Credibilidad
(teoría)

Marzo de 2024

Eva Boj del Val
Departamento de Matemática Económica, Financiera y Actuarial
Universidad de Barcelona

Credibilidad clásica
(teoría)

Eva Boj del Val

Departamento de Matemática Económica, Financiera y Actuarial
Universidad de Barcelona

Credibilidad

La credibilidad proporciona el marco analítico básico para fijar el precio de los productos de seguros. Destaca la importancia de combinar información sobre la experiencia reciente de los individuos versus la experiencia pasada agregada. El paradigma comenzó en la literatura con un tratamiento analítico riguroso del tema por Hans Bühlmann. El enfoque de Bühlmann proporciona una solución simple a la metodología bayesiana y logra la optimización dentro del subconjunto de predictores lineales de un modelo de predicción.

Partes:

- Enfoque de credibilidad de la fluctuación limitada (clásico)
- Credibilidad de Bühlmann
- Enfoque Bayesiano (asignatura AMAM con detalle)
- Implementación empírica de estas técnicas

Referencia básica (manual examen SOA): TSE, Y.K. *Nonlife Actuarial Models. Theory, Methods and Evaluation*. Cambridge University Press, New York, 2009.

1. Enfoque de credibilidad de fluctuación limitada – CLÁSICO

Los modelos de credibilidad en sus orígenes propusieron actualizar la predicción de la experiencia de siniestralidad como un promedio ponderado de la predicción basada únicamente en los datos recientes, D, y la tasa del "manual de seguros", M.

La **teoría de la credibilidad** se refiere a la actualización de la predicción del próximo período usando la experiencia de siniestralidad reciente, D, y la del manual de seguros, M. La predicción revisada, V, determina la prima de seguro del próximo período para el grupo de riesgo.

<u>Grupo de riesgo</u>: Es un bloque de pólizas de seguro. El grupo de riesgo está cubierto durante un período de tiempo (un año) mediante el pago de una prima. La prima se basa parcialmente en una tasa especificada del manual, denominada tasa del manual y parcialmente en las características de riesgo específicas del grupo. Con base en la experiencia reciente de siniestros del grupo de riesgo, se revisa la prima para el próximo período.

- **Credibilidad total**: se logra si la cantidad de datos recientes es suficiente, en cuyo caso la predicción actualizada se basará únicamente en los datos recientes.

- **Credibilidad parcial**: si la cantidad de datos recientes es insuficiente, sólo se atribuye credibilidad parcial a los datos y la predicción actualizada también depende de la tasa del manual.

Medidas de la experiencia de siniestralidad:

- Número de siniestros en el periodo, N.

- Cuantía total de los siniestros en el periodo, S.

  Es igual a $$S = X_1 + X_2 + \cdots + X_N$$

  donde X_i es la cuantía del siniestro i-ésimo.

- Cuantía media de un siniestro, $\bar{X}$.

  Es la media muestral de las cuantías de los siniestros de X_1, X_2, …, X_N,

$$\bar{X} = \frac{S}{N}.$$

- Prima pura, P.

  Si E es el número de unidades de exposición del grupo de riesgo, entonces,

$$P = \frac{S}{E}.$$

Realizamos la predicción actualizada, V, de:

$$N, S, \bar{X}, P.$$

Fórmula:

$$V = ZD + (1 - Z)M.$$

El peso Z se refiere a D y se denomina "**Factor de credibilidad**".

Toma valores entre 0 y 1 ($0 \leq Z \leq 1$).

1.1 Credibilidad total

El enfoque de credibilidad clásico determina el tamaño mínimo de datos requerido para que los datos de experiencia tengan credibilidad total ($Z = 1$). El tamaño de credibilidad total se denomina estándar de credibilidad total. Intuitivamente, un conjunto de datos grande justificaría una Z elevada.

- **Estándar de credibilidad total para el número de siniestros**. Supongamos que N es una variable aleatoria con media μ_N y varianza σ_N^2. Para evaluar la probabilidad de que un valor observado de N sea "representativo" de la media verdadera hacemos la siguiente pregunta:

  ¿Cuál es la probabilidad de observar el número de siniestros dentro del $100K\%$ de la media?
  Esta probabilidad viene dada por:

  Caso Normal, $N \sim N(\mu_N, \sigma_N^2)$:

$$P(\mu_N - K\mu_N \leq N \leq \mu_N + K\mu_N) =$$

$$= P\left(-\frac{K\mu_N}{\sigma_N} \leq \frac{N-\mu_N}{\sigma_N} \leq \frac{K\mu_N}{\sigma_N}\right) =$$

$$= \phi\left(\frac{K\mu_N}{\sigma_N}\right) - \phi\left(-\frac{K\mu_N}{\sigma_N}\right) = \phi\left(\frac{K\mu_N}{\sigma_N}\right) - \left(1 - \phi\left(\frac{K\mu_N}{\sigma_N}\right)\right) = 2\phi\left(\frac{K\mu_N}{\sigma_N}\right) - 1$$

Consideremos que la probabilidad anterior es igual a $1 - \alpha$:

$$2\phi\left(\frac{K\mu_N}{\sigma_N}\right) - 1 \quad = 1 - \alpha$$

$$\phi\left(\frac{K\mu_N}{\sigma_N}\right) = \frac{2-\alpha}{2} = 1 - \frac{\alpha}{2}$$

$$\left(\frac{K\mu_N}{\sigma_N}\right) = Z_{1-\frac{\alpha}{2}} \ \ (*)$$

Siendo ϕ la función de distribución de la Normal estandarizada.

Por lo tanto, existe una probabilidad de $1 - \alpha$ de que el número de siniestros observado esté dentro del $100K\%$ de la media verdadera, donde α satisface $(*)$.

Caso Poisson, $N \sim Poisson(\lambda_N)$:

Suponemos que λ_N es lo suficientemente grande como para justificar la aproximación normal. Se refiere a:

$$N \sim P(\lambda_N) \rightarrow \lim_{\lambda_N \to \infty} \left(\frac{N - \lambda_N}{\sqrt{\lambda_N}} \leq z \right) = P\ (Z \leq z)\ , \qquad Z \sim N(0,1).$$

Si $N \sim P_0(\lambda_N)$, $\sigma_N^2 = \lambda_N = \mu_N$, entonces la ecuación $(*)$ es

$$\frac{K\mu_N}{\sigma_N} = \frac{K\lambda_N}{\sqrt{\lambda_N}} = K\sqrt{\lambda_N} = Z_{1-\frac{\alpha}{2}}.$$

Se atribuye credibilidad total a los datos si la probabilidad es al menos 1 – α y por tanto:

$$\lambda_N \geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2 .$$

Definimos

$$\lambda_F = \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2$$

y damos a los datos credibilidad total si la probabilidad es al menos 1 – α, y por tanto:

$$\lambda_N \geq \lambda_F .$$

Como el número de siniestros λ_N se desconoce en la práctica, la implementación del modelo de credibilidad consiste en comparar el valor observado de N en el período actual con λ_F:

se alcanza credibilidad total si

$$N \geq \lambda_F.$$

- **Estándar de credibilidad total para la cuantía media de un siniestro.** Supongamos una muestra de N cuantías de siniestros (fijos, no aleatorios) $X_1, X_2, \ldots, X_N$. Asumiendo $\{X_i\}$ independientes e idénticamente distribuidas con respecto a μ_X y σ_X^2 y usando la media muestral $\bar{X}$ para estimar μ_X, y asumiendo también que N es suficientemente grande para que $\bar{X} \sim N\left(\mu_X \,, \sqrt{\frac{\sigma_X^2}{N}}\right)$.

