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,

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

La obsevación anterior nos permite dar una defineión 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

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

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. □

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. □

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 $x \in \{0, 1, \ldots\}$ 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,

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

$\square$

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

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

□

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,
$$X_i|P_i \sim \mathrm{Ber}(P_i), \; i = 1, 2, \ldots, n,$$
donde $P_i \sim \mathrm{Beta}(\alpha, \beta)$. Defina $Y_n = \sum_{i=1}^{n} X_i$. Encuentre $\mathbb{E}(Y_n)$ y $\mathrm{Var}(Y_n)$.

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

1.5. Caminatas aleatorias simples

El esta sección veremos nuestro primer ejemplo de proceso estocástico llamado *caminata aleatoria*. La camina aleatoria es un proceso simple de describir. Sin embargo, eso no quiere decir que sea sencillo estudiarla. Por otro lado, nos permite presentar algunos problemas que son de interés en procesos mucho más complicados. Comencemos con la definición de proceso estocástico.

Definición 1.5.1 *Un proceso estocástico es una colección de variables aleatorias,* $\{Z_t, t \in T\}$*, definidas sobre un mismo espacio de probabilidad* $(\Omega, \mathcal{F}, \mathbb{P})$*, donde* T *es un conjunto de indices.*

Para propósitos del presente curso, $T = \mathbb{Z}_+$ (ó $\mathbb{R}_+$) y todas las variables aleatorias Z_t, $t \in T$, toman valores en $\mathbb{Z}$.

Antes de definir la caminata aleatoria conviene recordar la definición de independencia de una colección de variables aleatorias. Se dice que $(X_n, n \geq 1)$ es una sucesión de variables indendientes, si para cada n-éada de enteros $(k_1, k_2, \cdots, k_n)$ distintos se cumple que las variables aleatorias $X_{k_1}, X_{k_2}, \cdots, X_{k_n}$ son indenpendientes.

Ahora ya tenemos todos los elementos necesarios para definir nuestro proceso estocástico de interés:

Definición 1.5.2 *Sea* $(X_n, n \geq 1)$ *una sucesión de variables aleatorias i.i.d. con distribución común dada por*

$$\mathbb{P}(X_1 = 1) = 1 - \mathbb{P}(X_1 = -1) = 1 - q = p.$$

La sucesión $(S_n, n \geq 0)$*, donde*

$$S_n = S_0 + \sum_{i=1}^{n} X_i,$$

es llamada caminata aleatoria simple. En general, S_0 *puede ser constante o una variable aleatoria, se dice que la caminata inicia en* S_0*. Si* $p = q = \frac{1}{2}$ *es llamada caminata aleatoria simple simétrica.*

Las caminatas aleatorias son útiles para modelar varios fenómenos: podemos usarlo para modelar la posición de una partícula que se mueve en los enteros, a cada paso la partícula puede avanzar o retroceder por un paso con probabilidad p y $1-p$, respectivamente. Además, la dirección (subir o bajar) es independiente de los pasos anteriores. Asimismo pueden servir para modelar un juego de apuestas donde en cada jugada se pierde o se gana una unidad.

Las caminatas aleatorias simples se gráfican en el plano cartesiano con los puntos $(n, S_n)_{n=0}^{\infty}$ uniendo los puntos vecinos con lineas rectas con pendiente 1 ó -1. A la gráfica resultante se le llama *trayectoria o realización*, y dado que es trata de una sucesión de variables aleatorias, para $\omega \in \Omega$ se tiene una trayectoria o realización.

1.5.1. Propiedades de las caminatas aleatorias simples

Lema 1.5.3 *Toda caminata aleatoria simple* $\{S_n, n \geq 0\}$*, con* $S_0 = a$*, poseé las siguientes propiedades:*

(i) Homogeneidad espacial:

$$\mathbb{P}(S_n = j | S_0 = a) = \mathbb{P}(S_n = j + b | S_0 = a + b).$$

(ii) Homogeneidad temporal, para todo $n, m \in \mathbb{Z}$:

$$\mathbb{P}(S_n = j | S_0 = a) = \mathbb{P}(S_{n+m} = j | S_m = a).$$

(iii) Propiedad de Markov, para todo $n, m \in \mathbb{Z}$,

$$\mathbb{P}(S_{n+m} = j | S_0, S_1, \cdots, S_n) = \mathbb{P}(S_{n+m} = j | S_n). \tag{1.9}$$

Demostración: (i) Veamos el lado izquierdo

$$\mathbb{P}(S_n = j | S_0 = a) = \mathbb{P}\left(\sum_{i=1}^{n} X_i = j - a\right).$$

Análogamente, el lado derecho satisface

$$\mathbb{P}\left(\sum_{i=1}^{n} X_i = j - a\right) = \mathbb{P}(S_n = j + b, S_0 = a + b) = \mathbb{P}\left(\sum_{i=1}^{n} X_i = j + b - (a + b)\right).$$

□

(ii) Procederemos como en (i). El lado derecho es igual a

$$\begin{aligned}
&\frac{\mathbb{P}\left(S_0 + \sum_{i=1}^{n+m} X_i = j, S_0 + \sum_{i=1}^{m} X_i = a\right)}{\mathbb{P}\left(S_0 + \sum_{i=1}^{m} X_i = a\right)} \\
&= \frac{\mathbb{P}\left(\sum_{i=m+1}^{n+m} X_i = j - a, S_0 + \sum_{i=1}^{m} X_i = j\right)}{\mathbb{P}\left(S_0 + \sum_{i=1}^{m} X_i = j\right)} \\
&= \mathbb{P}\left(\sum_{i=m+1}^{n+m} X_i = j - a\right) \text{ (independencia)} \\
&= \mathbb{P}\left(\sum_{i=1}^{n} X_i = j - a\right),
\end{aligned}$$

la última igualdad es debido al hecho que el vector $(X_1, \cdots, X_n)$ tiene la misma distribución que el vector $(X_{m+1}, X_{m+1}, \cdots, X_{n+m})$. Un cálculo similar, pero más simple, demuestra la igualdad deseada. □

