

Teoría de la credibilidad

Introducción

La teoría de la credibilidad es el conjunto de técnicas actuariales que permiten al asegurador ajustar de modelo sistemático las primas de los seguros en función de la experiencia de la siniestralidad ocurrida.

En la teoría de la credibilidad tienen roles primordiales los dos tipos de riesgo ya considerados: el riesgo *individual* y el riesgo *colectivo*, y se da una solución rigurosa al problema de cómo analizar la información proveniente de estas dos fuentes, para calcular la prima de seguros y obtener una tarifa justa.

La teoría de la credibilidad como disciplina matemática, utiliza diversas herramientas de varios campos de las matemáticas: Estadística Bayesiana, análisis funcional, mínimos cuadrados, modelos de espacio de estados, entre muchos otros. Varios autores, Bailey, Longley-Cook, Mayerson, Bühlmann, Straub, Jewell, entre otros, se han dado a la tarea de dar una fundamentación matemática rigurosa a esta teoría, que la ha convertido en una de las ramas más atractiva y estudiada de la ciencia actuarial. Uno de sus principales usos aparece en el seguro de automóviles, en el que las primas se van transformando paulatinamente a medida que se incorpora información sobre la siniestralidad, dando origen a los denominados sistemas de tarificación *bonus-malus*.

El término *credibilidad* se introdujo por primera vez en USA antes de la primera guerra mundial, en relación con los sistemas de ajuste de primas en seguros de compensación obrera o seguros de accidentes. Por ese entonces, numerosas empresas ejercieron una fuerte presión a las aseguradoras dada la baja siniestralidad laboral y la elevada tasa de actividad, para que se les reconociera este hecho en los importes de primas a pagar.

Withney (1918) publicó los primeros trabajos en esta materia con la aparición en los *Proceedings de la Casualty Actuarial Society*, de una forma simple, a través de una matemática elemental, propone que la prima que debe pagar un asegurado considere tanto la experiencia individual (del asegurado) y la del colectivo (la carta de seguros). De esta manera, la estimación del monto de la prima, se calculará como:

$$P = Z \cdot X + (1 - Z) \cdot C \quad (1)$$

Con X la experiencia individual, C es la información disponible del colectivo y Z es un factor que pondera estas dos informaciones, conocido como *factor de credibilidad*. Esta expresión dio respuesta a la idea que rondaba la mente de muchos actuarios de la época. Encontrar un

mecanismo que permitiera asignar a estos dos tipos de información, la individual y la colectiva, un peso o ponderación que las complementara para la determinación de la prima a cobrar

Intuitivamente, este factor de credibilidad, Z , debería satisfacer las siguientes condiciones

- Debe ser una función del tiempo de vigencia de la póliza, n , i.e., $Z = Z(n)$
- Debe ser una función creciente de n , de tal manera que converja a *uno* si $n \rightarrow \infty$ y tienda a *cero* cuando $n \rightarrow 0$. Este ultimo caso, ($n = 0$), implicaría que no se tiene información sobre el asegurado (sería un contrato nuevo), y la prima a cobrar sería, C , la que se basa en la información del colectivo. En la medida que se incremente la información del asegurado (que n crezca), entonces esta información empezaría a tener más peso en el cálculo de la prima a cobrar, i.e., la experiencia de la siniestralidad del asegurado tendría mayor verosimilitud o credibilidad. En el caso extremo, ($n \rightarrow \infty$), el valor de la prima debería ser X , esto es, la prima debería basarse únicamente en la experiencia individual de la siniestralidad del asegurado.
- El factor de la credibilidad, Z , debería ser también una función creciente de la varianza de las primas teóricas, con límite *uno* cuando esta varianza tienda a infinito, y *cero* cuando tienda a cero. La lógica de esta cuestión es que si la cartera no es *heterogénea*, i.e., es *homogénea* entonces la prima basada en la información del colectivo sería el mejor estimador de la prima individual. Por el contrario, una mayor heterogeneidad de la cartera, debería propiciar un mayor peso a la información individual del asegurado.

A mediados del *siglo XX* empezaba a tomar forma un nuevo enfoque de la estadística, la *Estadística Bayesiana*. No pasó mucho tiempo para que se constatará que muchos estimadores de Bayes, obtenidos para ciertas verosimilitudes (distribución conjunta de los datos) y la distribución *A priori* o inicial natural conjugada del parámetro o parámetros que determinan esta verosimilitud, correspondían a la expresión (1). De hecho, Whetney (1918) ya señalaba que el problema de credibilidad era un caso de cálculo de probabilidades inversas (teorema de Bayes). En el trabajo de Mayerson (1964) se utilizan por primera vez los términos de credibilidad y estadística Bayesiana.

Bajo el enfoque Bayesiano, la fórmula de credibilidad (1) puede interpretarse también de la siguiente manera. Puede verse a C como la información a priori (basada, por ejemplo, en contratos similares) y X la nueva información obtenida mediante la observación de la siniestralidad de los últimos años. Finalmente, la prima, P , es el resultado de combinar la información a priori con la información adquirida para obtener un *estimador actualizado* de la prima. Por lo tanto, la teoría de la credibilidad es un proceso Bayesiano que combina la información inicial o apriori con la información muestral para lograr una actualización del estimador de la prima.

Consideremos un riesgo determinado que proveniente de un conjunto de asegurados vigentes por un periodo determinado. Si este grupo de asegurados es homogéneo, en el sentido de que todos sus miembros tienen la misma probabilidad de realizar una reclamación, entonces es razonable aplicar una misma prima para todos ellos. Sin embargo, cuando el grupo no es homogéneo, o bien, al paso del tiempo aparecen factores de heterogeneidad dentro del mismo, habrá subgrupos de bajo riesgo y otros de alto riesgo. Cobrar una misma prima a todos resultaría injusto, y no sería conveniente para la aseguradora, pues, eventualmente, los asegurados de bajo riesgo buscarían un mejor trato con otra aseguradora. La idea fundamental es aplicar primas menores a los asegurados de bajo riesgo y primas mayores a los de alto riesgo, con base en el historial de reclamaciones que cada uno de los asegurados o subgrupos hayan realizado durante los periodos anteriores. En la teoría de la credibilidad se estudian métodos para el cálculo de primas a través de la combinación de la experiencia individual (historial de reclamaciones, datos propios) y la experiencia de grupo (datos del mercado, contratos similares, experiencia propia acumulada, datos colaterales).

Con base en lo dicho anteriormente, podemos decir que la finalidad de la teoría de la credibilidad es ajustar el valor de una prima con base en el historial/experiencia que tiene la aseguradora con cierto siniestro. Para lograr esto, nosotros en este caso vamos a trabajar ajustando la prima de riesgo, en general se puede hacer esto para que la prima de tarifa pueda ser calculada tomando como referencia a la de riesgo.

Tomaremos para esta sección S un riesgo arbitrario que busca absorber una aseguradora, correspondiente a un asegurado o grupo de asegurados con características homogéneas y válido por un periodo determinado. Denotaremos como $\{S_i\}_{i=1}^m$ los montos registrados de las reclamaciones efectuadas por el asegurado o portafolio de asegurados durante m periodos consecutivos.

Con base en la estadística clásica y Bayesiana, las metodologías que hay para la teoría de la credibilidad son dos principales ramas:

$$\text{Teoría de la credibilidad} \begin{cases} \text{Clásica} \begin{cases} \text{Completa} \\ \text{Parcial} \end{cases} \\ \text{Bayesiana} \end{cases}$$

Al igual que se hizo a lo largo de la historia, exploraremos la forma clásica, pues tiene sus fundamentos en resultados asintóticos que ya se han trabajado. Posteriormente veremos la perspectiva Bayesiana que es otra manera de atacar el problema pero con el apoyo de los fundamentos de la parte clásica.

Credibilidad Completa/Total

Para lograr el ajuste a la **prima de riesgo**, la cual estamos modelando como $\mathbb{E}[S]$, con base en el histórico, vamos a recurrir a su versión muestral.

