Funciones de masa y densidad

<u>*Definición I.3.-*</u> La *función de probabilidad* de la variable aleatoria X está dada por:

$$p(x) = P(X = x) \quad \text{Si } X \text{ es discreta } \forall x \in R \quad \text{(función de masa)}$$

$$f(x) = P(x_i \leq X \leq x_j) \quad \text{Si } X \text{ es continua } \forall x_i, x_j \in R \quad \text{(función de densidad)}$$

Propiedades

<u>Función Discreta</u>	<u>Función Continua</u>
(1) $p(x) \geq 0 \quad \forall x \in R$	(1) $f(x) \geq 0, \quad -\infty \leq x \leq \infty$
(2) $\sum_{\forall x} p(x) = 1$	(2) $\int_{-\infty}^{\infty} f(x)dx = 1$
	(3) $P(a \leq X \leq b) = \int_a^b f(x)dx \quad \forall a, b \in R$

Funciones de distribución

<u>*Definición I.4.-*</u> La *función de distribución de probabilidad* o simplemente *función de distribución*, está dada por:

$$F_X(x) = F(x) = \sum_{x_i \leq x} p(x_i) = P(X \leq x) \qquad \text{Si } X \text{ es discreta}$$

$$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)dt \qquad \text{Si } X \text{ es continua}$$

Propiedades

<u>Función Discreta</u>	<u>Función Continua</u>
1) $0 \leq F(x) \leq 1 \quad \forall x \in R$	(1) $F(-\infty) = 0, F(\infty) = 1$
(2) $F(x_i) \geq F(x_j)$ Si $x_i \geq x_j$	(2) $P(a < X < b) = F(b) - F(a)$
(3) $P(X > x) = 1 - F(x)$	(3) $\frac{d}{dx}F(x) = f(x)$

Esperanza y varianza

<u>*Definición I.5.*-</u> El *valor esperado* de una variable aleatoria X se define como:

$$E(X) = \sum_x xp(x) \quad \text{Si } X \text{ es discreta, } p(x) \text{ función de masa}$$

$$= \int_{-\infty}^{\infty} xf(x)dx \quad \text{Si } X \text{ es continua, } f(x) \text{ función de densidad}$$

El valor esperado de una función g(x) de la variable aleatoria X , se define como:

$$E(g(X)) = \sum_x g(x)p(x) \quad \text{Si } X \text{ es discreta}$$

$$= \int_{-\infty}^{\infty} g(x)f(x)dx \quad \text{Si } X \text{ es continua}$$

<u>*Definición I.6.*-</u> El *momento* de una variable aleatoria X se define como:

$$\mu_n = m_n = E(X^n)$$

Esto es, como los valores esperados de las sucesivas potencias de la variable aleatoria X, por lo tanto, el r-eximo momento de X alrededor del cero se define como:

$$\mu_r^{'} = E(X^r) = \sum_x x^r p(x) \qquad \text{(Caso discreto)}$$

$$= \int_{-\infty}^{\infty} x^r f(x)dx \qquad \text{(Caso continuo)}$$

Entonces, el r-eximo momento de X alrededor de la media se define como:

$$\mu_r = E(X-\mu)^r = \sum_x (x-\mu)^r p(x) \qquad \text{(Caso discreto)}$$

$$= \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx \qquad \text{(Caso continuo)}$$

Definición I.7.- La *varianza* de una variable aleatoria X se define como:

$$\sigma_x^2 = Var(X) = E(X^2) - E^2(X) = \mu_2$$

Por lo tanto, la desviación stándard es:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{Var(X)}$$

Funciones conjuntas, marginales y condicionales

Definición I.8.- La *función de probabilidad conjunta* de las variables aleatorias (X,Y) está dada por:

$p(x,y) = P(X = x, Y = y)$ *(caso discreto)*

$f(x,y) \geq 0 \ \ \forall (x,y) \in R$ *(caso continuo, función de densidad conjunta)*