(iii) Sean $s_0, s_1, \cdots, s_n$ enteros tales que

$$\begin{aligned}
&\mathbb{P}(S_{n+m} = j | S_0 = s_0, S_1 = s_1, \cdots, S_n = s_n) \\
&= \frac{\mathbb{P}(S_0 = s_0, S_1 = s_1, \cdots, S_n = s_n, S_{n+m} = s_{n+m})}{\mathbb{P}(S_0 = s_0, S_1 = s_1, \cdots, S_n = s_n)} \\
&= \frac{\mathbb{P}(S_0 = s_0, X_1 = s_1 - s_0, X_2 = s_2 - s_1, \cdots, X_n = s_n - s_{n-1}, \sum_{i=n+1}^{n+m} X_i = s_{n+m} - s_n)}{\mathbb{P}(S_0 = s_0, X_1 = s_1 - s_0, X_2 = s_2 - s_1, \cdots, X_{n-1} = s_{n-1} - s_{n-2}, X_n = s_n - s_{n-1})} \\
&= \mathbb{P}(\sum_{i=n+1}^{n+m} X_i = s_{n+m} - s_n).
\end{aligned}$$

Por otro lado, tenemos que

$$\begin{aligned}\mathbb{P}(S_{n+m} = s_{n+m}|S_n = s_n) &= \frac{\mathbb{P}(S_{n+m} = s_{n+m}, S_n = s_n)}{\mathbb{P}(S_n = s_n)} \\ &= \mathbb{P}(\sum_{i=n+1}^{n+m} X_i = s_{n+m} - s_n).\end{aligned}$$

□

Observación 1.5.4 *(i) Cualquier proceso estocástico, $\{Z_n, n \geq 0\}$, que cumpla la propiedad (1.9) es llamdado cadena de Markov a tiempo discreto, es decir,*

$$\mathbb{P}(Z_{n+m} = j|Z_0, Z_1, \cdots, Z_n) = \mathbb{P}(Z_{n+m} = j|Z_n).$$

(ii) Se dirá que la probabilidad $\mathbb{P}(S_n = j|S_0 = a)$ es la probabilidad de transición del estado a al estado j en n pasos.

En el siguiente resultado calcularemos las probabilidades de transición para la caminata aleatoria simple.

Lema 1.5.5 *Para todo $a, b \in \mathbb{Z}$ y $n \geq 0$, se tiene que*

$$\mathbb{P}(S_n = b|S_0 = a) = \begin{cases} \binom{n}{(n+b-a)/2} p^{(n+b-a)/2} q^{(n-b+a)/2} & si\ (n+b-a)/2 \in \mathbb{Z}, \\ 0 & en\ otro\ caso. \end{cases}$$

Demostración: Se tien que, una realización que lleva del punto $(0, a)$ al punto $(0, b)$ en n pasos tiene r pasos hacia arriba $(+1)$ y l pasos hacia abajo (-1), donde r, l son tales que $l + r = n$ y $r - l = b - a$. Lo anterior es debido a que, $S_n = r(+1) + l(-1) = b - a$. Resolviendo las ecuaciones antariores obtenemos que,

$$r = \frac{n+b-a}{2} \text{ y } l = \frac{n-b+a}{2}.$$

Ahora bien, cada realización que lleva de a a b en n pasos tiene probabilidad $p^r q^l$, y hay $\binom{n}{(n+b-a)/2}$ realizaciones posibles. Por lo que el resultado se sigue. □

Note que, la prueba del resultado anterior se basa en el conteo de trayectorias, i.e., "casos favorables"/"casos posibles". Esta es una propiedad de muy interesante y que ha llamado la atención no sólo de la comunidad probabilista sino que también es explotada en teoría combinatoria, teoría de juegos, entre otras.

En lo que sigue procederemos a calcular probabilidades asociadas a la caminata aleatoria simple mediante las herramientas estudiadas hasta el momento. A saber, por medio de condicionamientos.

Definición 1.5.6 *Para cada $j \in \mathbb{Z}$, el primer tiempo de llegada al estado j se define por*

$$T_j = \min\{n \geq 0 : S_n = j\}.$$

Proposición 1.5.7 *Para cada $j \in \mathbb{Z}$, sea h_j la probabilidad de que una caminata aleatoria que parte del estado j llegue al estado 0 antes de llegar al estado N, i.e., $h_j = \mathbb{P}(T_0 < T_n | S_0 = j)$. Entonces,*

$$h_j = \begin{cases} \frac{(\frac{q}{p})^j - (\frac{q}{p})^N}{1-(\frac{q}{p})^N} & p \neq q, \\ 1 - \frac{j}{N} & p = \frac{1}{2}. \end{cases}$$

Demostración: Condicionando en la primera transición obtenemos la ecuación,

$$h_j = p h_{j+1} + q h_{j-1},$$

para $j \in \{1, 2, \cdots, N-1\}$. Además, notemos que $h_0 = 1$ y $h_N = 0$. Reescribiendo la encuación anterior se obtiene

$$h_n = p h_{n+1} + q h_{n-1} \Longleftrightarrow q(h_{n+1} - h_n) = p(h_{n+1} - h_n), \; n \geq 1. \tag{1.10}$$

El caso simétrico: $p = q = 1/2$. En este caso, se tiene la ecuación

$$h_n - h_{n-1} = h_{n+1} - h_n, \;\; n \geq 1.$$

Por lo tanto, la recta h_n tiene una pendiente constante $c := h_{n+1} - h_n$, en consecuencia

$$h_n = 1 + \sum_{j=1}^{n} (h_j - h_{j-1}) = 1 + nc, \; 1 \leq n \leq N.$$

Ahora bien, recordando que $h_N = 1$ obtemos que $c = -1/N$, i.e., $h_n = 1 - n/N$.
El caso general: $p \neq q$. Definamos la sucesión $(x_n, n \geq 0)$ como sigue, $x_0 \in \mathbb{R}$ (se determinará más adelante) y $x_n = h_n - h_{n-1}$, para $1 \leq n \leq N$. De la ecuación en el lado derecho de (1.10) obtenemos que la sucesión $(x_n, n \geq 0)$ satisface la relación $x_{n+1} = \frac{q}{p} x_n$, para $1 \leq n \leq N$. Por lo tanto,

$$x_{n+1} = \left(\frac{q}{p}\right)^n x_0, \; 0 \leq n \leq N. \tag{1.11}$$

Luego, dado que

$$h_n = h_0 + \sum_{j=1}^{n} (h_j - h_{j-1}),$$

la ecuación (1.11) implica

$$\begin{aligned} h_n &= h_0 + x_0 \sum_{j=1}^{n} \left(\frac{q}{p}\right)^j \\ &= h_0 + x_0 \left(\frac{q}{p}\right) \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \left(\frac{q}{p}\right)}. \end{aligned} \tag{1.12}$$

