Notas[1]: Curso de Probabilidad I

Antonio Murillo Salas
Departamento de Matemáticas
Universidad de Guanajuato
amurillos@ugto.mx

14 de octubre de 2011

[1]Versión preliminar. No distribuirlas.

Índice general

Capítulo 1

Esperanza condicional

Esperanza condicional es una herramienta fundemental en la Teoría de Procesos Estocásticos. El propósito del presente capítulo es definir dicho concepto y estudiar algunas de sus propiedades más importantes. Trabajaremos en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ fijo, i.e., todas las variables aletorias estarán definidas en dicho espacio de probabilidad sin necesidad de hacer mención explícita de ello.

1.1. Definición de esperanza condicional

Sabemos que si X es una variable aleatoria discreta entonces $\mathbb{E}(X) = \sum_k x_k \mathbb{P}(X = k)$. Sea A un evento tal que $\mathbb{P}(A) \neq 0$, entonces podemos definir

$$\begin{aligned}
\mathbb{E}[X|A] &= \sum_k x_k \mathbb{P}(X = x_k|A) \\
&= \sum_k x_k \frac{\mathbb{P}(X = x_k, A)}{\mathbb{P}(A)} \\
&= \sum_k x_k \frac{\mathbb{P}(X\mathbf{1}_A = x_k)}{\mathbb{P}(A)} \\
&= \frac{1}{\mathbb{P}(A)} \mathbb{E}(X\mathbf{1}_A).
\end{aligned}$$

Lo anterior no da la esperanza condicional de la variable aleatoria X dado el evento A. Tal concepto se puede extender de la siguiente manera:

Definición 1.1.1 *Sean X y Y variables aleatorias discretas. La esperanza condicional de X dado que $Y = y$, donde $f_Y(y) > 0$, se define por*

$$\mathbb{E}(X|\{Y = y\}) = \sum_x x f_{X,Y}(x, y)/f_Y(y), \tag{1.1}$$

siempre y cuando la suma sea absolutamente convergente.

Notese que conforme y varia (sobre todos los posibles valores de Y) en la ecuación (1.1), se obtiene una función de Y, la cual denotarémos por $\mathbb{E}(X|Y)$. Entonces, $\mathbb{E}(X|Y)$ es una variable aleatoria tal que

$$\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|\{Y = y_n\}), \text{ si } Y(\omega) = y_n, \tag{1.2}$$

donde $y_1, y_2, \cdots$ son los posibles valores de Y. A la variable aleatoria $\mathbb{E}(X|Y)$ le llamarémos *esperanza condicional* de X dado Y.

Ejemplo 1.1.2 *Considere el lanzamiento de 3 monedas con denominación de 1, 5 y 10 pesos, respectivamente. Sea X la suma de las monedas que caen águila.*
(i) ¿Cual es el valor esperado de X dado que dos monedas caen águila?
(ii) Sea Y la suma de las monedas que caen águila, y que además, tienen denominación de 1 ó 5 pesos. ¿Cual es la esperanza condicional de X dado Y?
Solución: *(i) El espacio muestral está dado por*

$$\Omega = \{AAA, AAS, ASA, SAA, ASS, SAS, SSA, SSS\}.$$

Sea B el evento que dos monedas caen águila, i.e.,

$$B = \{AAS, ASA, SAA\}$$

Nos interesa determinar $\mathbb{E}(X|B)$. Notemos que, cada punto en B ocurre con probabilidad $1/8$. Luego,

$$\begin{aligned} X(AAS) &= 1 + 5 = 6, \\ X(ASA) &= 1 + 10 = 11, \\ X(SAA) &= 5 + 10 = 15. \end{aligned}$$

Por lo tanto,

$$\mathbb{E}(X|B) = \frac{1}{\frac{3}{8}}\left(6\frac{1}{8} + 11\frac{1}{8} + 15\frac{1}{8}\right) = \frac{32}{3}.$$

(ii) Ahora observamos que, $Y \in \{0, 1, 5, 6\}$ con probabilidades

$$\mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \mathbb{P}(Y = 5) = \mathbb{P}(Y = 6) = \frac{1}{4}.$$

Finalmente, siguiendo el mismo procedimiento que en (i) obtenemos

$$\begin{aligned} \mathbb{E}(X|\{Y = 0\}) = 5, \qquad \mathbb{E}(X|\{Y = 1\}) = 6, \\ \mathbb{E}(X|\{Y = 5\}) = 10, \qquad \mathbb{E}(X|\{Y = 6\}) = 11. \end{aligned}$$

