1 Esperanza para variables aleatorias multivariadas conjuntas

Cuando X e Y son variables aleatorias distribuidas conjuntamente y Z es una variable aleatoria que es una función de X y Y, esto es : $Z = h(X, Y)$, entonces la esperanza de Z, puede ser encontrada directamente desde la distribución conjunta. Cuando X y Y son variables aleatorias discretas, se tiene:

$$\mathbb{E}(Z) = \sum_{\text{todo x}} \sum_{\text{todo y}} h(x, y) p_{X,Y}(x, y),$$

Y cuando las variables son continuas se tiene:

$$\mathbb{E}(Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f_{X,Y}(x, y) dx dy.$$

Ejemplo 1.1 Sean X y Y dos variables aleatorias continuas, cuya función densidad de probabilidad conjunta es:

$$f_{X,Y}(x, y) = \begin{cases} 1 - (x + y)/3, & 0 \leq x \leq 1, 0 \leq y \leq 2, \\ 0 & \text{en otros casos} \end{cases}$$

Deseamos encontrar $\mathbb{E}(XY)$. Tenemos:

$$\begin{aligned} \mathbb{E}(XY) &= \int_{y=0}^{2} \int_{x=0}^{1} xy[1 - (x + y)/3] dx dy = \int_{y=0}^{2} \int_{x=0}^{1} \left(xy - \frac{x^2 y}{3} - \frac{xy^2}{3} \right) dx dy \\ &= \int_{y=0}^{2} \left(\frac{x^2 y}{2} - \frac{x^3 y}{9} - \frac{x^2 y^2}{6} \right) \Bigg|_{x=0}^{1} dy = \int_{y=0}^{2} \left(\frac{y}{2} - \frac{y}{9} - \frac{y^2}{9} \right) dy \\ &= \left(\frac{7y^2}{36} - \frac{y^3}{18} \right) \Bigg|_{0}^{2} = \frac{12}{36} = 1/3 \end{aligned}$$

Si X e Y son variables aleatorias discretas y sea $g : \mathbb{R}^2 \to \mathbb{R}$ una función, entonces $Z = g(X, Y)$ es una variable aleatoria también, definida como $Z(\omega) = g(X(\omega), Y(\omega))$ para $\omega \in \mathbb{R}$. La esperanza de Z puede ser calculada directamente desde la función $p_{X,Y}(x, y) = \mathbb{P}(X = X, Y = y)$ como lo indica el siguiente teorema:

Teorema 1.1 Tenemos que:

$$\mathbb{E}(g(X, Y)) = \sum_{x \in \text{todo x}} \sum_{y \in \text{todo y}} g(x, y) \mathbb{P}(X = x, Y = y)$$

siempre que esta suma converga absolutamente.

Una consecuencia importantes de este teorema, es que el operador $\mathbb{E}$ actua de manera lineal sobre el conjunto de las variables discretas. Es decir si X e Y son variables aleatorias discretas y $a, b \in \mathbb{R}$, entonces:

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y),$$

siempre que $\mathbb{E}(X)$ y $\mathbb{E}(Y)$ existan.

En efecto si $g(x,y) = ax + by$, por propiedad se tiene que :

$$\begin{aligned}\mathbb{E}(aX + bY) &= \sum_x \sum_y (ax + by)\mathbb{P}(X = x, Y = y)\\ &= a\sum_x x \sum_y \mathbb{P}(X = x, Y = y) + b\sum_y y \sum_x \mathbb{P}(X = x, Y = y)\\ &= a\sum_x x\mathbb{P}(X = x) + b\sum_y y\mathbb{P}(Y = y)\\ &= a\mathbb{E}(X) + b\mathbb{E}(Y).\end{aligned}$$

Dos variables aleatorias discretas X e Y son independientes si el par de eventos $\{X = x\}$ y $\{Y = y\}$ son independientes para todo $x, y \in \mathbb{R}$. Se escribe esta condición como:

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y) \quad \text{para} \;\; x, y \in \mathbb{R}.$$

Las variables aleatorias que no son independientes, se llaman `dependientes`.

Teorema 1.2 Las variables aleatorias discretas X e Y son independientes si y sólo si existen las funciones $f, g : \mathbb{R} \to \mathbb{R}$ tal que se cumple que las funciones de masa de probabilidad conjunta de X e Y satisface:

$$p_{X,Y}(x,y) = f(x)g(y) \quad \text{para} \;\; x, y \in \mathbb{R}. \tag{1}$$

Probemos que si se cumple la condición (1) para las funciones f y g y $x \in X$ y $y \in Y$ tenemos por definición de las funciones marginales de X e Y:

$$p_X(x) = f(x)\sum_y g(y), \quad p_Y(y) = g(y)\sum_x f(x)$$

Por propiedad:

$$\begin{aligned}1 = \sum_{x,y} p_{X,Y}(x,y) &= \sum_{x,y} f(x)g(y)\\ &= \sum_x f(x) \sum_y g(y)\end{aligned}$$

Por tanto :

$$\begin{aligned}p_{X,Y}(x,y) = f(x)g(y) &= f(x)g(y)\sum_x f(x)\sum_y g(y)\\ &= p_X(x)p_Y(y).\end{aligned}$$

Ejemplo 1.2 Supongamos que X e Y son variables aleatorias tomando valores enteros no negativos, con función de masa de probabilidad conjunta:

$$p_{X,Y}(i,j) = \mathbb{P}(X = i, Y = j) = \frac{1}{i!j!}\lambda^i \mu^j e^{-(\mu+\lambda)} \;\; \text{para } i, j = 0, 1, 2, \ldots$$

Por el teorema anterior, X e Y son independientes.