  La probabilidad de observar la cuantía media de un siniestro dentro del $100K\%$ de la media es:

$$P(\mu_X - K\mu_X \leq \bar{X} \leq \mu_X + K\mu_X) =$$

$$= P\left(\frac{-K\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{\bar{X} - \mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{K\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right).$$

$$\phi\left(\frac{K\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) - 1 = 1 - \alpha \quad \rightarrow \quad \frac{K\mu_X}{\frac{\sigma_X}{\sqrt{N}}} = Z_{1-\frac{\alpha}{2}}$$

Para que

$$\sqrt{N} = \frac{Z_{1-\frac{\alpha}{2}}}{K}\left(\frac{\sigma_X}{\mu_X}\right)$$

ha de pasar

$$N = \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2\left(\frac{\sigma_X}{\mu_X}\right)^2$$

y puesto que

$$C_X = \frac{\sigma_X}{\mu_X}$$

es el coeficiente de variación de X, entonces para que la probabilidad sea al menos $1 - \alpha$, debemos tener:

$$N \geq \lambda_F\, C_X^2.$$

Por lo tanto $\lambda_F\, C_X^2$ es el estándar para la credibilidad total de la cuantía media de un siniestro. Así, si el número de siniestros excede de $\lambda_F C_X^2$, $\bar{X}$ será la predicción para la cuantía media para el próximo período. Debemos tener en cuenta que N se trata como una constante en lugar de una variable aleatoria. Para implementar la metodología en la práctica, μ_X y σ_x^2 deben estimarse a partir de la muestra.

- **Estándar de credibilidad total para la cuantía total de los siniestros.** Para derivar el estándar de credibilidad total para la cuantía total de los siniestros determinamos la frecuencia mínima (esperada) de siniestros tal que la probabilidad de que la pérdida total observada S esté dentro del 100K% de la pérdida total esperada y sea al menos 1 – α,

$$P(\mu_S - K\mu_S \leq S \leq \mu_S + K\mu_S) =$$

$$= P\left(-\frac{K\mu_S}{\sigma_S} \leq \frac{S - \mu_S}{\sigma_S} \leq \frac{K\mu_S}{\sigma_S}\right).$$

Caso S Poisson compuesto:

$$\mu_S = \mu_X \lambda_N$$
$$N \sim P(\lambda_N)$$
$$\sigma_S^2 = \mu_N \sigma_X^2 + \mu_X^2 \sigma_N^2$$

y como $\mu_N = \sigma_N^2 = \lambda_N$,

$$\sigma_S^2 = \lambda_N (\mu_X^2 + \sigma_X^2),$$

de este modo,

$$\frac{\mu_S}{\sigma_S} = \frac{\mu_X \lambda_N}{\sqrt{\lambda_N(\mu_X^2 + \sigma_X^2)}} = \frac{\mu_X \sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}} .$$

Usando la aproximación normal,

la probabilidad de observar la cuantía total de los siniestros dentro del 100K% de la media es:

$$P\left(-\frac{K\mu_S}{\sigma_S} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{K\mu_S}{\sigma_S}\right) = 2\,\phi\left(\frac{K\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2+\sigma_X^2}}\right) - 1.$$

Para que la probabilidad anterior sea al menos 1 – α, debemos tener:

$$\frac{K\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2+\sigma_X^2}} \geq Z_{1-\frac{\alpha}{2}}.$$

Para que,

$$\sqrt{\lambda_N} \geq \frac{Z_{1-\frac{\alpha}{2}}\sqrt{\mu_X^2+\sigma_X^2}}{K\mu_X}$$

$$\lambda_N \geq \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right)$$

$$\left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right) = \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2 (1 + C_X^2) = \lambda_F (1 + C_X^2) =$$

$$= \lambda_F + \lambda_F C_X^2 .$$

Así,

$$\lambda_N \geq \lambda_F + \lambda_F C_X^2$$

siendo

$$C_X = \left(\frac{\sigma_X}{\mu_X}\right).$$

Estándar de credibilidad total de la cuantía total de los siniestros =

Estándar de credibilidad total del número de siniestros
+
Estándar de credibilidad total de la cuantía media de un siniestro

- **Estándar de credibilidad total para la prima pura**:

$$P = \frac{S}{E}$$

siendo E las unidades de exposición. La probabilidad de observar la prima pura dentro del $100K\%$ de la media es:

$$P(\mu_P - K\mu_P \leq P \leq \mu_P + K\mu_P) = 2\,\phi\left(\frac{K\mu_P}{\sigma_P}\right) - 1,$$

siendo μ_P y σ_P^2 la media y varianza de P. Se observa que:

$$\left(\frac{\mu_P}{\sigma_P}\right)$$

y

$$\left(\frac{\mu_S}{\sigma_S}\right)$$

son iguales porque E es fijo. Entonces, el estándar de credibilidad total para la prima pura es el mismo que para la cuantía total de los siniestros.

Estándar de credibilidad total

Número de siniestros	$\lambda_F = \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2$	$\lambda_N \geq \lambda_F$
Cuantía media de un siniestro	$\lambda_F\, C_X^2$	$N \geq \lambda_F\, C_X^2$
Cuantía total de los siniestros	$\lambda_F(1 + C_X^2)$	$\lambda_N \geq \lambda_F(1 + C_X^2)$
Prima pura	$\lambda_F(1 + C_X^2)$	$\lambda_N \geq \lambda_F(1 + C_X^2)$

1.2 Credibilidad parcial

Cuando el grupo de riesgo no es lo suficientemente grande, no se puede asignar credibilidad total. En este caso el valor de $Z \leq 1$ se ha de determinar.

La suposición básica al derivar Z es que la probabilidad de que ZW se encuentre dentro del intervalo $[Z\mu_W - K\mu_W, Z\mu_W + K\mu_W]$ y que sea igual a $1 - \alpha$ para un valor dado de K. Siendo W la variable de siniestralidad de interés.