Por lo tanto, la esperaza condicional de X dado Y resulta ser

$$\mathbb{E}(X|Y)(\omega) = \begin{cases} 5 & si\ Y(\omega) = 0, \\ 6 & si\ Y(\omega) = 1, \\ 10 & si\ Y(\omega) = 5, \\ 11 & si\ Y(\omega) = 6. \end{cases} \tag{1.3}$$

□

Notemos que en el ejemplo anterior $\mathbb{E}(X|Y)$ toma cada valor con la misma probabilidad, es decir, 1/4. Por lo tanto, $\mathbb{E}\left(\mathbb{E}(X|Y)\right) = 8 = \mathbb{E}(X)$. La propiedad anterior no es particular de este ejemplo. Más adelante veremos que tal propiedad se cumple en general.

Ejemplo 1.1.3 *Sean (X, Y) un vector aleatorio con función de probabilidad conjunta*

$$f(x,y) = \begin{cases} \frac{2}{N(N+1)}, & si\ x \leq y,\ x, y \in \{1, 2, \ldots, N\} \\ 0, & en\ otro\ caso, \end{cases}$$

donde N es un entero positivo. Encuentre (i) $\mathbb{E}(X|Y)$ y (ii) $\mathbb{E}(Y|X)$.
Solución: *(i) Notemos que*

$$f_Y(y) = \sum_{x=1}^{y} f(x,y) = \frac{2}{N(N+1)} y,\ \ y = 1, 2, \ldots, N.$$

Luego,

$$\mathbb{E}[X|Y=y] = \sum_{x=1}^{n} x \frac{f(x,y)}{f_Y(y)} = \frac{1}{y} \sum_{x=1}^{y} x = \frac{y+1}{2}.$$

Por lo tanto, $\mathbb{E}[X|Y] = \frac{1}{2}(Y+1)$. □

(ii) Porcediendo de manera análogo al inciso anterior se tiene que

$$f_X(x) = \sum_{y=1}^{N} f(x,y) = \sum_{y=x}^{N} \frac{2}{N(N+1)} = \frac{2}{N(N+1)}(N+1-x),\ \ x = 1, \ldots, N.$$

Luego, para $x \in \{1, \ldots, N\}$, se tiene que

$$\begin{aligned} E[Y|X=x] &= \sum_{y=1}^{N} y \frac{f(x,y)}{f_X(x)} \\ &= \frac{1}{N+1-x} \sum_{y=x}^{N} y \\ &= \frac{x+N}{2}. \end{aligned}$$

Por lo tanto, $\mathbb{E}[Y|X] = \frac{1}{2}(X+N)$. □

Teorema 1.1.4 *Sean X una variable aletoria discreta con esperanza finita y Y cualquier variable aleatoria discreta. Entonces,*
(i)

$$\mathbb{E}\left(\mathbb{E}(X|Y)\right) = \mathbb{E}(X), \tag{1.4}$$

siempre que ambos lados existan.
(ii) Para toda función $g : \mathbb{R} \to \mathbb{R}$ medible y acotada, se tiene

$$\mathbb{E}[g(Y)\mathbb{E}[X|Y]] = \mathbb{E}[g(Y)X].$$

Demostración: (i) Siempre que las sumatorias sean absolutamente convergentes se tiene que,

$$\begin{aligned}
\mathbb{E}\left(\mathbb{E}(X|Y)\right) &= \sum_y \mathbb{E}(X|\{Y=y\}) f_Y(y) \\
&= \sum_y \left(\sum_x \frac{x f_{X,Y}(x,y)}{f_Y(y)} \right) f_Y(y) \\
&= \sum_x x f_X(x) \\
&= \mathbb{E}(X).
\end{aligned}$$

(ii) Sea $g : \mathbb{R} \to \mathbb{R}$ cualquier función medible y acotada, entonces

$$\begin{aligned}
\mathbb{E}[g(Y)\mathbb{E}[X|Y]] &= \sum_k g(y_k)\mathbb{E}[X|Y=y_k]\mathbb{P}(Y=y_k) \\
&= \sum_k g(y_k) \left(\sum_j x_j \frac{\mathbb{P}(X=x_j, Y=y_k)}{\mathbb{P}(Y=y_k)} \right) \mathbb{P}(Y=y_k) \\
&= \sum_k g(y_k) \sum_j x_j \mathbb{P}(X=x_j, Y=y_k) \\
&= \sum_{k,j} g(y_k) x_j \mathbb{P}(X=x_j, Y=y_k) \\
&\equiv \mathbb{E}[g(Y)X].
\end{aligned}$$