Teorema 1.3 Si X e Y son variables aleatorias discretas independientes con esperanza $\mathbb{E}(X)$ y $\mathbb{E}(Y)$ entonces:

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Por el teorema (1.1):

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{x,y} xy\mathbb{P}(X = x, Y = y) \\
&= \sum_{x,y} xy\mathbb{P}(X = x)\mathbb{P}(Y = y) \\
&= \sum_{x} x\mathbb{P}(X = x) \sum_{y} y\mathbb{P}(Y = y) = \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

El inverso de este teorema es falso, si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ entonces X e Y son independientes como indica el siguiente ejemplo:

Ejemplo 1.3 Supongamos que X tiene una distribución dada por:

$$\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{3}$$

y Y es dado por

$$Y = \begin{cases} 0 & \text{si} \quad X = 0 \\ 1 & \text{si} \quad X \neq 0 \end{cases}$$

Si tomamos $\Sigma = \{-1, 0, 1\}$ y $\mathbb{P}$ dado por $\mathbb{P}(-1) = \mathbb{P}(0) = \mathbb{P}(1) = \frac{1}{3}$ y sea $X(\omega) = \omega, Y(\omega) = |\omega|$. Entonces X e Y son dependientes:

$$\mathbb{P}(X = 0, Y = 1) = 0$$

Pero,

$$\mathbb{P}(X = 0)\mathbb{P}(Y = 1) = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}$$

Por otro lado:

$$\begin{aligned}
\mathbb{E}(XY) &= \sum_{x,y} xy\mathbb{P}(X = x, Y = y) \\
&= (-1) \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = 0
\end{aligned}$$

$$\mathbb{E}(X)\mathbb{E}(Y) = 0 \cdot \frac{2}{3} = 0$$

Teorema 1.4 Las variables aleatorias X e Y son independientes si y sólo si

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

para todas las funciones $g, h : \mathbb{R} \to \mathbb{R}$ en el cual las esperanzas existen.

Para el caso en el que $X = (X_1, X_2, \dots, X_n)$ de variables aleatorias con $n > 2$. La familia X se dice que es independiente si:

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \cdots \mathbb{P}(X_n = x_n)$$

Además $X_1, X_2, \dots, X_n$ son independientes si se cumple que:

$$\mathbb{E}(X_1, X_2, \cdots, X_n) = \mathbb{E}(X_1)\mathbb{E}(X_2) \cdots \mathbb{E}(X_n).$$

La familia X es llamada `independiente por pares` si X_i y Y_j son independientes siempre que $i \leq j$.

Observa que, cuando X e Y son variables independientes entonces $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$, en el sentido continuo:

$$\begin{aligned}
\mathbb{E}(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{XY}(x,y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_X(x) y f_Y(y) dx dy = \int_{-\infty}^{\infty} \mathbb{E}(X) y f_Y(y) dy = \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

En el caso, en el cuál una variable aleatoria es definida como la suma de dos variables aleatorias, merece una especial atención:

$$\begin{aligned}
\mathbb{E}(Z) = \mathbb{E}(X+Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+y) f_{XY}(x,y) dx dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{XY}(x,y) dx dy + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{XY}(x,y) dx dy \\
&= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f_{XY}(x,y) dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f_{XY}(x,y) dx \right] dy \\
&= \int_{-\infty}^{\infty} x f_X(x) dx + \int_{-\infty}^{\infty} y f_Y(y) dy = \mathbb{E}(X) + \mathbb{E}(Y).
\end{aligned}$$

Debes notar que esta propiedad, llamada propiedad de linealidad de la esperanza, determina si las variables aleatorias son independientes o no. En términos generales, la esperanza de una suma de variables aleatorias es la suma de las esperanzas de esas variables aleatorias. Esto es, sean $X_1, X_2, \dots, X_n$ un conjunto de n variables aleatorias. Entonces:

$$\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}(X_n).$$

El mismo resultado no se aplica a las varianzas: la varianza de una suma de dos variables aleatorias es igual a la suma de las varianzas sólo si las variables aleatorias son independientes. Para mostrar esto, sean $\mu_X = \mathbb{E}(X)$ y $\mu_Y = \mathbb{E}(Y)$. Entonces :

$$
\begin{aligned}
\mathbb{V}(X+Y) &= \mathbb{E}[(X+Y-\mu_X-\mu_Y)^2] \\
&= \mathbb{E}[(X-\mu_X)^2] + 2\mathbb{E}[(X-\mu_X)(Y-\mu_Y)] + \mathbb{E}[(Y-\mu_Y)^2] \\
&= \mathbb{E}[(X-\mu_X)^2] + \mathbb{E}[(Y-\mu_Y)^2] = \mathbb{V}(X) + \mathbb{V}(Y),
\end{aligned}
$$

Donde hemos usado el hecho, de que cuando X e Y son independientes, se cumple:

$$\mathbb{E}[(X-\mu_X)(Y-\mu_Y)] = \mathbb{E}(X-\mu_X)\mathbb{E}(Y-\mu_Y) = 0,$$

desde, por ejemplo $\mathbb{E}(X-\mu_X) = \mathbb{E}(X) - \mathbb{E}(\mu_X) = \mu_X - \mu_X = 0$.

Para calcular la varianza de la suma de dos variables aleatorias X e Y que no son necesariamente independientes, utilizamos la relación $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$ y procedemos de la siguiente manera:

$$
\begin{aligned}
\mathbb{V}(X+Y) &= \mathbb{E}(X+Y)^2 - (\mathbb{E}(X)+\mathbb{E}(Y))^2 \\
&= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}[X])^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - (E(Y))^2
\end{aligned}
$$

y así:

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)).$$

Así, la varianza de la suma de dos variables aleatorias es igual a la suma de las varianzas sólo si $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ es cero. Este es el caso, por ejemplo, cuando X e Y son independientes. La cantidad $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ se denomina covarianza de X e Y y se denota como:

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

La varianza de la suma de variables aleatorias $X_1, X_2, \ldots, X_n$ es dada por la siguiente fórmula:

$$\mathbb{V}(X_1 + X_2 + \cdots + X_n) = \sum_{k=1}^{n} \mathbb{V}(X_k) + 2\sum_{j<k} \mathrm{Cov}(X_i, X_j).$$

Agunas propiedades de la covarianza de dos variables aleatorias son:

- $\mathrm{Cov}(\alpha X, \beta Y) = \alpha\beta\mathrm{Cov}(X,Y)$,
- $\mathrm{Cov}(X+\alpha, Y+\beta) = \mathrm{Cov}(X,Y)$,
- $\mathrm{Cov}(X, \alpha X+\beta) = \alpha\mathbb{V}(X)$.