- **Credibilidad parcial para el número de siniestros**:

$$P(Z\mu_N - K\mu_N \leq ZN \leq Z\mu_N + K\mu_N) = 1 - \alpha$$

$$P\left(\frac{Z\mu_N - K\mu_N - Z\mu_N}{Z\sigma_N} \leq \frac{ZN - Z\mu_N}{Z\sigma_N} \leq \frac{Z\mu_N + K\mu_N - Z\mu_N}{Z\sigma_N}\right) =$$

$$P\left(-\frac{K\mu_N}{Z\sigma_N} \leq \frac{N - \mu_N}{\sigma_N} \leq \frac{K\mu_N}{Z\sigma_N}\right) = 1 - \alpha.$$

Caso Poisson, $N \sim Poisson(\lambda_N)$, y usando la aproximación normal:

$\mu_N = \lambda_N, \sigma_N = \sqrt{\lambda_N}$

$$P\left(-\frac{K\sqrt{\lambda_N}}{Z} \leq \frac{N-\lambda_N}{\sigma_N} \leq \frac{K\sqrt{\lambda_N}}{Z}\right) =$$

$$= 2\,\phi\left(\frac{K\sqrt{\lambda_N}}{Z}\right) - 1 = 1 - \alpha$$

$$\frac{K\sqrt{\lambda_N}}{Z} = Z_{1-\frac{\alpha}{2}},$$

para que,

$$Z = \left(\frac{K}{Z_{1-\frac{\alpha}{2}}}\right)\sqrt{\lambda_N} = \sqrt{\frac{\lambda_N}{\lambda_F}}\ .$$

Esta ecuación se llama regla de la raíz cuadrada para la credibilidad parcial.

- **Credibilidad parcial para la cuantía media de un siniestro**:

$$Z = \sqrt{\frac{N}{\lambda_F \, C_X^2}} \; .$$

- **Credibilidad parcial para la cuantía total de los siniestros (prima pura)**:

$$Z = \sqrt{\frac{N}{\lambda_F \, (1 + \, C_X^2)}} \; .$$

Caso Binomial, $N \sim$ Binomial (E,θ):

$$\lambda_N = E\theta,$$

$$\sigma_N^2 = E\theta(1 - \theta).$$

Utilizando la aproximación normal, el estándar de credibilidad total para el número de siniestros es:

$$\frac{K\mu_N}{\sigma_N} = \frac{KE\theta}{\sqrt{E\theta(1-\theta)}} = K\sqrt{\frac{E\theta}{(1-\theta)}} = Z_{1-\frac{\alpha}{2}}$$

$$E = \left(\frac{Z_{1-\frac{\alpha}{2}}}{K}\right)^2 \left(\frac{1-\theta}{\theta}\right) = \lambda_F \left(\frac{1-\theta}{\theta}\right)$$

donde θ es la probabilidad de un siniestro y E el número de unidades de exposición.

Esta expresión puede expresarse en términos del número esperado de siniestros:

$$E\theta = \lambda_F(1-\theta).$$

Observar que

$$\lambda_F(1-\theta) \leq \lambda_F.$$

El estándar para la credibilidad total bajo Binomial es menor que bajo Poisson. Sin embargo, como θ es pequeña, $(1-\theta)$ está cerca de 1 y la diferencia entre estos dos modelos es pequeña.

Credibilidad parcial asumiendo $N \sim \text{Poisson}(\lambda_N)$

Número de siniestros	$Z = \sqrt{\dfrac{\lambda_N}{\lambda_F}}$
Cuantía media de un siniestro	$Z = \sqrt{\dfrac{N}{\lambda_F\, C_X^2}}$
Cuantía total de los siniestros	$Z = \sqrt{\dfrac{N}{\lambda_F\,(1 + C_X^2)}}$
Prima pura	$Z = \sqrt{\dfrac{N}{\lambda_F\,(1 + C_X^2)}}$

COMENTARIOS

Aunque el enfoque clásico de credibilidad es de fácil aplicación, no se basa en un principio estadístico de predicción bien adoptado. En particular, hay varias deficiencias del enfoque, tales como:

1. Este enfoque enfatiza el papel de D. No otorga ninguna importancia a la precisión de la información previa M.

2. Los estándares de credibilidad total dependen de algunos valores de parámetros desconocidos. El enfoque no aborda el tema de cómo la elección de estos parámetros puede afectar la credibilidad.

3. Hay algunas limitaciones en los supuestos que se hacen con el fin de obtener resultados analíticos tratables.

Credibilidad de Bühlmann y Bühlmann-Straub
(teoría)

Eva Boj del Val

Departamento de Matemática Económica, Financiera y Actuarial
Universidad de Barcelona

Credibilidad de Bühlmann y Bühlmann-Straub

Aunque la teoría clásica de la credibilidad aborda el problema de combinar la experiencia de siniestralidad y la información previa para actualizar la predicción, no ofrece una solución muy satisfactoria. El método se basa en la selección arbitraria de la probabilidad de cobertura α y el parámetro de precisión K. Además, por razones de manejabilidad, hay que imponer algunos supuestos restrictivos sobre la distribución de la siniestralidad analizada.

La teoría de la credibilidad de Bühlmann sitúa el problema en un marco estadístico riguroso de predicción óptima, utilizando el criterio del mínimo error cuadrático medio (ECM) para el ajuste de los modelos. Es flexible como para incorporar diversos supuestos de distribución de las variables de siniestralidad. El enfoque se amplía además para permitir combinar la experiencia de siniestralidad de distintos bloques de pólizas con diferentes exposiciones para mejorar la predicción mediante el modelo Bühlmann-Straub. Los modelos de Bühlmann y Bühlmann-Straub reconocen la interacción de dos fuentes de variabilidad en los datos, la varianza debida a las diferencias entre grupos y la varianza debida a las fluctuaciones dentro de los grupos.

Referencia básica (manual examen SOA): TSE, Y.K. *Nonlife Actuarial Models. Theory, Methods and Evaluation*. Cambridge University Press, New York, 2009.

1 1. Marco teórico

Consideremos un grupo de riesgo con una medida de siniestralidad denominada X, que puede ser el número de siniestros, la cuantía media de un siniestro, la cuantía total de los siniestros o la prima pura.

Suponemos que los perfiles de riesgo del grupo se caracterizan por un parámetro θ, que determina la distribución de la medida de siniestralidad X. Denotamos la media condicional y la varianza de X dado θ por

$$\mathrm{E}(X|\theta) = \mu_X(\theta),$$

y

$$Var(X|\theta) = \sigma_X^2(\theta).$$

Suponemos que la compañía de seguros tiene bloques similares de pólizas con distintos perfiles de riesgo. Así, el parámetro θ varía con los distintos grupos de riesgo. Tratamos θ como la realización de una variable aleatoria Θ, cuya distribución se denomina **distribución a priori**.
Cuando θ varía sobre el soporte de Θ, la media condicional y la varianza de X se convierten en variables aleatorias en Θ, y se denotan por:

$$\mu_X(\Theta) \ = \ E(X|\Theta)$$

y

$$\sigma_X^2\ (\Theta) = Var(X|\Theta).$$

El modelo de Bühlmann supone que existen n observaciones de siniestralidad, denotadas por $\boldsymbol{X}=\{X_1, \ldots, X_n\}$. Las observaciones pueden ser siniestralidades registradas en n períodos que se suponen independientes e idénticamente distribuidas como X, dependiendo del parámetro θ.

El objetivo es actualizar la predicción de X para el siguiente periodo, es decir, X_{n+1}, basándonos en $\boldsymbol{X}$. En el enfoque de Bühlmann, la solución depende de la varianza entre las medias condicionales, así como de la media de las varianzas condicionales de los grupos de riesgo.