$\square$

Observación 1.1.5 *La esperanza condicional $\mathbb{E}[X|Y]$ está bien definida. En efecto, se $h : \mathbb{R} \to \mathbb{R}$ una función medible tal que $h(Y)$ tiene esperanza finita y $\mathbb{E}[g(Y)h(Y)] = \mathbb{E}[g(Y)\mathbb{E}[X|Y]]$ para cualquier función g medible y acotada. Luego,*

$$\sum_k g(y_k)h(y_k)\mathbb{P}(Y=y_k) = \sum_{k,j} g(y_k) x_j \mathbb{P}(X=x_j, Y=y_k).$$

Ahora bien, la identidad anterior se cumple para todo g, en particular para $f = \mathbf{1}_{\{y_k\}}$, se tiene que

$$h(y_k)\mathbb{P}(Y=y_k) = \sum_j x_j \mathbb{P}(X=x_j, Y=y_k),$$

es decir,

$$h(y_k) = \mathbb{E}[X|Y=y_k], \textit{ para todo } k.$$

La obsevación anterior nos permite dar una defínción de esperanza condicional de una variable aleatoria dada otra variable aleatoria sin el supuesto de que estas sean discretas.

Definición 1.1.6 *(***Esperanza condicional de una variable aleatoria dada otra variable aleatoria***) Sea X una variable aleatoria con esperanza finita y Y cualquier variable aleatoria. Si existe una función medible $h : \mathbb{R} \to \overline{\mathbb{R}}$ tal que $h(Y)$ tiene media finita y*

$$\mathbb{E}[g(Y)h(Y)] = \mathbb{E}[g(Y)X],$$

para cualquier función medible $g : \mathbb{R} \to \mathbb{R}$ medible y acotada, entonces se dice que $h(Y)$ es una versión de esperanza condicional $\mathbb{E}[X|Y]$ y se define

$$\mathbb{E}[X|Y] = h(Y) \text{ y } \mathbb{E}[X|Y = y] = h(y), \ y \in \mathbb{R}.$$

Ejemplo 1.1.7 *Una gallina pone X huevos, donde X es Poisson con parámetro λ. Cada huevo es fecundado con probabilidad p, independientemente de los otros, produciendo así Y pollos. Demuestre que $\rho(X, Y) = \sqrt{p}$.*

Solución: *Observemos que, condicional en $X = k$, Y tiene distrubución binomial $Bin(k, p)$. Por lo tanto, $\mathbb{E}(Y|\{X = k\} = kp$. Más generalmente,*

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

Entonces, por el teorema anterior se tiene que

$$\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|X)) = \mathbb{E}(X\mathbb{E}(Y|X)) = \mathbb{E}(X^2p) = (\lambda^2 + \lambda)p.$$

De manera similar, obtenemos

$$\begin{aligned} \mathbb{E}(Y^2) &= \mathbb{E}(\mathbb{E}(Y^2|X)) \\ &= \mathbb{E}(Xp(1-p) + X^2p^2) \\ &= \lambda p(1-p) + (\lambda^2 + \lambda)p^2 \\ &= \lambda p + \lambda^2 p^2. \end{aligned}$$

Finalmente, tenemos que

$$\begin{aligned} \rho(X, Y) &= \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{(Var(X)\,Var(Y))^{1/2}} \\ &= \frac{(\lambda^2 + \lambda)p - \lambda \cdot \lambda p}{(\lambda(\lambda p + \lambda^2p^2 - \lambda^2p^2))^{1/2}} \\ &= \sqrt{p}. \end{aligned}$$

□

Ejemplo 1.1.8 *Sea (X_i) una sucesión de v.a. i.i.d., y sea Y una v.a. con valores en los enteros no negativos independiente de la sucesión (X_i). Defina $S_Y = \sum_{i=1}^{Y} X_i$. Demuestre que*

