EYP 1025-1027 Métodos Probabilísticos

Clase 18

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Contenido I

Dado un un vector aleatorio (X, Y) en $(\Omega, \mathcal{A}, P)$, suponga que se desea conocer la distribución de probabilidad condicional de Y cuando se sabe que $X = x$ para algún x en el recorrido de X; es decir, se quiere cacular

$P(Y \in B | X = x)$ para cualquier subconjunto B de números reales.

Distribuciones condicionales

Ejemplo 1.1

Suponga que se lanza un dado justo dos veces. Si X_1 y X_2 son los puntajes del primer y segundo lanzamiento, respectivamente, defina $X = X_1 + X_2$ e $Y = \text{mín}\{X_1, X_2\}$. Note que X e Y son variables aleatorias discretas. Para calcular $P(Y = 2|X = 7)$, considere los eventos $A = \{X = 7\}$ y $B = \{Y = 2\}$. Es claro que $P(A) = 6/36$ y $P(A \cup B) = 2/36$. Entonces,

$$
\begin{aligned}
P(Y = 2|X = 7) = P(B|A) &= \frac{P(A \cap B)}{P(A)} \\
&= \frac{P(X = 7, Y = 2)}{P(X = 7)} \\
&= \frac{2/36}{6/36} = \frac{1}{3}.
\end{aligned}
$$

Distribuciones condicionales

En general, si (X, Y) es un vector aleaorio discreto, entonces

$$P(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)} \quad \text{si} \quad f_X(x) = P(X = x) > 0.$$

Cuando $f_X(x) = P(X = x) = 0$, esta probabilidad condicional se puede definir de forma arbitraria, digamos $P(Y = y|X = x) = 0$. Sea, $f_{Y|X=x}(y) = P(Y = y|X = x)$ definida como,

$$(*) \quad f_{Y|X=x}(y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)}, & \text{si } f_X(x) > 0, \\ 0, & \text{si no.} \end{cases}$$

Entonces,

(a) $0 \leq f_{Y|X=x}(y) \leq 1$ para todo (x, y),

(b) $\displaystyle\sum_{y \in \mathbb{R}} f_{Y|X=x}(y) = \sum_{y \in \mathbb{R}} \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{\sum_{y \in \mathbb{R}} f_{X,Y}(x, y)}{f_X(x)} = \frac{f_X(x)}{f_X(x)} = 1$

para todo x.

Distribuciones condicionales

Es decir, de (a) y (b) sigue que la función $f_{Y|X=x}(y) = P(Y = y|X = x)$ definida en $(*)$ constituye una fmp en $\mathbb{R}$, llamada fmp condicional de Y dado $X = x$.

De forma análoga se define $f_{X|Y=y}(x) = P(X = x|Y = y)$ como la fmp condicional de X dado $Y = x$.

Además, procediendo de forma relativamente similar, se pueden construir las fdp's condicionales $f_{Y|X=x}(y)$ y $f_{X|Y=y}(x)$ en el caso continuo.

Nota: Para la fmp o fdp condicional de Y dado $X = x$ (X dado $Y = y$), también se usa la notación $f(y|x)$ ($f(x|y)$).

Distribuciones condicionales

Definición 1.1

Sea (X, Y) un vector aleatorio discreto o continuo, con fmp (c.d.) o fdp (c.c.) conjunta $f_{X,Y}(x, y)$ y fmp's (c.d.) o fdp's (c.c.) marginales $f_X(x)$ y $f_Y(y)$. La fmp (c.d.) o fdp (c.c.) condicional de Y dado $X = x$ se define como,

$$f_{Y|X=x}(y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)}, & \text{si } f_X(x) > 0, \\ 0, & \text{si no.} \end{cases}$$

Análogamente, fmp (c.d.) o fdp (c.c.) condicional de X dado $Y = y$ se define como,

$$f_{X|Y=y}(x) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)}, & \text{si } f_Y(y) > 0, \\ 0, & \text{si no.} \end{cases}$$