Haciendo uso del hecho que, $h_N = 0$ y $h_0 = 1$, obtenemos que

$$0 = 1 + x_0 \left(\frac{q}{p}\right) \frac{1 - \left(\frac{q}{p}\right)^n}{1 - \left(\frac{q}{p}\right)},$$

de donde se sigue que

$$x_0 = -\frac{p}{q} \frac{1 - \frac{q}{p}}{1 - \left(\frac{q}{p}\right)^N}.$$

Finalmente, de (1.12) se concluye que

$$h_n = \frac{\left(\frac{q}{p}\right)^n - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N}.$$

□

Corolario 1.5.8 *Para cada $j \in \mathbb{N}$ se tiene que*

$$\mathbb{P}(T_0 < \infty | S_0 = j) = \begin{cases} 1 & si\ q \leq p, \\ \left(\frac{q}{p}\right)^j & si\ q > p. \end{cases}$$

Demostración: Para cada n, sea $A_n := \{T_0 < T_n\}$. Notemos que $A_n \subset A_{n+1}$, dado que $T_n \leq T_{n+1}$, para cada n. Además, observemos que

$$\{T_0 < \infty\} = \cup_{n=1}^{\infty} \{T_0 < T_n\}.$$

Por lo tanto, dado que (A_n) es una sucesión creciente, la continuidad de la medida de probababilidad implica

$$\lim_{n \to \infty} \mathbb{P}(A_n | S_0 = j) = \mathbb{P}(T_0 < \infty | S_0 = j).$$

Luego, el resultado se sigue de la proposición anterior. □

Sean $a, b \in \mathbb{Z}$ y $n \in \mathbb{N}$. Sea $N_{n(a,b)}$ el número de trayectorias que van de a a b en n pasos y $N^0_{n(a,b)}$ las trayectorias que unen a y b en n pasos; y que además, pasan por 0 al menos una vez.

Teorema 1.5.9 *(Principio de Reflexión) Para cada $a, b \in \mathbb{N}$ se tiene que*

$$N^0_{n(a,b)} = N_{n(-a,b)}.$$

Demostración: Haciendo bosquejo podemos ver que cada trayectoria que lleva de $(0, a)$ a (b, n) cruza el eje x por lo menos una vez, sea $(k, 0)$ el punto donde esto ocurre por primera vez. Reflejando el segmento de la trayectoria anterior se obtiene una trayectoria de $(0, a)$ a (b, n) y que pasa por el eje x por lo menos una vez. Luego, haciendo lo mismo en el sentido opuesto obtenemos el resultado. □

Lema 1.5.10 *Para todo $a, b \in \mathbb{Z}$, se cumple que*

$$N_{n(a,b)} = \binom{n}{\frac{1}{2}(n+b-a)}.$$

Veamos el siguiente resultado importante, el cual es una consecuencia del lema anterior.

Teorema 1.5.11 *(Teorema de las votaciones (Ballot Theorem) Sea $b \in \mathbb{N}$, entonces el número de realizaciones que van de $(0,0)$ a (n,b) y que no visitan al eje x después del primer paso está dado por*

$$\frac{b}{n} N_{n(0,b)}.$$

Demostración: Notemos que, las trayectorias que nos interesa contar en el primer paso se encuentran en $(1,1)$. Por lo tanto, en número de trayectorias de interés está dado por

$$\begin{aligned}
N_{n-1(1,b)} - N^0_{n-1(1,b)} &= N_{n-1(1,b)} - N_{n-1}(-1,b) \\
&= \frac{(n-1)!}{\left(\frac{n-b}{2}\right)!\left(\frac{n+b-2}{2}\right)!} - \frac{(n-1)!}{\left(\frac{n-b-2}{2}\right)!\left(\frac{n+b}{2}\right)!} \\
&= \frac{(n-1)!}{\left(\frac{n-b}{2}\right)!\left(\frac{n+b}{2}\right)!}\left(\frac{n+b}{2} - \frac{n-b}{2}\right) \\
&= \frac{b}{n} N_{n(0,b)}.
\end{aligned}$$

$\square$

Veamos ahora porque el resultado anterior se llama Teorema de las votaciones. Sunpongamos que tenemos dos candidatos A y B, y que A obtiene a votos y B obtiene b votos, donde $a > b$. Cual es la probabilidad de que A tenga la ventaja durante toda la votación?

Supongamos que $X_i = 1$ si el i-ésimo individuo vota por el candidato A y vale -1 si vota el canditato B. Supongamos que cualquier combinación de votos es igualmente probable, i.e., cada una tiene probabilidad $\binom{\alpha+\beta}{\alpha}$. La trayectoria que deben seguir las votaciones para que A tenga las preferencias durante toda la jornada de votaciones va del punto $(0,0)$ al punto $(\alpha+\beta, \alpha-\beta)$. Por lo tanto, por el Teorema 1.5.11 está dada por

$$\frac{\alpha-\beta}{\alpha+\beta} N_{\alpha+\beta(0,\alpha-\beta)} \frac{1}{\binom{\alpha+\beta}{\alpha}} = \frac{\alpha-\beta}{\alpha+\beta}.$$

El siguiente resultado es una aplicación del Principio de Reflexión (Teorema 1.5.9)

Teorema 1.5.12 *Supongamos que $S_0 = 0$, entonces para todo $n \geq 0$ se cumple que*

$$\mathbb{P}(S_1 S_2 \cdots S_n \neq 0, S_n = b) = \frac{|b|}{n} \mathbb{P}(S_n = b) \tag{1.13}$$