Propiedades

Función Discreta	Función Continua
(1) $p(x,y) \geq 0$	(1) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1$
(2) $\sum_x \sum_y p(x,y) = 1$	(2) $P(a < X < b, c < Y < d) = \int_a^b \int_c^d f(x,y)dydx$

$\forall (a,b,c,d) \in R$

Definición I.9.- La *función de distribución conjunta* de (X,Y), para el caso continuo, está dada por:

$$F(x,y) = P(X \leq X, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dvdu$$

<u>*Definición I.10.-*</u> Sean (X,Y) dos variables aleatorias con funciones de probabilidad conjunta $p(x,y)$ y $f(x,y)$. Las *funciones marginales* de (X,Y) , están dadas por:

<u>*Caso discreto*</u> <u>*Caso continuo*</u>

(funciones de probabilidad marginal) (funciones de densidad probabilidad marginal)

$$p_X(x) = \sum_y p(x,y) \qquad f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

$$p_Y(y) = \sum_x p(x,y) \qquad f_Y(x) = \int_{-\infty}^{\infty} f(x,y)dx$$

<u>*Definición I.11.-*</u> Sea $F(x,y)$ la función de distribución conocida. *Las distribuciones acumulativas marginales de (X,Y), son:*

$$P(X \leq x) = F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(t,y)dydt = \int_{-\infty}^{x} f_X(t)dt = F(x,\infty)$$

$$P(Y \leq y) = F_Y(y) = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f(x,t)dxdt = \int_{-\infty}^{y} f_Y(t)dt = F(\infty,y)$$

<u>*Definición I.12.-*</u> Sea (X,Y) una variable aleatoria bidimensional con funciones de probabilidad conjunta $p(x,y)$ y $f(x,y)$. Sean $p(x) = p_X(x)$, $p(y) = p_Y(y)$ y $f(x) = f_X(x)$, $f(y) = f_Y(y)$ funciones de probabilidad marginales, *las funciones de probabilidad condicionales* están dadas por:

<u>*Caso discreto*</u> <u>*Caso continuo*</u>

$$p(x/y) = \frac{p(x,y)}{p_Y(y)} \quad p_Y(y) > 0 \qquad f(x/y) = \frac{f(x,y)}{f_Y(y)} \quad f_Y(y) > 0$$

$$p(y/x) = \frac{p(x,y)}{p_X(x)} \quad p_X(x) > 0 \qquad f(y/x) = \frac{f(x,y)}{f_X(x)} \quad f_X(x) > 0$$

Independencia

<u>*Definición I.13.-*</u> Sean (X,Y) dos variables aleatorias con distribución conjunta. Las variables aleatorias X ,Y son independientes si y sólo si:

$p(x,y) = p_X(x)p_Y(y)$ Si X y Y son discretas

$f(x,y) = f_X(x)f_Y(y)$ Si X y Y son continuas

Covarianza y correlación

<u>*Definición I.14.-*</u> Sean (X,Y) dos variables aleatorias con distribución conjunta. Sean $E(X) = \mu_X$, $E(Y) = \mu_Y$, $E(Y) = \mu_Y$, $Var(X) = \sigma_x^2$, $Var(Y) = \sigma_Y^2$. La covarianza de X e Y, se define como:

$$Cov(X,Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

De aquí se desprende que:

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

El valor de la covarianza puede ser positivo, negativo o cero.