Distribuciones condicionales

Teorema 1.1

La distribución de probabilidad condicional de Y dado $X = x$, esta dada por,

$$P(Y \in B|X = x) = \begin{cases} \sum_{y\in B} f_{Y|X=x}(y) & \text{c.d.} \\ \int_{y\in B} f_{Y|X=x}(y)dy & \text{c.c.} \end{cases}$$

para todo $B \subset \mathbb{R}$. En particular, la fda condicional de de Y dado $X = x$, esta dada por,

$$F_{Y|X=x}(y) = P(Y \leq y|X = x) = \begin{cases} \sum_{z\leq y} f_{Y|X=x}(z) & \text{c.d.} \\ \int_{-\infty}^{y} f_{Y|X=x}(z)dz & \text{c.c.} \end{cases}$$

para todo y.

Nota: La definición y los resultados anteriores son análogos si $(\boldsymbol{X}, \boldsymbol{Y})$ es un vector aleatorio discreto o continuo, con $\boldsymbol{X} : \Omega \longrightarrow \mathcal{X} \subseteq \mathbb{R}^n$ e $\boldsymbol{Y} : \Omega \longrightarrow \mathcal{Y} \subseteq \mathbb{R}^m$.

Ejemplo 1.2

Se extrae al azar una bolita de una urna con N bolitas numeradas del 1 al N. Luego se lanza una moneda tantas veces como lo indica el número de la bolita seleccionada. Sea X el número de la bolita extraída. Si $X = x$, entonces se lanza la moneda x veces. Si Y es el número caras obtenidas en los x lanzamientos de la moneda, entonces, $Y|X = x \sim Bin(x, p)$, donde p es la probabilidad de obtener cara al lanzar la moneda. Es decir,

$$
\begin{aligned}
f_{Y|X=x}(y) &= P(Y = y|X = x) \\
&= \begin{cases} \left(\begin{array}{c} x \\ y \end{array} \right) p^y (1-p)^{x-y}, & \text{si } y = 0, 1, \ldots, x, \text{ para } x = 1, \ldots, N, \\ 0, & \text{si no.} \end{cases}
\end{aligned}
$$

Ejemplo 1.3

Sean X e Y variables aleatorias continuas con fdp conjunta dada por,

$$f_{X,Y}(x,y) = \begin{cases} 2, & \text{si } x > 0,\ y > 0,\ x + y < 1, \\ 0, & \text{si no.} \end{cases}$$

$$\Longrightarrow f_X(x) = \begin{cases} \int_0^{1-x} 2dy = 2(1-x), & \text{si } 0 < x < 1, \\ 0, & \text{si no.} \end{cases}$$

Luego,

$$f_{Y|X=x}(y) = \begin{cases} \frac{1}{1-x}, & \text{si } 0 < y < 1 - x,\ \text{para } 0 < x < 1, \\ 0, & \text{si no.} \end{cases}$$

Note que $Y|X = x \sim U(0, 1-x)$ para cada $x \in (0,1)$. Análogamente, se tiene que $X|Y = y \sim U(0, 1-y)$ para cada $y \in (0,1)$.

1) $f_{X,Y}(x,y) = f_{Y|X=x}(y)f_X(x) = f_{X|Y=y}(x)f_Y(y)$ para todo (x,y)

2) Si X e Y son variables aleatorias independientes, entonces $f_{Y|X=x}(y) = f_Y(y)$ y $f_{X|Y=y} = f_X(x)$ para todo (x,y)

3) Para cada y (fijo), se tiene que $f_{Y|X=x}(y) = g(x)$ es una función (no aleatoria) de x definida sobre el recorrido de x

Por ejemplo, si $Y|X = x \sim Bin(x,p)$, entonces,
$f_{Y|X=x} = \left(\begin{array}{c} x \\ y \end{array} \right) p^y(1-p)^{x-y} = g(x)$ para cada $y = 0, 1, \ldots, x$.

Similarmente, si $Y|X = x \sim U(0, 1-x)$, entonces
$f_{Y|X=x}(y) = 1/(1-x) = g(x)$ para cada $y \in (0, 1-x)$.