Donde α y β son escalares. Como regla general, la covarianza es positiva si valores superiores a la media de X están asociados con valores por encima de la media de Y o si valores por debajo de la media de X están asociados con valores por debajo de la media de Y. Si los valores superiores a la media de uno están asociados con valores inferiores a la media del otro, entonces la covarianza será negativa. En ocasiones se plantean dificultades para interpretar la magnitud de la covarianza y por esta razón, se utiliza una nueva cantidad, llamada `correlación`.

Así como la desviación estándar de una variable aleatoria a menudo da una sensación mejor que la varianza para la desviación de una variable aleatoria desde el valor medio, la correlación de dos variables aleatorias normaliza la covarianza para hacerla sin unidad. La correlación de dos variables aleatorias se define como:

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X \sigma_y}.$$

Donde σ_X y σ_Y son la desviación estándar de X e Y respectivamente. Para dos variables aleatorias, se tiene $-1 \leq \text{Corr}(X, Y) \leq 1$.

Ejemplo 1.4 Calculemos la covarianza y la correlación de dos variables aleatorias X e Y cuya función de densidad de probabilidad conjunta es:

$$f_{X,Y}(x, y) = \begin{cases} 6x, & 0 < x < y < 1 \\ 0 & \text{en otros casos.} \end{cases}$$

Hallemos $\mathbb{E}(XY), \text{Cov}(X, Y), \text{Corr}(X, Y), \mathbb{E}(X + Y)$ y $\mathbb{V}(X + Y)$.

Comenzamos por calcular la función de densidad marginal de X e Y y de estos resultados, calculamos la esperanza y la varianza de X y de Y. Con esta información podemos calcular los elementos solicitados:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) dy = \int_{x}^{1} 6x dy = x(x - 1),\ 0 \leq x \leq 1,$$

y cero en otros casos.

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) dx = \int_{0}^{y} 6x dx = 3y^2,\ 0 \leq y \leq 1,$$

y cero en otros casos.

$$\begin{aligned}
\mathbb{E}(X) &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_{0}^{1} 6x^2(1 - x) dx = 1/2. \\
\mathbb{E}(Y) &= \int_{-\infty}^{\infty} xy f_Y(y) dy = \int_{0}^{1} 3y^3 dy = 3/4. \\
\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{0}^{1} 6x^3(1 - x) dx = 3/10, \\
\mathbb{E}(Y^2) &= \int_{-\infty}^{\infty} y^2 f_Y(y) dy = \int_{0}^{1} 3y^4 dy = 3/5,
\end{aligned}$$

También:

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = 3/10 - 1/4 = 1/20, \\
\mathbb{V}(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 = 3/5 - 9/16 = 3/80,
\end{aligned}$$

Ahora estamos en posición para calcular las cantidades requeridas:

$$\mathbb{E}(XY) = \int_{x=0}^{1} \int_{y=x}^{1} 6x^2 y dy dx = \int_{0}^{1} 3x^2 y^2 \Big|_{y=x}^{1} dx = \int_{0}^{1} (3x^2 - 3x^4) dx = 2/5.$$

$$\text{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 2/5 - 1/2 \times 3/4 = 1/40$$

$$\text{Corr}(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{1/40}{\sqrt{1/20}\sqrt{3/80}} = \frac{1}{\sqrt{3}} = \sqrt{3}/3.$$

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y) = 1/2 + 3/4 = 5/4.$$

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{Cov}(X,Y) = 1/20 + 3/80 + 1/20 = 11/80.$$

La covarianza de dos variables aleatorias X e Y captura el grado en que están correlacionadas. X e Y no están correlacionados si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Si, por ejemplo, X e Y son independientes, entonces $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ y X e Y tampoco están correlacionados. Sin embargo, si X e Y no están correlacionados, no necesariamente se sigue que son también independientes.

Ejemplo 1.5 Sean X e Y dos variables aleatorias discretas con función de masa de probabilidad conjunta dada por:

	X = -1	X = 0	X=1	$p_Y(y)$
Y = -1	1/12	3/12	1/12	5/12
Y = 0	1/12	0/12	1/12	2/12
Y = 1	1/12	3/12	1/12	5/12
$p_X(x)$	3/12	6/12	3/12	1

Desde esta tabla tenemos:

$$\mathbb{E}(X) = -1 \times 3/12 + 1 \times 3/12 = 0, \ \ \mathbb{E}(Y) = -1 \times 5/12 + 1 \times 5/12 = 0.$$

$$\mathbb{E}(XY) = (-1)(-1)1/12 + (-1)(1)1/12 + (1)(-1)1/12 + (1)(1)1/12 = 0.$$

Se sigue entonces, que X y Y son no correlacionadas, pues $\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$, pero ellas son dependientes:

$$p_{X,Y}(0,0) = 0 \neq p_X(0)p_Y(0) = 6/12 \times 2/12.$$

En resumen, X y Y son no correlacionadas, si $\text{Corr}(X,Y) = 0$ o si $\text{Cov}(X,Y) = 0$ o si $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Las variables aleatorias independientes no están correlacionadas, pero las variables aleatorias no correlacionadas no son necesariamente son independientes.