Como ya lo hemos hecho antes, nos interesa conocer el comportamiento de

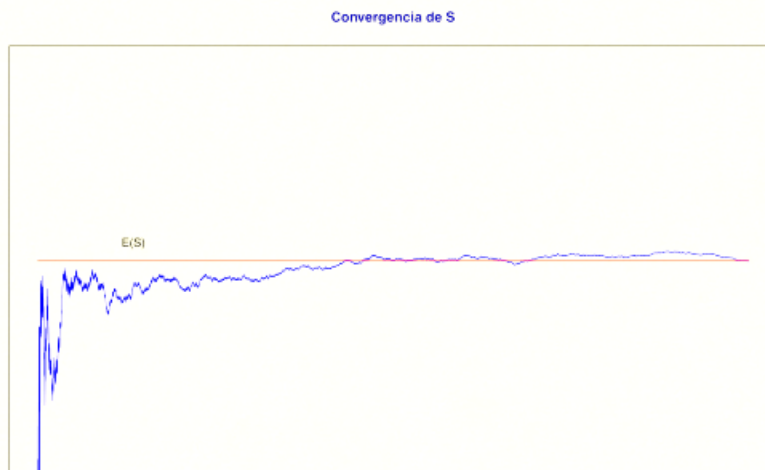
$$\bar{S} = \frac{1}{m} \sum_{i=1}^m S_i$$

Esto lo haremos así pues, si la **modelación del riesgo** que nosotros estamos proponiendo es correcta, entonces por las leyes de los grandes números

$$\begin{array}{ccc} \text{La prima de} & & \text{La prima de} \\ \text{riesgo que nos} & = \bar{S} = \frac{1}{m} \sum_{i=1}^m S_i & \xrightarrow[m \rightarrow \infty]{c.s.} \mathbb{E}[S] = \\ \text{dicen los datos} & & \text{riesgo según el} \\ & & \text{modelo propuesto} \end{array}$$

Por lo que si nuestro modelo es incorrecto, esta relación no se da.

Si estamos modelando de manera correcta, tiene sentido que, **mientras el portafolio se mantenga homogéneo** a medida que **m** crezca la información nueva sustente nuestra teoría. De lo contrario, aún consiguiendo **m lo suficientemente grande**, los datos no ajustarán el modelo.



Por lo que nos interesa encontrar el valor de **m** tal que según nuestro modelo, \bar{S} se encuentre “razonablemente” cercano a $\mathbb{E}[S]$. El que esto suceda es precisamente lo que le da credibilidad al modelo.

Definición Sean k , p dos números fijos, se dice que \bar{S} tiene **Credibilidad completa** (k,p) si

$$\mathbb{P} \left[|\bar{S} - \mathbb{E}[S]| \leq k\mathbb{E}[S] \right] \geq p$$

Nota: En general se asume que $\mathbb{E}[S] \neq 0$ por(espero) obvias razones. Así que para que lo dicho anteriormente tenga sentido, se toman valores para k cercanos a cero y P cercanos a uno. Así como se toma un nivel de significación $\alpha = 0.05$, lo más usual es tomar $k = 0.05$, $P = 0.90$.

Credibilidad Completa Bajo Hipótesis de Normalidad

Como usualmente es complicado obtener expresiones analíticas para las probabilidades de S , acudiremos a las aproximaciones.

Encontraremos la condición sobre el número de periodos de observación m , para obtener credibilidad completa usando la distribución normal.

Ejercicio: Arrastra el lápiz si no recuerdas que:

$$\mathbb{E}[\bar{S}] = \mathbb{E}[S] \quad \text{y} \quad \text{Var}(\bar{S}) = \frac{1}{m} \text{Var}(S)$$

Tenemos que

$$P \leq \mathbb{P} \left[|\bar{S} - \mathbb{E}[S]| \leq k\mathbb{E}[S] \right] = \mathbb{P} \left[\frac{|\bar{S} - \mathbb{E}[S]|}{\sqrt{\text{Var}(\bar{S})}} \leq \frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(\bar{S})}} \right]$$

Llamando

$$\gamma = \frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(\bar{S})}}$$

$$P \leq \mathbb{P} \left[-\gamma \leq \frac{\bar{S} - \mathbb{E}[\bar{S}]}{\sqrt{\text{Var}(\bar{S})}} \leq \gamma \right]$$

Usando el valor absoluto anterior y que $\mathbb{E}[\bar{S}] = \mathbb{E}[S]$

$$\approx \Phi(\gamma) - \Phi(-\gamma)$$

usando ley de los grandes números

$$\begin{aligned}
&= \underbrace{2\Phi(\gamma) - 1}_{\text{Por la simetría de la normal } \Phi(-x)=1-\Phi(x)} \\
&= 2\Phi\left(\frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) - 1 = 2\Phi\left(\frac{\sqrt{m}k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) - 1 \\
&\quad \underbrace{\text{Var}(\bar{S}) = \frac{1}{m}\text{Var}(S)} \\
&\therefore P \leq 2\Phi\left(\frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) - 1 = 2\Phi\left(\frac{\sqrt{m}k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) - 1 \\
&\iff \frac{P+1}{2} \leq \Phi\left(\frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) = \underbrace{\Phi\left(\frac{\sqrt{m}k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right)}_{\text{Las operaciones realizadas respetan la igualdad}}
\end{aligned}$$

Tomando $Z_x \doteq$ cuantil del $(\alpha - 100)\%$ de una $N(0, 1) = \Phi^{-1}(x)$

$$\iff Z_{(\frac{P+1}{2})} \leq \left(\frac{k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right) = \underbrace{\left(\frac{\sqrt{m}k\mathbb{E}[S]}{\sqrt{\text{Var}(S)}}\right)}$$

Aplicando Φ^{-1} de ambos lados respeta la desigualdad pues Φ es creciente $\implies \Phi^{-1}$ lo es

De esta última expresión obtenemos las dos siguientes:

- $\text{Var}(S) \left[\frac{Z_{(\frac{P+1}{2})}}{k\mathbb{E}[S]} \right]^2 \lesssim m \rightarrow$ Esta es una manera de obtener la cantidad mínima de periodos m a partir del modelo
- $\text{Var}(\bar{S}) \lesssim \left[\frac{k\mathbb{E}[S]}{Z_{(\frac{P+1}{2})}} \right]^2 \rightarrow$ si $m = 1$ vemos una cota teórica para la varianza del riesgo S .

Nota: Recordar que $\text{Var}(\bar{S}) = \frac{\text{Var}(S)}{m}$

Nota: Como $m \in \mathbb{N}$ si la cota inferior tiene decimales, la mínima m será el techo de la cota.

En concreto, ¿cómo se realiza usando una muestra? Lo primero que debemos realizar es, dado nuestro modelo, calcularemos:

$$m_{min} \doteq \left\lceil \frac{Var(S)}{\mathbb{E}^2[S]} \left\lceil \frac{Z_{(\frac{p+1}{2})}}{k} \right\rceil \right\rceil^2$$

Este cálculo se hará con el modelo teórico que nosotros vamos a probar

Nota: Como usualmente $k = 0.05$ y $p = 0.9 \Rightarrow \left\lceil \frac{Z_{(\frac{p+1}{2})}}{k} \right\rceil^2 \approx 1082.217$ salvo que se especifiquen otros valores para (k, p)

Donde $\lceil \cdot \rceil$ es la función "parte entera mayor o igual- o simplemente conocida como "techo". De tal manera que el modelo propuesto impone una cantidad de periodos necesaria para poder verificar credibilidad.