Distribuciones condicionales

Propiedades básicas

Considere la función aleatoria $g(X) = f_{Y|X}(y)$. Entonces,

$$\mathsf{E}\{g(X)\} = \mathsf{E}\{f_{Y|X}(y)\} = f_Y(y).$$

En efecto. Considere el caso continuo; entonces

$$\begin{aligned}
\mathsf{E}(g(X)) &= \int_{-\infty}^{\infty} g(x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} f_{Y|X=x}(y) f_X(x) dx \\
&= \int_{-\infty}^{\infty} f_{Y,X}(x,y) dx \quad \text{(por la propiedad 1))} \\
&= f_Y(y).
\end{aligned}$$

De aquí, también es inmediato que

$$\mathsf{E}\{P(Y \in B|X)\} = P(Y \in B) \quad \forall\ B \subseteq \mathbb{R}.$$

Definición 1.2

La esperanza condicional de Y dado $X = x$, provisto que exista, se define como,

$$E(Y|X = x) = \begin{cases} \sum_{y\in\mathbb{R}} y f_{Y|X=x}(y) & \text{c.d.} \\ \int_{-\infty}^{+\infty} y f_{Y|X=x}(y)dy & \text{c.c.} \end{cases}$$

La esperanza condicional de X dado $Y = y$ se define de forma análoga.

Nota: Si Y tiene esperanza finita, entonces la esperanza condicional de Y dado $X = x$ también es finita (con probabilidad 1).

Más generalmente, si $g(Y)$ tiene tiene esperanza finita, entonces la esperanza condicional de $g(Y)$ dado $X = x$, se define como,

$$\mathsf{E}\{g(Y)|X = x\} = \begin{cases} \sum_y g(y) f_{Y|X=x}(y) & \text{c.d.} \\ \int_{-\infty}^{+\infty} g(y) f_{Y|X=x}(y) dy & \text{c.c.} \end{cases}$$

Ejemplo 1.4

1) Si $Y|X = x \sim Bin(x, p)$, entonces

$$\mathrm{E}(Y|X = x) = \sum_{y=0}^{x} y \binom{x}{y} p^y (1-p)^{x-y} = xp.$$

2) Si $Y|X = x \sim U(0, 1-x)$, entonces

$$\mathrm{E}(Y|X = x) = \int_{y=0}^{1-x} y \frac{1}{1-x} dy = \frac{1-x}{2}.$$

Distribuciones condicionales

Propiedad importante

Se desprende de los ejemplos anteriores que

$$\mathsf{E}(Y|X=x) = h(x) \quad \text{(función no aleatoria de } x\text{)}$$

Sea

$$h(X) = \mathsf{E}(Y|X) \quad \text{(función aleatoria de } X\text{)}$$

Teorema 1.2

Ley de probabilidad total para esperanzas. Sean X e Y variables aleatorias definidas en el mismo espacio de probabilidad. Si Y tiene esperanza finita, entonces,

$$\mathrm{E}(Y) = \mathrm{E}\{\mathrm{E}(Y|X)\}.$$

Distribuciones condicionales

Demostración. Caso continuo: Ya que $\mathsf{E}(Y|X) = h(X)$, entonces,

$$
\begin{aligned}
\mathsf{E}\{\mathsf{E}(Y|X)\} &= \mathsf{E}\{h(X)\} \\
&= \int_{-\infty}^{\infty} h(x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} \mathsf{E}(Y|X = x) f_X(x) dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{Y|X=x}(y) dy f_X(x) dx \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{Y|X=x}(y) f_X(x) dy dx \\
&= \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy \quad (f_{Y|X=x}(y) f_X(x) = f_{X,Y}(x, y)) \\
&= \int_{-\infty}^{\infty} y f_Y(y) dy = \mathsf{E}(Y).
\end{aligned}
$$

Ejemplo 1.5

1) Si $Y|X = x \sim Bin(n, p)$, entonces, $\mathrm{E}(Y|X = x) = xp$, de modo que $\mathrm{E}(Y|X) = Xp$; luego,

$$\mathrm{E}(Y) = \mathrm{E}\{\mathrm{E}(Y|X)\} = \mathrm{E}(Xp) = \mathrm{E}(X)p.$$

Así, si $X \sim P(\lambda)$, entonces $\mathrm{E}(X) = \lambda$ y por tanto $\mathrm{E}(Y) = \lambda p$.