Demostración: Supongamos que $S_0 = 0$ y $S_n = b > 0$. Notemos que, $S_1 S_2 \cdots S_n \neq 0$ si y sólo si la caminata aleatoria no visita el eje x en el intervalo de tiempo $[1, n]$. Por lo tanto, por el Teorema 1.5.11 se tiene que el número de tales trayectorias es

$$\frac{b}{n} N_n(0,b)$$

y por argumentos similares a los del Lema 1.5.5 se sigue que hay $(n+b)/2$ pasos hacia arriba y $(n-b)/2$ pasos hacia abajo. Por lo tanto,

$$\begin{aligned}\mathbb{P}(S_1S_2\cdots S_n \neq 0, S_n = b) &= \frac{b}{n}N_n(0,b)p^{(n+b)/2}q^{(n-b)/2}\\ &= \frac{b}{n}\binom{n}{\frac{1}{2}(n+b)}p^{(n+b)/2}q^{(n-b)/2}\\ &= \frac{b}{n}\mathbb{P}(S_n = b|S_0 = 0).\end{aligned}$$

El caso $b < 0$ es similar, concluyendo así que

$$\mathbb{P}(S_1S_2\cdots S_n \neq 0, S_n = b) = \frac{|b|}{n}\mathbb{P}(S_n = b).$$

□

Observación 1.5.13 *Notemos que, la ecuación (1.13) implica*

$$\mathbb{P}(S_1S_2\cdots S_n \neq 0) = \frac{1}{n}\mathbb{E}(|S_n|).$$

Ahora vamos a analizar el comportamiento de los máximos de una caminata aleatoria. Sea $M_n := \{S_k : 1 \leq n\}$, $n \geq 1$, el máximo de (S_n) hasta el tiempo n. Tenemos el siguiente

Teorema 1.5.14 *Supongamos que $S_0 = 0$. Entonces, para cada $r \geq 1$, se cumple*

$$\mathbb{P}(M_n \geq r, S_n = b) = \begin{cases}\mathbb{P}(S_n = b), & si\ b \geq r,\\ \left(\frac{q}{p}\right)^{r-b}\mathbb{P}(S_n = 2r-b), & si\ b < r.\end{cases}$$

Demostración: Supongamos que $r \geq 1$ y que $b < r$, pues el caso $b \geq r$ es trivial. Sea $N_n^r(0,b)$ el número de realizaciones que van del $(0,0)$ a (n,b) y que pasan por el estado r al menos una vez. Sea i_r el primer tiempo al cual la caminata visita el estado r, reflejando la trayectoria entre i_r y n en la recta r se obtiene una trayectoria que va de $(0,0)$ a $(n, 2r-b)$. Ahora bien, a una de estas trayectorias le aplicamos la transformación inversa y obtenemos una que va de $(0,0)$ a (n,b) y que además pasa por r. Entonces, se obtiene que

$$N_n^r(0,b) = N_n(0, 2r-b),$$

y sabemos que cada una de tales realizaciones tiene probabilidad $p^{(n+b)/2}q^{(n-b)/2}$. Por lo tanto,

$$\begin{aligned}\mathbb{P}(M_n \geq r, S_n = b) &= N_n^r(0,b)p^{(n+b)/2}q^{(n-b)/2}\\ &= \left(\frac{q}{p}\right)^{r-b}N_n(0,2r-b)p^{(n+2r+b)/2}q^{(n-2r+b)/2}\\ &= \left(\frac{q}{p}\right)^{r-b}\mathbb{P}(S_n = 2r-b).\end{aligned}$$

$\square$

Una pregunta interesate que podemos hacerno es la siguiente, ¿Cuál es la probabilidad de que (S_n), $S_0 = 0$, alcance el nivel b por primera vez al tiempo n? Sea $f_b(n)$ tal probabilidad.

Teorema 1.5.15 *Para cada $n \geq 1$ se cumple que*

$$f_b(n) = \frac{|b|}{n}\mathbb{P}(S_n = b).$$

Demostración: Supongamos que $b > 0$. Notemos que,

$$\begin{aligned}
f_b(n) &\equiv \mathbb{P}(M_{n-1} = S_{n-1} = b-1, S_n = b) \\
&= (\mathbb{P}(M_{n-1} = S_{n-1} = b-1, S_n = b)\mathbb{P}(M_{n-1} = S_{n-1} = b-1) \\
&= p\mathbb{P}(M_{n-1} = S_{n-1} = b-1) \\
&= p\left[\mathbb{P}(M_{n-1} \geq b-1, S_{n-1} = b-1) - \mathbb{P}(M_{n-1} \geq b, S_{n-1} = b-1)\right] \\
&= p\left[\mathbb{P}(S_{n-1} = b-1) - \frac{q}{p}\mathbb{P}(S_{n-1} = b-1)\right] \\
&= \frac{b}{n}\mathbb{P}(S_n = b),
\end{aligned}$$

donde en la penúltima igualdad usamos el Teorema 1.5.14.

El caso $b < 0$ se obtiene de manera similar. $\square$

Ejemplo 1.5.16 *Sea $(S_n)_{n\geq 0}$ una caminata aleatoria simple con $S_0 = 0$. Para cada $r \neq 0$ definamos V_r como el número de visitas al estado r antes de que la cadena regrese a su estado inicial.*
(i) Demuestre que $\mathbb{E}(V_r) = 1$.
(ii) Dar un criterio para determinar si el número de visitas a 0 es finito o infinito.