2. Componentes de la varianza

La varianza de la media de siniestralidad X consta de dos componentes:

- La varianza entre grupos de riesgo: se debe a la aleatoriedad de los perfiles de riesgo de cada grupo y se capta mediante el parámetro Θ.
- La varianza dentro de los grupos de riesgo: se mide por la varianza condicional del grupo de riesgo.

Consideramos en primer lugar el cálculo de la media global de la medida de siniestralidad X. La **media incondicional** (o **media global**) de X mide la tendencia central global de X, promediada sobre todas las diferencias subyacentes en los grupos de riesgo:

$$E(X) = E[E(X|\Theta)] = E[\mu_X(\Theta)].$$

Así, la media incondicional es la media de las medias condicionales tomadas sobre la distribución de Θ. Para la **varianza incondicional** (o **varianza total**), el cálculo es más complicado. La varianza total de X se debe a la esperanza de la varianza en Θ así como a la varianza de X condicional en Θ:

$$Var(X) = E[Var(X|\Theta)] + Var[E(X|\Theta)].$$

La parte $Var(X|\Theta)$ mide la varianza de un determinado grupo de riesgo. Es una función de la variable aleatoria Θ y la denominamos **varianza del proceso**. Así, $E[Var(X|\Theta)]$ es el **valor esperado de la varianza del proceso (EPV)**.

La parte $E(X|\Theta)$ es la media de un grupo de riesgo determinado. A esta media condicional denominamos **media hipotética**. Así, $Var[E(X|\Theta)]$ es la **varianza de las medias hipotéticas** (**VHM**), ya que mide las variaciones de las medias de los grupos de riesgo.

varianza total =
valor esperado de la varianza del proceso + **varianza de las medias hipotéticas**

o

varianza total = **EPV** + **VHM**

o

varianza total = **media de la varianza condicional** + **varianza de la media condicional**

Notaciones:

$$E[Var(X|\Theta)] = E[\sigma_X^2(\Theta)] = \mu_{PV}$$

y

$$Var[E(X|\Theta)] = Var[\mu_X(\Theta)] = \sigma_{HM}^2$$

así

$$Var(X) = \mu_{PV} + \sigma_{HM}^2.$$

3. Credibilidad de Bühlmann

El enfoque de Bühlmann para actualizar la medida de siniestralidad se basa en un predictor lineal que utiliza observaciones pasadas. También se denomina enfoque de la mayor precisión o enfoque de mínimos cuadrados. Recordemos que para el enfoque clásico de la credibilidad, la predicción actualizada U viene dada por:

$$U = ZD + (1 - Z)M.$$

El método de credibilidad de Bühlmann tiene una ecuación básica similar, en la que D es la media muestral de los datos y M es la media global *a priori* $E(X)$. La credibilidad de Bühlmann, Z, depende del tamaño de la muestra n y de la relación entre EPV y VHM, k:

$$k = \frac{\mu_{PV}}{\sigma_{HM}^2} = \frac{EPV}{VHM}.$$

En particular, Z varía con n y k de la siguiente manera:

1. Z aumenta con el tamaño de la muestra n.
2. Z aumenta con la distinción de los grupos de riesgo. Los grupos de riesgo son más distinguibles cuando k es pequeño. Por lo tanto, Z aumenta a medida que k disminuye.

Supuestos del modelo de Bühlmann y fórmula de actualización del predictor lineal de mínimo ECM:

1. $\boldsymbol{X}$ = $\{X_1, \ldots, X_n\}$ son medidas de siniestralidad que se distribuyen independiente e idénticamente como la variable aleatoria X. La distribución de X depende del parámetro θ.

2. El parámetro θ es una realización de una variable aleatoria Θ. Dado θ, la media condicional y la varianza de X son:

$$E(X|\theta) = \mu_X(\theta)$$

y

$$Var(X|\theta) = \sigma_X^2(\theta).$$

3. La media incondicional de X es $E(X) = E[E(X|\Theta)] = \mu_x$. La media de la varianza condicional de X es

$$E[Var(X|\Theta)] = E[\sigma_X^2(\Theta)] = \mu_{PV} = EPV,$$

y la varianza de la media condicional es

$$Var[E(X|\Theta)] = Var[\mu_X(\Theta)] = \sigma_{HM}^2 = VHM.$$

Así,

$$Var(X) = E[Var(X|\Theta)] + Var[E(X|\Theta)] = \mu_{PV} + \sigma_{HM}^2 = EPV + VHM.$$

4. Se formula un predictor de X_{n+1} basado en una función lineal de X, donde se supone que X_{n+1} tiene la misma distribución que X. El predictor minimiza el ECM en la predicción de X_{n+1} sobre la distribución conjunta de Θ, X_{n+1} y X. En concreto, el predictor viene dado por

$$X_{n+1} = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$$

donde $\beta_0, \beta_1, \ldots, \beta_n$ se estiman minimizando el ECM definido como:

$$ECM = E\left[\left(X_{n+1} - \hat{X}_{n+1}\right)^2\right].$$

La solución a esta minimización nos da el predictor lineal de menor MSE de X_{n+1}, denominado **prima de Bühlmann**:

$$\hat{X}_{n+1} = \widehat{\beta_0} + \widehat{\boldsymbol{\beta}'}\boldsymbol{X} = \frac{n\bar{X}}{n+k} + \frac{k\mu_X}{n+k} = Z\bar{X} + (1-Z)\mu_X$$

donde el **factor de credibilidad de Bühlmann** o **credibilidad de Bühlmann** es:

$$Z = \frac{n}{n+k}$$

siendo el **parámetro de credibilidad de Bühlmann**:

$$k = \frac{\mu_{PV}}{\sigma_{HM}^2} .$$

Y la media, $\bar{X}$, se refiere a:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i .$$

La credibilidad de Bühlmann, Z, es función de k y del tamaño n de los datos.

- Para predecir el número de siniestros, N, el tamaño de la muestra n es el número de periodos en los que se agrega el número de siniestros.

- Para predecir la cuantía media de un siniestro, X, el tamaño de la muestra n es el número de siniestros.

- Como la cuantía total de los siniestros, S, se refiere al pago total de los siniestros por periodo, el tamaño de la muestra es el número de periodos de siniestralidad.

4. Credibilidad de Bühlmann-Straub

Una limitación importante de la teoría de la credibilidad de Bühlmann es que las observaciones de siniestralidad X_i se suponen idénticamente distribuidas. Este supuesto se incumple si los datos corresponden a distintos periodos con distintas exposiciones. El modelo de credibilidad de Bühlmann-Straub amplía la teoría de Bühlmann a los casos en los que los datos de siniestros X_i no están distribuidos idénticamente. En concreto, se supone que la varianza del proceso de la medida de la siniestralidad depende de la exposición. Denotamos la exposición por m_i, y la pérdida por unidad de exposición por X_i . La exposición no tiene por qué ser el número de asegurados, aunque a menudo puede ser el caso. A continuación, suponemos lo siguiente para la varianza condicional de X_i

$$Var(X_i|\Theta) = \frac{\sigma_X^2(\Theta)}{m_i}$$

para una $\sigma_X^2(\Theta)$ convenientemente definida.