2) Si $Y|X = x \sim U(0, 1-x)$, entonces, $\mathrm{E}(Y|X = x) = (1-x)/2$, de modo que $\mathrm{E}(Y|X) = (1-X)/2$; luego,

$$\mathrm{E}(Y) = \mathrm{E}\{\mathrm{E}(Y|X)\} = \mathrm{E}\{(1-X)/2\} = (1-\mathrm{E}(X))/2.$$

Así, si $f_X(x) = 2(1-x)I_{(0,1)}(x)$, entonces
$\mathrm{E}(X) = \int_0^1 2x(1-x)dx = 1/3$ y por tanto $\mathrm{E}(Y) = 1/3$.

Ejemplo 1.6

Encuesta de Hogares Sean X el número de miembros en un hogar seleccionado aleatoriamente en la encuesta, e Y el número de automóviles de propiedad de dicho hogar.

Los 250 hogares encuestados tienen la misma probabilidad de ser seleccionados, por lo que $P(X = x, Y = y)$ es igual al número de hogares con x miembros e y autos, dividido por 250; estas probabilidades se presentan en la Tabla 1 dada a continuación.

Suponga que el hogar seleccionado tiene $X = 4$ miembros.

La fmp condicional de Y dado $X = 4$ es $f_{Y|X=4}(y) = f_{X,Y}(4, y)/f_X(4)$, y corresponde a los valores de la columna $x = 4$ de la Tabla 1 dividido por $f_X(4) = 0.208$, es decir,

$$f_{Y|X=4}(0) = 0.0385, \quad f_{Y|X=4}(1) = 0.5769,$$
$$f_{Y|X=4}(2) = 0.2885, \quad f_{Y|X=4}(3) = 0.0962.$$

Distribuciones condicionales

Tabla 1

fmp's conjunta, $f_{X,Y}(x,y)$, y marginales, $f_X(x)$ y $f_Y(y)$, de X e Y.

	x								
y	1	2	3	4	5	6	7	8	$f_Y(y)$
0	0.040	0.028	0.012	0.008	0.008	0.004	0	0	0.100
1	0.048	0.084	0.100	0.120	0.100	0.060	0.020	0.004	0.536
2	0.004	0.020	0.040	0.060	0.080	0.044	0.020	0.012	0.280
3	0	0.008	0.012	0.020	0.020	0.012	0.008	0.004	0.084
$f_X(x)$	0.092	0.140	0.164	0.208	0.208	0.120	0.048	0.020	1.000

La media condicional de Y dado $X = 4$ es,

$$\mathrm{E}(Y|X = 4) = 0 \times 0.0385 + 1 \times 0.5769 + 2 \times 0.2885 + 3 \times 0.0962 = 1.442$$

Similarmente, podemos calcular $\mathrm{E}(Y|X=x)$ para los ocho valores de x; estos son,

x	1	2	3	4	5	6	7	8
$\mathrm{E}(Y\|X=x)$	0.609	1.057	1.317	1.442	1.538	1.533	1.75	2

La variable aleatoria $h(X)$ que toma el valor 0.609 cuando el hogar muestreado tiene un miembro, toma el valor 1.057 cuando el hogar muestreado tiene dos miembros, y así sucesivamente, es $h(X)=\mathrm{E}(Y|X)$, es decir, la esperanza condicional de Y dado la variable aleatoria X.

Distribuciones condicionales

Además de la propiedad importante de que $\mathsf{E}\{\mathsf{E}(Y|X)\} = \mathsf{E}(Y)$ (ver Teorema 1.2), la esperanza condicional posee (condicionalmente) todas las propiedades de la esperanza ordinaria, ya que es la media de la distribución condicional. A cotinuación se enuncian sólo algunas de estas propiedades.

