecho se obtiene sumando $f_s(x - ki)$ sobre todo i y k tal que $1 \le ik \le x.$ Luego:

$$
xf_s(x) = \sum_{i=1}^{x}\sum_{k=1}^{[x/i]} f_s(x-ik) A(i,k), \text{ de donde:}
$$

$$
xf_s(x) = \frac{1}{x}\sum_{i=1}^{x}\sum_{k=1}^{[x/i]} f_s(x-ik) A(i,k), \quad x = 1, 2, \ldots
$$

Finalmente, puesto que $A(i,k) = 0 \ para\ i \succ a$ resulta:

$$x f_s(x) = \frac{1}{x} \sum_{i=1}^{\min\{a,x\}} \sum_{i=1}^{\left[x/i \right]} f_s(x - ik) A(i,k), \quad x = 1, 2, \ldots$$

CAPITULO 3

ANÁLISIS PROSPECTIVO DEL NÚMERO DE RECLAMACIONES Y LA CANTIDAD TOTAL RECLAMADA

3.1. INTRODUCCIÓN

En este capítulo se realizará una aplicación de la teoría desarrollada en los capítulos 1 y 2 a una base de datos que corresponden a un grupo de asegurados de la Caja Mutual de los empleados del Ministerio de Educación correspondiente al período 2004 - 2009.
Para efectos expositivos se han ajustados los datos, planteando la siguiente situación:
Una compañía de seguros cuenta con un número de asegurados que conforman una póliza en una cartera de seguros, identificando las siguientes variables:

- *Año*
- *Número de reclamaciones durante un año determinado.*
- *Cantidad total reclamada durante un año determinado.*

En primer lugar se hará un análisis de los datos correspondiente al número de reclamaciones, siendo que la distribución Poisson es una de las distribuciones que habitualmente se utiliza para el ajuste de datos relativos al número de reclamaciones, será dicha distribución la que supondremos para los datos y posteriormente se hará una prueba de bondad de ajuste de la distribución supuesta. Seguidamente, de manera análoga se hará el análisis de los datos correspondiente a la cantidad total reclamada.

De la teoría desarrollada en el capítulo 1, se aplicará el análisis bayesiano para determinar *la distribución a priori, la distribución a posteriori, la distribución predictiva, los intervalos bayesianos de credibilidad para un parámetro desconocido θ, que para el caso en estudio corresponde al número medio de reclamaciones y a la cantidad media reclamada.* Así como también se realizará *las pruebas de hipótesis bayesianos para el parámetro desconocido θ.*
Lo ideal en estadística actuarial es trabajar con las distribuciones del número de reclamaciones y de la cantidad total reclamada para obtener la función de distribución total reclamada, para ello

se aplicará *la teoría de la credibilidad para calcular la prima neta, sistemas bonus malus*, *el modelo de riesgo colectivo y el modelo de riesgo individual*, teoría desarrollada en el capítulo 2.

3.2. DESCRIPCIÓN DEL PROGRAMA FIRSTBAYES. Versión 1.3

El programa ***firstbayes*** ha sido diseñado para facilitar la enseñanza y aprendizaje de la estadística bayesiana básica. Permite la implementación de varios modelos estadísticos simples pero de gran aplicación.

El programa incluye cuatro de los principales modelos conjugados en el análisis bayesiano, estos son: beta-binomial, gamma-Poisson y normal con varianza conocida. Tanto la distribución a priori, como la de verosimilitud pueden ser representadas simultáneamente a través de la herramienta triplot.

El programa también permite el análisis descriptivo de trece familias de distribuciones. Dichas distribuciones son: chi-cuadrado x^2, $\mathcal{F}$ de Snedecor, geométrica, hipergeométrica, inversa x^2, binomial negativa, normal, Poisson y *t* de Student.

Del análisis bayesiano se obtienen las distribuciones predictivas, que permiten aproximar los valores futuros de la variable de interés.

A.1. TRATAMIENTO DE LOS DATOS.

Firstbayes permite cargar, manipular y crear nuevos datos. La siguiente figura A.1 muestra la pantalla inicial del programa.

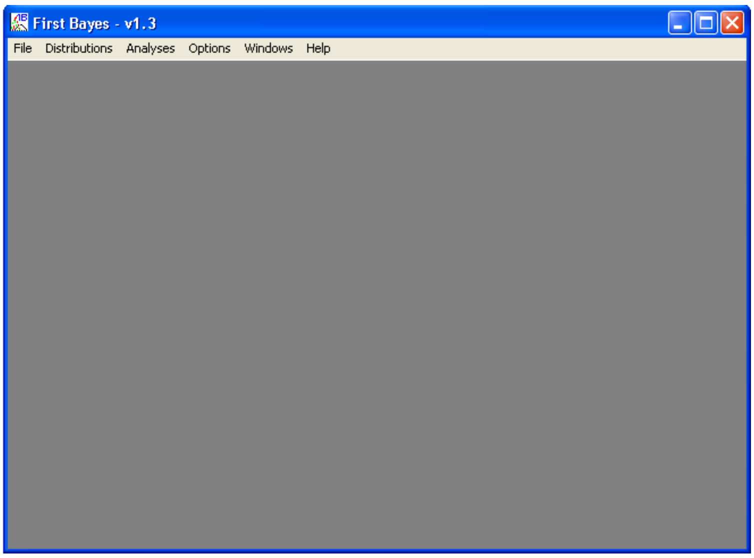

Figura A.1. *Pantalla inicial del programa.*

A la ventana de datos se accede a través del menú File de la barra de herramientas: Los datos son cargados por defecto de la carpeta C:\1b\. Se puede modificar dicho directorio sin más que seleccionar la opción change. La figura A.2 muestra el menú correspondiente al tratamiento de datos.

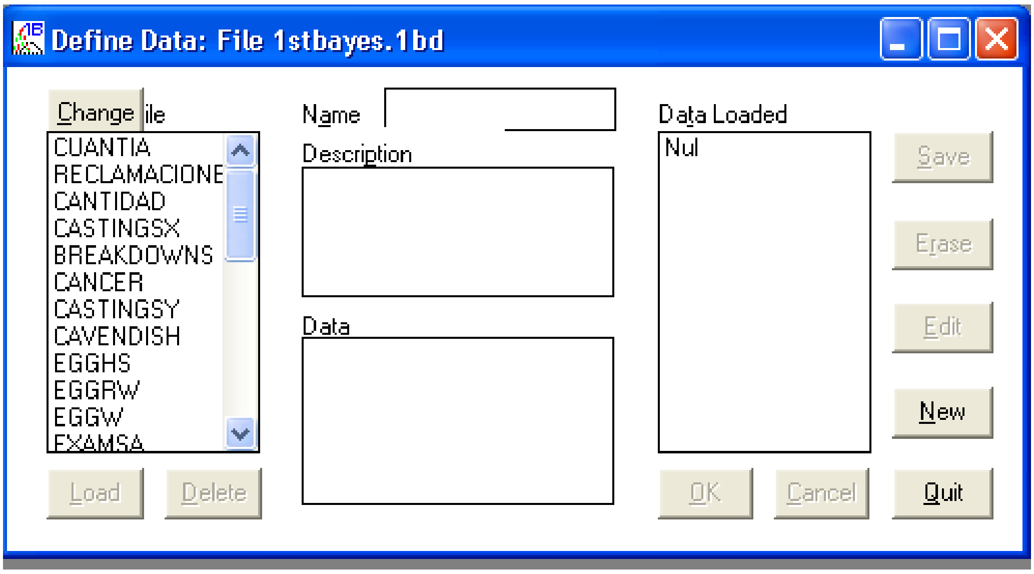

Figura A.2. *Menú para el tratamiento de datos.*

A.2. CARGAR DATOS.

El programa ***Firstbayes*** incluye un listado de bases de datos que pueden ser utilizadas como ejemplos ilustrativos de los diferentes modelos propuestos. Para cargar una de las bases de datos se debe seguir el siguiente procedimiento:

1. *En primer lugar debemos marcar la base de datos que queremos cargar, por ejemplo "CASTINGX". Al marcar dicha base de datos se activa la pantalla central de datos. Ella muestra la siguiente información: el nombre, una breve descripción de los datos y los propios datos.*
2. *Para poder utilizar los datos en los modelos posteriores debemos cargar dichos datos en la memoria. Para ello debemos seleccionar la opción Load. En la parte izquierda de la ventana se muestra el listado de los datos cargados en memoria.*

La figura A.3 muestra los datos cargados que corresponden a “CASTINGX”.