<u>*Definición I.15.-*</u> Si $0 < \sigma_X^2 < \infty$ y $0 < \sigma_Y^2 < \infty$, entonces *la correlación o coeficiente de correlación* de X e Y, se define como:

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Nótese que si X y Y son independientes $Cov(X,Y) = 0$

Esperanza, varianza y covarianza condicional

<u>*Definición I.16.-*</u> Sean (X,Y) dos variables aleatorias distribuidas conjuntamente. La esperanza condicional, se define como:

$E(X/Y=y)=E(X/y)=\int_{-\infty}^{\infty} xf(x/y)dx$ Esperanza condicional de X dado que $Y=y$

$E(Y/X=x)=E(Y/x)=\int_{-\infty}^{\infty} yf(y/x)dy$ Esperanza condicional de Y dado que $X=x$

Nótese que si X y Y son independientes, se tiene que:

$$E(X/y)=E(X) \text{ y } E(Y/x)=E(Y)$$

Ahora bien, definiremos la esperanza condicional de $g(X,Y)$, dado que $X=x$, la cual se establece como:

$$E[g(X,Y)/X=x]=\int_{-\infty}^{\infty} g(x,y)f_{Y/X}(y/x)dy$$

Se deja para el lector, la deducción de la esperanza condicional de $g(X,Y)$, dado que $Y=y$.

<u>*Definición I.17.-*</u> Sean (X,Y) dos variables aleatorias distribuidas conjuntamente. La varianza y la covarianza de Y dado que $X=x$, se definen como:

$$Var[Y/X=x]=E[Y^2/x]-E^2[Y/x]$$

$$Cov[X,Y/\Theta]=E[(X-E[X/\Theta])(Y-E[Y/\Theta])]$$

Es para el lector, la deducción de la varianza y la covarianza de X dado que $Y=y$. *Resultados importantes*

$E[X]=E[E[X/Y]]$

$$Var[Y] = E[Var[Y/X]] + Var[E[Y/X]]$$

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

<u>*Definición I.18.-*</u> Sean $X_1, X_2, ..., X_n$ n variables aleatorias con una función de densidad conjunta de probabilidad $f(x_1, x_2, ..., x_n)$. Entonces:

$$E\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n}\left[a_i E(X_i)\right]$$

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var(X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ i\neq j}}^{n} a_i a_j Cov(X_i, X_j)$$

Para cualquier constante $a_i, i = 1,2,...,n$

Conjugadas

<u>*Definición I.19-*</u> Sea Y una *variable aleatoria* (discreta o continua) tal que sus valores representen las posibles opciones en que puede ocurrir un fenómeno aleatorio antes de llevar a cabo un experimento. ***El grado de creencia*** del investigador con respecto a estas posibilidades se encuentra expresado por una función de probabilidad la cual se denomina como:

$p_Y(y)$ *función de probabilidad a priori* de Y (si es discreta)

$f_Y(y)$ *función de densidad de probabilidad a priori* Y (si es continua).

<u>*Definición I.20-*</u> Sea $f(x/y)$ la función de probabilidad condicional de cualquier variable aleatoria X (discreta o continua), la cual representa evidencia muestral en función de una alternativa fija y de Y. La función $f(x/y)$ recibe el nombre de

función de verosimilitud, dado que representa el grado de concordancia del resultado muestral x, dado el valor y de Y.

Definición I.21- Sea $p_Y(y)$ o $f_Y(y)$ la función de probabilidad o función de densidad de probabilidad *a priori* de Y, respectivamente, y sea $f(x/y)$ la función de verosimilitud. Entonces la *probabilidad a posteriori* o *función de densidad de probabilidad a posteriori* de Y dada la evidencia muestral x, es:

$$p(y/x) = \frac{f(x/y)p_Y(y)}{\sum_Y f(x/y)p_Y(y)} \quad \text{Si } Y \text{ es discreta}$$

$$f(y/x) = \frac{f(x/y)f_Y(y)}{\int_Y f(x/y)f_Y(y)dy} \quad \text{Si } Y \text{ es continua}$$

La función de probabilidad a posteriori $p(x/y)$ o la función de densidad de probabilidad a posteriori $f(x/y)$ reflejan el grado de creencia modificado del investigador con respecto a la variable aleatoria Y después de obtener información muestral. Esta información puede verificarse de forma periódica. Lo anterior significa que en un futuro la distribución *a posteriori* puede convertirse en una distribución *a priori*.