1) $\mathsf{E}(aY + b|X = x) = a\mathsf{E}(Y|X = x) + b$

2) $\mathsf{E}\{g(X,Y)|X = x\} = \mathsf{E}\{g(x,Y)|X = x\}$ (principio de sustitución para la esperanza condicional); en particular, $\mathsf{E}(XY|X = x) = x\mathsf{E}(Y|X)$. Además, $\mathsf{E}\{g(X)h(Y)\} = \mathsf{E}[g(X)\mathsf{E}\{h(Y)|X\}]$; por ejemplo, $\mathsf{E}(XY) = \mathsf{E}\{X\mathsf{E}(Y|X)\}$.

3) Si X e Y son independientes, entonces la distribución condicional de Y dado $X = x$ coincide con la distribución marginal de Y para todo x, es decir, $P(Y \in B|X = x) = P(Y \in B)$ para todo x y todo B, luego $\mathsf{E}(Y|X = x) = \mathsf{E}(Y)$; del mismo modo se tiene que $\mathsf{E}(X|Y = y) = \mathsf{E}(X)$.

Tal como la media condicional, la varianza condicional es simplemente la varianza de la distribución condicional como se define a continuación; por ende también satisface todas las propiedades de la varianza ordinaria.

Definición 1.3

La varianza condicional de Y dado $X = x$, se define como

$$\begin{aligned}\mathrm{Var}(Y|X=x) &= \mathrm{E}\{(Y - E(Y|X=x))^2|x\} \\ &= \begin{cases} \sum_{y\in\mathbb{R}}(y - E(Y|X=x))^2 f_{Y|X=x}(y) & \text{c.d.} \\ \int_{-\infty}^{\infty}(y - E(Y|X=x))^2 f_{Y|X=x}(y)dy & \text{c.c.} \end{cases}\end{aligned}$$

provisto que la esperanza exista.

Tarea: Pruebe que la varianza condicional de Y dado $X = x$, también puede calcularse como $\mathrm{Var}(Y|X=x) = \mathrm{E}(Y^2|X=x) - \{\mathrm{E}(Y|X=x)\}^2$.

Ejemplo 1.7

1) Si $Y|X = x \sim Bin(n,p)$, entonces, $\mathrm{E}(Y|X = x) = xp$, de modo que $\mathrm{E}(Y|X) = Xp$; luego,

$$\mathrm{Var}(Y|X = x) = \sum_{y=0}^{x} (y - xp)^2 \left(\begin{array}{c} x \\ y \end{array} \right) p^y (1-p)^{x-y} = xp(1-p).$$

2) Si $Y|X = x \sim U(0, 1-x)$, entonces, $\mathrm{E}(Y|X = x) = (1-x)/2$; luego,

$$\mathrm{Var}(Y|X = x) = \int_{y=0}^{1-x} \left(y - \frac{1-x}{2} \right)^2 \frac{1}{1-x} dy = \frac{(1-x)^2}{12}.$$

De los ejemplos anteriores se desprende que $\text{Var}(Y|X=x)=v(x)$ (función no aleatoria), mientras que $\text{Var}(Y|X)=v(X)$ (función aleatoria)

Teorema 1.3

Ley de probabilidad total para varianzas Sean X e Y variables aleatorias definidas en un mismo espacio de probabilidad. Si $\mathrm{E}(Y^2)$ es finita, entonces,

$$\mathrm{Var}(Y)=\mathrm{Var}\{\mathrm{E}(Y|X)\}+\mathrm{E}\{\mathrm{Var}(Y|X)\}.$$

Demostración 1.1

Se deja como ejercicio para el lector.

Ejemplo 1.8

Si $Y|X = x \sim Bin(n, p)$, entonces, $\mathrm{E}(Y|X = x) = xp$ y $\mathrm{Var}(Y|X = x) = xp(1 - p)$; luego,

$$\begin{aligned} \mathrm{Var}(Y) &= \mathrm{Var}\{Xp\} + \mathrm{E}\{Xp(1-p)\} \\ &= p^2 \mathrm{Var}(Xp) + p(1-p)\mathrm{E}(X) \end{aligned}$$

Si $X \sim P(\lambda)$ entonces $\mathrm{Var}(X) = \mathrm{E}(X) = \lambda$, de modo que $\mathrm{Var}(Y) = \mathrm{E}(Y) = \lambda p$.