$$Var(S_Y) = \mathbb{E}(X_1^2)\,Var(Y) + \mathbb{E}(Y)\,Var(X_1).$$

Solución: *Por el ejercicio 2 de la tarea 2 tenemos que,*

$$\mathbb{E}(S_Y) = \mathbb{E}(X_1)\mathbb{E}(Y).$$

Ahora bien, por el Teorema 1.1.4 se obtiene

$$\begin{aligned}
\mathbb{E}(S_Y^2) &= \mathbb{E}(\mathbb{E}(S_Y^2|Y)) \\
&= \sum_{k=0}^{\infty} \mathbb{E}(S_k^2)p_k \;\; p_k = \mathbb{P}(Y=k) \\
&= \sum_{k=0}^{\infty} p_k \left(\sum_{i=1}^{k} \mathbb{E}(X_i^2) + \sum_{i=1}^{k} \sum_{j\neq i}^{k} \mathbb{E}(X_i X_j) \right) \\
&= \sum_{k=0}^{\infty} p_k \left(k\mathbb{E}(X_1^2) + k(k-1)[\mathbb{E}(X_1)]^2 \right) \\
&= \mathbb{E}(Y\mathbb{E}(X_1^2) + Y(Y-1)[\mathbb{E}(X_1)]^2) \\
&= \mathbb{E}(Y)\left(\mathbb{E}(X_1^2) - [\mathbb{E}(X_1)]^2\right) + \mathbb{E}(Y^2)[\mathbb{E}(X_1)]^2.
\end{aligned}$$

El resultado se sigue usando la identidad $Var(S_Y) = \mathbb{E}(S_Y^2) - [\mathbb{E}(S_Y)]^2$.

1.2. Propiedades de la esperanza condicional

Teorema 1.2.1 *Sean a y b constantes, g una función de valor real, y suponga que X, Y y Z son conjuntamente distribuidas. Entonces,*

1. $\mathbb{E}(a|Y) = a$.
2. $\mathbb{E}(aX + bZ|Y) = a\mathbb{E}(X|Y) + b\mathbb{E}(Z|Y)$.
3. $\mathbb{E}(X|Y) \geq 0$ *si* $X \geq 0$.
4. $\mathbb{E}(X|Y) = \mathbb{E}(X)$ *si* X *e* Y *son independientes.*
5. $\mathbb{E}(Xg(Y)|Y) = g(Y)\mathbb{E}(X|Y)$.
6. $\mathbb{E}(X|Y, g(Y)) = \mathbb{E}(X|Y)$.
7. $\mathbb{E}(\mathbb{E}(X|Y,Z)|Y) = \mathbb{E}(X|Y)$.

Demostración: (1) Sabemos que $f_{a,Y}(a,y) = f_Y(y)$. Entonces,

$$\mathbb{E}(a|Y) = a\frac{f_{a,X}(a,y)}{f_Y(y)} = a.$$

$\square$

(2) Tarea.

(3) Si $X \geq 0$, entonces cada sumando en la defición de esperanza condicional será no-negativo. Por lo tanto, $\mathbb{E}(X|Y) \geq 0$. $\square$

(4) Para cada y en el rango de Y, tenemos que

$$\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X|Y = y), \; Y(\omega) = y.$$

Luego, por definición de esperanza condicional se tiene que

$$\begin{aligned}
\mathbb{E}(X|Y = y) &= \sum_x x \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \sum_x x \frac{f_X(x) f_Y(y)}{f_Y(y)} \\
&= \mathbb{E}(X).
\end{aligned}$$

Entonces, $\mathbb{E}(X|Y)(\omega) = \mathbb{E}(X)$, para todo $\omega \in \Omega$. $\square$

(5) Notemos que, conjunto y perteneciente al rango de Y se tiene

$$\begin{aligned}
\mathbb{E}(Xg(Y)|\{Y = y\}) &= \frac{\sum_x x g(y) f_{X,Y}(x,y)}{f_Y(y)} \\
&= g(y) \frac{\sum_x x f_{X,Y}(x,y)}{f_Y(y)} \\
&= g(y)\mathbb{E}(X|\{Y = y\}.
\end{aligned}$$

Entonces, $\mathbb{E}(Xg(Y)|Y) = g(Y)\mathbb{E}(X|Y)$.

El resultado anterior se puede interpretar de la siguiente manera: esperanza condicional es una manera de "medir" la información que aporta Y sobre la variable aleatoria $Xg(Y)$. En consecuencia, al menos intuitivamente, se tiene que conociendo el valor de Y automáticamente conocemos el valor de $g(Y)$, y por lo tanto, puede tratarse como una constante dentro de la esperanza condicional. $\square$

(6) Tarea.