Nótese que el denominador de las ecuaciones anteriores es la función de densidad de probabilidad marginal o no condicional de X; esto es,

$$f_X(x) = \sum_Y f(x/y)p_Y(y) \quad \text{Caso discreto}$$

$$f_X(x) = \int_Y f(x/y)f_Y(y)dy \quad \text{Caso continuo}$$

Herramientas estadísticas

Definición I.22 – Si las variables aleatorias $X_1, X_2,\ldots,X_n$ tienen la misma función de densidad de probabilidad que la de la distribución de la población y su función de distribución conjunta de probabilidad es igual al producto de las marginales, entonces $X_1, X_2,\ldots,X_n$ forman un conjunto de n variables aleatorias independientes e idénticamente distribuidas a esto se le denomina una *muestra aleatoria* de la población.

Definición I.23 – *Un parámetro* es una caracterización numérica de la distribución de la población de manera que describe, parcial o completamente, la función de densidad de probabilidad de la característica de interés. Por ejemplo, cuando se específica el valor del parámetro de escala exponencial θ, se describe de manera completa la función de densidad de probabilidad:

$$f(x;\theta) = \frac{1}{\theta}\exp(-x/\theta)$$

Definición I.24 – Sea λ el parámetro de la variable aleatoria X, si X puede ser generada bajo el supuesto de que el parámetro Λ es una variable aleatoria con función de densidad $U(\lambda)$, a la distribución $U(\lambda)$ se denomina *función estructural*

Definición I.25 – Una *estadística* es cualquier función de las variables aleatorias que se observaron en la muestra de manera que esta función no contiene cantidades desconocidas, denótese a una estadística como $T = \mu(\underline{X})$.

Definición I.26 – Una *estimador* del parámetro θ, basado en las variables aleatorias $X_1, X_2,\ldots,X_n$, es una función $\delta(X_1, X_2,\ldots,X_n)$ que especifica el valor estimado de θ para cada conjunto de valores posibles de $X_1, X_2,\ldots,X_n$. En otras palabras, si los valores observados de $X_1, X_2, ..., X_n$ son $x_1, x_2, ..., x_n$ entonces el valor estimado de θ es $\delta(X_1, X_2, ..., X_n)$.

Las propiedades de un estimador son las siguientes:

a) Insesgado

Siendo $X_1, X_2, ..., X_n$ una muestra aleatoria de $f(x|\theta)$ y $t(\underline{x})$ un estimador de θ, entonces

$E(t(\underline{x})) = \theta$

b) Eficiente(varianza mínima)

Sea $t(\underline{x})$ un estimador, si $h(\underline{x})$ es otro estimador entonces $Var(t(\underline{x})) \leq Var(h(\underline{x}))$

c) Suficiente

Si la distribución de la muestra aleatoria dado el estimador $t(\underline{x})$, no depende del parámetro, entonces: $f(\underline{x}\,|t(\underline{x})) = h(\underline{x})$

d) Consistente

Si el estimador $t(\underline{x})$, tiende al parámetro, se interpreta como:

$\lim_{n\to\infty} P = \{|t(\underline{x}) - \theta| > \varepsilon\} = 0$

e) Invarianza

Si $t(\underline{x})$ es un estimador de θ y $g(*)$ es una función continua, entonces:

$g(t(\underline{x})) = g(\hat{\theta})$

f) Robusto

Si $t(\underline{x})$ es un estimador que no se afecta por valores extremos, entonces $t(\underline{x})$ es un estimador robusto

g) Completo

Si $E(t(\underline{x})) \equiv 0$, se dice que $t(\underline{x})$ es un estimador completo.

En estadística existen métodos de estimación, tales como *método de los momentos*, el de *máxima verosimilitud* o por *mínimos cuadrados*, dependiendo de las características que se deseen en la estimación, se cumplirán las propiedades generales descritas.