Tarea: Pruebe que si $Y|X = x \sim Bin(n, p)$ y $X \sim P(\lambda)$, entonces $Y \sim P(\lambda p)$.

Ejemplo 1.9

Si $Y|X = x \sim U(0, 1-x)$, entonces, $\mathrm{E}(Y|X = x) = (1-x)/2$ y $\mathrm{Var}(Y|X = x) = (1-x)^2/12$; luego,

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{Var}\left(\frac{1-X}{2}\right) + \mathrm{E}\left\{\frac{(1-X)^2}{12}\right\}\\ &= \frac{1}{4}\mathrm{Var}(X) + \frac{1}{2}\mathrm{E}\{(1-X)^2\}.\end{aligned}$$

Tarea: Termine el ejemplo para $X \sim f_X(x) = 2(1-x)I_{(0,1)}(x)$.

Ejemplo 1.10

Sea N el número de personas por día que entra a un supermercado.

Sean $X_1, \ldots, X_N$ las cantidades gastadas por cada una de las N personas que ingreso al supermercado durante un determinado día. Suponga que N y $X_1, \ldots, X_N$ son variables aleatorias independientes.

Encuentre la media y la varianza del ingreso total del supermercado durante un día.

Sea $Y = \sum_{i=1}^{N} X_i$ el ingreso diario total del supermercado, donde es N una variable aleatoria con valores en los enteros positivos, mientras que los X_i son variables aletorias continuas con valores positivos. Asuma también que el gasto de cada persona que entra al supermacado durante un día tiene la misma distribución.

Distribuciones condicionales

i) Esperanza de Y. Se tiene que $\mathrm{E}(Y) = \mathrm{E}(\mathrm{E}(Y|N))$, donde $\mathrm{E}(Y|N=n) = \mathrm{E}(\sum_{i=1}^{N} X_i | N=n) = \mathrm{E}(\sum_{i=1}^{n} X_i) = n\mathrm{E}(X_1)$; luego,

$$\begin{aligned}\mathrm{E}(Y) &= \mathrm{E}(\mathrm{E}(Y|N)) \\ &= \mathrm{E}(N\mathrm{E}(X_1)) = \mathrm{E}(N)\mathrm{E}(X_1).\end{aligned}$$

ii) Varianza de Y. Sabemos que $\mathrm{Var}(Y) = \mathrm{E}[\mathrm{Var}(Y|N)] + \mathrm{Var}[\mathrm{E}(Y|N)]$. Ahora $\mathrm{Var}(Y|N=n) = \sum_{i=1}^{n} \mathrm{Var}(X_i) = n\mathrm{Var}(X_1)$; luego,

$$\begin{aligned}\mathrm{Var}(Y) &= \mathrm{E}\{\mathrm{Var}(Y|N)\} + \mathrm{Var}\{\mathrm{E}(Y|N)\} \\ &= \mathrm{E}(N)\mathrm{Var}(X_1) + \mathrm{Var}\{NE(X_1)\} \\ &= \mathrm{E}(N)\mathrm{Var}(X_1) + \mathrm{Var}(N)\{E(X_1)\}^2.\end{aligned}$$

Tarea: Concluya el ejemplo asumiendo que $N \sim P(\lambda)$ y $X_1 \sim U(0, \theta)$.

Considere dos variables aleatorias X e Y con fdp (o fmp) conjunta $f_{X,Y}(x,y)$.

Suponga que después de que se haya observado el valor de X, se debe predecir el valor de Y.

En otras palabras, el valor predicho de Y puede depender del valor de X.

Suponga que este valor predicho $h(X)$ debe elegirse de modo que minimice el error cuadrático medio $\mathrm{E}\{(Y-h(X))^2\}$.

Teorema 1.4

El predictor $h(X)$ que minimiza $\mathrm{E}\{(Y-h(X))^2\}$ es $h(X)=\mathrm{E}(Y|X)$.

Tarea: Obtenga $h(x)=\mathrm{E}(Y|X=x)$ cuando (X,Y) tiene distribucion normal bivariada $NB(\mu_X,\mu_Y,\sigma_X^2,\sigma_Y^2,\rho_{XY})$, con $|\rho_{XY}|<1$.