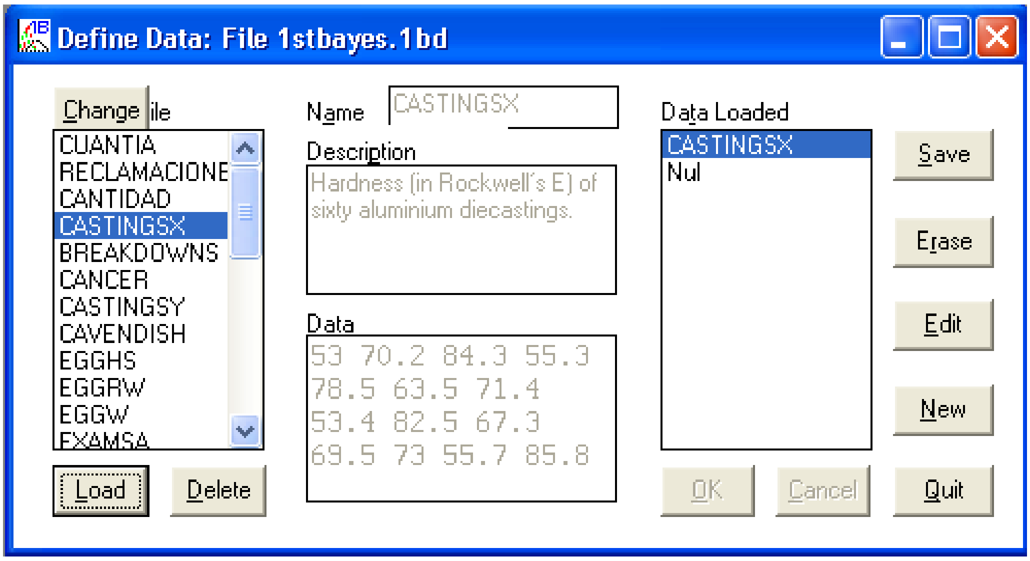

Figura A.3. *Datos cargados.*

A.3. CREAR UNA NUEVA BASE DE DATOS

El programa permite también la incorporación de nuevos datos. Para ello se debe de seguir los siguientes pasos:

1. *Para crear nuevos datos debemos señalar el icono* New *situado en la esquina inferior derecha de la ventana de datos.*
2. *De esta forma se activa la ventana central. Esto nos permitirá escribir el nombre de la base de datos, una breve descripción y los datos. Los datos deben estar separados por un espacio. Al finalizar debemos presionar* Ok.
3. *Los datos se guardan momentáneamente en memoria. Para poder utilizar los datos se debe confirmar su grabación a través del icono* Save. *Los datos quedarán guardados para poder ser utilizados posteriormente.*

3.3. ANÁLISIS BAYESIANO.

La siguiente tabla muestra el número de reclamaciones de una cartera de seguros y la cuantía de seguros pagados por una aseguradora en intervalos de 1 año, es decir, entre el año 2004 y el año 2009.

AÑO	NÚMERO DE RECLAMACIONES	MONTO PAGADO (MILES DE DÓLARES)
2004	686	760,967
2005	748	965,224
2006	666	781,384
2007	707	907,371
2008	658	838,145
2009	423	955,211
TOTAL	3,888	5,208,302

TABLA 1. *Número de reclamaciones y monto pagado en el período 2004 - 2009.*

ANÁLISIS DESCRIPTIVO DE LOS DATOS

Primeramente se hará un análisis de los datos correspondiente al número de reclamaciones.

En el gráfico de barras se puede observar también que el número de reclamaciones entre el año 2004 y 2009 tiende a permanecer constante, por lo que podemos suponer que los datos se ajustan a la distribución Poisson.

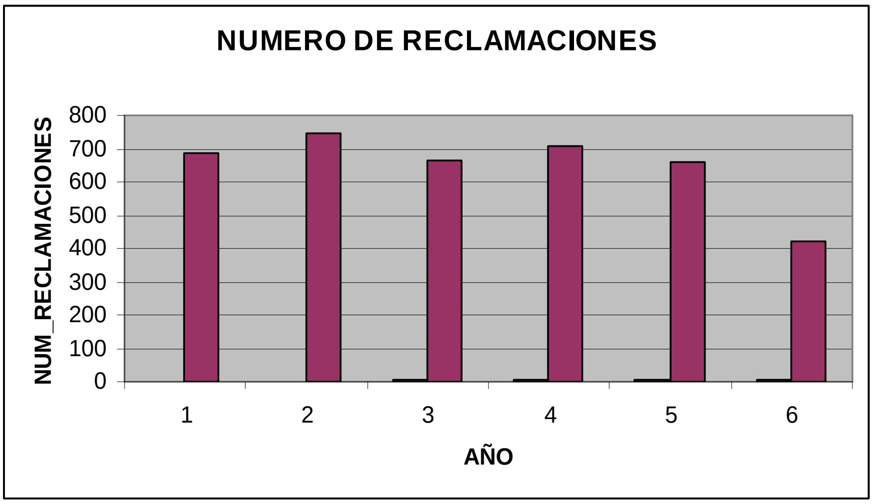

GRAFICO 1. *Gráfico correspondiente al número de reclamaciones*

Para determinar que los datos se ajustan a la distribución Poisson se contrastará las hipótesis siguientes:

$$H_0 : \mathbf{X} \sim p(\lambda)$$
$$H_1 : \mathbf{X} \nsim p(\lambda)$$

Para realizar el contraste de hipótesis se realizará la *prueba de bondad de ajuste de Kolmogorov-Smirnov* para los datos de la muestra. El siguiente resultado se obtiene con el programa estadístico SPSS

One-Sample Kolmogorov-Smirnov Test		
		NUM_RECLAMACIONES
N		6
Poisson Parameter[a,,b]	Mean	648.00
Most Extreme Differences	Absolute	0.481
	Positive	0.167
	Negative	-0.481
Kolmogorov-Smirnov Z		1.178
Asymp. Sig. (2-tailed)		0.125
a. Test distribution is Poisson.		
b. Calculated from data.		

TABLA 2. *Prueba de bondad de ajuste Kolmogorov-Smirnov*

De los resultados mostrados en la tabla 2, se observa que el nivel de significancia para el estadístico de prueba es 0.125 > 0.05; por lo tanto, se acepta la hipótesis nula, es decir, los datos que corresponden al número de reclamaciones se ajustan a la distribución Poisson.

Como el número de reclamaciones se ajustan a la distribución Poisson se hará un análisis bayesiano de los datos con el programa ***FirstBayes***. Para ello, se creará un fichero en dicho programa que permitirá realizar diferentes cálculos e interpretaciones importantes de los datos.

3.3.1. ANÁLISIS BAYESIANO PARA EL NÚMERO DE RECLAMACIONES.

Con base a lo desarrollado en los apartados 1.11, 1.13, 1.14, 1.15, 1.16, 1.17 del capítulo 1, se realizará el análisis bayesiano para el número de reclamaciones con el objetivo de:

1. *La distribución a posteriori para el número de reclamaciones a partir de la distribución a priori. La distribución a priori se especifica a partir de los datos de la muestra.*
2. *Obtener Intervalos bayesiano de credibilidad para el parámetro λ.*
3. *Realizar el test de hipótesis bayesianos para el parámetro λ*
4. *Realizar predicciones para el número de reclamaciones.*

En la figura 1 se puede observar el fichero creado en el programa ***FirstBayes*** correspondiente al número de reclamaciones, el fichero ha sido guardado con el nombre de RECLAMACIONES.

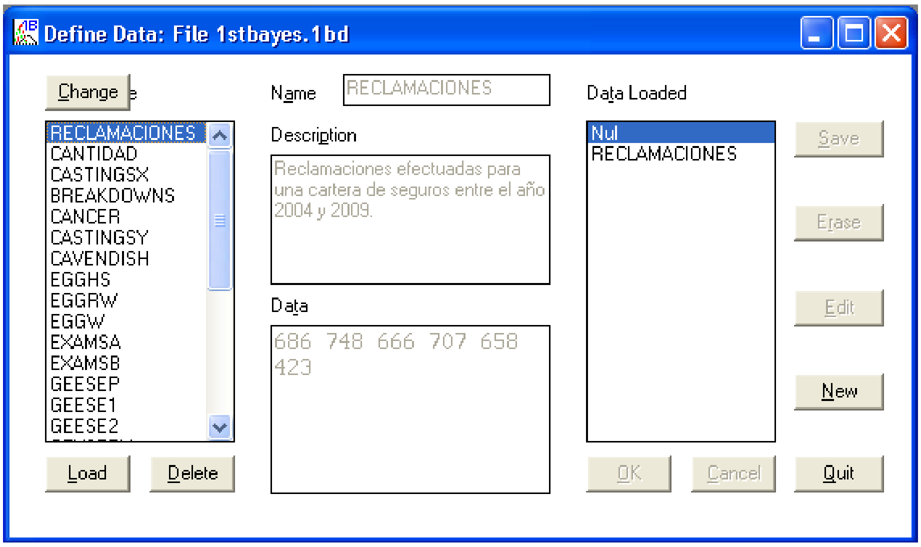

FIGURA 1. *Fichero creado en FirstBayes para el número de reclamaciones.*

Primeramente se obtendrá una distribución a posteriori de los datos bajo un modelo de Poisson $P(\lambda)$.