Solución: (i) Sea A_n el evento que "al tiempo n la caminata visita el estado r y no ha visitado el estado 0 hasta ese instante". Entonces, $V_r \equiv \sum_{\geq 1} \mathbf{1}_{A_n}$. Por otro lado, se tiene que

$$\mathbb{E}(\mathbf{1}_{A_n}) = \mathbb{P}(A_n) = \mathbb{P}(S_n = r, S_1 S_2 \cdots S_n \neq 0) = \frac{|r|}{n}\mathbb{P}(S_n = r) \equiv f_r(n).$$

Por lo tanto,

$$\mathbb{E}(V_r) = \mathbb{E}(\sum_{\geq 1} \mathbf{1}_{A_n}) = \sum_{n\geq 1} f_r(n) = 1.$$

Vamos a ver demostrar la última igualdad. Note que,

$$\sum_{n\geq 1} f_r(n) \equiv \mathbb{P}(S_n = r, \text{para algún } n) := f_r.$$

Condicionando en el primer salto, i.e., en S_1 obtemos la ecuación

$$f_r = \frac{1}{2}(f_{b+1} - f_{b-1}),\ b > 0,$$

con condición inicial $f_0 = 1$. Resolviendo la ecuación obtenemos que $f_b = 1$. Lo mismo se puede hacer para el caso $b < 0$.
(ii) Sea R el número total de visitas al estado 0. Notemos que, $R = \sum_{n\geq 0} \mathbf{1}_{\{S_n=0\}}$, entonces

$$\mathbb{E}(R) = \sum_{n\geq 0} \mathbb{P}(S_n = 0) = \sum_{k\geq 0} \binom{2k}{k} p^k q^k. \tag{1.14}$$

Notemos que, $pq \leq 1/4$ y $pq = 1/4$ si y sólo si $p = 1/2$. Luego, usando la identidad de Stirling $n! \sim n^{n+\frac{1}{2}} e^{-n}\sqrt{2\pi}$, se tiene que para k suficientemente grande

$$\binom{2k}{k} \sim \frac{(2k)^{2k+1/2}e^{-2k}}{\sqrt{2\pi}(k^{k+1/2}e^{-2})^2} = (\sqrt{2\pi})^{-1}2^{2k+1/2}k^{-1/2},$$

es decir, el término general en la seríe está dado por la aproximación

$$\binom{2k}{k} p^k q^k \sim (\sqrt{\pi})^{-1} k^{-1/2}(4pq)^k.$$

Por lo tanto, la seríe que aparece en (1.14) no es es convergente para $p = 1/2$ ya que el término general es de orden de $k^{-1/2}$. Por otro lado, si $p \neq 1/2$, se tiene que $4pq < 1$, en consecuencia (1.14) es convergente.

□

1.6. Ejercicios

1. Sea $\{S_n, n \geq 0\}$ la caminata aleatoria simple simétrica con $S_0 = 0$, y defina $T =: \{n \geq 1 : S_n = 0\}$ el primer tiempo de regreso al punto de inicio. Demuestre que

   $$\mathbb{P}(T = 2n) = \frac{1}{2n-1}\binom{2n}{n}2^{-2n}.$$

   Deduzca de lo anterior que $\mathbb{E}(T^\alpha) < \infty$ si, y sólo si, $\alpha < \frac{1}{2}$. *Sugerencia:* recuerde la fórmula de Stirling, $n! \sim n^{n+\frac{1}{2}} e^{-n}\sqrt{2\pi}$.

2. Sea $\{S_n, n \geq 0\}$ la caminata aleatoria simple simétrica con $S_0 = 0$ y sea $M_n = \max_{n\geq 0} S_n$. Demuestre que

   $$\mathbb{P}(M_n = r) = \mathbb{P}(S_n = r) + \mathbb{P}(S_n = r + 1),\ \ r \geq 0.$$

3. Sea $\{S_n, n \geq 0\}$ la caminata aleatoria simple simétrica con $S_0 = 0$.
   a) Demuestre que

   $$\mathbb{P}(S_1 S_2 \cdots S_{2m} \neq 0) = \mathbb{P}(S_{2m} = 0),\ \ m \geq 1.$$

   b) Sea $\alpha_{2n}(2k)$ la probabilidad de que la última visita a 0 antes del tiempo $2n$ ocurrió en el tiempo $2k$. Justifique que

   $$\alpha_{2n}(2k) = \mathbb{P}(S_{2k} = 0)\mathbb{P}(S_1 S_2 \cdots S_{2n-2k} \neq 0).$$

c) Pruebe que

$$\alpha_{2n}(2k) = \mathbb{P}(S_{2k} = 0)\mathbb{P}(S_{2n-2k} = 0).$$

Capítulo 2

Funciones generadoras

2.1. Funciones generadoras de probabilidades

Definición 2.1.1 *Sea* $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ *un vector aleatorio discreto con función de probabilidades conjuntada* $f_{\mathbf{X}}(x_1, x_2, \ldots, x_n)$*, definimos la función generadora de probabilidades del vector* $\mathbf{X}$ *por*

$$\begin{aligned} G_{\mathbf{X}}(s_1, s_2, \ldots, s_n) &= \sum_{x_1, x_2, \ldots, x_n} s_1^{x_1} s_2^{x_2} \cdots s_n^{x_n} f_{\mathbf{X}}(x_1, x_2, \ldots, x_n),\ |s_i| \leq 1,\ i = 1, \ldots, n, \\ &\equiv \mathbb{E}\left[s_1^{X_1} s_2^{X_2} \cdots s_n^{X_n}\right]. \end{aligned}$$

De la definición anterior obtenemos que, la función generadora de probabilidades (f.g.p.) de X_i está dada por

$$G_{X_i}(s) = G_{\mathbf{X}}(1, \ldots, 1, s, 1, \ldots, 1) = \mathbb{E}\left[s^{X_i}\right],\ |s| \leq 1,$$

donde s aparece en la i-ésima entrada.

Notemos que, en general $G_X(s)$ está bien definida para todo $|s| \leq 1$. En efecto,

$$|G_X(s)| \leq |\sum_x s^x \mathbb{P}(X = x)| \leq \sum_x |s|^x \mathbb{P}(X = x) \leq \sum_x \mathbb{P}(X = x) = 1.$$

Sin embargo, puede extenderse el rango de definición de G_X. Al número $R > 0$ tal que $|G_X(s)| < \infty$, $|s| < R$, se le llama *radio de convergencia*.

Ejemplo 2.1.2 *(i) Supongamos que* $X \sim Bin(n, p)$*. Entonces,*

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{n} s^k \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (ps)^k (1-p)^{n-k} \\ &= (q + ps)^n,\ q := 1 - p. \end{aligned}$$

(ii) Si $X \sim Poisson(\lambda)$, se tiene

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} e^{-\lambda s} = e^{-\lambda(1-s)} \end{aligned}$$

Note que, en ambos ejemplos $R = \infty$.