(7) Supongamos que $Y = y$ y que $Z = z$, entonces

$$\mathbb{E}(X|\{Y = y\}, \{Z = z\}) = \frac{\sum_x x f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)}.$$

Entonces, por definición tenemos que, para cada ω tal que $Y(\omega) = y$

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(X|Y,Z)|Y) &= \sum_z \mathbb{E}(X|Y=y, Z=z)\frac{f_{Y,Z}(y,z)}{f_Y(y)} \\
&= \sum_z \sum_x x \frac{f_{X,Y,Z}(x,y,z)}{f_{Y,Z}(y,z)} \frac{f_{Y,Z}(y,z)}{f_Y(y)} \\
&= \sum_x x \frac{\sum_z f_{X,Y,Z}(x,y,z)}{f_Y(y)} \\
&= \sum_x x \frac{f_{X,Y}(x,y)}{f_Y(y)} \\
&= \mathbb{E}(X|\{Y=y\}).
\end{aligned}$$

De lo anterior se sigue el resultado, ya que se cumple para cada y en el rango de Y. $\square$

Concluimos la presente sección con un resultado que nos dice que, $\mathbb{E}(X|Y)$ es la variable aletorias que se encuentra a una menor distancia de X de todas la variables aleatoria que se pueden determinar apartir de Y.

Teorema 1.2.2 *Sea h cualquier función de valor real tal que $\mathbb{E}(h(Y)^2) < \infty$. Entonces,*

$$\mathbb{E}\left[(X - h(Y))^2\right] \geq \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2\right]. \tag{1.5}$$

Más aún, si h es tal que

$$\mathbb{E}\left[(X - h(Y))^2\right] = \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2\right],$$

entonces

$$\mathbb{E}\left[(h(Y) - \mathbb{E}(X|Y))^2\right] = 0.$$

Demostración:

$$\begin{aligned}
\mathbb{E}\left[(X - h(Y))^2\right] &= \mathbb{E}\left[(X - \mathbb{E}(X|Y) + \mathbb{E}(X|Y) - h(Y))^2\right] \\
&= \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2\right] + \mathbb{E}\left[(\mathbb{E}(X|Y) - h(Y))^2\right] \\
&\quad +2\mathbb{E}\left[(X - \mathbb{E}(X|Y))(\mathbb{E}(X|Y) - h(Y))\right].
\end{aligned}$$

Ahora bien, por Teorema 1.1.4 tenemos que

$$\begin{aligned}
\mathbb{E}\left[(X - \mathbb{E}(X|Y))(\mathbb{E}(X|Y) - h(Y))\right] &= \mathbb{E}\left[\mathbb{E}\left((X - \mathbb{E}(X|Y))(\mathbb{E}(X|Y) - h(Y))\right)|Y\right] \\
&= \mathbb{E}\left[(\mathbb{E}(X|Y) - h(Y))\right]\mathbb{E}\left[(X - \mathbb{E}(X|Y))|Y\right],
\end{aligned}$$

donde la segunda igualdad es debido al Teorema 1.2.1 (5). Luego, obsevemos que $\mathbb{E}\left[(X - \mathbb{E}(X|Y))|Y\right] = 0$. Entonces,

$$\mathbb{E}\left[(X - h(Y))^2\right] = \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2\right] + \mathbb{E}\left[(\mathbb{E}(X|Y) - h(Y))^2\right]. \tag{1.6}$$

Para terminar la prueba note que $\mathbb{E}\left[(\mathbb{E}(X|Y) - h(Y))^2\right] \geq 0$, y por lo tanto, obtenemos (1.5).

Por último, si $\mathbb{E}\left[(X - h(Y))^2\right] = \mathbb{E}\left[(X - \mathbb{E}(X|Y))^2\right]$ de (1.6) se concluye que $\mathbb{E}\left[(\mathbb{E}(X|Y) - h(Y))^2\right] = 0$, es decir, $h(Y) = \mathbb{E}(X|Y)$ salvo en un conjunto de probabilidad cero. $\square$

Ejemplo 1.2.3 *En una fiesta n de los asistentes se quita el sombrero. Se ponen los n sombreros en un contenedor, y cada persona selecciona uno al azar. Decimos que ocurre un "coincidencia" si una persona selecciona su propio sombrero. ¿Cuál es la probabilidad de que no ocurra ninguna coincidencia? ¿Cuál es la probabilidad de que ocurran exactamente k coincidencias?*
Solución: *Sea E el evento que no ocurra ninguna coincidencia, sea $p_n := \mathbb{P}(E)$. Definamos el evento $M :=$"la primera persona selecciona su propio sombrero". Entonces,*

$$p_n = \mathbb{P}(E) = \mathbb{P}(E|M)\mathbb{P}(M) + \mathbb{P}(E|M^c)\mathbb{P}(M^c).$$