Consideremos ahora la varianza de la diferencia de dos variables aleatorias X e Y:

$$\mathbb{V}(X-Y) = \mathbb{V}(X) + \mathbb{V}(Y) - 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))$$

que se obtiene sustituyendo $X+Y$ por $X-Y$ en la fórmula de la varianza de una suma. Así, la varianza de la suma de dos variables aleatorias también es igual a la varianza de su diferencia, si no están correlacionadas.

Sean $X_1, X_2, \ldots, X_n$, n variables aleatorias independientes cuyas funciones de distribución acumulativa están dadas por $F_{X_1}(x), F_{X_2}(x), \ldots, F_{X_n}(x)$, respectivamente. Deseamos calcular las funciones de distribución que se definen como el máximo y el mínimo de estas n variables aleatorias. En otras palabras, buscamos la función de distribución de las variables aleatorias $X_{\max}$ y $X_{\min}$ definidas como:

$$X_{\max} = \max(X_1, X_2, \ldots, X_n) \ \text{ y } \ X_{\min} = \min(X_1, X_2, \ldots, X_n).$$

Debes observar que $X_{\max} \leq x$ si y sólo si $X_i \leq x_i$ para todo $i = 1, 2, \ldots, n$ y por tanto:

$$F_{\max}(x) = \mathbb{P}(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x).$$

Así, desde nuestra aseveración de que las n variables son independientes, se sigue que :

$$F_{\max}(x) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) = \prod_{i=1}^{n} F_{X_i}(x).$$

Ejemplo 1.6 Comparemos el tiempo esperado para realizar tres tareas dependientes de datos en un sistema informático multiprocesador y el tiempo esperado para realizar estas tareas en un solo procesador. Sean X_1, X_2 y X_3 las variables aleatorias que describen los tiempos necesarios para realizar las tres tareas y debido a la dependencia de los datos, supongamos además que cada una tiene una distribución acumulativa (exponencial) idéntica dada por:

$$F_{X_i} = 1 - e^{-t/4}, \;\; i = 1, 2, 3.$$

Así, el tiempo medio de procesamiento para cada una de las tres tareas es dado por $\mathbb{E}(X_i) = 4$. Sea X la variable aleatoria que describe el tiempo necesario para completar las tres tareas. Dado que las tres tareas del multiprocesador finalizarán tan pronto como se complete el último de ellos, tenemos $X = \max\{X_1, X_2, X_3\}$. Así, la función de distribución acumulativa de X es:

$$F_X(t) = (1 - e^{-t/4})^3$$

y la función densidad, obtenida diferenciando $F_X(t)$, es dado por:

$$f_X(t) = 3(1 - e^{-t/4})^2(1/4)e^{-t/4} = \frac{3}{4}e^{-t/4}(1 - e^{-t/4})^2.$$

Notemos que la distribución del máximo de un conjunto de variables aleatorias distribuidas exponencialmente no es exponencial. La esperanza de X puede ahora ser calculada. Teniendo en cuenta el resultado previamente calculado,

$$\int_0^{\infty} \lambda t e^{-\lambda t} dt = \frac{1}{\lambda}, \; t \geq 0,$$

Obtenemos:

$$\begin{aligned}
\mathbb{E}(X) &= \frac{3}{4} \int_0^{\infty} e^{-t/4}(1 - e^{-t/4})^2 dt \\
&= 3 \int_0^{\infty} \frac{t}{4} e^{-t/4}(1 - 2e^{-t/4} + e^{-t/2}) dt \\
&= 3\left(4 - 2 + \frac{4}{9}\right) = 7.3333.
\end{aligned}$$

Observa que el cálculo del valor esperado para la última de las tres tareas a completar, es decir, 7.3333, es bastante mayor que el tiempo esperado para completar cualquiera de las tareas, que son 4.

En un solo procesador, la variable aleatoria que describe el tiempo total requerido ahora es dada por $X = X_1 + X_2 + X_3$ y calculamos la esperanza simplemente como:

$$\mathbb{E}(X) = \sum_{i=1}^{3} \mathbb{E}(X_i) = 12.$$

Ahora consideramos, la variable aleatoria $X_{\min} = \min(X_1, X_2, \ldots, X_n)$. En este caso es fácil, primero calcular $\mathbb{P}(X_{\min}) > x$ y formamos $F_{\min}(x)$, desde la relación:

$$F_{\min}(x) = 1 - \mathbb{P}(X_{\min} > x).$$

Si $X_{\min} > x$ si y sólo si $X_i > x$ para todo $i = 1, 2, \ldots, n$ y por tanto:

$$\mathbb{P}(X_{\min} > x) = \mathbb{P}(X_1 > x, X_2 > x, \ldots X_n > x).$$

Una vez más, volviendo a la independencia de las n variables aleatorias $X_1, X_2, \ldots, X_n$, tenemos:

$$\mathbb{P}(X_{\min} > x) = \prod_{i=1}^{n} \mathbb{P}(X_i > x) = \prod_{i=1}^{n} (1 - F_i(x))$$

y así:

$$F_{\min}(x) = 1 - \prod_{i=1}^{n} (1 - F_i(x)).$$

Ejemplo 1.7 Encuentra la distribución del tiempo hasta que termine la primera de las tres tareas en el sistema informático multiprocesador del ejemplo anterior. La variable aleatoria X que describe esta situación estÃ¡ dada por $X = \min\{X_1, X_2, X_3\}$ y su función de distribución acumulativa está dada por:

$$F_X(t) = 1 - e^{-t/4}e^{-t/4}e^{-t/4} = 1 - e^{-3t/4},\ t \geq 0.$$

Así, el tiempo medio para que la primera tarea se termine es igual a $4/3$, que es considerablemente menor que la media de cualquiera de las tareas individuales.

Obsérvese, en el ejemplo anterior, que la distribución obtenida como mínimo del conjunto de tres distribuciones exponenciales es en si misma una distribución exponencial. Esto es válido en general: a diferencia del máximo de un conjunto de distribuciones exponenciales, el mínimo de un conjunto de distribuciones exponenciales es en si misma una distribución exponencial.