Algunos ejemplos:

1. X_i es el número medio de siniestros por asegurado en el año i, $\sigma_X^2(\Theta)$ es la varianza de la frecuencia de siniestralidad de un asegurado en un año y la exposición m_i es el número de asegurados cubiertos en el año i.

2. X_i es la pérdida agregada media por mes del i-ésimo bloque de pólizas, $\sigma_X^2(\Theta)$ es la varianza de la pérdida agregada del bloque en un mes, y la exposición m_i es el número de meses del seguro para el i-ésimo bloque de pólizas.

3. X_i es la pérdida media por unidad de prima en el año i, $\sigma_X^2(\Theta)$ es la varianza de la cuantía de un asegurado por año dividida por la prima por asegurado, y la exposición m_i es el importe de las primas percibidas en el año i. Para verlo, supongamos que hay N_i asegurados en el año i, cada uno de los cuales paga una prima *P*. Así pues, $m_i = N_i P$,

$$
\begin{aligned}
Var(X_i|\Theta) &= Var(\frac{perdida\ por\ asegurado}{P}) \\
&= \frac{1}{P^2}\frac{Var(cuantia\ de\ un\ asegurado)}{N_i} \\
&= \frac{1}{m_i}\frac{Var(cuantia\ de\ un\ asegurado)}{P} \\
&= \frac{\sigma_X^2(\Theta)}{m_i} \ .
\end{aligned}
$$

En cada uno de los ejemplos anteriores, las distribuciones de X_i no son idénticas. La varianza condicional de X_i varía con m_i .

Formalmente los supuestos del modelo de credibilidad de Bühlmann-Straub y la fórmula de predicción son:

1. Sea m_i la exposición en el periodo i y X_i la siniestralidad por unidad de exposición, para $i = 1, \dots, n$. Supongamos que $\boldsymbol{X} = \{X_1, \dots, X_n\}$ se distribuyen de forma independiente (pero no idéntica) y que la distribución de X_i depende del parámetro θ.

2. El parámetro θ es una realización de una variable aleatoria Θ. Dado θ, la media condicional y la varianza de $\boldsymbol{X}$ son:

$$E(X_i|\theta) = \mu_X(\theta)$$

y

$$Var(X_i|\theta) = \frac{\sigma_X^2(\theta)}{m_i}$$

para $i = 1, \dots, n$, donde $\sigma_X^2(\theta)$ se define adecuadamente como en los ejemplos anteriores.

3. La media de la varianza condicional de X_i es

$$E\left[Var(X_i|\Theta)\right] = E\left[\frac{\sigma_X^2(\Theta)}{m_i}\right] = \frac{\mu_{PV}}{m_i}.$$

para $i = 1, \dots, n$, donde $\mu_{PV} = E[\sigma_X^2(\Theta)]$, y la varianza de la media condicional es

$$Var[E(X_i|\Theta)] = Var[\mu_X(\Theta)] = \sigma_{HM}^2.$$

4. El predictor de Bühlmann-Straub minimiza el ECM de todos los predictores de X_{n+1} que son lineales en $\boldsymbol{X}$ sobre la distribución conjunta de Θ, X_{n+1} y X. El predictor viene dado por

$$\hat{X}_{n+1} = \beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n$$

donde $\beta_0, \beta_1, \dots, \beta_n$ se eligen tales que minimizan el MSE.

La varianza total de X_i :

$$Var(X_i) = E[Var(X_i|\Theta)] + Var[E(X_i|\Theta)] = \frac{\mu_{PV}}{m_i} + \sigma_{HM}^2.$$

Denotamos:

$$m = \sum_{i=1}^{n} m_i$$

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{n} m_i X_i$$

y,

$$Z = \frac{m}{m+k}$$

$$k = \frac{\mu_{PV}}{\sigma_{HM}^2}.$$

Sea

$$\mu_{PV} = m_i E[Var(X_i|\Theta)]$$

y

$$\sigma_{HM}^2 = Var[E(X_i|\Theta)],$$

entonces la predicción de Bühlmann-Straub de X_{n+1} es:

$$\hat{X}_{n+1} = Z\bar{X} + (1 - Z)\mu_X$$

donde

$$\mu_X = E(X_i) \text{ para } i = 1, \ldots, n$$

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{n} m_i X_i$$

$$Z = \frac{m}{m + k}$$

$$k = \frac{\mu_{PV}}{\sigma_{HM}^2} .$$

En el caso particular en que las exposiciones de todos los periodos sean iguales, es decir, $m_i = \overline{m}$ para $i = 1, \dots, n$, entonces, a partir del predictor de Bühlmann-Straub se obtiene como caso particular el predictor de Bühlmann.

En los ejemplos de este tema suponemos que se conocen las componentes de la varianza. En la práctica, deben estimarse a partir de los datos. Para ello se puede considerar la estimación empírica de los modelos de credibilidad de Bühlmann y Bühlmann-Straub en los casos en que se desconocen EPV y VHM.

Método de Bayes empírico aplicado a credibilidad clásica y a credibilidad de Bühlmann y Bühlmann-Straub

(teoría)

Eva Boj del Val

Departamento de Matemática Económica, Financiera y Actuarial
Universidad de Barcelona

La aplicación de los métodos de predicción de la siniestralidad futura:

- Credibilidad clásica
- Credibilidad de Bühlmann y Bühlmann-Straub
- Método bayesiano

requieren el conocimiento o la suposición de algunos parámetros desconocidos del modelo.

Método empírico de Bayes:

- no paramétrico
- semiparamétrico
- paramétrico

dependiendo de los supuestos relativos a la distribución *a priori* y a la verosimilitud.

Nos centramos principalmente en los modelos de credibilidad de Bühlmann y Bühlmann-Straub, cuya aplicación no paramétrica es relativamente sencilla.

Referencia básica (manual examen SOA): TSE, Y.K. *Nonlife Actuarial Models. Theory, Methods and Evaluation*. Cambridge University Press, New York, 2009.

Para el **método de fluctuación limitada o credibilidad clásica**, dependiendo de la variable de pérdida de interés, se requiere la media y/o la varianza de la variable de pérdida. Por ejemplo, para determinar si se alcanza la plena credibilidad en la predicción de la frecuencia de siniestralidad, necesitamos conocer λ_N, que puede estimarse mediante la media muestral de la frecuencia de los siniestros. Para predecir la cuantía de los siniestros y la pérdida agregada/prima pura, también es necesario conocer el coeficiente de variación de la variable de siniestralidad, C_X, que puede estimarse de la siguiente manera

$$\hat{C}_X = \frac{S_X}{\bar{X}}$$

donde S_X, y $\bar{X}$ son la desviación típica muestral y la media muestral de X, respectivamente.

Para los **métodos de Bühlmann y Bühlmann-Straub**, las cantidades clave necesarias son el valor esperado de la varianza del proceso, μ_{PV}, y la varianza de las medias hipotéticas, σ^2_{HM}. Estas cantidades dependen de la hipótesis de la distribución *a priori* de los parámetros de riesgo y de la distribución condicional de la variable aleatoria de siniestralidad.