Si m es la cantidad de datos disponibles una condición para poder verificar credibilidad completa (k, p) (Bajo hipótesis de normalidad) es:

$$m_{min} \leq m$$

Nota: Observando la definición de credibilidad completa y de m_{min} se vislumbra que cuando $k \downarrow 0$ o bien $p \uparrow 1$ estamos siendo más estrictos con el modelo, y de hecho $m_{min} \uparrow \infty$ lo que significa que necesitaremos más periodos. Recíprocamente cuando $k \uparrow 1$ o bien $p \downarrow 0$ necesitamos menos periodos.

De tal manera que si tenemos m datos, usando el modelo teórico m_{min} nos dirá si tenemos la cantidad suficiente de datos para verificar si nuestro modelo tiene credibilidad completa (k, p) o no.

Ahora, con base en la teoría desarrollada por la SOA para el examen STAM, notemos que:

$$\frac{Var(S)}{\mathbb{E}^2[S]} \left\lceil \frac{Z_{(\frac{p+1}{2})}}{k} \right\rceil^2 \approx \frac{Var(S)}{\mathbb{E}[S]^2} \left\lceil \frac{Z_{(\frac{p+1}{2})}}{k} \right\rceil^2 = Var(S) \left\lceil \frac{Z_{(\frac{p+1}{2})}}{k \mathbb{E}[S]} \right\rceil^2 \leq m$$

De aquí usaremos la muestra para verificar si se **puede** cumplir la **credibilidad completa** (k,p). Esto es:

Si $m_{min} \leq m$ entonces nuestro modelo **puede** satisfacer credibilidad completa (k,p)

(bajo hipótesis de normalidad) si:

$$\underbrace{\frac{Var(S)}{E[S]}}_{\text{Modelo teórico}} \underbrace{\left[\frac{Z_{(\frac{p+1}{2})}}{k} \right]^2}_{\text{C.C. (k,p) normalidad}} \leq \underbrace{m\bar{S} = \sum_{i=1}^m S_i}_{\text{experiencia de los siniestros}}$$

Esto nos da otra condición que deberían satisfacer los datos si deseamos verificar la credibilidad completa (k,p)

De tal manera que para poder si quiera preguntarnos si nuestros datos **pueden** satisfacer credibilidad completa (k,p) deberían satisfacerse las siguientes **condiciones**:

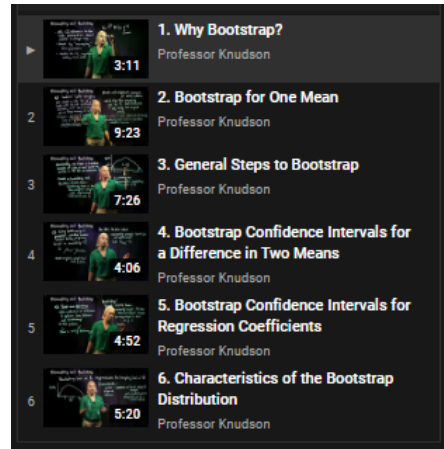
1. $m_{min} \leq m \rightarrow$ cantidad **mínima** de periodos requeridos **SOA**
2. $\frac{Var(S)}{E[S]} \left[\frac{Z_{(\frac{p+1}{2})}}{k} \right]^2 \leq \sum_{i=1}^m S_i \rightarrow$ Monto total mínimo de reclamaciones **experimentadas** del riesgo **SOA**
3. $Var(\bar{S}) \lesssim \left[\frac{kE[S]}{Z_{(\frac{p+1}{2})}} \right]^2 \rightarrow$ Cota superior de la volatilidad de la media de las reclamaciones experimentadas

Si no se satisfacen 2 y 3 son indicios de replantear el modelo. Más adelante hablamos de qué hacer si falla. Ahora estas **condiciones** las ponemos así porque estamos esperando que $\bar{S} \rightarrow E[S]$ es decir que nuestros datos en efecto sean descritos por la teoría. Pero al final \bar{S} se aproximará a su propia media. De tal manera que **aún cumpliéndose las condiciones anteriores no significa que nuestro modelo explique los datos**. Más bien de cumplirse aún falta verificar que en efecto nuestros datos explican el modelo mediante la **definición** de **credibilidad completa** (k,p). Podemos pensar que las condiciones anteriores son una **puerta** que necesito **abrir** para poder ver si mi modelo explica los datos.

En resumen una vez satisfechas las condiciones hay que ver que en efecto $|\bar{S} - E[S]| \leq kE[S]$ con probabilidad $\geq p$.

Una posible verificación decisiva puede ser usando **bootstrap** para mostrar que se cumpla la definición de **credibilidad completa** (k,p)

Lo ideal es primero verificar las **condiciones** y posteriormente esto.



Equivalentemente decimos que \bar{S} cumple **credibilidad completa** (k,p) bajo hipótesis de normalidad con nuestro modelo teórico.

La idea es sencilla. Pensemos entonces que tenemos una muestra aleatoria $\underline{S}_m = \{S_1, S_2, \dots, S_m\}$. Ahora, si esta muestra cumpliera la **definición** de **credibilidad completa** con (K, P) entonces:

$$\mathbb{P} \left[|\bar{S} - \mathbb{E}[S]| \leq k \quad \mathbb{E}[S] \right] \geq p$$

lo que haremos es verificar que esto sucede en n ensayos realizados a partir de la primera muestra.

Lo que haremos será muestrear con reemplazo de \underline{S} y de cada una de estas muestras veremos si $|\bar{S} - \mathbb{E}[S]| \leq k \quad \mathbb{E}[S]$ o no. Es decir, tendremos $\{\underline{S}_m^{(i)}\}_{i=1}^n$ muestras con reemplazo de \underline{S}_m y luego:

$$\begin{aligned} \underline{S}^{(1)} &\Rightarrow \text{Calculamos su promedio y lo llamamos } \bar{S}_m^{(1)} \\ \underline{S}^{(2)} &\Rightarrow \text{Calculamos su promedio y lo llamamos } \bar{S}_m^{(2)} \\ &\vdots \\ \underline{S}^{(n)} &\Rightarrow \text{Calculamos su promedio y lo llamamos } \bar{S}_m^{(n)} \end{aligned}$$

Teniendo así, una muestra de tamaño n de $\{\underline{S}_m^{(i)}\}_{i=1}^n$ para cada una de estas estadísticas podemos calcular: $x_i = \mathbb{I}\{|\bar{S}_m^{(i)} - \mathbb{E}[S]| \leq k \quad \mathbb{E}[S]\}$ y entonces si $\bar{x} = \frac{\text{éxitos}}{\text{ensayos}} \geq p$, tendremos que el modelo satisface **c.c** (k,p) por **bootstrap**.

- Cuando no hay credibilidad completa hay que replantear el modelo
- Si $m_{\min} > m(1)$ entonces podemos ver la siguiente alternativa

Credibilidad Parcial

Cuando no se obtiene credibilidad completa, por ejemplo cuando no tenemos la cantidad de periodos mínima para obtener este criterio obtenemos **credibilidad parcial**. Para esto, tomamos $\alpha \in (0, 1]$ y proponemos la combinación convexa del estimador de $\mathbb{E}[S]$:

$$\alpha \bar{S} + (1 - \alpha) \mathbb{E}[S] \quad \text{con Factor de credibilidad} \quad \alpha \in (0, 1)$$

Mediante esta expresión se le otorga **credibilidad completa** a una parte de la media muestral \bar{S} y el complemento a $\mathbb{E}[S]$. Es decir:

Definición:

Sean $k, p, \alpha \in (0, 1)$ tres números fijos. Diremos que \bar{S} tiene **credibilidad parcial** (k, p, α) si:

$$\mathbb{P} \left[|\alpha \bar{S} + (1 - \alpha) \mathbb{E}[S] - \mathbb{E}[S]| \leq k \mathbb{E}[S] \right] \geq p$$