En el apartado 1.17 del capítulo 1, se presenta la siguiente tabla que resume algunas distribuciones a priori conjugadas respecto a una verosimilitud dada.

Verosimilitud Distribución a priori	Distribución a posteriori
$X \sim \mathcal{P}(\theta)$ $\theta \sim \mathcal{G}\left(a, \frac{1}{b}\right)$	$\mathcal{G}\left(a + n\bar{x}, \frac{1}{(b+n)}\right)$
$X \sim \mathcal{P}(\theta)$ $\theta \sim \mathcal{IG}(\mu, \beta)$	$\mathcal{GIG}\left(n\bar{x} - \frac{1}{2}, \mu\sqrt{\frac{1}{2\beta n+1}}, \frac{\beta}{2\beta n+1}\right)$
$X \sim \mathcal{BN}(r, \theta)$ $\theta \sim \mathcal{B}e(a, b)$	$\mathcal{B}(a + nr, b + n\bar{x})$
$X \sim \mathcal{B}(m, \theta)$ $\theta \sim \mathcal{B}e(a, b)$	$\mathcal{B}(a + n\bar{x}, b + mn - n\bar{x})$
$X \sim \mathcal{G}\left(\alpha, \frac{1}{\theta}\right)$ $\theta \sim \mathcal{G}\left(a, \frac{1}{b}\right)$	$\mathcal{G}\left(n\alpha + a, \frac{1}{(n\bar{x} + b)}\right)$
$X \sim \mathcal{N}(\theta, \sigma^2)$ $\theta \sim \mathcal{N}(a, r^2)$	$\mathcal{N}\left(\frac{a\sigma^2 + n\bar{x}r^2}{\sigma^2 + tr^2}, \frac{\sigma^2 r^2}{\sigma^2 + tr^2}\right)$

TABLA 3. *Distribuciones a priori conjugadas*

De acuerdo a la información presentada en la tabla 3 se tiene que para una distribución Poisson de parámetro λ, λ tiene una distribución a priori gamma $G(\alpha,\beta)$ para los datos proporcionados, es decir, se tiene que $G(648,1)$.

Se tiene que $n=6$ con $n\bar{x}=3888, \alpha=648$ y $\beta=1$ de donde la densidad a posteriori es $G\left(a+n\bar{x}, \frac{1}{(b+n)}\right)$, es decir, $G\left(4536, \frac{1}{7}\right)$. Utilizando la opción *Poisson simple* del menú *Análisis* de *FirstBayes* se obtienen los resultados siguientes.

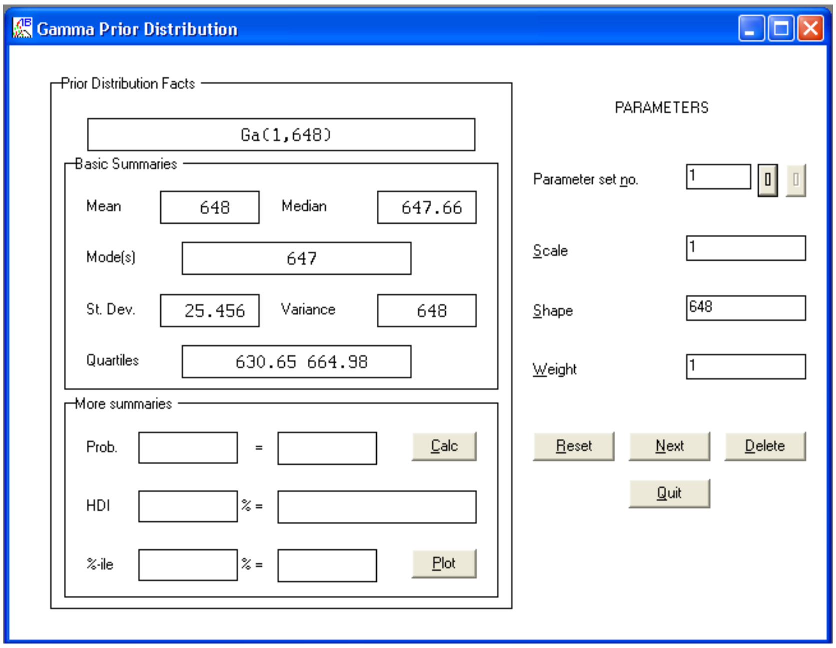

FIGURA 2. *Densidad a Priori para el número de reclamaciones.*

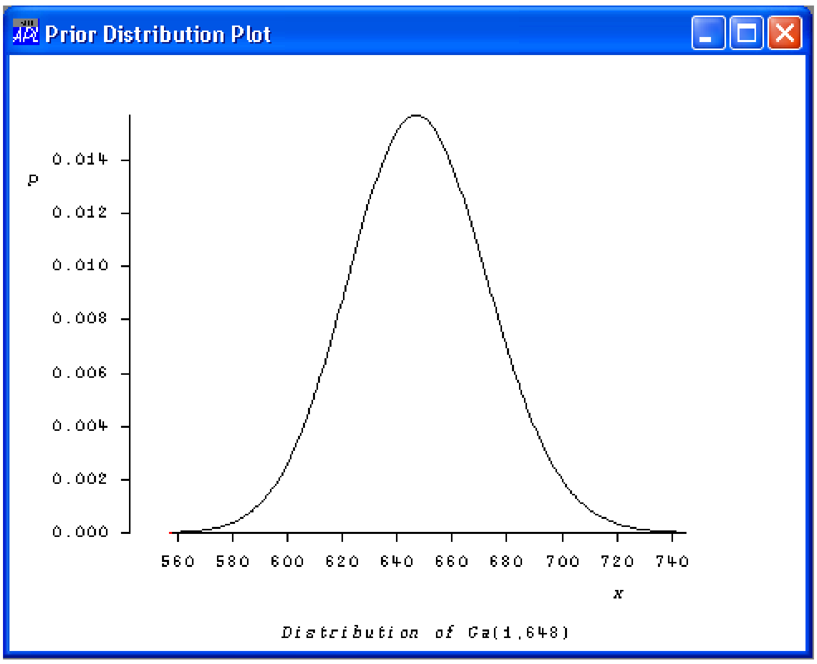

FIGURA 3. *Gráfico de la densidad a priori del número de reclamaciones.*

La figura 2 muestra la distribución a priori del número de reclamaciones, estos valores producen unos valores a priori para la media, moda y desviación típica. La figura 3 muestra el gráfico de la densidad a priori del número de reclamaciones.

La función de densidad a posteriori se obtiene al elegir la opción Quit de la ventana mostrada en la figura 2. El resultado se observa en la figura 4.