Para el **método bayesiano**, la siniestralidad prevista puede obtenerse con relativa facilidad si la distribución *a priori* es conjugada con la verosimilitud. Sin embargo, la media *a posteriori*, que es el predictor bayesiano de la siniestralidad futura, depende de los hiperparámetros de la distribución *a posteriori*. Por lo tanto, para la aplicación empírica del método bayesiano, hay que estimar los hiperparámetros.

Estas cantidades pueden deducirse de la hipótesis bayesiana y dependen tanto de la distribución *a priori* como de la probabilidad plasmada en la función de verosimilitud. En un enfoque estrictamente bayesiano, la distribución *a priori* viene dada y la inferencia se basa en ella. En aplicaciones prácticas, cuando los investigadores no pueden establecer la distribución *a priori*, se pueden aplicar métodos empíricos para estimar los hiperparámetros. Esta metodología se denomina **método de Bayes empírico**. El método Bayes empírico adopta el enfoque bayesiano de análisis, pero trata los hiperparámetros como cantidades que deben obtenerse a partir de los datos.

Dependiendo de los supuestos sobre la distribución *a priori* y la verosimilitud, la estimación empírica de Bayes puede adoptar uno de los siguientes enfoques:

1 **Enfoque no paramétrico**: En este enfoque, no se hacen suposiciones sobre las formas particulares de la densidad *a priori* de los parámetros de riesgo $f_{\Theta}\,(\theta)$ y la densidad condicional de la variable de siniestralidad $f_{X|\Theta}\,(x|\theta)$. El método es muy general y se aplica a una amplia gama de modelos.

2 **Enfoque semiparamétrico**: En algunas aplicaciones prácticas la experiencia previa puede sugerir una distribución particular para la variable de siniestralidad X, mientras que la especificación de la distribución *a priori* sigue siendo elusiva. En tales casos, pueden hacerse suposiciones paramétricas sobre $f_{X|\Theta}\,(x|\theta)$, mientras que la distribución *a priori* de los parámetros de riesgo $f_{\Theta}\,(\theta)$ permanece sin especificar.

3 **Enfoque paramétrico**: Cuando el investigador hace suposiciones específicas sobre $f_{X|\Theta}\,(x|\theta)$ y $f_{\Theta}\,(\theta)$, la estimación de los parámetros del modelo puede llevarse a cabo mediante el método de estimación de máxima verosimilitud. Aunque en algunos casos la estimación por máxima verosimilitud puede derivarse analíticamente, en muchas situaciones debe calcularse numéricamente.

Estimación no paramétrica

Para aplicar la predicción de **credibilidad clásica** se utiliza una estimación del coeficiente de variación $\hat{C}_X$. Bajo el supuesto de una muestra aleatoria, S_X, y $\bar{X}$ son estimadores consistentes para la desviación típica de la población y la media, respectivamente, independientemente de la distribución real de la variable aleatoria de siniestralidad X. Así pues, $\hat{C}_X$ es un estimador de C_X, aunque generalmente no es insesgado.

Para la aplicación de los **modelos de credibilidad de Bühlmann y Bühlmann-Straub** las cantidades clave necesarias son el valor esperado del proceso varianza, μ_{PV}, y la varianza de las medias hipotéticas, σ_{HM}^2, que en conjunto determinan el parámetro de credibilidad k de Bühlmann.

Veamos estimaciones insesgadas no paramétricas de estas cantidades, en la medida en que la insesgadez se mantiene bajo el leve supuesto de que las observaciones de siniestralidad son estadísticamente independientes, y que no se hace ningún supuesto específico sobre la probabilidad de las variables aleatorias de siniestralidad y la distribución *a priori* de los parámetros de riesgo.

Consideremos múltiples grupos de riesgo, cada uno con múltiples muestras de observaciones de siniestralidad a lo largo de (posiblemente) períodos diferentes:

1) Sea X_{ij} variable aleatoria siniestralidad por unidad de exposición y m_{ij} la cantidad de exposición. El índice i denota el i-ésimo grupo de riesgo, para $i = 1, \dots, r$, con $r > 1$. Dado i, el índice j denota la j-ésima observación de siniestralidad en el i-ésimo grupo, para $j = 1, \dots, n_i$, donde $n_i > 1$ para $i = 1, \dots, r$. El número de observaciones n_i en cada grupo de riesgo puede diferir. Podemos pensar en j como un índice de un individuo dentro del grupo de riesgo o un período del grupo de riesgo. Así, para el i-ésimo grupo de riesgo tenemos observaciones de siniestralidad de n_i individuos o periodos.
2) Sea X_{ij} se supone que se distribuyen de forma independiente. El parámetro de riesgo del grupo i-ésimo se denota por θ_i, que es una realización de la variable aleatoria Θ_i. Suponemos que Θ_i se distribuye de forma independiente e idéntica como Θ.
3) Media de la varianza del proceso y varianza de las medias hipotéticas:
   Esperanza y varianza condicional

$$E\left(X_{ij} \,\middle|\, \Theta = \theta_i\right) = \mu_X(\theta_i) \quad \text{para} \quad i = 1, \dots, r \quad j = 1, \dots, n_i$$

$$Var\left(X_{ij} \middle| \theta_i\right) = \frac{\sigma_X^2(\theta_i)}{m_{ij}} \quad \text{para} \quad i = 1, \dots, r \quad j = 1, \dots, n_i$$

Media global

$$\mu_X = E[\mu_X(\Theta_i)] = E[\mu_X(\Theta)]$$

Media de la varianza del proceso

$$\mu_{PV} = E[\sigma_X^2(\Theta_i)] = E[\sigma_X^2(\Theta)]$$

Varianza de las medias hipotéticas

$$\sigma_{HM}^2 = Var\,[\mu_X(\Theta_i)] = Var\,[\mu_X(\Theta)]$$

Exposición total para el i-ésimo grupo de riesgo

$$m_i = \sum_{j=1}^{n_i} m_{ij} \qquad i = 1, \dots, r$$

Exposición total sobre todos los grupos de riesgo

$$m = \sum_{i=1}^{r} m_i$$

Media ponderada por exposición del i-ésimo grupo de riesgo

$$\bar{X}_i = \frac{1}{m_i}\sum_{j=1}^{ni} m_{ij}X_{ij} \qquad i = 1, \ldots, r$$

Media ponderada global

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{r} m_i\bar{X}_i$$

Predictor de credibilidad de Bühlmann-Straub:

$$Z_i\bar{X}_i + (1 - Z_i)\mu_X$$

donde

$$Z_i = \frac{m_i}{m_i + k}$$

con

$$k = \frac{\mu_{PV}}{\sigma_{HM}^2}$$

Para aplicar la predicción de credibilidad, necesitamos estimar μ_X, μ_{PV} , y σ_{HM}^2. Es natural estimar μ_X mediante $\bar{X}$.