Vamos a ver ahora la utilidad de la f.g.p.

Teorema 2.1.3 *Sea X una variable aleatoria no negativa tal que $\mathbb{P}(X = n) \equiv p_n$, $n = 0, 1, \ldots$ y con función generadora de probabilidades G. Entonces,*
(i) $G(s)$ es diferenciable en todo $|s| < 1$, y su derivada está dada por

$$G'(s) = \sum_{n=1}^{\infty} n p_n s^{n-1}.$$

Para $s = 1$,

$$G'(1) := \lim_{s \uparrow 1} \sum_{n=1}^{\infty} n p_n s^{n-1} \text{ finito o infinito.}$$

(ii) Para cada $k \geq 1$, se tiene la derivada k-ésima está dada por

$$G^{(k)} = \sum_{n=1}^{\infty} \frac{n!}{(n-k)!} p_n s^{n-k}.$$

(iii) G determina la distribución de X, es decir, $(p_n)_{\geq 0}$.

Demostración: Sólo vamos a demostrar parte (iii). Por definición se tiene que

$$G(0) = \mathbb{P}(X = 0) = p_k.$$

Ahora bien, para cada $k \geq 1$,

$$G^{(k)}(0) = \lim_{s \downarrow 0} \sum_{n=1}^{\infty} \frac{n!}{(n-k)!} p_n s^{n-k} = k p_k,$$

entonces

$$p_k = \frac{1}{k!} G^{(k)}(0).$$

Por lo tanto, G determina $(p_k)_{k \geq 0}$. □

La parte (iii) del teorema anterior nos dice que, para conocer (su distribución) a una variable aleatoria es suficiente con conocer su f.g.p. Por otro lado, conociendo la distribución de una variable aleatoria se determina su f.g.p.

Como un corolario del teorema anterior temenos el siguiente resultado.

Corolario 2.1.4 *Sea X una variable aleatoria con función generadora de probabilidades G. Entonces,*

$$\mathbb{E}(X) = G'(1).$$

Más generalmente, el k-ésimo momento factorial, $\mu^{(k)}$, de X está dado por

$$\begin{aligned} \mu^{(k)} &:= \mathbb{E}[X(X-1)(X-2)\cdots(X-k+1)] \\ &= G^{(k)}(1). \end{aligned}$$

En particular, del corolario anterio se sigue que

$$\mathrm{Var}(X) = G^{(2)}(1) + G'(1) - (G'(1))^2 \tag{2.1}$$

El siguiente resultado nos habla de la función generadora de probabilidades conjunta cuando hay independencia.

Teorema 2.1.5 *Suponga que X y Y tienen función generadora de probabilidades conjunta $G(s,t)$. Entonces, X y Y son independientes si y sólo si $G(s,t) = G(s,1)G(1,t)$.*

Demostración: $\Rightarrow$) Por definición tenemos que

$$\begin{aligned} G(s,t) &= \mathbb{E}(s^X t^Y) \\ &= \mathbb{E}(s^X)\mathbb{E}(t^Y), \text{ (independencia)} \\ &= G(s,1)G(1,t). \end{aligned}$$

$\Leftarrow$) Notemos que,

$$\begin{aligned} G(s,1)G(1,t) &= \left(\sum_x s^x \mathbb{P}(X=x)\right)\left(\sum_y t^y \mathbb{P}(Y=y)\right) \\ &= \sum_x \sum_y s^x t^y \mathbb{P}(X=x)\mathbb{P}(Y=y). \end{aligned}$$

Por otro lado,

$$G(s,t) = \sum_{x,y} s^x t^y \mathbb{P}(X=x, Y=y).$$

Luego, para que se cumpla $G(s,t) = G(s,1)G(1,t)$ se debe tener que

$$\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\mathbb{P}(Y=y), \text{ para todo } x, y,$$

la última relación en justamente la definición de independencia. $\square$

Ejemplo 2.1.6 *(Continuando con el 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. Sea Y en número de huevos fecundados y Z los restantes. produciendo así Y pollos. Demuestre que Y y Z son independientes.*

Solución: *Condicionalmente en $X = x$, $Y \sim Bin(x, p)$. Entonces,*

$$\mathbb{E}[s^Y | X = x] = (ps + q)^x.$$

Luego,

$$\begin{aligned}
\mathbb{E}(s^Y t^Z) &= \mathbb{E}(s^Y t^{X-Y}) \\
&= \mathbb{E}\left\{\mathbb{E}[(s/t)^Y t^X | X]\right\} \\
&= \mathbb{E}\left\{t^X \mathbb{E}[(s/t)^Y | X]\right\} \\
&= \mathbb{E}[t^X (p(s/t) + q)^X] \\
&= \mathbb{E}[(ps + qt)^X].
\end{aligned}$$

Recordando que X es tiene distribución Poisson temenos que

$$\mathbb{E}(s^Y t^Z) = \exp\{\lambda(ps + qt - 1)\} = \exp\{\lambda p(s-1)\} \exp\{\lambda q(t-1)\},$$

usando el teorema anterior obtemos que Y y Z son independientes. Además, se observa que $Y \sim$ Poisson(λp) y $Z \sim$ Poisson(λq).

Vamos a concluir la sección con un resultado que será muy últil mas a adelante.

Proposición 2.1.7 *Sean N y $(X_i)_{i \geq 1}$ variables aleatorias, supongamos que N es no negativa y que, para cada $i \geq 1$*

$$\mathbb{E}(s^{X_i}) = G(s),$$

es decir, las X_i's tienen la misma distribución. Entonces, $Z := \sum_{i=1}^{N} X_i$ tiene función generadora de probabilidades

$$G_Z(s) = G_n(G(s)).$$

Demostración: Condicionando tenemos que,

$$\begin{aligned}
\mathbb{E}(s^Z) &= \mathbb{E}[\mathbb{E}(s^Z | N)] \\
&= \mathbb{E}[\mathbb{E}(s^{X_1}) \cdots \mathbb{E}(s^{X_N]})] \\
&= \mathbb{E}[G(s)^N] \\
&= G_N(G(s)).
\end{aligned}$$

□

2.2. Una breve introducción a procesos de Galtoa-Watson

Supongamos que una población de partículas (moléculas, virus, etc.) evoluciona de la siguiente manera. La población inicial al tiempo $n = 0$ con una partícula, al tiempo $n = 1$ dicha partícula muere y da oringen a un número aleatorio (X) de partículas indénticas entre si y su progenitora y, a tiempos subsecuentes $n = 2, 3, \ldots$, cada individuo evoluciona de la misma manera (muriendo y ramificandose) produciendo así X partículas. Supondremos que X tiene función de probabilidades $f_X(k)$, $k = 0, 1, 2, \ldots$. Vamos a suponer que el número de partículas que produce cada individuo es independiente de los demás y que tiene la misma distribución que X.