Sin embargo, esta definición se vuelve irrelevante cuando se vislumbra que:

$$\begin{aligned} p &\leq \mathbb{P} \left[|\alpha \bar{S} + (1 - \alpha) \mathbb{E}[S] - \mathbb{E}[S]| \leq k \mathbb{E}[S] \right] \quad \text{¿Qué pasa si } \alpha \equiv 0? \\ &= \mathbb{P} \left[|\alpha \bar{S} - \mathbb{E}[S]| \leq k \mathbb{E}[S] \right] \\ &= \mathbb{P} \left[|\bar{S} - \mathbb{E}[S]| \leq \frac{k}{\alpha} \mathbb{E}[S] \right] \end{aligned}$$

Por lo tanto

$$\begin{aligned} p &\leq \mathbb{P} \left[|\alpha \bar{S} + (1 - \alpha) \mathbb{E}[S] - \mathbb{E}[S]| \leq k \mathbb{E}[S] \right] \\ &= \mathbb{P} \left[|\bar{S} - \mathbb{E}[S]| \leq \frac{k}{\alpha} \mathbb{E}[S] \right], \text{ es decir:} \end{aligned}$$

$$\text{credibilidad parcial } (k, p, \alpha) \Leftrightarrow \text{credibilidad completa } \left(\frac{k}{\alpha}, p \right)$$

Como $\alpha \in (0, 1] \Rightarrow \frac{k}{\alpha} \geq k$ y esto permite un rango de error mayor con credibilidad parcial que con credibilidad completa con la misma (k, p) .

Credibilidad Parcial bajo Hipótesis de Normalidad

Como ya vimos, credibilidad parcial no es más que tomar credibilidad total pero modificando una de sus componentes. Por lo que bajo el supuesto de normalidad asintótica, para credibilidad parcial tenemos que:

$$\text{Var}(S) = \left[\frac{z\left(\frac{p+1}{2}\right)}{\frac{k}{\alpha} \mathbb{E}[S]} \right]^2 \lesssim m \quad ; \quad \text{Var}(\bar{S}) \lesssim \left[\frac{\frac{k}{\alpha} \mathbb{E}[S]}{z\left(\frac{p+1}{2}\right)} \right]^2$$

Nota: Observemos que el factor de credibilidad justo hace menos estricta la condición de credibilidad completa

Las interpretaciones son idénticas al caso que ya vimos con credibilidad total. El punto interesante aquí es que si el tamaño de muestra (m) NO fuese suficientemente grande pero fijo y conocido, entonces podemos calcular el factor de credibilidad de la información que tenemos y el modelo propuesto como:

$$\alpha = \left[\frac{k \cdot \mathbb{E}[S] \sqrt{m}}{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(S)}} \right] = \left[\frac{k \cdot \mathbb{E}[S]}{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(\bar{S})}} \right]$$

Nota: Si m es lo suficientemente grande $\Rightarrow \alpha$ puede ser mayor a uno. Esto pues precisamente va alcanzando **credibilidad completa**. Con la finalidad de dar una versión generalizada del **factor de credibilidad** tenemos que:

$$\alpha = \min \left\{ \underbrace{\frac{k \cdot \mathbb{E}[S] \sqrt{m}}{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(S)}}}_{\text{Usaremos este}} = \frac{k \cdot \mathbb{E}[S]}{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(\bar{S})}}, 1 \right\}$$

De aquí notamos que:

$$\frac{k \cdot \mathbb{E}[S] \sqrt{m}}{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(S)}} = \frac{\sqrt{m}}{\frac{z_{\left(\frac{p+1}{2}\right)} \sqrt{\text{Var}(S)}}{k \cdot \mathbb{E}[S]}} = \sqrt{\frac{m}{\frac{\text{Var}(S)}{\mathbb{E}^2[S]} \left[\frac{z_{\left(\frac{p+1}{2}\right)}}{k} \right]^2}}$$

Con base en la teoría desarrollada por la SoA para el examen STAM, tendremos entonces lo siguiente:

Dependiendo de la cantidad de **información** que se tenga, el número anterior se calcula de la siguiente manera:

$$\sqrt{\frac{m}{\frac{Var(S)}{E^2[S]} \left[\frac{z\left(\frac{p+1}{2}\right)}{k} \right]^2}} \approx \sqrt{\frac{m \cdot \bar{S}}{\frac{Var(S)}{E[S]} \left[\frac{z\left(\frac{p+1}{2}\right)}{k} \right]^2}}$$

Nota: En esta expresión el numerador se calcula con la **muestra** y el denominador con el **modelo teórico**

Nota: Si no se tiene muestra, de hecho no tiene sentido invocar credibilidad. Sin embargo en ejercicios **100 % teóricos** donde se solicita el cálculo del **factor de credibilidad** **pero no se da una muestra**, se debe tomar $\alpha = 1$

Ahora, existen diversas propuestas de diferentes autores de cómo estimar/obtener el **factor de credibilidad**; Nosotros nos centraremos en la metodología propuesta por la **SoA**.

Usando la teoría desarrollada por la SoA tenemos que:

Para invocar credibilidad parcial (k, p, α) (Bajo el supuesto de normalidad), el cálculo del factor de credibilidad se hace de la siguiente manera:

$$\alpha = \min \left\{ \frac{\text{“Información disponible”}}{\text{“Información necesaria para credibilidad completa”}} \right\}$$

Donde:

- “Información disponible” Es el numerador de las expresiones anteriores y se obtiene dependiendo de la muestra dada (o bien si no se cuenta con muestra, se hace teórica).
- “Información necesaria para credibilidad completa” Es el denominador de las expresiones anteriores y se obtiene con el modelo propuesto.

Las cosas pueden ponerse más interesantes dependiendo si S es un **modelo colectivo**, algunos de los cálculos con los que se verifican o se obtienen ciertos factores tanto para credibilidad completa o parcial se modifican dependiendo de la “Información disponible”.

Esto se hace con la finalidad de explotar los datos y la experiencia obtenida de la muestra de la forma más adecuada posible. Aunque de momento, vamos a tratar la credibilidad como hasta este punto.

Ajuste de primas con credibilidad clásica

Finalmente llegamos al objetivo principal de esta sección; recordemos que la finalidad de la teoría de la credibilidad es dar verosimilitud al modelo que estamos proponiendo con base en el histórico que tenemos.

En el caso en el que nuestro modelo cumpla **credibilidad completa**, como su nombre lo indica, nuestro modelo funciona correctamente a los niveles (k, p) , por lo que **no es necesario hacer un ajuste al modelo de prima de riesgo**. Sin embargo, se acostumbra dar más peso a la información recolectada y con base en esto tener la nueva prima. Esto significa:

$$\text{Credibilidad completa } (k, p) \Rightarrow \text{Prima de riesgo ajustada} \doteq \underbrace{\bar{S}}_{(k,p)} \approx E[S]$$

Nota: Recuerda que buscamos $k \approx 0$ y $p \approx 1$. Sin ser exactamente iguales.

Por otro lado, en el caso en que tengamos credibilidad parcial (k, p, α) , entonces realizamos un ajuste a la prima y de hecho podemos considerar modificar el modelo propuesto, si $\alpha \approx 0$ o si el número de periodos con los que contamos es aún muy pequeño con respecto a la cota inferior de m . El ajuste de la prima en este caso será:

$$\text{Credibilidad Parcial } (k, p, \alpha) \Rightarrow \text{Prima de riesgo ajustada } \alpha \bar{S} + (1 - \alpha) \mathbb{E}[S] \underbrace{\approx}_{(k, p, \alpha)} \mathbb{E}[S]$$

Nota: No queremos que $\alpha \approx 0$ pues hacemos que el modelo ajuste a los datos no a sí mismo

Enfoque bayesiano

En estadística tradicional, enfoque clásico, uno de los problemas inferenciales más importantes es la estimación del parámetro θ , de una distribución de probabilidad $f(x; \theta)$.

Para realizar dicha estimación seleccionamos una muestra aleatoria de esta distribución y se tienen distintos métodos para estimar θ , considerando siempre que este parámetro tiene un valor desconocido y fijo. En el enfoque Bayesiano, θ se trata como una variable aleatoria para la que se supone una distribución de probabilidad $p(\theta)$, llama distribución inicial o distribución a priori. Esta distribución refleja la información subjetiva o cuantitativa que el observador pueda tener sobre este parámetro θ , antes de observar la muestra.