Notemos que, $\mathbb{P}(E|M) = 0$, (al menos ocurre una coincidencia). Luego,

$$p_n = \mathbb{P}(E|M^c)\frac{n-1}{n}. \tag{1.7}$$

Ahora bien, note que $\mathbb{P}(E|M^c)$ es la probabilidad de que no ocurra ninguna coincidencia cuando $n-1$ personas seleccionan de $n-1$ sombreros, y que además, hay una persona cuyo sombrero no está dentro de los $n-1$. Lo anterior puede pasar de dos formas mutuamente excluyentes: (1) no ocurre ninguna coincidencia y la persona extra no selecciona el sombrero extra (el sombrero perteneciente a la primera persona en seleccionar); ó (2) no ocurre ninguna coincidencia y la persona extra selecciona el sombrero extra. La probabilidad de (1) es exactamente p_{n-1}, lo anterior es considerando que el sombrero extra pertenece a la persona extra. Por otro lado, la probabilidad de (2) está dada por $\frac{1}{n-1}p_{n-2}$. Entonces,

$$\mathbb{P}(E|M^c) = p_{n-1} + \frac{1}{n-1}p_{n-2}.$$

Por lo tanto, combinando la ecuación anterior con (1.7) tenemos,

$$p_n = \frac{n-1}{n}p_{n-1} + \frac{1}{n}p_{n-2},$$

equivalentemente,

$$p_n - p_{n-1} = -\frac{1}{n}(p_{n-1} - p_{n-2}). \tag{1.8}$$

Ahora bien, dado que p_n es la probabilidad de que no ocurra ninguna coincidencia cuando n personas seleccionan un sombrero, tenemos

$$p_1 = 0,\ p_2 = \frac{1}{2},$$

y por lo tanto, de (1.8) resulta que

$$p_3 - p_2 = -\frac{p_2 - p_1}{3} = -\frac{1}{3!},$$

es decir,

$$p_3 = \frac{1}{2!} - \frac{1}{3!}.$$

De manera análoga, obtenemos

$$p_4 - p_3 = -\frac{p_3 - p_2}{4} = \frac{1}{4!},$$

es decir,

$$p_4 = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}.$$

Procediendo de manera similar, obtenemos que

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^n}{n!}.$$

Para responder la segunda pregunta considere un grupo fijo de k personas. La probabilidad que ellos, y solamente ellos, seleccionen sus propios sombreros está dada por

$$\frac{1}{n}\frac{1}{n-1}\cdots\frac{1}{n-(k-1)}p_{n-k} = \frac{(n-k)!}{n!}p_{n-k},$$

donde $p_n - k$ es la probabilidad que las $n-k$ personas restantes, que seleccionan dentro de sus propios sombreros, no haya ninguna coincidencia. Ahora bien, dado que hay exactamente $\binom{n}{k}$ formas diferentes de seleccionar un grupo de k personas, la probabilidad de que haya exactamente k coincidencias está dada por

$$\begin{aligned} \binom{n}{k}\frac{(n-k)!}{n!}p_{n-k} &= \frac{p_{n-k}}{k!} \\ &= \frac{\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots + \frac{(-1)^{n-k}}{(n-k)!}}{k!}. \end{aligned}$$

Por lo tanto, para n suficientemente grande, se tiene que la probabilidad de que haya exactamente k coincidencias es aproximadamente $\frac{e^{-1}}{k!}$. □

Caso absolutamente continuo.

Hasta ahora, en los ejemplos, hemos puesto mucho enfásis en caso de variables (vectores) aleatorias discretas. Sin embargo, hemos dado la definición general de esperanza condicional de una variable aleatoria dado otra variable aleatoria. En el caso absolutamente continuo se tiene lo siguiente: sea (X, Y) un vector aleatorio con función de densidad conjunta $f_{X,Y}$ tal que $\mathbb{E}(X) < \infty$. Entonces, una función de densidad para la esperanza condicional $\mathbb{E}(X|Y)$ esta dada por

$$\frac{f_{X,Y}(x,y)}{f_Y(y)}, \text{ siempre que } f_Y(y) > 0.$$

Comunmente se usa la notación

$$f_{X|Y}(x|y) \equiv \frac{f_{X,Y}(x,y)}{f_Y(y)}.$$

1.3. Más ejemplos

Ejemplo 1.3.1 *Supongamos que el número de accidentes que tiene una persona en un año tiene distribución Poisson con parámetro Y, de modo que, para cada $y > 0$, el porcentaje de personas para las cuales $Y > y$ es igual a $\lambda e^{-\lambda y}$, donde λ es una constante positiva. Si X es el número de accidentes en un año de una persona seleccionada al azar, encuentre i) la distribución de X y $\mathbb{E}(X)$, ii) la distribución condicional de Y dado que $X = x$, para $x \in \{0, 1, \dots\}$ y iii) $\mathbb{E}(Y|X)$.*