2 Distribuciones condicionales

Si X y Y son discretas, con función de masa de probabilidad conjunta $p_{XY}(x, y)$, entonces la `función de masa de probabilidad condicional` de X dado que el evento $Y = y$ se define como:

$$p_{Y|X}(y|x) = \mathbb{P}(Y = y|X = x) = \frac{\mathbb{P}(Y = y, X = x)}{\mathbb{P}(X = x)} = \frac{p_{X,Y}(x, y)}{p_Y(y)} \tag{2}$$

si se cumple que $p_X(x) > 0$. En otras palabras, elegimos un valor específico para y y encontramos $\mathbb{P}(X = x|Y = y)$ como función de x.

De la ecuación anterior, se tiene el siguiente resultado:

$$p_{X|Y}(x, y) = p_Y(y)p_{X|Y}(x|y) \text{ para } p_Y(y) > 0.$$

De manera similar se cumple:

$$p_{X|Y}(x, y) = p_X(x)p_{Y|X}(y|x) \text{ para } p_X(x) > 0.$$

De esta manera, podemos calcular la función de masa de probabilidad conjunta de X e Y, a partir de sus distribuciones marginales y sus funciones de masa de probabilidad condicional. El procedimiento se aplica si las variables aleatorias son independientes o no. Si son independientes, entonces:

$$p_{X|Y}(x|y) = p_X(x) \text{ y } p_{Y|X}(y|x) = p_Y(y)$$

y $p_{X|Y}(x,y) = p_X(x)p_Y(y)$.

También las probabilidades marginales, se pueden formar como:

$$p_X(x) = \sum_{\text{todo y}} p_{XY}(x,y) = \sum_{\text{todo y}} p_{X|Y}(x|y)p_Y(y),$$

que es otra forma del teorema de probabilidad total. Continuando con el caso discreto de dos variables aleatorias discretas, podemos formar la función de distribución acumulativa condicional como:

$$F_{X|Y}(x|y) = \mathbb{P}(X \leq x|Y = y) = \frac{\mathbb{P}(X \leq x, Y = y)}{\mathbb{P}(Y = y)}$$

para todos los posibles valores de x e y para el cual $\mathbb{P}(Y = y) > 0$. Esta función también puede ser obtenida, desde la función de masa de probabilidad condicional como:

$$F_{X|Y}(x|y) = \frac{\sum_{u \leq x} p_{XY}(u,y)}{p_Y(y)} = \sum_{u \leq x} p_{X|Y}(u|y).$$

Ejemplo 2.1 Sean X e Y dos variables aleatorias discretas con función de masa de probabilidad conjunta:

	X = -1	X = 0	X=1	$p_Y(y)$
Y = -1	1/12	3/12	1/12	5/12
Y = 0	1/12	0/12	1/12	2/12
Y = 1	1/12	3/12	1/12	5/12
$p_X(x)$	3/12	6/12	3/12	

Calculemos $\mathbb{P}(X = x|Y = 1)$ tenemos:

$$\begin{aligned}
\mathbb{P}(X = -1|Y = 1) &= \mathbb{P}(X = -1, Y = 1)/\mathbb{P}(Y = 1) = (1/12)/(5/12) = 1/5, \\
\mathbb{P}(X = 0|Y = 1) &= \mathbb{P}(X = 0, Y = 1)/\mathbb{P}(Y = 1) = (3/12)/(5/12) = 3/5, \\
\mathbb{P}(X = 1|Y = 1) &= \mathbb{P}(X = 1, Y = 1)/\mathbb{P}(Y = 1) = (1/12)/(5/12) = 1/5.
\end{aligned}$$

Obsérvese que para cada valor diferente que puede asumir Y, obtenemos una función de masa de probabilidad diferente $p_{XY}(x|y)$. En el ejemplo (2.1), se pueden obtener tres funciones de densidad de probabilidad diferentes correspondientes a $y = -1, 0, 1$, sólo la última $(Y = 1)$ fue calculada.

Calculamos la función de distribución acumulativa condicional de X, dado $Y = 1$:

$$F_{X|Y}(x|1) = \begin{cases} 0 & x < -1 \\ 12/5 \times 1/12 = 1/5 & -1 \leq x < 0 \\ 12/5 \times (1/12 + 3/12) = 4/5 & 0 \leq x < 1 \\ 12/5 \times (1/12 + 3/12 + 1/12) = 1, & x \geq 1. \end{cases}$$

Para distribuciones continuas, no podemos calcular $\mathbb{P}(Y \leq y|X = x)$, de la fórmula $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B$, desde que $\mathbb{P}(B) = 0$(estamos condicionando el evento $\{X = x\}$ que tiene probabilidad 0). En lugar de condicionar el evento $X = x$, condicionamos el evento $x \leq X \leq x + \delta$ y tomamos límite $\delta x \to 0^+$. Así:

$$\mathbb{P}(Y \leq y|x \leq X \leq x+\delta) = \frac{\mathbb{P}(Y \leq y, x \leq X \leq x+\delta)}{\mathbb{P}(x \leq X \leq x+\delta)}$$
$$= \frac{\displaystyle\int_{u=x}^{x+\delta x} \int_{v=-\infty}^{y} f_{X,Y}(u,v)dudv}{\displaystyle\int_{x}^{x+\delta x} f_X(u)du}$$

Dividiendo el numerador y el denominados entre δx y tomando límite cuando $\delta x \to 0^+$, obtenemos:

$$\mathbb{P}(Y \leq y|x \leq X \leq x+\delta) \to \int_{-\infty}^{y} \frac{f_{X,Y}(x,y)}{f_X(x)}dv$$
$$= G(y)$$

G es una función de distribución con una función densidad:

$$g(y) = \frac{f_{X,Y}(x,y)}{f_X(x)} \quad \text{para } y \in \mathbb{R}.$$

Llamamos a G y g la `función de distribución condicional` y la `función densidad condicional` de Y dado que X es igual a x, para valores de x tal que $f_X(x) > 0$.