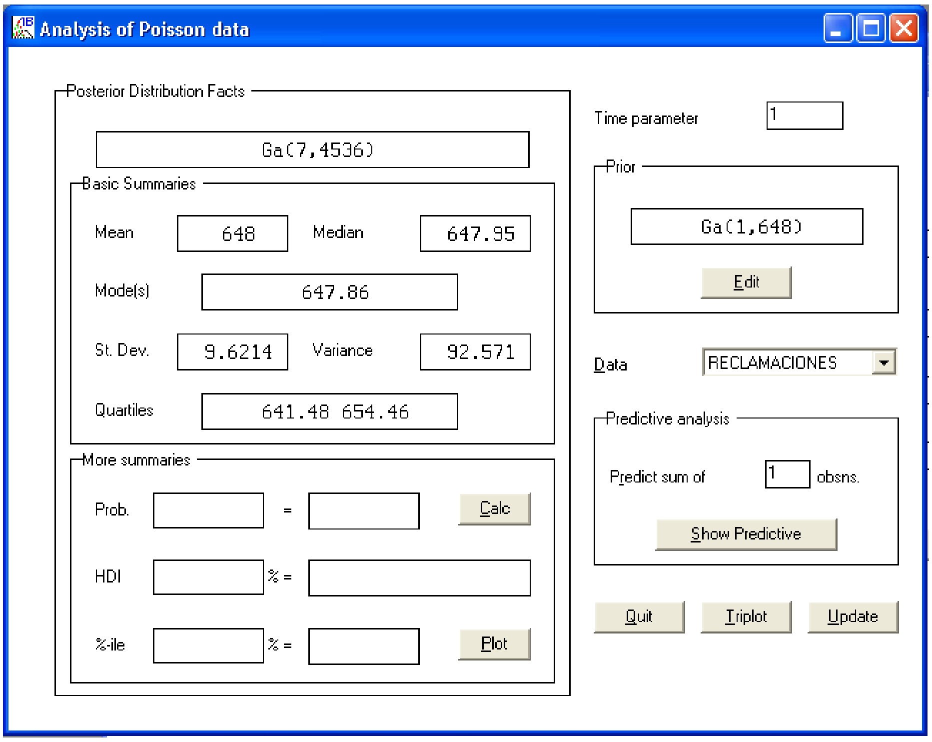

FIGURA 4. *Información de la densidad a posteriori del número de reclamaciones.*

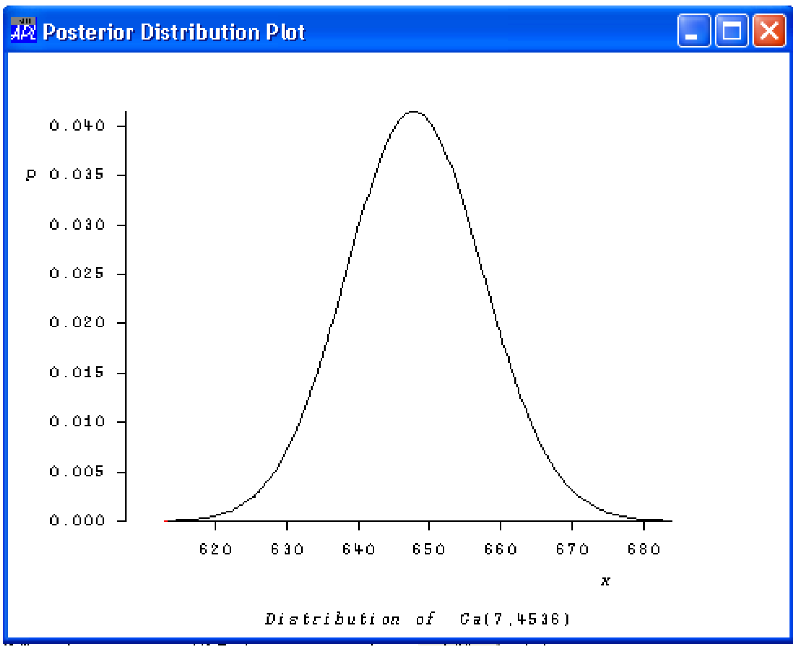

FIGURA 5. *Gráfico de la densidad a posteriori del número de reclamaciones.*

.En la figura 4 y 5 puede apreciarse la información y el gráfico, respectivamente de la densidad a posteriori del número de reclamaciones, es decir la distribución gamma $G\left(4536, \frac{1}{7}\right)$.

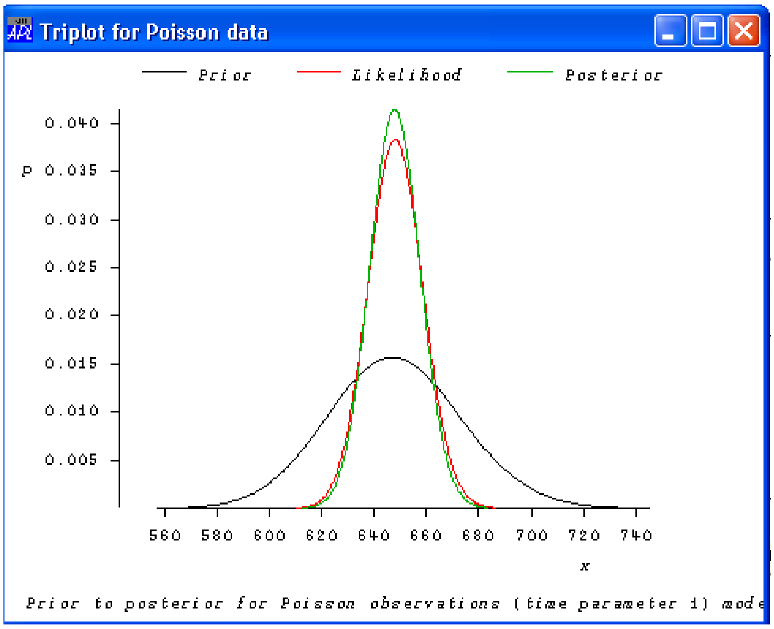

FIGURA 6. *Gráfico triplot del número de reclamaciones.*

En la figura 6 se puede observar que la información a priori no ha pesado prácticamente nada en el análisis a posteriori ya que la verosimilitud o densidad a priori y la distribución a posteriori del número de reclamaciones son bastante coincidentes. Por lo tanto, la modelización a priori no aporta información y son los datos los que producen la información a posteriori.

El intervalo de credibilidad al 95% para el número medio de reclamaciones, se puede observar en la figura siguiente:

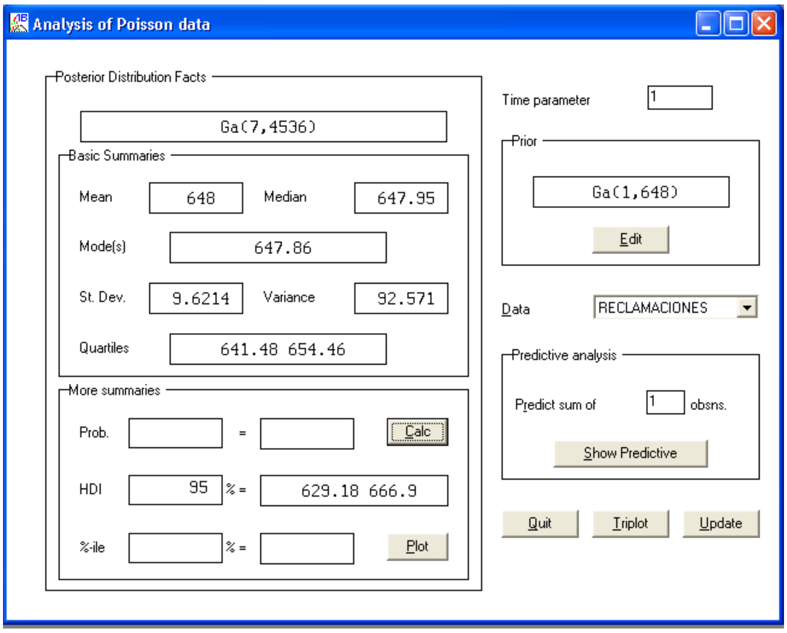

FIGURA 7. *Intervalo de credibilidad para el número medio de reclamaciones.*

La figura 7 muestra los resultados obtenidos $\left(P_r\left(629.18 \leq \lambda \leq 666.9\right) = 0.95\right)$ y permiten afirmar que con una probabilidad de 0.95 el número medio de reclamaciones entre el año 2004 y 2009, se sitúa entre 629.18 y 666.9.

El comportamiento de los próximos 5 años, se obtiene mediante la distribución predictiva, la figura 8 muestra los resultados.

Puesto que para los datos en estudio $X/\lambda \rightarrow P(\lambda)$ y $\lambda \rightarrow G\left(\alpha, \frac{1}{\beta}\right)$ la distribución incondicional o predictiva es la binomial negativa $BN\left(\alpha, p = \frac{1}{(1+\beta)}\right)$, es decir, $BN(4536{,}0.5833)$

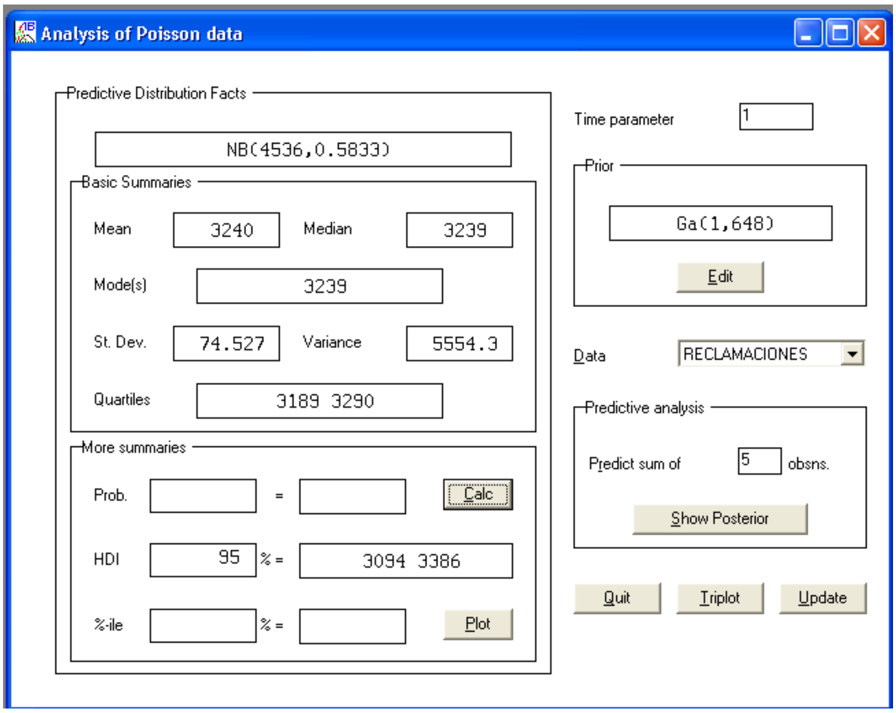

FIGURA 8. *Medidas predictivas para el número de reclamaciones.*

De acuerdo a los resultados anteriores se puede afirmar que por término medio en los próximos 5 años se esperará atender 3240 reclamaciones y que con una probabilidad del 95% para el año 2014 se atenderá entre 3094 y 3386 reclamaciones.