Solución: i) Sabemos que $X|Y$ tiene distribución Poisson de parámetro Y. Entonces, por Ley de Probabilidad Total, para cada $x \in \{0, 1, \dots\}$, se tiene que

$$\begin{aligned}\mathbb{P}(X = x) &= \int_0^\infty \mathbb{P}(X = x|Y = y) f_Y(y)\, dy \\ &= \int_0^\infty \frac{e^{-y} y^x}{x!} \lambda e^{-\lambda y}\, dy \\ &= \frac{\lambda}{x!} \int_0^\infty y^x e^{-(\lambda+1)y}\, dy,\end{aligned}$$

vamos a completar la integral anterior para que sea una Gama$(x, \lambda + 1)$, entonces

$$\begin{aligned}\mathbb{P}(X = x) &= \frac{\lambda}{(\lambda+1)^{x+1}} \int_0^\infty \frac{(\lambda+1)^{x+1} y^x e^{-(\lambda+1)y}}{x!}\, dy \\ &= \frac{\lambda}{\lambda+1} \left(\frac{1}{\lambda+1}\right)^x.\end{aligned}$$

Por lo tanto, X tiene distribución geométrica de parámetro $p = \lambda/(\lambda + 1)$, puesto que

$$\mathbb{P}(X = x) = p(1 - p)^x,\ x = 0, 1, \dots.$$

Entonces,

$$\mathbb{E}(X) = \frac{1-p}{p} = \frac{1}{\lambda} \equiv \int_0^\infty \mathbb{E}(X|Y = y) \lambda e^{-\lambda y}\, dy = \int_0^\infty y \lambda e^{-\lambda y}\, dy.$$

ii) Para $x \in \{0, 1, \dots\}$ y $y > 0$ se tiene que

$$f_{Y|X}(x|y) = \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{f_{X|Y}(x|y) f_Y(y)}{f_X(x)},$$

notemos que no conocemos $f_{X,Y}$. Sin embargo, en el último término si conocemos todos los factores. Por lo tanto,

$$f_{Y|X}(x|y) = \frac{\frac{y^x e^{-y}}{x!} \lambda e^{-\lambda y}}{\frac{\lambda}{\lambda+1}\left(\frac{1}{\lambda+1}\right)^x} = \frac{(\lambda+1)^{x+1} y^x e^{-(\lambda+1)y}}{x!}.$$

iii) De la parte ii) observamos que $Y|X$ tiene distribución gama con parámetros $x + 1$ y $\lambda + 1$. Luego, recordando que si $Z \sim$ Gamma(α, β), entonces $\mathbb{E}(Z) = \alpha/\beta$. Por lo tanto,

$$\mathbb{E}(Y|X) = \frac{X+1}{\lambda+1}.$$

□

Ejemplo 1.3.2 *Sean X y Y dos variables aleatorias con esperanza finita tales que $\mathbb{E}(X|Y) = \mathbb{E}(Y)$. Supongamos que XY también tiene esperanza finita, demuestre que $Cov(X, Y) = 0$.*

Solución: Sabemos que

$$\begin{aligned}
\mathrm{Cov}(X,Y) &\equiv \mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)\\
&= \mathbb{E}[\mathbb{E}(XY|X)]-\mathbb{E}(X)\mathbb{E}(Y),\ \text{(por Teorema 1.1.4) (i))}\\
&= \mathbb{E}[X\mathbb{E}(Y|X)]-\mathbb{E}(X)\mathbb{E}(Y),\ \text{(por Teorema 1.2.1 5)}\\
&= \mathbb{E}[X\mathbb{E}(Y)]-\mathbb{E}(X)\mathbb{E}(Y)\\
&= 0.
\end{aligned}$$

□

Ejemplo 1.3.3 *Supongamos que el número de personas que suben a un elevador, en la planta baja de un edificio de N pisos, tiene distribución Poisson con parámetro λ. Supongamos que cada persona deja el elevador al azar, en cualquiera de los N pisos, independientemente de donde bajen los demás. Encuentre el número esperado de paradas que hace el elevador hasta que bajan todas las personas.*