Actualización de la información a través del teorema de Bayes

La forma en la que se actualiza la información inicial sobre nuestro parámetros de interés θ , mediante la información contenida en la muestra $p(x|\theta) = L(\theta, \underline{X})$ (la verosimilitud), es a través del teorema de Bayes.

Teorema de Bayes

Teorema de Bayes

Dados dos eventos A y B tales que $\mathbb{P}(B) > 0$, la probabilidad condicional de A dado B, $\mathbb{P}(A|B)$, se define como:

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}\end{aligned}$$

Teorema de Bayes

Si $\{A_i : i = 1, 2, \dots, M\}$ en un conjunto exhaustivo de eventos mutuamente excluyentes, entonces:

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i)\mathbb{P}(A_i)}{\sum_{j=1}^M \mathbb{P}(B|A_j)\mathbb{P}(A_j)}$$

La forma esquemática de actualizar esta información a través del teorema de Bayes es la siguiente:

1. θ debe tener una distribución de probabilidad $p(\theta)$, que refleje nuestro conocimiento inicial acerca de su valor.
2. La información sobre θ que contiene la muestra aleatoria seleccionada, está resumida en la verosimilitud $p(x|\theta) = L(\theta; \underline{X})$.

Por lo tanto, nuestro conocimiento acerca del valor θ queda descrito a través de su distribución final.

$$p(\theta|\underline{X})$$

El teorema de Bayes nos dice cómo encontrarla:

$$p(\theta|\underline{X}) = \frac{p(x|\theta)p(\theta)}{\int p(x|\theta)p(\theta)d\theta}$$

Este proceso se conoce como el “proceso de actualización de la información sobre θ ”, y es la manera de combinar las dos fuentes de información que tenemos: La inicial dada a través de $p(\theta)$, y la de la muestra, dada por medio de la verosimilitud: $p(x|\theta)$, para obtener la distribución final: $p(\theta; \underline{X})$, que contiene la suma de estas dos fuentes de información.

Obsérvese que el denominador, $p(x) = \int p(x|\theta)p(\theta)d\theta$, no depende de θ , por lo que es común escribir esta distribución final como:

$$p(\theta|\underline{X}) \propto p(x|\theta)p(\theta)$$

En la práctica, el cálculo de la distribución final puede ser un asunto complicado, especialmente si la dimensión del parámetro es grande.

Sin embargo, para ciertas combinaciones de distribuciones iniciales y verosimilitudes es posible simplificar el análisis.

En otros casos se requieren aproximaciones analíticas y/o técnicas computacionales relativamente sofisticadas.

Como hemos visto, la prima de riesgo puede cambiar según el comportamiento de la muestra y que a medida que tuviéramos más información, se nos indica que el modelo va en dirección correcta.

Todo esto es algo que nos hace pensar en estadística bayesiana, por lo que veremos un pequeño repaso de cómo hacer estimaciones con este tipo de estadística.

Comencemos dejando claro un poco la notación en términos probabilísticos. Supongamos que $X_1, X_2, \dots, X_n \sim \text{Algo}(\Theta)$ (Θ podría ser un vector de parámetros). $\Theta \sim \text{Algo}_2(\Delta)$, donde Δ asumimos totalmente conocido (propuesto).

Verosimilitud

$$p(\underline{x}|\Theta) \doteq f_{x_1, x_2, \dots, x_n}(x_1, x_2, \dots, x_n)$$

Es la **función de densidad conjunta** de la muestra, cuya distribución depende del parámetro de interés Θ

Distribución inicial (a priori)

$$p(\theta) \doteq f_{\Theta}(\theta)$$

Es la **función de densidad** del parámetro de interés, recordando que este se piensa como v.a

Distribución final (a posteriori)

$$p(\theta|\underline{x}) \doteq f_{\Theta|\underline{X}=\underline{x}}(\theta)$$

Es la **función de densidad** que depende únicamente de θ . **Tomando en cuenta valores ya observados y fijos de la muestra**

El siguiente resultado **NO** viene directamente del teorema de Bayes, ya que, aunque es muy común usar la notación “ $p(\theta)$ ”, esto no necesariamente es probabilidad, pues de hecho es una **función de densidad**. Del Teorema de Bayes se puede encontrar que:

$$p(\theta|\underline{x}) = \frac{p(\underline{x}|\theta)p(\theta)}{\int_{\Omega_{\Theta}} p(\underline{x}|\theta)p(\theta)d\theta} \underbrace{\propto}_{\text{proporcional}} p(\underline{x}|\theta)p(\theta)$$

Donde \propto Kernel de la v.a $\Theta|\underline{X}$

La razón de porque $p(\theta|\underline{x}) \propto p(\underline{x}|\theta)p(\theta)$ es porque $\int_{\Omega_\theta} p(\underline{x}|\theta)p(\theta)d\theta$ no depende de θ (Se integran) y quedan simplemente constantes para θ , como lo son \underline{X} y Δ .

Nota: En general se asume una muestra aleatoria en \underline{X} entonces la independencia queda implícita. Así: $p(\underline{x}|\theta) = \prod_{i=1}^n p(x_i|\theta)$.

Definición

Dada una **distribución final** y una **muestra fija**, decimos que el **estimador bayesiano** para Θ es:

$$\hat{\Theta} \doteq \mathbb{E}[\Theta|\underline{X} = \underline{x}] = \int_{\Omega_\Theta} \theta p(\theta|\underline{x}) d\theta$$

Nota: “Muestra fija” es la clave, ya que de otro modo $\hat{\Theta}$ sería v.a.

Ejemplo

Suponga $\{X_i\}_{i=1}^n$ una muestra aleatoria con distribución Bernoulli(Θ) y consideremos $\Theta \sim \text{Beta}(\Delta = (a, b))$ entonces:

$$p(\theta) = \frac{1}{\beta(a, b)} \theta^{a-1} (1-\theta)^{b-1} \mathbb{I}_{(0,1)}(\theta) \quad ; \quad \mathbb{E}[\Theta] = \frac{a}{b+a} \quad \beta(a, b) \doteq \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$p(x_i|\theta) = \theta^{x_i} (1-\theta)^{1-x_i} \mathbb{I}_{\{0,1\}}(x_i) \quad \text{Noten que está bien definido pues } \theta \in (0, 1)$$

Kernel

$$\begin{aligned}
p(\underline{x}|\theta)p(\theta) &= \prod_{i=1}^n p(x_i|\theta)p(\theta) \\
&= \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} \frac{1}{\beta(a,b)} \theta^{a-1} (1-\theta)^{b-1} \\
&= \frac{1}{\beta(a,b)} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1} \\
&\propto \underbrace{\theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1}}_{\text{Un kernel conocido}}
\end{aligned}$$

Por lo tanto

Como $p(\theta|\underline{x}) \propto p(\underline{x}|\theta)p(\theta)$ entonces $\Theta|\underline{X} = \underline{x} \sim \text{Beta}(n\bar{x} + a, n(1 - \bar{x}) + b)$

Sin embargo, es común que la gente sea escéptica a “mandar a la ... constante de integración” las cosas. Entonces, sin tomar proporciones, tenemos:

Kernel:

$$\blacksquare P(\underline{x}|\theta)P(\theta) = \frac{1}{\beta(a,b)} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1}$$