Definición 2.1 La `función densidad condicional` de Y dado que $X = x$, denotada por $f_{Y|X}(\cdot|x)$ es definida como:

$$f_{Y|X}(y|x) = \frac{f_{X|Y}(x,y)}{f_X(x)} \tag{3}$$

para $y \in \mathbb{R}$ y x que cumplen que $f_X(x) > 0$.

Enfatizamos que expresiones como $\mathbb{P}(Y \leq y|X = x)$ no se puede ser interpretada de la manera usual, usando la fórmula para $\mathbb{P}(A|B)$. La única forma de darle significado a esta cantidad es hacer una nueva definición de $\mathbb{P}(Y \leq y|X = x)$ como la función distribución condicional $G(y)$ de Y dado $X = x$.

Si X e Y son independientes, entonces $f_{X,Y}(x,y) = f_X(x)f_Y(y)$. Por definición $f_{Y|X}(y|x) = f_Y(y)$, lo que dice que es irrelevante la información de X, cuando se estudia Y.

Ejemplo 2.2 Sean X y Y variables aleatorias, con distribución uniforme sobre un cuadrado unidad. Así se cumple que $f_{X|Y}(x|y) = 1$, para $0 \leq x \leq 1$ y 0 en otros casos. Dado $Y = y$, X tienen una distribución uniforme en (0.1), se puede escribir como $X|Y = y \sim$ Uniforme$(0,1)$.

Ejemplo 2.3 Supongamos que $X \sim$ Uniforme$(0,1)$. Después de obtener un valor de X, generamos $Y|X = x \sim$ Uniforme$(x,1)$. La distribución marginal de Y se puede calcular a partir de:

$$f_X(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otros casos.} \end{cases}$$

luego por definición tenemos:

$$f_{Y|X}(y|x) = \begin{cases} \dfrac{1}{1-x} & \text{si } 0 < x < y < 1 \\ 0 & \text{en otros casos.} \end{cases}$$

Así :

$$f_{X,Y}(x,y) = f_{Y|X}(y|x)f_X(x) = \begin{cases} \dfrac{1}{1-x} & \text{si } \ 0 < x < y < 1 \\ 0 & \text{en otros casos.} \end{cases}$$

La distribución marginal para Y es:

$$f_Y(y) = \int_0^y f_{X,Y}(x,y)dx = \int_0^y \frac{dx}{1-x} = -\int_1^{1-y} \frac{du}{u} = -\log(1-y)$$

para $0 < y < 1$.

Ejemplo 2.4 Sea:

$$f(x,y) = \begin{cases} x+y & \text{si } \ 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en otros casos.} \end{cases}$$

Encontremos $\mathbb{P}(X < 1/4|Y = 1/3)$. Como $f_Y(y) = y + (1/2)$, tenemos:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{x+y}{y+\frac{1}{2}}$$

Así:

$$\begin{aligned} \mathbb{P}\Big(X < \frac{1}{4}\Big|Y = \frac{1}{3}\Big) &= \int_0^{1/4} f_{X|Y}\Big(x \ \Big| \ \frac{1}{3}\Big)dx \\ &= \int_0^{1/4} \frac{x+\frac{1}{3}}{\frac{1}{3}+\frac{1}{2}}dx = \frac{\frac{1}{32}+\frac{1}{12}}{\frac{1}{3}+\frac{1}{2}} = \frac{11}{80}. \end{aligned}$$

Ejemplo 2.5 Sean X e Y variables aleatorias con una función densidad conjunta dada por:

$$f_{XY}(x,y) = \begin{cases} 2e^{-x-y} & \text{si } \ 0 < x < y < \infty \\ 0 & \text{en otros casos} \end{cases}$$

Las funciones marginales, para cada variable aleatoria son:

$$\begin{aligned} f_X(x) &= 2e^{-2x} \qquad \text{para } x > 0 \\ f_Y(y) &= 2e^{-y}(1-e^{-y}) \quad \text{para } y > 0 \end{aligned}$$

La función densidad condicional de Y dado $X = x(> 0)$ es dada por:

$$f_{Y|X}(y|x) = \frac{2e^{-x-y}}{2e^{-2x}} = e^{x-y} \ \text{ para } \ y > x$$

La función densidad condicional de X dado $Y = y$, se puede expresar como:

$$f_{X|Y}(x|y) = \frac{e^{-x}}{1-e^{-y}} \ \text{ para } \ 0 < x < y.$$

En ambos casos, la funciones densidades son iguales a 0 si $x > y$.

Ejemplo 2.6 Calcula la función densidad de probabilidad marginal condicional $f_{X|Y}(x|Y=y)$ desde la siguiente función de densidad conjunta de dos variables aleatorias X e Y:

$$f_{X,Y}(x,y) = \begin{cases} e^{-y}, & 0 < x < y < \infty \\ 0 & \text{en otros casos.} \end{cases}$$

Primero calculemos la función de densidad de probabilidad marginal. Tenemos:

$$f_X(x) = \int_x^{\infty} e^{-y} dy = -e^{-y}\Big|_x^{\infty} = e^{-x}.$$

desde que $f_X(x) > 0$, sólo cuando $y > x$. La otra función marginal es más simple. Tenemos:

$$f_Y(y) = \int_0^{y} e^{-y} dx = ye^{-y}.$$

Formemos ahora la distribución marginal condicional $f_{Y|X}(x|Y=y)$. Desde la ecuación (3),

$$f_{Y|X}(x|Y=y) = \frac{e^{-y}}{ye^{-y}} = \frac{1}{y}, \text{ para } 0 < x < y.$$

Esta función densidad condicional es uniformemente distribuida en $(0.y)$.