Solución: sea Y en número de personas que sube al elevador en la planta baja y X en número de paradad necesarias para que el elevador quede vacio. Definamos las variables aleatorias $X_1, X_2, \ldots, X_N$ como sigue:

$$X_i = \begin{cases} 1, & \text{si el elevador para en el piso } i,\\ 0, & \text{en otro caso.}\end{cases}$$

Entonces, $X = \sum_{i=1}^{N} X_i$, y para cada $k \in \{0, 1, \ldots, \}$ se tiene que

$$\mathbb{P}(X_1 = 0|Y = k) = \mathbb{P}(\text{ninguna persona baja en el piso } i) = (1 - \frac{1}{N})^k,$$

es decir,

$$\mathbb{E}(X_i|Y = k) = 1 - (1 - \frac{1}{N})^k.$$

Luego,

$$\mathbb{E}(X|Y = k) = N\left[1 - (1 - 1/N)^k\right].$$

Por lo tanto, por la Ley de Probabilidad Total,

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=0}^{\infty} \mathbb{E}(X|Y=k)\mathbb{P}(Y=k)\\
&= \sum_{k=0}^{\infty} N\left[1-(1-1/N)^k\right]\frac{\lambda^k e^{-\lambda}}{k!}\\
&= N\left(1-e^{-\lambda/N}\right).
\end{aligned}$$

□

1.4. Ejercicios

Los ejercicios marcados con * son para entregar.

1. *Sean X y Y dos variables aleatorias discretas, y g una función de real-valuada. Demuestre lo siguiente
$$\mathbb{E}(X|Y, g(Y)) = \mathbb{E}(X|Y).$$

2. *Sea (X, Y) un vector aleatorio con función de densidad conjunta dada por
$$f_{X,Y}(x, y) = \begin{cases} \frac{6}{N(N^2-1)}(y - x), & \text{si } x < y \text{ y } x, y \in \{1, \ldots, N\}, \\ 0, & \text{en otro caso.} \end{cases}$$
Encuentre la función de densidad condicional de: (i) X dado Y; (ii) Y dado X.

3. *Sea X y Y variables aleatorias independientes con distribución Poisson. Sea $Z := X + Y$. Encuentre la distribución condicional de X dado Z.

4. *Una moneda muestra águila con probabilidad p. Sea X_n el número de lanzamientos necesarios para obtener un corrida de n águilas consecutivas. Demuestre que
$$\mathbb{E}(X_n) = \sum_{k=1}^{n} p^{-k}.$$
Sugerencia: recuerde que $\mathbb{E}(X_n) = \mathbb{E}[\mathbb{E}(X_n|X_{n-1})]$.

5. Sea $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas, con esperanza finita. Demuestre que, para cualquier $k \in \{1, \ldots, n\}$, se cumple
$$\mathbb{E}\left(X_k \middle| \sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

6. *Se eligen, al azar y sin reemplazo, dos tarjetas de una urna que contiene N tarjetas numeradas del 1 al N, con $N \geq 1$. Sean X y Y el menor y mayor, respectivamente, de los números en la tarjetas seleccionadas. Encuentre $\mathbb{E}(X|Y)$ y $\mathbb{E}(Y|X)$.

7. Supongamos que (X_i) es una suseción de variables aleatorias idependientes tales que,
$$X_i|P_i \sim \text{Ber}(P_i),\ i = 1, 2, \ldots, n,$$
donde $P_i \sim \text{Beta}(\alpha, \beta)$. Defina $Y_n = \sum_{i=1}^{n} X_i$. Encuentre $\mathbb{E}(Y_n)$ y $\text{Var}(Y_n)$.

8. *Dos jugadores A y B tienen n monedas. Se las reparten de la siguiente manera: lanzan cada moneda y A obtine las que resultan "águila", digamos X, entonces B obtine las restantes $n - X$ monedas.

   Luego, A y B juegan volados independientes y justos, cada vez que A gana (la moneda cae águila) B le da una moneda al jugador A; y cada vez que pierde le da una moneda a B. El juego termina cuando uno de ellos se queda sin monedas.

   Sea D_X el número de volados jugados. Encuentre $\mathbb{E}(D_X)$, y demuestre que $\rho(X, D_X) = 0$.

9. Sean X y Y dos variables aleatorias independientes, ambas con distribución uniforme en el conjunto $\{1, \ldots, N\}$. Encuentre $\mathbb{E}(X|Y-X)$ y $\mathbb{E}(Y|Y-X)$.