Estimadores insesgados de μ_{PV} y de σ_{HM}^2

$$\hat{\mu}_{PV} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}(X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}$$

$$\hat{\sigma}_{HM}^2 = \frac{\left[\sum_{i=1}^{r} m_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\hat{\mu}_{PV}}{m - \frac{1}{m}\sum_{i=1}^{r} m_i^2}$$

El modelo de Bühlmann es un caso especial del modelo de Bühlmann-Straub, $m_{ij} \equiv 1$. Estimadores insesgados:

$$\tilde{\mu}_{PV} = \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2}{\sum_{i=1}^{r}(n_i - 1)}$$

$$\tilde{\sigma}_{HM}^2 = \frac{\left[\sum_{i=1}^{r} m_i(\bar{X}_i - \bar{X})^2\right] - (r-1)\,\tilde{\mu}_{PV}}{m - \frac{1}{m}\sum_{i=1}^{r} m_i^2}$$

donde $n = \sum_{i=1}^{r} n_i$. En particular, si todos los grupos de riesgo tienen el mismo tamaño de muestra, de modo que $n_i = n^*$ para $i = 1, \ldots, r$ entonces:

$$\tilde{\mu}_{PV} = \frac{1}{r(n^* - 1)}\left[\sum_{i=1}^{r}\sum_{j=1}^{n^*}\left(X_{ij} - \bar{X}_i\right)^2\right] = \frac{1}{r}\left[\sum_{i=1}^{r} S_i^2\right]$$

donde S_i^2 es la varianza muestral del grupo i-ésimo y

$$\tilde{\sigma}_{HM}^2 = \frac{1}{(r-1)}\left[\sum_{i=1}^{r}(\bar{X}_i - \bar{X})^2\right] - \frac{\tilde{\mu}_{PV}}{n^*} = S^2 - \frac{\tilde{\mu}_{PV}}{n^*}$$

donde S^2 es la varianza muestral entre grupos.

Predictor de credibilidad de Bühlmann-Straub:

$$\widehat{Z}_i \bar{X}_i + \left(1 - \widehat{Z}_i\right)\bar{X}$$

donde

$$\widehat{Z}_i = \frac{m_i}{m_i + \hat{k}}$$

con

$$\hat{k} = \frac{\hat{\mu}_{PV}}{\hat{\sigma}_{HM}^2}$$

Predictor de credibilidad de Bühlmann:

$$\widetilde{Z}_i \bar{X}_i + \left(1 - \widetilde{Z}_i\right) \bar{X}$$

donde

$$\widetilde{Z}_i = \frac{n_i}{n_i + \tilde{k}}$$

con

$$\tilde{k} = \frac{\tilde{\mu}_{PV}}{\tilde{\sigma}^2_{HM}}$$

Nótese que $\hat{\sigma}^2_{HM}$ y $\tilde{\sigma}^2_{HM}$ pueden ser negativas en aplicaciones empíricas. En tales circunstancias, pueden fijarse en cero, lo que implica que $\hat{k}$ y $\tilde{k}$ serán infinito, y que $\widehat{Z}_i$ y $\widetilde{Z}_i$ serán cero para todos los grupos de riesgo. En efecto, si las medias hipotéticas no tienen variación, los grupos de riesgo son homogéneos y no debería haber ponderación diferencial. En resumen, la pérdida prevista es la media general.

Pérdida total experimentada igual *Pérdida total predicha*

La pérdida total experimentada es $m\bar{X} = \sum_{i=1}^{r} m_i \ X_i$. La pérdida total predicha será en general diferente de la pérdida total experimentada. Si se desea equiparar la pérdida total prevista a la pérdida total experimentada, es necesario realizar algún reajuste. Para ello se puede utilizar una estimación alternativa de la siniestralidad media, denotada por $\hat{\mu}_X$, y se sustituye en lugar de $\bar{X}$.

Pérdida total prevista

$$\sum_{i=1}^{r} m_i \left[\widehat{Z}_i \overline{X}_i + \left(1 - \hat{Z}_i\right)\hat{\mu}_X\right] = \sum_{i=1}^{r} m_i \left\{\left[1 - \left(1 - \widehat{Z}_i\right)\right] \overline{X}_i + \left(1 - \hat{Z}_i\right)\hat{\mu}_X\right\}$$

como $m_i\left(1 - \widehat{Z}_i\right) = \widehat{Z}_i \hat{k}$, la ecuación anterior puede escribirse como

$$Perdidas\ totales\ predecidas = \sum_{i=1}^{r} m_i \left\{\left[1 - \left(1 - \widehat{Z}_i\right)\right] \overline{X}_i + \left(1 - \widehat{Z}_i\right)\hat{\mu}_X\right\}$$
$$= \sum_{i=1}^{r} m_i X_i + \sum_{i=1}^{r} m_i \left(1 - \widehat{Z}_i\right)\left(\hat{\mu}_X - \overline{X}_i\right)$$
$$= Perdidas\ totales\ experimentadas + \hat{k} \textstyle\sum_{i=1}^{r} \widehat{Z}_i \left(\hat{\mu}_X - \overline{X}_i\right)$$

Para equilibrar la pérdida total prevista y la pérdida total experimentada se ha de cumplir que $\hat{k} \sum_{i=1}^{r} \widehat{Z}_i \left(\hat{\mu}_X - \overline{X}_i\right) = 0$, lo que implica que:

$$\hat{\mu}_X = \frac{\sum_{i=1}^{r} \widehat{Z}_i \overline{X}_i}{\sum_{i=1}^{r} \widehat{Z}_i}$$

y la pérdida prevista para el grupo i-ésimo es

$$\widehat{Z}_i \overline{X}_i + \left(1 - \widehat{Z}_i\right)\hat{\mu}_X$$

Método Bayesiano:

Los modelos de credibilidad clásico y de Bühlmann actualizan la predicción de la siniestralidad futura basándose en la experiencia reciente de siniestralidad y en la información previa (*a priori*) existente. En estos modelos, la variable aleatoria de siniestralidad X tiene una distribución que varía según los distintos grupos de riesgo. A partir de una muestra de n observaciones de siniestralidad, se actualiza el valor predicho para el siguiente periodo. La predicción es una media ponderada de la media muestral de X y la media *a priori*, donde las ponderaciones dependen de la distribución de X entre los distintos grupos de riesgo.

Formulado como un problema estadístico adecuado para el enfoque bayesiano:

1) Sea X variable aleatoria (frecuencia de siniestralidad, cuantía de los siniestros y cuantía agregada) de un grupo de riesgo. La distribución de X depende de un parámetro θ, que varía con los distintos grupos de riesgo y, por tanto, se trata como la realización de una variable aleatoria Θ.

2) Sea Θ una variable aleatoria con una distribución estadística llamada distribución *a priori*. La función de densidad de probabilidad *a priori* de Θ se denota por $f_\Theta(\theta|\gamma)$ (o simplemente $f_\Theta(\theta)$), que depende del parámetro γ, llamado hiperparámetro.

3) La función de densidad de probabilidad condicional de X dado el parámetro θ se denota por $f_{X|\Theta}(X|\theta)$. Supongamos que $X = \{X_1, X_2 \dots, X_n\}$ es una muestra aleatoria de X, y $x = (x_1, x_2 \dots, x_n)$ es una realización de X. La función de densidad de probabilidad condicional de X es:

$$f_{X|\Theta}(x|\theta) = \prod_{i=1}^{n} f_{X|\Theta}(x_i|\theta)$$

Llamamos $f_{X|\Theta}(X|\theta)$ a la función de verosimilitud.