Constante de integración:

$$\begin{aligned}
\blacksquare \int_{\Omega_\Theta} P(\underline{x}|\theta)P(\theta)d\theta &= \int_0^1 \frac{1}{\beta(a,b)} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1} d\theta \\
&= \frac{1}{\beta(a,b)} \beta(a+n\bar{x}, n(1-\bar{x})+b) \underbrace{\int_0^1 \frac{1}{\beta(a+n\bar{x}, n(1-\bar{x})+b)} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1} d\theta}_{\text{Una densidad Beta sobre todo su soporte}} \\
&= \frac{1}{\beta(a,b)} \beta(a+n\bar{x}, n(1-\bar{x})+b) \\
\therefore P(\theta|\underline{x}) &= \frac{P(\underline{x}|\theta)P(\theta)}{\int_{\Omega_\Theta} P(\underline{x}|\theta)P(\theta)d\theta} = \frac{\cancel{\frac{1}{\beta(a,b)}} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1}}{\cancel{\frac{1}{\beta(a,b)}} \beta(a+n\bar{x}, n(1-\bar{x})+b)} \\
&= \frac{1}{\beta(a+n\bar{x}, n(1-\bar{x})+b)} \theta^{n\bar{x}+a-1} (1-\theta)^{n(1-\bar{x})+b-1}
\end{aligned}$$

$\therefore \Theta | \underline{X} = \underline{x} \sim \text{Beta}(n\bar{x} + a, n(1 - \bar{x}) + b)$, también haciendo más cuentas.

$$\Rightarrow \hat{\Theta} = E[\Theta | \underline{X} = \underline{x}] = \frac{n\bar{x} + a}{(n\bar{x} + a) + (n(1 - \bar{x}) + b)} = \frac{n\bar{x} + a}{n + a + b} \quad \therefore \hat{\Theta} = \frac{n\bar{x} + a}{n + a + b}$$

En este caso, la distribución final tuvo una distribución conocida, sin embargo, esto no sucede necesariamente. Afortunadamente, existen distribuciones que si las utilizamos para la **verosimilitud** y la **inicial**, tendremos una **final conocida**. Esto se conoce como **familias conjugadas**.

Familias conjugadas

Comentamos en uno de los puntos anteriores, que existen ciertas combinaciones de distribuciones y verosimilitudes, que simplifican el análisis Bayesiano, esencialmente, porque el modelo de la distribución final de θ , pertenece a la misma familia que el de la inicial.

Definición. Sea $\mathcal{P} = \{p(x|\theta) : \theta \in \Theta\}$ una familia paramétrica. Una clase (o colección) de distribuciones de probabilidad \mathcal{F} es una **familia conjugada** para \mathcal{P} si para toda $p(x|\theta) \in \mathcal{P}$ y $p(\theta) \in \mathcal{F}$ se tiene que $p(\theta|x) \in \mathcal{F}$

Algunos modelos paramétricos univariados con sus respectivas familias conjugadas:

Familia paramétrica $P(\underline{x} \theta)$	Familia Conjugada $P(\theta)$	
<i>Bernoulli</i> ($x \theta$)	<i>Beta</i> ($\theta \alpha, \beta$)	
<i>Poisson</i> ($x \lambda$)	<i>Gamma</i> ($\lambda \alpha, \beta$)	→ Modelo Poisson- Gamma
<i>Geométrica</i> ($x \theta$)	<i>Beta</i> ($\theta \alpha, \beta$)	
<i>Exponencial</i> ($x \lambda$)	<i>Gamma</i> ($\lambda \alpha, \beta$)	
<i>Uniforme</i> ($x 0, \theta$)	<i>Pareto</i> ($\lambda \alpha, \beta$)	
<i>Normal</i> ($x \mu$)	<i>Normal</i> ($\mu \mu_0, \lambda_0$)	→ Modelo Normal-Normal
<i>Normal</i> ($x \lambda$)	<i>Gamma</i> ($\lambda \alpha, \beta$)	
<i>Normal</i> ($x \mu, \lambda$)	<i>Normal - Gamma</i> ($\mu, \lambda \mu_0, \lambda_0, \alpha, \beta$)	

Ejemplo. Consideremos la familia paramétrica $\mathcal{P} = \{\text{Poisson}(x|\lambda) : \lambda \in \mathbb{R}^+\}$. Si utilizamos como distribución inicial $p(\lambda) \in \mathcal{F} = \{\text{Gamma}(\lambda|\alpha, \beta) : \alpha, \beta \in \mathbb{R}^+\}$. Entonces, si se tiene una muestra aleatoria $\underline{x} = (x_1, \dots, x_n)$, la distribución final es

$$p(\lambda|\underline{X}) = \text{Gamma}(\lambda|\alpha + r, n + \beta) \quad \text{con } r = \sum_{i=1}^n x_i$$

Dem. Solo demostraremos que el *kernel* de la distribución final, pertenece a la distribución $\text{Gamma}(\lambda|\alpha + r, n + \beta)$.

Sabemos, por Bayes, que

$$p(\lambda|X) = \frac{p(x|\lambda)p(\lambda)}{\int p(x|\lambda)p(\lambda)d\lambda}$$

con

$$p(x|\lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \propto \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \text{ y}$$

$$p(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}, \text{ por lo que}$$

$$p(\lambda|X) \propto \lambda^{\sum_{i=1}^n x_i} e^{-n\lambda} \lambda^{\alpha-1} e^{-\beta\lambda} = \lambda^{\alpha + \sum_{i=1}^n x_i - 1} e^{-\lambda(\beta+n)} \propto \text{Gamma}\left(\alpha + \sum_{i=1}^n x_i, \beta + n\right)$$

Cálculo Bayesiano de primas de seguros

El uso de las distribuciones iniciales que tienen un carácter evidentemente subjetivo, resulta de utilidad en el mercado de seguros, sobre todo si se tiene en cuenta que cuando se quiere tarifar un riesgo nuevo no se dispone de información para ello.

La visión Bayesiana se incorporó rápidamente a la disciplina actuarial, demostrando que algunas primas que se obtienen a través de la metodología Bayesiana pueden escribirse como fórmulas de credibilidad.

En estos términos actuariales, la cuestión básica de credibilidad es determinar una *prima* establecida como una combinación lineal convexa entre la experiencia particular de un asegurado y la experiencia del colectivo, esto es, toda la cartera. Es decir

$$\mathbb{P}_j = Z\hat{\mathbb{P}} + (1 - Z)\mathbb{P}_0$$

Con

- \mathbb{P}_j prima a aplicar a los asegurados por el riesgo j . (final-posterior)
- \mathbb{P}_0 prima a aplicar a un colectivo al que pertenece el asegurado j . (inicial-propuesta)
- $\hat{\mathbb{P}}$ Prima calculada con base en la experiencia del asegurado j . (muestral)
- Z Factor de credibilidad. Que debe verificar las condiciones: $\lim_{m \rightarrow \infty} Z = 1$, con m el número de sujetos expuestos al riesgo j o el periodo de observación de la póliza j .

Entonces, si $Z = 1$ la experiencia del asegurado recibe credibilidad total o del 100 %, mientras que si $Z = 0$, $\mathbb{P}_j = \mathbb{P}_0$ y la prima del asegurado j coincide con la del colectivo a la que pertenece dicha póliza, o la experiencia del colectivo recibe credibilidad total o del 100 %

Entonces, desde el punto de vista Bayesiano, esta fórmula de credibilidad puede interpretarse como: Podemos considerar \mathbb{P}_0 como la información inicial o *a priori*. A \mathbb{P}_0 como la nueva información que se obtiene mediante la observación de la siniestralidad del riesgo j (los datos recabados; la información recabada) y \mathbb{P}_j la actualización del cálculo de la póliza (prima a posteriori), resultado de combinar la información inicial con la información recabada. Por lo tanto

$$\text{Prima(a posteriori)} = (1 - Z) * \text{Prima a priori} + Z * \text{Experiencia dada por los datos}$$

De esta manera, la teoría de la credibilidad Bayesiana, sigue un esquema donde la información a priori sobre el cálculo de las primas, se actualiza con la información dada por la observación del siniestro (muestra), dando como resultado la actualización de la prima, mediante el cálculo de la prima a posteriori.

¿Cómo modelar un riesgo para obtener una prima de credibilidad Bayesiana?