Vamos a denotar por Z_n el tamaño de la población al tiempo n, entonces $Z = \{Z_n : n \geq 0\}$ es un proceso estocástico en cual, a cada tiempo, no da el total de la población. Sea $(X_i^n, n \geq 0, i \geq 1)$ un colección de variables aleatorias independientes todas con función de probabilidades f_X. Entonces, el proceso Z se puede describir de la siguiente manera, $Z_0 = 1$ y

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_i^n, \;\; n = 0, 1, \ldots, \tag{2.2}$$

donde X_i^n representa el número de descendientes que produce el i-ésimo individuo prensete en la generación n. Una consecuencia de la independencia de la colección (X_i^n) es que el proceso Z es una cadena de Markov a tiempo discreto, ver la Observación 1.5.4 (i). El proceso Z es llamado *proceso de Galton-Watson*[1]. El proceso de Galton-Watson ha sido fuente de inspiración para procesos de ramificación mucho más generales, los cuales conforman un área de investigación dentro de la teoría de probabilidad y procesos estocásticos por su riqueza en la variedad de modelos y su interacción con otras áreas de las matemáticas.

A principios del presente capítulo vimos que la función generadora de probabilidades es muy útil cuando se trabaja con variables aleatorias que toman valores en los enteros no negativos, que es caso del proceso de Galton-Watson. Supongamos que X tiene función generadora de probabilidades $G(s)$, $|s| < 1$. ... Entonces,

$$G_n(s) := \mathbb{E}(s^{Z_n}).$$

Notemos que $G_1(s) = G(s)$. Luego, por identidad (2.2) y la Proposición 2.1.7 se tiene que para $|s| \leq 1$

$$G_{n+1}(s) = G_n(G(s)), \quad \text{para todo } \; n \geq 1, \tag{2.3}$$

es decir, G_n es la convolución de G consigo misma n veces.

Proposición 2.2.1 *Supongamos que $\mu = \mathbb{E}(X)$ y $\sigma^2 = Var(X)$. Entonces, para cada $n \geq 1$,*

$$\mathbb{E}(Z_n) = \mu^n,$$

y

$$Var(Z_n) = \begin{cases} n\sigma^2, & si\ \mu = 1, \\ \sigma^2 \mu^{n-1} \frac{\mu^n - 1}{\mu - 1}, & si\ \mu \neq 1. \end{cases}$$

Demostración: Sabemos que $\mathbb{E}(Z_n) = G_n'(s)|_{s=1}$. Entonces, de (2.2) se sigue que

$$\mathbb{E}(Z_n) = \mu \mathbb{E}(Z_{n-1}) = \mu G_{n-1}'(1),$$

el resultado se sigue por iteración.

[1] Francis Galton propuso la pregunta sobre la probabilidad de extinción de apellídos aristocráticos en Inglaterra en 1873 y Henry William Watson lo resolvió; y el 1874 escribieron el paper "On the probability of extinction of families".

Ahora vamos a demostrar la segunda afirmación. Diferenciando dos veces en (2.3) y recordando que $G(1) = 1$ y se obtiene

$$G_n''(1) = G^{(2)}(1)\left[G_{n-1}^{(2)}(1)\right]^2 + G'(1)G_{n-1}^{(2)}(1).$$

Luego, el resultado se concluye usando la fórmula (2.1). En efecto, para $\mu = 1$ se tiene

$$G^{(2)}(1) = \sigma^2,$$

y

$$G_n^{(2)}(1) = \sigma^2 + G_{n-1}^{(2)}(1),$$

lo cual junto con el hecho $G_1'(1) = G'(1)$ implica

$$G_n^{(2)}(1) = \sigma^2 n,\ n \geq 0.$$

... □

En general hay muy pocos casos en los que se puede encontrar un expresión explícita para G_n. Uno de ellos es caso en que la ramificación sigue una ley geométrica como lo muestra el siguiente

Ejemplo 2.2.2 *(Ramificación geométrica) Supongamos que $G(s) = q(1-ps)^{-1}$ $(p+q=1)$, $|s| < \frac{1}{p}$, es decir, $\mathbb{P}(X = k) = qp^k,\ k \geq 0$. En tal caso se tiene (Ver ejercicio 3)*

$$G_n(s) = \begin{cases} \frac{n-(n-1)s}{n+1-ns}, & p = \frac{1}{2}, \\ \frac{q[p^n - q^n - ps(p^{n-1}-q^{n-1})]}{p^{n+1}-q^{n+1}-ps(p^n-q^n)}, & p \neq \frac{1}{2}. \end{cases} \tag{2.4}$$

Una de las preguntas importantes acerca del proceso Z es conocer la probabilidad de extinción al tiempo n, es decir, conocer $\mathbb{P}(Z_n = 0) = G_n(0)$ así como también $\lim \mathbb{P}(Z_n = 0)$. En el presente ejemplo se pueden encontrar de manera explícita. En efecto, de (2.4) tenemos que

$$\mathbb{P}(Z_n = 0) = \begin{cases} \frac{n}{n+1}, & p = q, \\ \frac{q(p^n-q^n)}{p^{n+1}-q^{n+1}}, & p \neq q. \end{cases}$$

Por lo tanto,

$$\lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \begin{cases} 1, & p \leq q, \\ \frac{q}{p}, & p > q. \end{cases}$$