4) A partir de los datos de la muestra x, se actualiza la distribución de Θ. La función de densidad de probabilidad condicional de Θ dado x se denomina función de densidad de probabilidad *a posteriori* y se denota por $f_{\Theta|X}(\theta|x)$.

5) Se calcula una estimación de la media, que es una función de Θ, utilizando la función de densidad de probabilidad *a posteriori* Θ. Esta estimación se denomina estimación de Bayes y es la predicción de la siniestralidad futura.

La función de verosimilitud y la función de probabilidad *a priori* determinan conjuntamente la función de probabilidad *a posteriori*, que se utiliza para la inferencia estadística.

Distribución *a posteriori* del parámetro

Dada la distribución *a priori* de Θ y la función de verosimilitud de X, la función de densidad de probabilidad conjunta de Θ y X se puede obtener de la siguiente manera:

$$P(A|B) = P(A|B)P(B) = P(B|A)P(A)$$
$$P(B|A) = \frac{P(B|A)P(B)}{P(A)}$$

$$f_{\Theta X}(\theta, x) = f_{X|\Theta}(x|\theta) f_{\Theta}(\theta)$$

La función de densidad de probabilidad marginal de x es:

$$f_X(x) = \int_{\theta \in \Omega_\Theta} f_{X|\Theta}(x|\theta)\, f_\Theta(\theta) d\theta$$

donde Ω_Θ es el soporte de f_Θ. La función de densidad de probabilidad condicional de Θ dado x:

$$f_{\Theta|X}(\theta|x) = \frac{f_{\Theta X}(\theta, x)}{f_X(x)} = \frac{f_{X|\Theta}(x|\theta) f_\Theta(\theta)}{\int_{\theta \in \Omega_\Theta} f_{X|\Theta}(x|\theta)\, f_\Theta(\theta) d\theta}$$

$$k(x) = \frac{1}{\int_{\theta \in \Omega_\Theta} f_{X|\Theta}(x|\theta)\, f_\Theta(\theta) d\theta}$$

$k(x)$ es una constante de proporcionalidad que integra hasta 1 y que no depende de θ. Re-escribimos:

$$f_{\Theta|X}(\theta|x) = k(x) f_{X|\Theta}(x|\theta) f_\Theta(\theta)$$

$$f_{\Theta|X}(\theta|x) \propto f_{X|\Theta}(x|\theta) f_\Theta(\theta)$$

La expresión $f_{X|\Theta}(x|\theta) f_\Theta(\theta)$ nos permite identificar la forma funcional de la función de densidad de probabilidad *a posteriori* en términos de θ sin calcular la función de densidad de probabilidad marginal de X.

Para la función de pérdida de error cuadrática,

$$L[\mu_x(\Theta), w(x)] = [\mu_x(\Theta) - w(x)]^2$$

la regla de decisión que minimiza la pérdida esperada es $w^*(x) = E[\mu_x(\Theta)|x]$ que es la media *a posteriori* denotada por $\mu_x(\mathrm{x})$, de modo que la prima bayesiana:

$$\hat{\mu}_X(x) = E[\mu_X(\Theta), x] = \int_{\theta \in \Omega_\Theta} \mu_x(\Theta)\, f_{\Theta|\mathrm{X}}(\theta|x)d\theta$$

La prima bayesiana también puede interpretarse como la esperanza condicional de X_{n+1} dado X:

$$E(X_{n+1}|x) = \int_0^\infty X_{n+1}\, f_{X_{n+1}\,|\,X}(X_{n+1}|\, X)dX_{n+1}$$

$$E(X_{n+1}|x) = \int_0^\infty X_{n+1}\, [\int_{\theta \in \Omega_\Theta} f_{X_{n+1}\,|\,\Theta}(X_{n+1}|\,\theta) f_{\,\Theta\,|\,\mathrm{X}}(\theta|\,x)d\theta]dX_{n+1}$$

$$= \int_{\theta \in \Omega_\Theta} [\int_0^\infty X_{n+1}\, f_{X_{n+1}\,|\,\Theta}(X_{n+1}|\,\theta)dX_{n+1}] f_{\Theta|\mathrm{X}}(\theta, x)d\theta = \int_{\theta \in \Omega_\Theta} E\,(X_{n+1}|\,\theta) f_{\Theta|\mathrm{X}}(\theta|x)d\theta$$

$$= \int_{\theta \in \Omega_\Theta} \mu_x(\Theta)\, f_{\Theta|\mathrm{X}}(\theta\mid x)d\theta = E[\mu_X(\Theta)|x]$$

Distribuciones conjugadas

Sea $f_\Theta(\theta\ ,\gamma)$ la distribución *a priori* de Θ, donde γ es el hiperparámetro. La distribución *a priori* de $f_\Theta(\theta\ ,\gamma)$ es conjugada con la función de verosimilitud $f_{X|\Theta}(x|\theta)$ si la distribución *a posteriori* es igual a $f_\Theta(\theta,\gamma^*)$, que tiene la misma forma funcional que la distribución *a priori* pero, generalmente, un hiperparámetro diferente γ^*. La *a priori* y la *a posteriori* pertenecen a la misma familia de distribuciones.

Algunos ejemplos de aplicación actuarial:

A priori $f_\Theta(\theta,\gamma)$	**Verosimilitud** $f_{X\|\Theta}(x\|\theta)$	***A posteriori*** $f_\Theta(\theta,\gamma^*)$
Gamma	Poisson	Gamma
Beta	Negative Binomial	Beta
Beta	Bernoulli	Beta
Beta	Bernoulli	Beta
Beta	Geometric	Beta
Gamma	Gamma	Gamma
Normal	Normal	Normal
Pareto	Uniform	Pareto
Gamma	Normal	Gamma
Gamma	Exponential	Gamma

Número de siniestros

Cuantía de siniestros

Se obtiene <u>credibilidad exacta</u>, pues las estimaciones de la predicción con el método bayesiano son las mismas que con credibilidad de Bühlmann.

Relacionado con la idea de un <u>sistema de tarificación *a posteriori*</u>.

Por ejemplo, <u>Gamma – Poisson – Gamma</u>:

Sea $X = \{X_1, \ldots, X_{n+1}\}$ Poisson independientes e idénticamente distribuidas $PN(\lambda)$. Suponemos que $\Lambda \sim \mathrm{Gamma}\ (\alpha, \beta)$. La distribución *a posteriori* de Λ es Gamma (α^*, β^*), donde $\alpha^* = \alpha + \mathrm{n}\bar{x}$ y $\beta^* = [\,n + \frac{1}{n}]^{-1} = \frac{\beta}{n\beta+1}$. La prima bayesiana es:

$$\hat{\mu}_X(x) = E[\mu_X(\Theta), x] = \alpha^*\,\beta^* = \alpha + \mathrm{n}\bar{x}\;\frac{\beta}{n\beta+1} = \frac{\alpha\beta}{n\beta+1} + \frac{\mathrm{n}\bar{x}\beta}{n\beta+1}$$

$$= \frac{\mathrm{n}\beta}{n\beta+1}\,\bar{x} + \left(1 - \frac{\mathrm{n}\beta}{n\beta+1}\right)\alpha\beta$$