Ahora aplicaremos estas ideas al problema del cálculo de primas tomando en cuenta la experiencia de un riesgo. Suponga que las variables S_1, \dots, S_m representan el historial de reclamaciones en m años o periodos que se han registrado de un riesgo dado. Suponga además que estas variables son independientes y todas ellas tienen una distribución común dependiente de un parámetro desconocido θ , y esta distribución es tal que $E(S) = \theta$. Bajo el enfoque Bayesiano se considera que el parámetro θ es una variable aleatoria para la cual se asume una distribución de probabilidad a priori. La esperanza a posteriori de θ , es decir, $E(\theta|S_1, \dots, S_m)$, representa una estimación para $E(S) = \theta$ tomando en cuenta el historial S_1, \dots, S_m . A esta esperanza posteriori se le llama prima de credibilidad Bayesiana. En los casos que analizaremos esta prima tiene la forma de la credibilidad parcial mencionada antes. Los casos que consideraremos para la distribución de S son: la distribución Poisson con media λ , y la distribución normal con media θ .

■ Modelo Poisson-Gamma

Este modelo adquiere su nombre a partir de las siguientes hipótesis: se postula que cada una de las variables aleatorias independientes S_1, \dots, S_m tiene distribución Poisson con parámetro λ , el cual se considera aleatorio con una distribución a priori gamma(γ, α), con γ y α parámetros conocidos. Observe que en este modelo se considera que los montos de las reclamaciones

toman valores enteros. La función de densidad a posteriori de λ es, para $x > 0$,

$$\begin{aligned}
 g(\lambda|S_1, \dots, S_m) &= \frac{f(S_1, \dots, S_m|\lambda)h(\lambda)}{\int_0^\infty f(S_1, \dots, S_m|\lambda)h(\lambda)d\lambda} \\
 &= \frac{\prod_{j=1}^m \left(\frac{\lambda^{S_j}}{S_j!} e^{-\lambda} \right) \frac{\alpha^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\alpha\lambda}}{\int_0^\infty \prod_{j=1}^m \left(\frac{\lambda^{S_j}}{S_j!} \right) \frac{\alpha^\gamma}{\Gamma(\gamma)} \lambda^{\gamma-1} e^{-\alpha\lambda} d\lambda} \\
 &= \frac{\lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda}}{\int_0^\infty \lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda} d\lambda} \\
 &= \frac{(m+\alpha)^{m\bar{S}+\gamma}}{\Gamma(m\bar{S}+\gamma)} \lambda^{m\bar{S}+\gamma-1} e^{-(m+\alpha)\lambda}.
 \end{aligned}$$

Es decir, la densidad a posteriori es gamma($m\bar{S} + \gamma, m + \alpha$). Por lo tanto, la prima por credibilidad, esperanza de esta densidad, es

$$\begin{aligned}
 \text{prima} &= E(\lambda|S_1, \dots, S_m) \\
 &= \frac{m\bar{S} + \gamma}{m + \alpha} \\
 &= \frac{m}{m + \alpha} \bar{S} + \frac{\alpha}{m + \alpha} \frac{\gamma}{\alpha} \\
 &= z\bar{S} + (1 - z) \frac{\gamma}{\alpha},
 \end{aligned}$$

en donde $z = \frac{m}{(m+\alpha)}$ es llamado factor de credibilidad. Esta cantidad crece monótonamente a uno cuando m crece a infinito, dando cada vez más credibilidad a la media muestral \bar{S} y favoreciendo cada vez menos a la media teórica $\frac{\gamma}{\alpha}$. Observe además que cuando m crece a infinito, la media de la distribución a posteriori converge a la media muestral límite dado por el historial de reclamaciones, y que la varianza de λ dada por $Var(\lambda|\underline{S}) = \frac{(m\bar{S}+\gamma)}{(m+\alpha)^2}$ converge a cero, lo cual indica que la distribución posteriori se concentra cada vez más alrededor de su media.

■ Modelo Normal-Normal

En este modelo se postula que cada una de las reclamaciones S_1, \dots, S_m tiene distribución $N(\theta, \sigma^2)$, en donde el parámetro σ^2 es conocido y la media θ es una variable aleatoria con distribución $N(\mu, \eta^2)$, con μ y η^2 conocidos. La primera hipótesis puede ser justificada en el caso cuando los montos anuales se componen de un gran número de reclamaciones

individuales, para ello no es necesario suponer que las reclamaciones individuales tienen la misma distribución. La segunda hipótesis podría ser razonable si es que los parámetros μ y η^2 son tales que la probabilidad asignada a la parte negativa del eje es muy pequeña. La función de densidad a posteriori de θ es

$$\begin{aligned} g(\theta|S_1, \dots, S_m) &= \frac{f(S_1, \dots, S_m|\theta)h(\theta)}{\int_{-\infty}^{\infty} f(S_1, \dots, S_m|\theta)h(\theta)d\theta} \\ &= \frac{\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\sum_{j=1}^m (S_j-\theta)^2/2\sigma^2} \frac{1}{\sqrt{2\pi\eta^2}}e^{-(\theta-\mu)^2/2\eta^2}}{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\sum_{j=1}^m (S_j-\theta)^2/2\sigma^2} \frac{1}{\sqrt{2\pi\eta^2}}e^{-(\theta-\mu)^2/2\eta^2} d\theta} \end{aligned}$$

Nos concentramos en analizar únicamente el exponente y tenemos que

$$-\sum_{j=1}^m \frac{(S_j - \theta)^2}{2\sigma^2} - \frac{(\theta - \mu)^2}{2\eta^2} = -\theta^2 \left(\frac{m}{2\sigma^2} + \frac{1}{2\eta^2} \right) + 2\theta \left(\frac{m\bar{S}}{2\sigma^2} + \frac{\mu}{2\eta^2} \right) - \left(\sum_{j=1}^m \frac{S_j^2}{2\sigma^2} \right) - \frac{\mu^2}{2\eta^2}.$$

Completando el cuadrado en θ , este exponente se puede escribir como sigue

$$-\frac{\left[\theta - \left(\frac{m\bar{S}}{\sigma^2} + \frac{\mu}{\eta^2} \right) / \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right) \right]^2}{2 \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right)^{-1}} + \frac{\left(\frac{m\bar{S}}{\sigma^2} + \frac{\mu}{\eta^2} \right)^2}{2 \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right)} - \left(\sum_{j=1}^m \frac{S_j^2}{2\sigma^2} \right) - \frac{\mu^2}{2\eta^2}.$$

Este exponente aparece tanto en el numerador como en el denominador, y como los últimos dos sumandos no dependen de θ éstos desaparecen completamente al hacer el cociente. Lo que resulta es nuevamente la distribución normal

$$g(\theta|S_1, \dots, S_m) = \frac{1}{\sqrt{2\pi \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right)^{-1}}} \exp \left\{ -\frac{\left[\theta - \left(\frac{m\bar{S}}{\sigma^2} + \frac{\mu}{\eta^2} \right) / \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right) \right]^2}{2 \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right)^{-1}} \right\}.$$

La media de esta distribución es entonces la prima por credibilidad, es decir,

$$\begin{aligned} \text{prima} &= E(\theta|S_1, \dots, S_m) \\ &= \left(\frac{m\bar{S}}{\sigma^2} + \frac{\mu}{\eta^2} \right) / \left(\frac{m}{\sigma^2} + \frac{1}{\eta^2} \right) \\ &= \frac{m\eta^2}{m\eta^2 + \sigma^2} \bar{S} + \frac{\sigma^2}{m\eta^2 + \sigma^2} \mu \\ &= z\bar{S} + (1-z)\mu, \end{aligned}$$

en donde $z = \frac{m\eta^2}{(m\eta^2 + \sigma^2)}$ es el factor de credibilidad, el cual tiene nuevamente comportamiento monótono creciente a uno conforme m crece a infinito.