Suponiendo que la compañía de seguros destina $1,500.00 para cada reclamación, se puede obtener un intervalo al 95% de previsión de costes para el siguiente año, es decir, para el año 2010.

Para la previsión de gastos, se debe obtener la distribución predictiva para el próximo año (1 observación), La figura 9 muestra que con una probabilidad de 0.95, no se esperará para el año 2010 más de 701 reclamaciones.

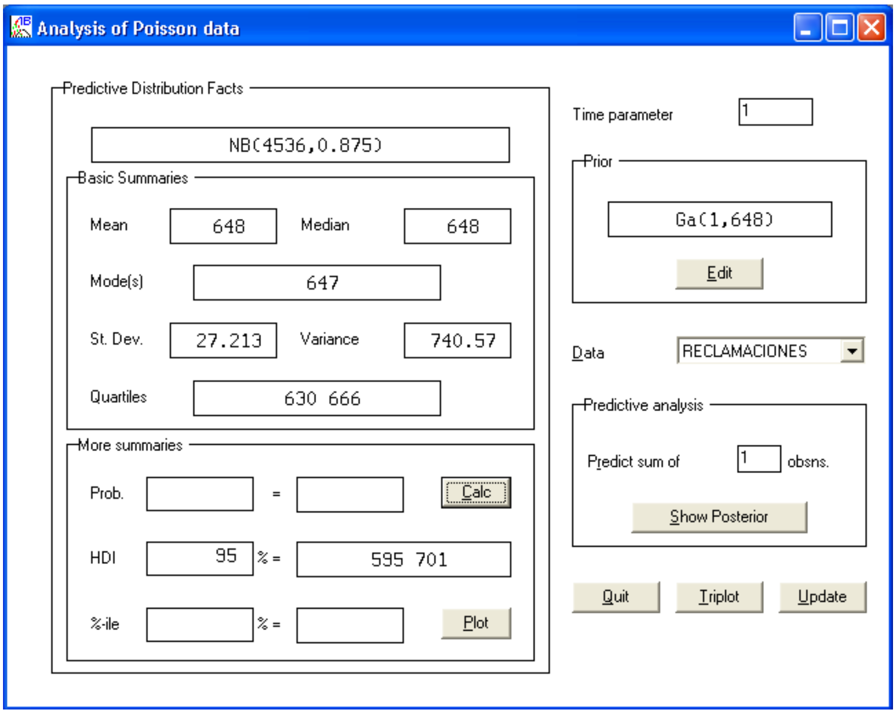

FIGURA 9. *Medidas predictivas para el número de reclamaciones para el año 2010*

La previsión para el próximo año deberá ser $701 \times 1500 = 1{,}051{,}500$, *es decir, con probabilidad de 0.95 la previsión de gastos para el año 2010 es* $\$1{,}051{,}500.00$

TEST DE HIPÓTESIS BAYESIANO

Se supondrá ahora que la compañía de seguros estima que el número medio de reclamaciones al año no es mayor de 750.

El test de hipótesis planteado es:

$$H_0 : \lambda \leq 750 \quad \text{vs} \quad H_1 : \lambda > 750$$

Desde el punto de vista bayesiano, cada una de esas hipótesis tiene una probabilidad a priori de ser cierta que puede obtenerse sin más que calcular mediante la distribución a posteriori la probabilidad de cada intervalo:

$$\pi_0 = P_r\left(H_0 \ cierta\right) = P_r\left(0 \leq \lambda \leq 750\right) = 0.9996$$

La figura 10 muestra la probabilidad del intervalo que se obtiene en la ventana de probabilidades en *FirstBayes*. Por tanto la probabilidad de H_1 denotada por π_1 será 0.0004

El "odds" a priori es $\dfrac{\pi_0}{\pi_1} = \dfrac{0.9996}{0.0004} = 2499$

"A priori, la hipótesis nula H_0 ***es 2,499 veces más creíble que*** H_1 ***"***

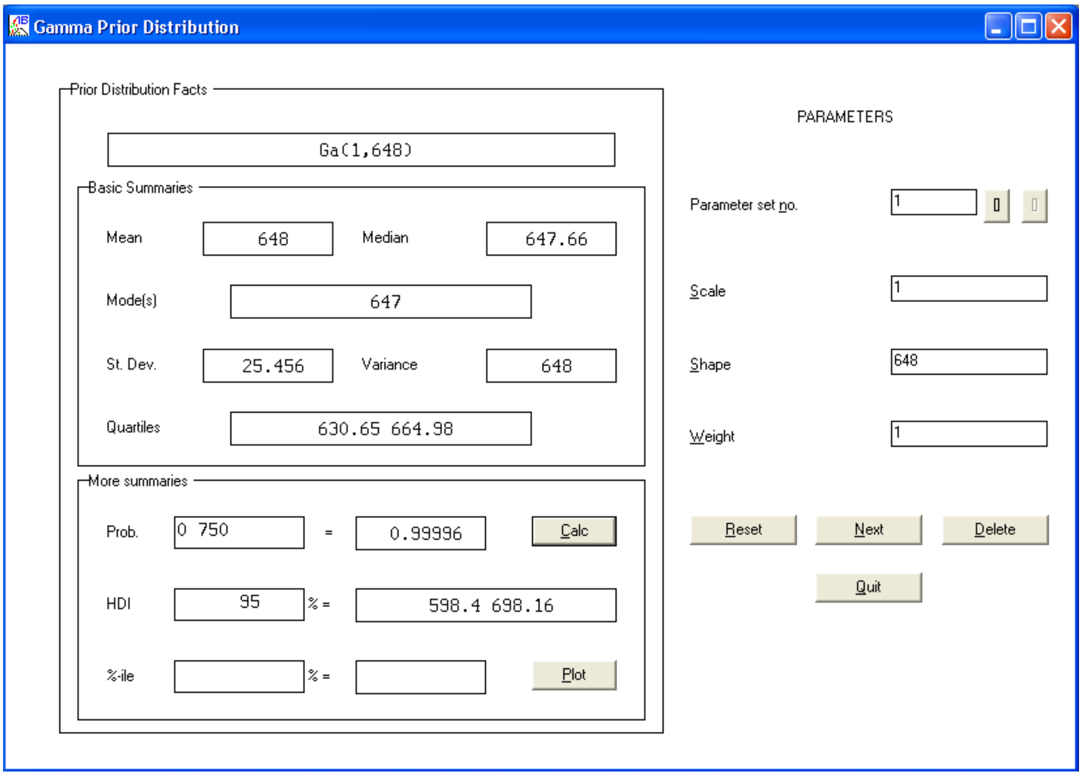

FIGURA 10. *Cantidades a priori para el número de reclamaciones.*

Ahora se actualizan las probabilidades utilizando la muestra, obteniendo las probabilidades de cada hipótesis mediante la densidad a posteriori. La figura 11 muestra el resultado.

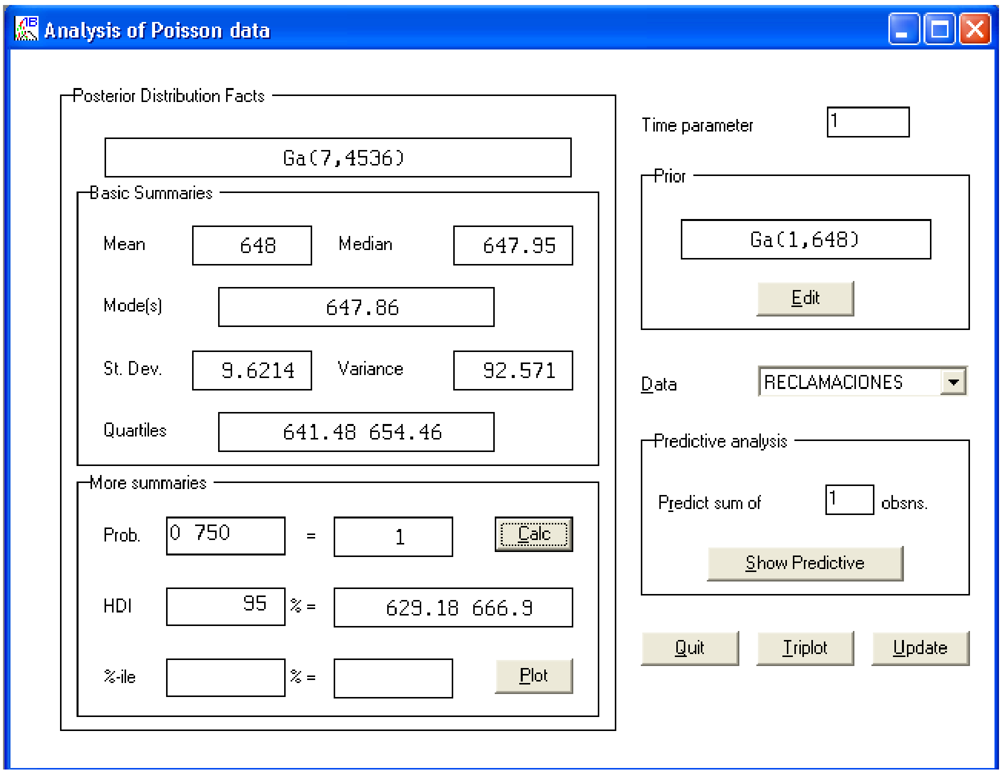

FIGURA 11. *Cantidades a posteriori para el número de reclamaciones.*

Se tiene entonces que:

$p_0 = P_r\left(H_0 \ cierta / datos\right) = P_r\left(0 \le \lambda \le 750 / datos\right) = 1.$

En consecuencia, $p_1 = P_r\left(H_1 \ cierta / datos\right) = 0.$

De donde se acepta H_0*. Por tanto, el número medio de reclamaciones al año no es mayor de 750.*

PREDICCIONES

La siguiente tabla muestra las predicciones del número de reclamaciones para los próximos 5 años, es decir, 2010 al 2014.

AÑO	2010	2011	2012	2013	2014
NUMERO DE RECLAMACIONES	648	1296	1944	2592	3240

TABLA 4. *Predicciones para el número de reclamaciones entre el año 2010 hasta el año 2014*

3.3.2. ANÁLISIS BAYESIANO PARA LA CANTIDAD RECLAMADA.

En este apartado se realizará el análisis bayesiano para la cantidad reclamada, conocida también como *cuantía de las reclamaciones* realizadas para el año 2004 y 2009.

Se realizará el análisis Bayesiano para la cuantía de las reclamaciones, se obtendrá una distribución a posteriori para la cuantía. En este caso no se incorporará la información a priori ya que no se cuenta con dicha información.

ANÁLISIS ESCRIPTIVO DE LOS DATOS.

Supongamos que la cuantía de las reclamaciones de la cartera de seguros se distribuye de forma normal con media μ y varianza σ^2 desconocidas, y además μ y σ^2 son independientes. En el gráfico 2 se puede observar la cantidad total reclamada durante los 12 meses del año 2004.

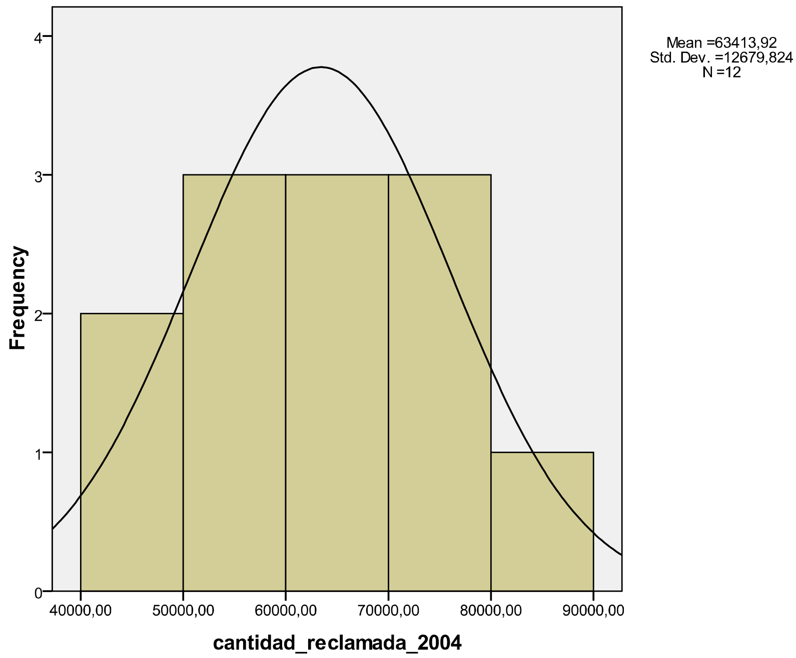

GRAFICO 2. *Gráfico correspondiente a la cantidad total reclamada para el año 2004.*

Para determinar que los datos se ajustan a la distribución normal se contrastará las hipótesis siguientes:

$$H_0 : \mathbf{X} \sim N(\mu, \sigma^2)$$
$$H_1 : \mathbf{X} \nsim N(\mu, \sigma^2)$$

Para realizar el contraste de hipótesis se realizará la *prueba de bondad de ajuste de Kolmogorov-Smirnov* para los datos de la muestra. El siguiente resultado se obtiene con el programa estadístico SPSS.

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Cantidad_Reclamada_2004	,169	12	,200*	,965	12	,853
a. Lilliefors Significance Correction						
*. This is a lower bound of the true significance.						

TABLA 5. *Prueba de bondad de ajuste Kolmogorov-Smirnov*

Al 5% de nivel de significación se llega a la conclusión de que la distribución de la CUANTÏA se adapta a la distribución normal con media μ y varianza σ^2, ya que el nivel de significación del estadístico de prueba es 0.20 > 0.05.

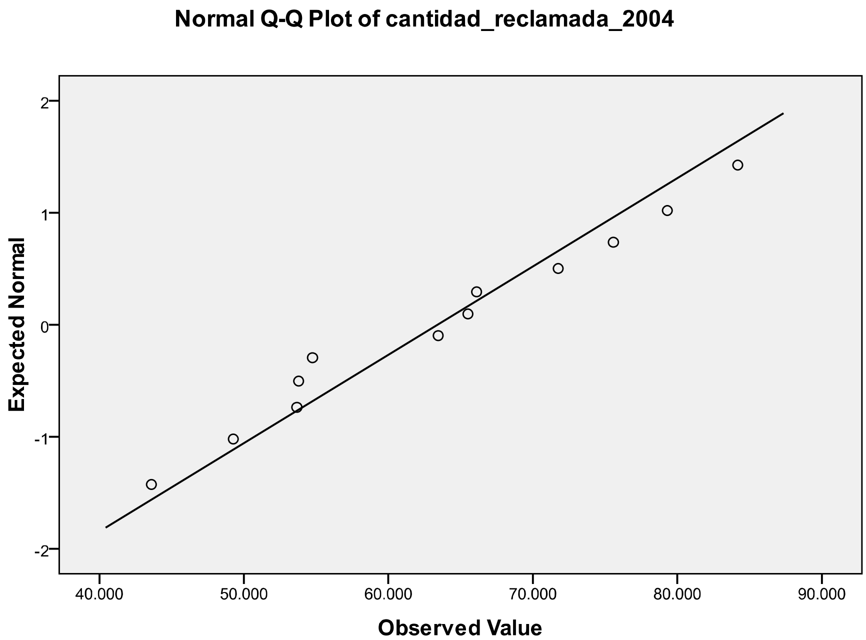

GRÁFICO 3. *Gráfico Q-Q Plot para la cantidad total reclamada para el año 2004.*

El gráfico Q-Q normal confirma la conclusión anterior ya que se observa que los valores observados para el año 2004 se sitúan sobre la recta esperada bajo el supuesto de normalidad. Gráfica 3.

De forma análoga, con base a lo desarrollado en los apartados 1.11, 1.13, 1.14, 1.15, 1.16, 1.17, 1.18 del capítulo 1, se realizará el análisis bayesiano para la cantidad total reclamada, con el objetivo de:

1. *La distribución a posteriori para la cantidad total reclamada a partir de la distribución a priori. La distribución a priori se especifica a partir de los datos de la muestra.*
2. *Obtener Intervalos bayesiano de credibilidad para el parámetro* μ
3. *Realizar el test de hipótesis bayesianos para el parámetro* μ
4. *Realizar predicciones para la cantidad total reclamada.*

En la figura 11 se puede observar el fichero creado en el programa ***FirstBayes*** correspondiente a la cuantía de las reclamaciones, el fichero ha sido guardado con el nombre de CUANTÍA.