Correspondiendo a la ecuación (2), podemos reescribir a ecuación (3) como:

$$f_{X,Y}(x,y) = f_Y(y) f_{X|Y}(x|y) = f_X(x) f_{Y|X}(y|x) \text{ para } 0 < f_Y(y) < \infty,\ 0 < f_X(x) < \infty.$$

Cuando X e Y son variables aleatorias independientes, entonces:

$$f_{X,Y}(x,y) = f_X(x) f_Y(y).$$

y de esto se sigue que:

$$f_{Y|X}(y|x) = f_Y(y).$$

Ejemplo 2.7 Sea X uniformemente distribuida en $[0,1]$ y sea Y uniformemente distribuida en $[0,X]$. Calculamos la función de densidad de probabilidad conjunta de X e Y. La declaración del problema nos permite escribir:

$$f_X(x) = 1{,}0 \leq x \leq 1,$$
$$f_{Y|X}(y|x) = 1/x{,}0 \leq y \leq x.$$

Por tanto, usando (4), obtenemos la función de densidad de probabilidad condicional, como:

$$f_{X,Y}(x,y) = f_X(x) f_{Y|X}(y|x) = 1 \times 1/x = 1/x, \quad 0 \leq y \leq x \leq 1.$$

Hay una serie de otros resultados que pueden obtenerse de la definición de la función de densidad conjunta. Por ejemplo, podemos calcular la densidad marginal de X en términos de la densidad condicional como:

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y) dy,$$

lo cuál es el análogo al teorema de la probabilidad total. Podemos también obtener, una analogía a la Regla de Bayes como:

$$f_{Y|X}(y|x) = \frac{f_Y(y) f_{X|Y}(x|y)}{\displaystyle\int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x|y) dy}.$$

Finalmente,desde la ecuación (3), definimos la función de distribución acumulativa condicional como:

$$F_{X|Y}(x|y) = \mathbb{P}(X \leq x | Y = y) = \frac{\displaystyle\int_{-\infty}^{x} f(u,y) du}{f_Y(y)} = \int_{-\infty}^{x} f_{X|Y}(u|y) du.$$

Ejemplo 2.8 Retornamos el ejemplo (2.7), calculamos la distribución marginal de Y y la función de densidad probabilidad condicional de X dado Y. En efecto:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_{y}^{1} \frac{1}{x} dx = -\ln y.$$

y

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \frac{1/x}{-\ln y},\ y \leq x \leq 1.$$

2.1 Esperanza condicional y varianza

En el caso de esperanza condicional, supongamos que X e Y son variables aleatorias continuas conjuntas, con densidad $f_{X,Y}$y que se nos dá $X = x$. A la luz de esta información , la nueva función densidad de Y es la función densidad condicional $f_{Y|X}(\cdot|x)$.

Definición 2.2 La `esperanza condicional de` Y `dado` $X = x$ escrita como $\mathbb{E}(Y|X = x)$ es de la media de la función densidad condicional:

$$\mathbb{E}(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_{-\infty}^{\infty} y \frac{f_{X,Y}(x,y)}{f_X(x)} dy,$$

válido para algún valor de x que cumple $f_X(x) > 0$. De manera analóga, dada una función de densidad condicional conjunta $f_{X|Y}(x|y)$ de las variables aleatorias X e Y, podemos formar la esperanza condicional de X dado Y como:

$$\mathbb{E}(X|Y = y) = \int_{-\infty}^{\infty} y f_{X|Y}(x|y) dx = \int_{-\infty}^{\infty} x \frac{f_{X,Y}(x,y)}{f_Y(y)} dy,\ \ 0 < f_Y(y) < \infty.$$

Más generalmente, la esperanza condicional de una función h de X e Y viene dada por:

$$\mathbb{E}[h(X,Y)|Y = y] = \int_{-\infty}^{\infty} h(x,y) f_{X|Y}(x|y) dx.$$

Cuando X e Y son variables aleatorias discretas, tenemos:

$$\mathbb{E}(X|Y = y) = \sum_{\text{todo x}} x \mathbb{P}(X = x | Y = y) = \sum_{\text{todo x}} x p_{X|Y}(x|y),$$

y si h es una función de X e Y, entonces:

$$\mathbb{E}[h(X|Y)|Y=y] = \sum_{\text{todo x}} h(x,y)p_{X|Y}(x|y).$$

Posiblemente la aplicación más útil de la esperanza condicional sea el siguiente teorema, una variante del teorema de la partición que nos permite calcular $\mathbb{E}(Y)$ en situaciones donde la esperanza condicional $\mathbb{E}(Y|X=x)$ son fácilmente calculables.

Teorema 2.1 Si X e Y son variables aleatorias continuas conjuntas, entonces:

$$\mathbb{E}(Y) = \int \mathbb{E}(Y|X=x)f_X(x)dx$$

donde la integral está definida para los valores de x, tal que $f_X(x) > 0$. En otras palabras, al calcular $\mathbb{E}(Y)$ podemos fijar primero el valor de X y luego promediar sobre este valor más tarde.

En efecto:

$$\begin{aligned}
\mathbb{E}(Y) = \int yf_Y(y)dy &= \int\int yf_{X,Y}(x,y)dxdy \\
&= \int\int yf_{Y|X}(y|x)f_X(x)dxdy \\
&= \int\Big(\int yf_{Y|X}(y|x)dy\Big)f_X(x)dx
\end{aligned}$$

Desde que $\mathbb{E}(Y|X)$ es una variable aleatoria, tiene una media $\mathbb{E}[\mathbb{E}(Y|X)]$ que puede ser calculada usando las siguientes propiedades:

$$\mathbb{E}[\mathbb{E}(Y|X)] = \begin{cases} \sum_x \mathbb{E}(Y|X=x)p_X(x) & \text{X discreta} \\ \displaystyle\int_{-\infty}^{\infty} \mathbb{E}(Y|X=x)f_X(x) & \text{X continua.} \end{cases}$$