Observe además que $Var(\theta|\underline{S}) = (\frac{m}{\sigma^2} + \frac{1}{\eta^2})^{-1}$ y que esta cantidad converge a cero cuando el tamaño de muestra m crece a infinito, indicando nuevamente que la distribución a posteriori se concentra cada vez más alrededor de su media.

Ejemplo:

Suponga $\{X_i\}_{i=1}^n$ una muestra aleatoria con distribución *Bernoulli*(Θ) y consideremos $\Theta \sim \text{Beta}(\Lambda = (a, b))$.

Con base en esto, ya vimos que: $\hat{\Theta} = \frac{n\bar{X} + a}{n + a + b} = \frac{n}{n + a + b}\bar{X} + \frac{1}{n + a + b}a$

Sea Z tal que: $\frac{1}{n + a + b} = Z \frac{1}{a + b} \Leftrightarrow Z = \frac{a + b}{n + a + b} \in (0, 1)$, además notemos que:

$1 - Z = 1 - \frac{a + b}{n + a + b} = \frac{n}{n + a + b}$. De tal manera que:

$$\hat{\Theta} = \frac{n}{n + a + b}\bar{X} + \frac{1}{n + a + b}a = (1 - Z)\bar{X} + Z\frac{a}{a + b} = (1 - Z)\bar{X} + Z\mathbb{E}[\theta]; \text{ si } Z' \doteq 1 - Z$$

$$\therefore \begin{array}{l} \text{Prima final} \\ \text{(Posterior)} \end{array} = \hat{\Theta} = Z'\bar{X} + (1 - Z')\mathbb{E}[\Theta] = Z' \left(\begin{array}{l} \text{Experiencia} \\ \text{de los datos} \end{array} \right) + (1 - Z') \left(\begin{array}{l} \text{Prima inicial} \\ \text{(a priori)} \end{array} \right)$$

\hookrightarrow De credibilidad Bayesiana

Nota: En esta metodología, debemos tener en cuenta que pensamos a la prima de riesgo como una v.a. que cambia con base en la muestra y $\mathbb{E}[S|\Theta = \theta] = \theta$. Pero puede ser que $\mathbb{E}[S|\Theta = \theta] = \tau(\theta)$

Un par de últimas observaciones de lo que se está haciendo. Noten que la **prima de credibilidad bayesiana** se puede obtener sin calcular el **factor de credibilidad** (Z), más bien, se obtiene simplemente de actualizar la distribución inicial.

La aparición de Z se hace para observar qué tanto se otorga en la transformación lineal convexa, a la **experiencia muestral** y a la propuesta **teórica**.

$$Z\bar{X} + (1 - Z)\mathbb{E}[\Theta]$$

Por eso en **credibilidad Bayesiana** es fundamental proponer una distribución inicial que tenga como media (al menos de forma numérica) la **Prima de Riesgo teórica** que se propone en el **modelo inicial**.

También se debe recalcar que **el modelo que se propone inicialmente es para $p(\underline{x}|\theta)$** (la verosimilitud) **no para $p(\theta)$** (la distribución inicial). Esto significa que todo lo que vimos en la primera parte del curso es para $p(\underline{x}|\theta)$. Mientras que $p(\theta)$ se propone “a colmillo” y tal **teóricamente**:

$$\text{si } \mathbb{E}[S] = \mathbb{E}[S|\Theta = \theta] = \tau(\theta) \Rightarrow \tau(\mathbb{E}[\Theta]) = \tau(\theta).$$

Más aún, la **prima de riesgo de Credibilidad Bayesiana** y el **factor de credibilidad Bayesiana** son respectivamente:

$$\tau(\hat{\Theta}) = \tau(\mathbb{E}[\Theta|\underline{S}]) \quad \text{y} \quad Z \cdot \tau(\hat{\Theta}) = Z\bar{S} + (1 - Z)\tau(\mathbb{E}[\Theta])$$

Todo esto significa que debemos plantear la distribución inicial con base en nuestro modelo y luego simplemente actualizarlo. Vía estadística Bayesiana.

Ejemplo: Por muchos años un investigador ha estudiado una muestra de observaciones S_i . Él ha dicho en diversos artículos que $S_i \sim \text{Exp}(\lambda)$, $\mathbb{E}[S_i] = 2 \forall i$. Hoy en día existe un fenómeno natural que hace dudar que $\lambda = \frac{1}{2}$.

Pero no se sabe qué valor pueda tomar λ , así que los expertos en las S_i 's han dicho que λ , podría tener un comportamiento $\Lambda \sim \text{Gamma}(\alpha, \beta)$ con $\alpha = 2$. Hoy se tomó una muestra aleatoria de S la cual es la siguiente: $\{2.71, 11.04, 0.53, 0.88, 0.14, 7.13, 5.35, 2.82, 1.14, 5.09\}$.

Si S está asociado a un monto de siniestros. ¿Cuál sería una prima de riesgo adecuada dada esta información?

Solución:

Como el investigador decía que antes $\mathbb{E}[S_i] = 2$ y se ha propuesto una distribución inicial $\Lambda \sim \text{Gamma}(2, \beta)$ entonces:

$$\mathbb{E}[S|\Lambda = \lambda] = \boxed{\tau(\lambda) = \frac{1}{\lambda}} = 2 \Rightarrow \frac{\beta}{2} = \frac{1}{\mathbb{E}[\Lambda]} = \tau(\mathbb{E}[\Lambda]) = \tau(\lambda) = 2, i.e.,$$

Los expertos proponen una distribución inicial $\Lambda \sim \text{Gamma}(\alpha = 2, \beta = 4)$.

Trabajando con la **nueva información**:

- Distribución inicial: $P(\lambda) \sim \text{Gamma}(\alpha = 2, \beta = 4)$
- Verosimilitud: $P(\underline{S}|\lambda) \sim \text{Exp}(\Lambda)$.

$$\Rightarrow \text{Distribución final: } p(\lambda|\underline{S}) \sim \text{Gamma}(\alpha + n = 12, \beta + \sum_{i=1}^n S_i = 40.83)$$

ejercicio para el lector

$$\Rightarrow \hat{\Lambda} \doteq \mathbb{E}[\Lambda|\underline{S}] = \frac{12}{40.83} = \frac{400}{1361} \approx 0.293901543$$

$$\Rightarrow \text{Prima de Riesgo de C.B.} = \tau(\hat{\Lambda}) = \frac{1361}{400}$$

	Antes:		Ahora:
λ	$\mathbb{E}[\Lambda] = \frac{1}{2}$	\rightarrow	$\hat{\Lambda} = \frac{400}{1361}$
Prima de Riesgo $\mathbb{E}[S] = \frac{1}{\lambda}$	$\mathbb{E}[S \Lambda = \lambda] = 2$	\rightarrow	$\tau(\hat{\Lambda}) \frac{1361}{400} \approx 3.4025$

\therefore La prima de riesgo actualizada es 3.4025.

Ejemplo (Parte 2): Del ejemplo anterior ¿cuál sería el factor de credibilidad?

Solución: Buscamos Z tal que:

$$\begin{aligned}
 Z \left(\frac{36.83}{10} \right) + (1 - Z)(2) &= Z\bar{S} + (1 - Z)\tau(\mathbb{E}[\Lambda]) = \text{Prima de Riesgo} = \tau(\hat{\Lambda}) = 3.4025 \\
 \Rightarrow 3.683Z + 2.2Z &= 2 + 1.6837 = 3.4025 \Leftrightarrow Z = \frac{3.4025 - 2}{1.683} \\
 \therefore \text{El factor de credibilidad es } Z &= \frac{3.4025 - 2}{1.683} = \frac{5}{6} \approx 83.3\bar{3} \%
 \end{aligned}$$

Moraleja: Recuerda siempre lo que se está haciendo y cómo se está haciendo.

Nota: Todo esto tiene sentido pues el autor obtuvo los datos de una *Exponencial* ($\lambda = \frac{1}{4}$)