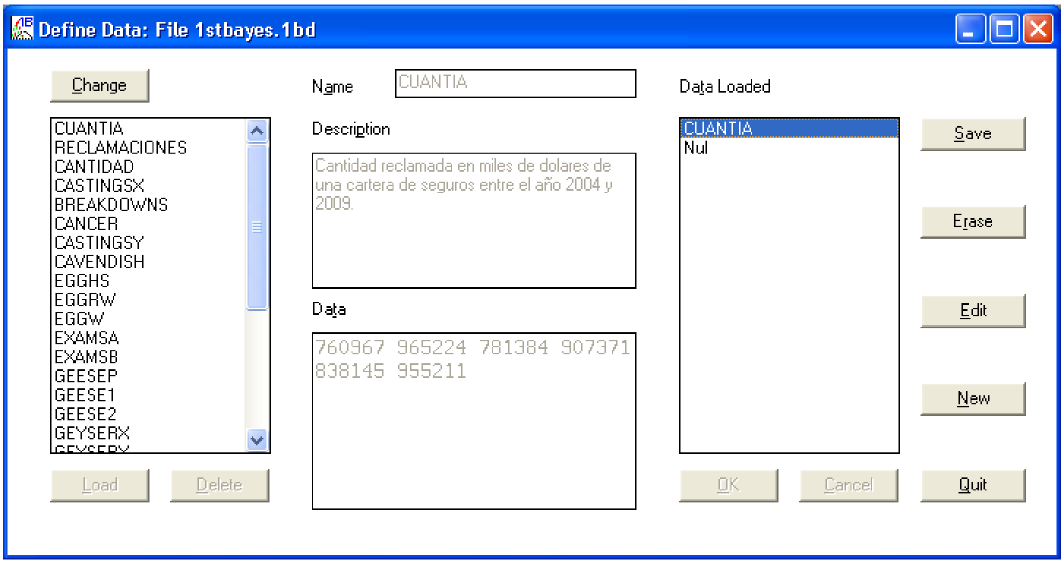

FIGURA 11. *Fichero creado en FirstBayes para la cuantía de las reclamaciones.*

Se efectuará la inferencia bayesiana sobre μ a partir de la densidad a posteriori obtenida de los datos de la muestra que corresponde entre el año 2004 y el año 2009.

Utilizando el programa ***Firstbayes*** mediante el análisis *Normal simple, unknown variance* se obtienen los resultados sobre la cuantía de las reclamaciones. La figura 12 muestra los estimadores bayesianos, asumiendo que μ y varianza σ^2 son desconocidas.

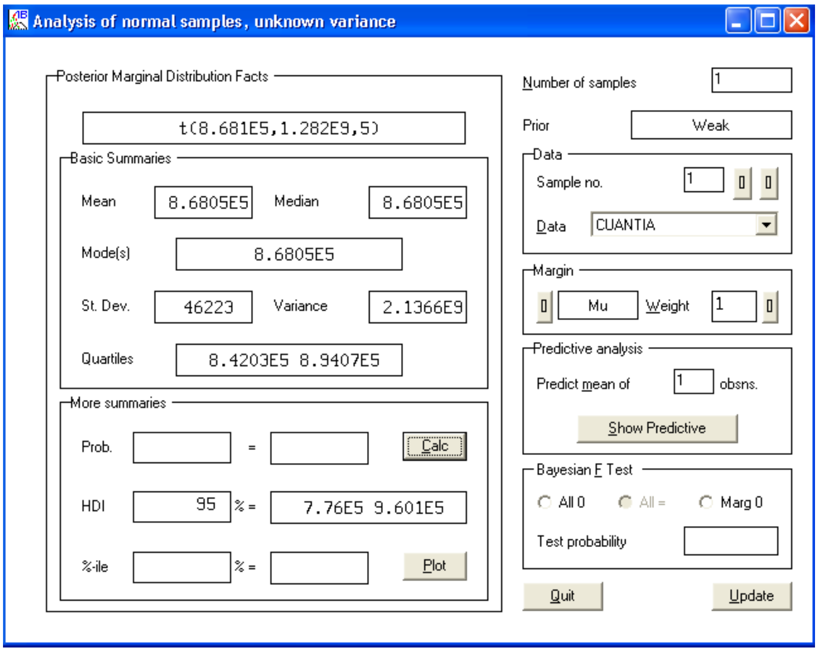

FIGURA 12. *Cantidades a posteriori para la cuantía con media y varianza desconocidas.*

En la figura 12 se puede observar la distribución a posteriori de μ es un t de Student generalizada cuya media (a posteriori) coincide con la mediana y la moda y valen $868{,}050$ dólares.

Con una probabilidad de 0.95 la cuantía media de reclamaciones entre el año 2004 y 2009, se sitúa entre 776,000 y 960,100 dólares.

La figura 13 muestra el gráfico de densidad a posteriori para la cuantía de las reclamaciones.

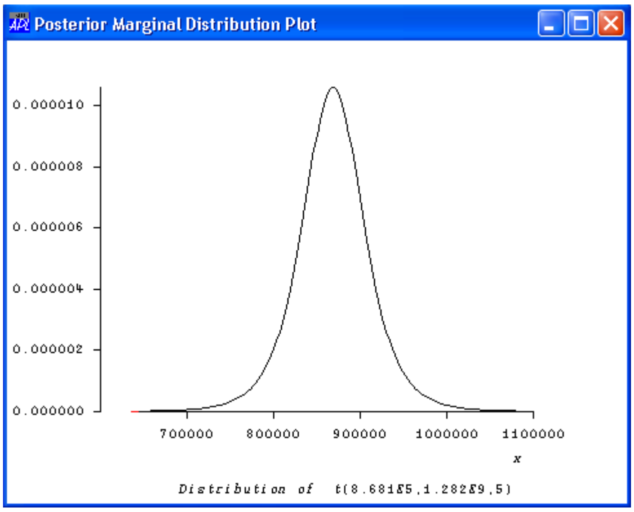

FIGURA 13. *Gráfico de la densidad a posteriori de la cuantía de las reclamaciones.*

El intervalo bayesiano de credibilidad con probabilidad de 0.95 para una nueva observación futura sobre la cuantía de reclamación para el próximo año, es decir, año 2010, se obtiene a partir de la distribución predictiva.

La distribución predictiva para el caso de datos normales donde μ y varianza σ^2 son desconocidas es una distribución t de Student con parámetro de localización $\bar{x}$, parámetro de escala $\sqrt{1+\frac{1}{n}}\,s$, $n-1$ grados de libertad, para el caso en estudio se tiene una distribución $t_5\left(8.681\times10^5, 8.974\times10^9\right)$. La figura 14 muestra los resultados.

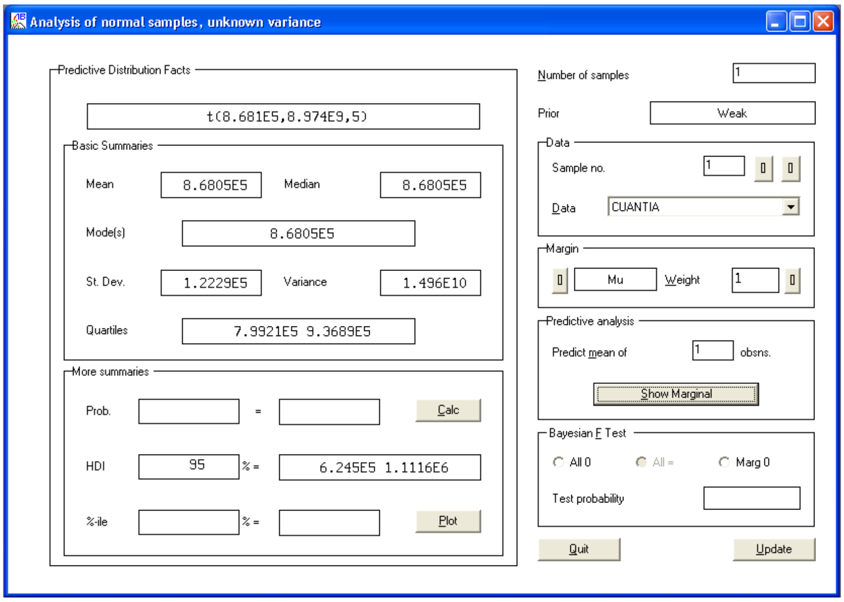

FIGURA 14. *Distribución predictiva para la cuantía de reclamaciones.*

El intervalo bayesiano al 95% es $\left(6.245\times10^5, 1.1116\times10^6\right)$, *e indica que con una probabilidad de 0.95 la cuantía de las reclamaciones para el año 2010 estará comprendida entre 624,500.00 y 1,111,600.00 dólares.*

A continuación se contrastará la hipótesis de que la cuantía media de reclamaciones que paga la compañía de seguro cada año no es mayor de 825,000 dólares.
El test de hipótesis planteado es:

$$H_0 : \mu \leq 825{,}000 \quad \text{vs} \quad H_1 : \mu > 825{,}000$$

Cada una de esas hipótesis tiene una probabilidad a posteriori de ser cierta que puede obtenerse sin más que calcular mediante la distribución a posteriori la probabilidad de cada intervalo:

$$p_0 = P_r\left(H_0 \ cierta\right) = P_r\left(0 \leq \mu \leq 825{,}000\right) = 0.14164$$

La figura 15 muestra la probabilidad del intervalo que se obtiene en la ventana de probabilidades en *FirstBayes*. Por tanto la probabilidad de H_1 denotada por p_1 será 0.85836.