Teorema 2.2 Este resultado llamado `ley de iteración de esperanzas` se cumple para toda variable aleatoria:

$$\mathbb{E}[\mathbb{E}(Y|X)] = \mathbb{E}(Y) \text{ y } \mathbb{E}[\mathbb{E}(X|Y)] = \mathbb{E}(X)$$

De forma general, para alguna función $r(X,Y)$, tenemos:

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

Probemos la primera ecuación:

$$\begin{aligned}
\mathbb{E}(\mathbb{E}(Y|X)) = \int \mathbb{E}(Y|X=x)f_X(x)dx &= \int\int yf(y|x)dyf(x)dx \\
&= \int\int yf(y|x)f(x)dxdy = \int\int yf(x,y)dxdy = \mathbb{E}(Y)
\end{aligned}$$

Ejemplo 2.9 Sea una barra de longitud l. Dividimos la barra en un punto que se elige al azar y de manera uniforme en su longitud, y mantenemos la pieza que contiene el extremo izquierdo de la barra. A continuación, repetimos el mismo proceso con la pieza que nos quedó. ¿Cuál es la esperanza de la longitud prevista de las piezas que quedan después de romper dos veces la pieza?

Sea Y la longitud de la pieza de la barra, después de que se rompe por primera vez. Sea X la longitud después de que se rompe la barra por segunda vez. Asi tenemos $\mathbb{E}(X|Y) = Y/2$, desde que el punto de ruptura es escogido de manera uniforme sobre una pieza de longitud Y. De manera similar, tenemos que $\mathbb{E}(Y) = l/2$. Así:

$$\mathbb{E}(X) = \mathbb{E}[(X|Y)] = \mathbb{E}(Y/2) = \frac{\mathbb{E}(Y)}{2} = \frac{l}{4}.$$

Definición 2.3 Para dos variables aleatorias X e Y, la varianza condicional de Y dado $X = x$ es definida como:

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

o por la siguiente fórmula equivalente, que es algunas veces más útil:

$$\mathbb{E}(Y|X) = \mathbb{E}(Y^2|X) - (\mathbb{E}(Y|X))^2.$$

También se tiene:

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

Ejemplo 2.10 Sean X y Y dos variables aleatorias continuas, cuya función de probabilidad de densidad conjunta es dada por:

$$f_{X,Y}(x,y) = \begin{cases} 1 - (x+y)/3, & 0 \le x \le 1,\ 0 \le y \le 2 \\ 0 & \text{en otros casos.} \end{cases}$$

Encontremos $\mathbb{E}(X|Y = 0.5)$ y $\mathbb{E}(Y|X = 0.5)$. Esto requiere un conocimiento de la distribución marginal de X e Y, como también las distribuciones marginales, así que empezamos calculando esas cantidades. Tenemos :

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy = \int_{y=0}^{2} [1 - (x+y)/3]dy = 4/3 - 2x/3, \quad 0 \le x \le 1,$$

y

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int_{x=0}^{1} [1 - (x+y)/3]dx = 5/6 - y/3, \quad 0 \le y \le 2.$$

También:

$$f_{X|Y}(x|0.5) = \frac{f_{X,Y}(x,0.5)}{f_Y(0.5)} = \frac{1 - x/3 - 0.5/3}{5/6 - 0.5/3} = 5/4 - x/2,$$
$$f_{Y|X}(y|0.5) = \frac{f_{X,Y}(0.5,y)}{f_X(0.5)} = \frac{1 - 0.5/3 - y/3}{4/3 - 1/3} = 5/6 - y/3.$$

Los resultados requeridos pueden formarse ahora:

$$\mathbb{E}(X|Y=0.5) = \int_{-\infty}^{\infty} x f_{X|Y}(x|0.5)dx = \int_0^1 x\left(\frac{5}{4} - \frac{x}{2}\right)dx \left(\frac{5x^2}{8} - \frac{x^3}{6}\right)\Bigg|_0^1 = 11/24,$$

$$\mathbb{E}(Y|X=0.5) = \int_{-\infty}^{\infty} y f_{Y|X}(y|0.5)dy = \int_0^2 x\left(\frac{5}{6} - \frac{y}{3}\right)dy \left(\frac{5y^2}{12} - \frac{y^3}{9}\right)\Bigg|_0^2 = 7/9.$$

Ejemplo 2.11 Retornamos al ejemplo (1.4) y encontremos $\mathbb{E}(Y|x)$. Por conveniencia, reescribimos la función densidad conjunta y los cálculos anteriores de la densidad marginal de X. Estas eran:

$$f_{X,Y}(x,y) = \begin{cases} 6x, & 0 < x < y < 1, \\ 0 & \text{en otros casos.} \end{cases}$$

y

$$f_X(x) = 6x(x-1),\ 0 \leq x \leq 1,$$

respectivamente. Ahora calculamos $\mathbb{E}(Y|x)$ como:

$$\mathbb{E}(Y|x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy = \int_{y=x}^{1} \frac{6xy}{6x(1-x)}dy = \frac{1}{1-x}\int_{y=x}^{1} ydy = \frac{1+x}{2}.$$

En general, para un `vector aleatorio`, $X = (X_1, \ldots, X_n)$ donde $X_1, \ldots, X_n$ son variables aleatorias, es posible definir las funciones marginales, condicionales etc, de la misma manera que las distribuciones bivariadas. En este caso se denota la densidad del vector aleatorio como $f(x_1, \ldots, x_n)$.

Decimos que $X_1, \ldots, X_n$ son independientes si, para cada $A_1, \ldots, A_n$

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i).$$

Es suficiente verificar que $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$.

Si $X_1, \ldots X_n$ son independientes y cada uno de ellos tiene la misma distribución marginal con una función de distribución cumulativa F, decimos que $X_1, \ldots, X_n$ son idénticamente distribuidas e independientes y escribimos como:

$$X_1, \ldots, X_n \sim F$$

Si F tiene densidad f, escribimos también $X_1, \ldots, X_n \sim f$.