Tabla I.1 Propiedades de los estimadores

Tipos de estimación	Propiedades cumplidas de los estimadores
Método de los momentos	Consistencia
Máxima verosimilitud	Consistencia Eficiencia Suficiencia Invarianza
Mínimos cuadrados	Insesgados Eficientes

Definición I.27 - Sea $\hat{\theta}$ un estimador de θ, la función $E\left[(\hat{\theta}-\theta)^2\right] = Var(\hat{\theta}) + sesgo^2(\hat{\theta})$, es llamado el error cuadrático medio del estimador $\hat{\theta}$, denotado por: $ECM(\hat{\theta})$

Apéndice C.

Modelos de Credibilidad Bayesiana.

Tomado de de MORENO MUÑOZ, M. y RAMOS BURGOA, L (2003) [7].

C.1. Modelo de Bühlmann

Con el objetivo de obtener la prima de riesgo de una cartera o conglomerado, se determina un estimador lineal que permita ponderar la experiencia individual con la de toda la cartera. Esta es la idea esencial del modelo original planteado por Bühlmann.

La cartera involucrada en el modelo se encuentra expuesta a un riesgo fijo y desconocido $\boldsymbol{\Theta} = \boldsymbol{\theta}$, durante el período de t años. Sean $\mathbf{X_1}$, $\mathbf{X_2}$,..., $\mathbf{X_t}$ los siniestros individuales en los períodos **1,...,** $\boldsymbol{t}$ respectivamente y sea $\boldsymbol{\theta}$ que se distribuye como la función estructural $\boldsymbol{\Pi_0(\theta)}$. Conocido el parámetro de riesgo $\boldsymbol{\Theta}$, las reclamaciones son condicionalmente independientes e idénticamente distribuidas con una función de distribución $\mathbf{F_{x/\theta}(x, \theta)}$.

A continuación se definen las siguientes variables con la finalidad de desarrollar el modelo:

$$\mu(\theta) = E\left[X_r / \Theta = \theta\right] \tag{2.4.1}$$

la prima por el principio de prima neta o equivalente.

$$m = E\left[X_r\right] = E\left[\mu(\theta)\right] \tag{2.4.2}$$

la esperanza de la prima teórica.

$$a = Var\left[E\left[X_r / \Theta = \theta\right]\right] = Var\left[\mu(\theta)\right] \tag{2.4.3}$$

la varianza del parámetro

$$\sigma^2(\theta) = Var[X_r / \Theta = \theta] \tag{2.4.4}$$

la varianza de los siniestros que pertenecen a la cartera

$$s^2 = E[Var[X_r / \Theta]] = E[\sigma^2(\Theta)] \tag{2.4.5}$$

heterogeneidad promedio en el tiempo de los montos de siniestros de las carteras.
Ahora bien, se debe determinar

$$\mu(\theta) = E(X_r / \Theta = \theta) \tag{2.4.6}$$

que es la prima de cobro que se estima según Bühlmann a través de una función ***g*(*)**, que depende de las reclamaciones observadas $\underline{\mathbf{X}} = (\mathbf{X_1, X_2, ..., X_t})$, es decir de la experiencia propia de la cartera.

Si se tiene una variable aleatoria ***X*** , con función de densidad $f(x,\theta)$ y si $\mathbf{T = \Pi_0(X_1, X_2, ..., X_t)}$ es cualquier estadística, mediante el cálculo del error cuadrático medio de ***T*** , denotado como ***ECM*(*T*)**, se encontrará una función $\boldsymbol{\mu}$ que nos proporcione la mejor estimación del parámetro $\boldsymbol{\theta}$; el estimador incluye dos cantidades mayores a cero, varianza y cuadrado del sesgo.

Se puede verificar que si ***X*** y ***Y*** son dos variables aleatorias, la función ***g*(*)** de ***X* / *Y*** que minimiza el ***ECM*** es:

$$g(X) = E[Y / X] \tag{2.4.7}$$

De aquí