El "odds" a posteriori es $\dfrac{p_0}{p_1} = \dfrac{0.14164}{0.85836} = 0.1650$

"A posteriori, H_1 ***es más creíble que*** H_0***."***

De donde se concluye que los datos no aportan ninguna evidencia en favor de H_0, por tanto, la hipótesis nula debe ser rechazada, es decir, la cuantía media de las reclamaciones que cubre la compañía de seguros cada año es mayor de 825,000 dólares.

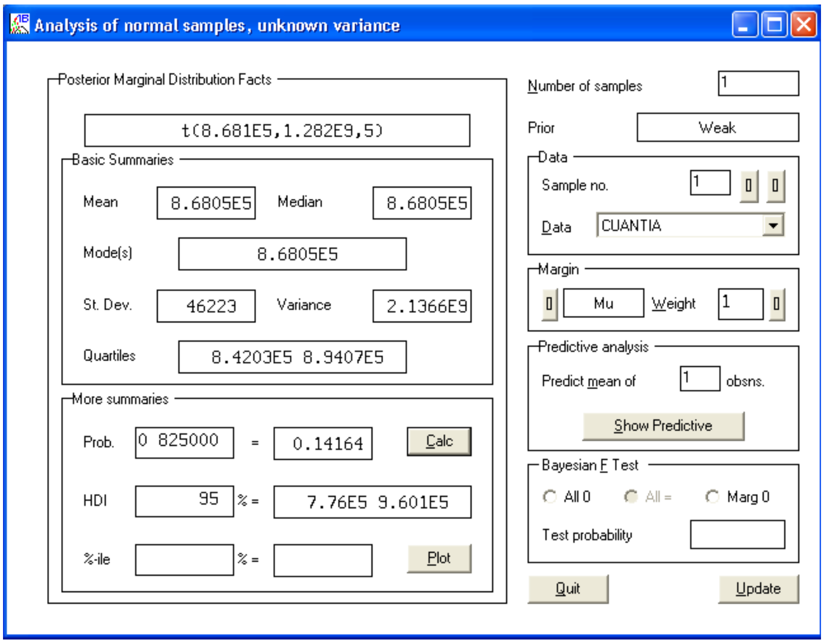

Figura 15. *Cantidades a posteriori para la cuantía de reclamaciones.*

3.3.3. LA TEORÍA DE LA CREDIBILIDAD.

En los siguientes casos se aplicará la teoría de la credibilidad para dar solución a las situaciones planteadas.

- Se cuenta con la experiencia X_j, $j = 1,\ldots,t$ de un contrato de seguro perteneciente a una cartera de seguros, los datos son los correspondientes al número de reclamaciones efectuadas por los asegurados entre el año 2004 y 2009 de Caja Mutual de los empleados del Ministerio de Educación. Sabiendo que $X_1,\ldots,X_t$ son el número de reclamaciones y son variables aleatorias independientes e idénticamente distribuidas Poisson de parámetro $\theta = 648$. Se obtendrá el valor más pequeño para t para suponer credibilidad total a la experiencia observada.

La condición para suponer credibilidad total es $\dfrac{c\varepsilon\sqrt{t}}{\sigma} \ge x_p$, donde $0 < p < 1$ y $c > 0$: En la práctica lo razonable es elegir un valor de p cercano a 1 y un valor de c cercano a 0.

Para dar solución al problema se supondrá que $c = 0.04$, $p = 0.95$ y que la aseguradora tarifica atendiendo sólo al número de reclamaciones.

En este caso $\varepsilon = \sigma^2 = 480$. Dados los valores de $c = 0.04$ y $p = 0.95$ resulta que $x_p = 0.95$, luego:

$$\lambda_0 = \left(\frac{x_p}{c^2}\right)^2 = 2401$$

Por tanto

$$t \geq \lambda_0 \left(\frac{\sigma}{\varepsilon}\right)^2$$

$$t \geq 2401 \frac{648}{648^2} = 3.71$$

El tiempo mínimo para suponer credibilidad total en la prima que la empresa aseguradora asigna es 3.7 años.

- Para efectos solamente expositivo se plantea la siguiente situación:

La compañía tiene 2 grupos de asegurados en una póliza que cubre determinado riesgo, supondremos que ambos grupos se formaron de aleatoriamente. La cantidad reclamada en miles de dólares para ambos grupos se muestra en la siguiente tabla.

A continuación se calculará la prima neta para el séptimo año, es decir, año 2010.

AÑO	MONTO PAGADO (MILES DE DÓLARES) GRUPO 1	MONTO PAGADO (MILES DE DÓLARES) GRUPO 2
2004	216,719	499,676
2005	222, 398	689,016
2006	199,714	521,099
2007	267,599	646,056
2008	212,434	541,139
2009	253,748	631,748
TOTAL	1,372,612	3,528,734

TABLA 6. Cantidad total pagada para 2 grupos de asegurados.

Para ello se utilizará el modelo de Bilman de distribución libre para estimar la prima neta para el séptimo año, mediante la fórmula:

$$\text{Pr}\,ima \;\; neta = \hat{Z}(t)\bar{x} + \left[1 - \hat{Z}(t)\right] \hat{m}$$

Donde:

$\hat{Z}(t)$ es el factor de credibilidad

$\bar{x}$ es la cantidad total reclamada media de cada grupo

$\hat{m}$ es la cantidad total reclamada media de ambos grupos

Para determinar la prima neta, se calculará primeramente el factor de credibilidad de ambos grupos, a partir de la fórmula:

$$\hat{Z}(t) = \frac{\hat{a}t}{\hat{s}^2 + \hat{a}t}$$

Además:

$$\widehat{m} = \frac{1}{k}\sum_{j=1}^{k}\overline{x}_j = \frac{1}{k}\sum_{j=1}^{k}\sum_{s=1}^{t}\frac{x_{js}}{t},$$

$$\widehat{s}^2 = \frac{1}{k}\sum_{j=1}^{k}\widehat{s}_j^2, \ \widehat{s}_j^2 = \frac{1}{t-1}\sum_{s=1}^{t}(x_{js} - \overline{x}_j)^2,$$

$$\widehat{a} = \frac{1}{k-1}\sum_{j=1}^{k}(\overline{x}_j - \overline{x})^2 - \frac{1}{t}\widehat{s}^2$$

Siendo $k = 2$, el número de grupos de asegurados y $t = 6$, los primeros 6 años de vigencia de la póliza.

El programa SPSS, muestra los valores descriptivos para ambos grupos. Ver tabla 7.

Descriptive Statistics							
	N	Minimum	Maximum	Sum	Mean	Std. Deviation	Variance
Grupo_1	6	199,714	267,599	1372,612	228,76867	26,186014	685,707
Grupo_2	6	499,676	689,016	3528,734	588,12233	77,409761	5992,271
Valid N (listwise)	6						

TABLA 7. Valores descriptivos para los dos grupos asegurados

Del cuadro anterior se puede observar que:

$\hat{m}_1 = 228{,}768.67,$ $\quad \hat{m}_1 = 588{,}122.33,$ $\quad m = 408{,}445.5$

$\hat{s}_1^2 = 685{,}70$ $\quad \hat{s}_2^2 = 5{,}992.27$ $\quad \hat{s}^2 = 3{,}338.99$

$\hat{a} = 64{,}567{,}525{,}364.7$

$\hat{z}(t) = 0.99999999 \approx 1$

$\hat{z}(t) = 1$

Se puede observar de acuerdo al cálculo anterior que $\hat{z}(t) = 1$, lo que indica que la experiencia individual contribuye a la prima en un 100%, es decir, que la aseguradora calcula la prima tomando en cuenta únicamente la experiencia de siniestralidad de los asegurados asignando un 100% de credibilidad.

La prima neta para el séptimo año para ambos grupos respectivamente es:

Grupo 1:
$$\text{Pr}\,ima \;\; neta = \hat{Z}(t)\bar{x} + \left[1 - \hat{Z}(t)\right]\; \hat{m}$$
$$\text{Pr}\,ima \;\; neta = 1 \times 228{,}768.67 + 0 \times 408{,}445.5 = 228{,}768.67$$

Grupo 2:
$$\text{Pr}\,ima \;\; neta = \hat{Z}(t)\bar{x} + \left[1 - \hat{Z}(t)\right]\; \hat{m}$$
$$\text{Pr}\,ima \;\; neta = 1 \times 588{,}122.33 + 0 \times 408{,}445.5 = 588{,}122.33$$

La prima neta para el año 2010, para el grupo 1 y el grupo 2 son respectivamente 228,768.67 y 588,122.33.

BIBLIOGRAFÍA.

1. Sarabia Alegría, José María; Gómez Déniz, Emilio y Vásquez Polo, Francisco José. Estadística Actuarial. Editorial Pearson Educación S.A, España, 2007.

2. Lareshas Sanz, Antonio. Matemática del Seguro. Editorial Dossat S.A, Madrid-Buenos Aires, 1948.