Observación 2.2.3 *En el ejemplo anterior sabemos que $\mathbb{E}(Z_1) = p/q \leq 1$ si y sólo si $p \leq q$. Luego, en tal caso, como $\lim_{n\to\infty} \mathbb{E}(Z_n) = 0$ ya que $\mathbb{E}(Z_n) = [\mathbb{E}(Z_n)]^n$. Lo cual nos indica que, de alguna manera, Z_n tiende a cero cuando n tiene a infinito.*

La propiedad anterior no es propia de caso geométrico como lo muestra en siguiente

Teorema 2.2.4 *Se cumple que $\lim_{n\to\infty} \mathbb{P}(Z_n = 0) = \eta$ existe. Más aún, η es la menor raíz no negativa de la ecuación $G(s) = s$, $0 \leq s \leq 1$.*

Demostración: Sea $\eta_n := \mathbb{P}(Z_n = 0)$, y sea η la menor raíz no negativa de de la ecuación $G(s) = s$, $0 \leq s \leq 1$. Vamos a demostrar que $\eta_n \to \eta$.

Consideremos los siguientes casos

1. Si $f_X(0) = 0$, entonces
$$\eta_n = G_n(0) = 0 = \eta.$$

2. Si $f_X(0) = 1$, entonces
$$\eta_n = G_n(1) = 1 = \eta.$$

3. Supongamos que $f_X(0) + f_X(1) = 1$ y $f_X(0) f_X(1) \neq 0$. Entonces,
$$\begin{aligned} \eta_n = G_n(0) &= 1 - \mathbb{P}(Z_n > 0) \\ &= 1 - [f(1)]^n \\ &\to 1, \text{ cuando } n \to \infty, \end{aligned}$$
y en este caso $\eta = 1$.

4. Finalmente, supongamos que $0 < f_X(0) < f_X(0) + f_X(1) < 1$. Notemos que
$$\{Z_n = 0\} \subset \{Z_{n+1} = 0\},$$
entonces
$$\eta_n = \mathbb{P}(Z_n = 0) \leq \mathbb{P}(Z_{n+1} = 0) = \eta_{n+1}.$$
Luego, (η_n) es una sucesión creciente y acotada, y por lo tanto es convergente. Sea $\lambda := \lim_{n\to\infty} \eta_n$.
Por otro lado, sabemos que $G_{n+1}(0) = G(G_n(0))$, entonces
$$\eta_n = G(\eta_n),$$
en consecuencia si hacemos n tender a infinito, por continuidad se sigue que
$$\lambda = G(\lambda).$$
Para concluir la prueba debemos demostrar que $\lambda = \eta$. Notemos que,
$$\eta_1 = G_1(0) = G(0) \leq G(\eta),$$
y
$$\eta_2 = G_2(0) = G(G(0)) = G(\eta_1) \leq G(\eta) = \eta.$$
Procediendo inductivamente obtenemos que $\eta_n \leq \eta$, entonces $\lambda \leq \eta$. Ahora bien, por hipótesis η es la menor raíz no negativa de la ecuación $G(s) = s$, $0 \leq s \leq 1$. Entonces, $\lambda \geq \eta$. Concluyendo así que
$$\lambda = \eta.$$

□

Puede demostraser que $\eta = 1$ si $\mathbb{E}(X) < 1$ y $\eta < 1$ si $\mathbb{E}(X) > 1$. Si $\mathbb{E}(X) = 1$, entonces $\eta = 1$ siempre que $\mathrm{Var}(X) > 0$, es decir, siempre que X no sea constanta con probabilidad 1.

2.3. Ejercicios

Nota: El fin de semana voy a agregar más ejecicios...

1. Sea X una variable aleatoria con distribución de Poisson de parámetro Y, donde Y tiene distribución de Poisson con parámetro μ. Demuestre que

$$G_{X+Y}(s) = \exp\{\mu[se^{s-1} - 1]\}.$$

2. Sean $X_0, X_1.X_2, \cdots$ una variables aleatorias independientes e idénticamente distribuidas todas con distribución logarítmica, es decir,

$$\mathbb{P}(X = k) = \frac{(1-p)^k}{k \log(1-p)},\ k \geq 1,$$

   donde $0 < p < 1$. Supong que N es independiente de las X_i's y tiene distribución de Poisson con parámetro λ. Demuestre que $Y := \sum_{i=1}^{N} X_i$ tiene distribución binomial negativa.

   Sedice que Z tiende distribución bonomial negativa con parámetros $r \in \mathbb{N}$ y $p \in (0,1)$ si

$$\mathbb{P}(Z = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r},\ k = r+1, r+2, \ldots.$$

   Sugerencia: recuerde que $\log(x) = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} + \cdots$.

3. Verifique la identidad (2.4) de Ejemplo 2.2.2.

4. Sea X una v.a. no-negativa con función generadora de probabilidades $G_X(s)$ tal que $G'_X(1) < \infty$. Demuestre que

$$G(s) = \frac{1}{\mathbb{E}(X)} \frac{1 - G_X(s)}{1-s}$$

   es la función generadora de probabilidades de alguna variable aleatoria Y. ¿Cuando se tiene que $G(s) = G_X(s)$?

5. Sea (Z_n) un proceso de Galton-Watson con $Z_0 = 1$, $\mathbb{E}(Z_1) = \mu > 0$ y $\mathrm{Var}(Z_1) > 0$. Demuestre que $\mathbb{E}(Z_n Z_m) = \mu^{n-m}\mathbb{E}(Z_m^2)$, $m \leq n$. Luego, encuentre $\rho(Z_n, Z_m)$ en términos de μ.

6. Sea (Z_n) con en ejercicio anterior. Sea G_n la función generadora de probabilidades de Z_n.

   (a) Encuentre una expresión para G_n cuando la función generadora de Z_1 está dada por $G_1(s) \equiv G(s) = 1 - \alpha(1-s)^\beta$, $0 < \alpha, \beta < 1$.

   (b)Encuentre $\mathbb{P}(Z_1 = k)$, $k = 0, 1, \ldots$.

7. Sea Z un proceso de Galton-Watson donde X (el número de descendientes) es tal que $\mathbb{P}(X = 0) = 2/5 = 1 - \mathbb{P}(X = 2)$. Encuentre la probabilidad de extinción de Z.