INTRODUCCIÓN A LA ESTADÍSTICA
EJERCICIOS RESUELTOS
PRÁCTICO 7

Andrea Rotnitzky — arotnitzky@utdt.edu
Daniela Cuesta — dcuesta@utdt.edu
Pablo M. Escobar — escobarpm@gmail.com

28 de junio de 2022

1. Cierto supermercado tiene una caja normal y una rápida. Llamá X_1 a la cantidad de clientes en la fila de la caja normal y X_2 a la cantidad de clientes en la fila rápida en cualquier momento a primera hora de la tarde. Suponé que la función de probabilidad conjunta de (X_1, X_2) está dada por:

		X_2			
		0	1	2	3
X_1	0	0.08	0.07	0.04	0
	1	0.06	0.15	0.05	0.04
	2	0.05	0.04	0.10	0.06
	3	0	0.03	0.04	0.07
	4	0	0.01	0.05	0.06

(a) Calculá las siguientes probabilidades, en cada caso expresando la fórmula de lo que estás calculando:
- que haya exactamente un cliente en cada línea de espera.
- que los números de clientes de las dos líneas de espera sean iguales.
- que haya por lo menos dos clientes más en una línea de espera que en la otra.
- que el número total de clientes de las dos líneas de espera sea exactamente cuatro.
- que el número total de clientes de las dos líneas de espera sea por lo menos cuatro.
- que el número de clientes que esperan en la caja normal sea mayor que dos si en la caja rápida hay un cliente.
- que el número de clientes que esperan en la caja rápida sea a lo sumo 2.

(b) Hallá la función de probabilidad marginal de la variable que registra la cantidad de clientes en espera en la caja rápida y separadamente halla la función de probabilidad marginal de la variable que registra la cantidad de clientes en la línea de espera de la caja lenta.

(c) Suponé que atender a un cliente en la línea de espera rápida lleva 3 minutos y atender a uno en la línea de espera lenta lleva 8 minutos. Hallá la función de probabilidad de masa del tiempo que llevará atender a todos los clientes que se encuentran en línea de espera en un momento dado a primera hora de la tarde y calculá su esperanza.

(d) Hallá la función de probabilidad de masa condicional $p_{X_1|X_2}(x|X_2 = 2)$ y expresá con palabras simples el significado de esta función.

Solution:

(a)

- $p_{X_1X_2}(1;1) = 0{,}15$
- $P(X_1 = X_2) = \sum_{k=0}^{3} p_{X_1X_2}(k;k) = 0{,}08 + 0{,}15 + 0{,}1 + 0{,}07 = 0{,}4$
- $$\begin{aligned} P(|X_1 - X_2| \geq 2) &= P(X_1 \geq X_2 + 2 \quad ó \quad X_1 + 2 \leq X_2) \\ &= 0{,}05 + 0 + 0 + 0{,}03 + 0{,}01 + 0{,}05 + 0{,}04 + 0 + 0{,}04 = 0{,}22 \end{aligned}$$
- $P(X_1 + X_2 = 4) = 0 + 0{,}03 + +0{,}1 + 0{,}04 = 0{,}17$
- $P(X_1 + X_2 \geq 4) = 0{,}17 + P(X_1 + X_2 \geq 5) = 0{,}17 + \sum_{i=5}^{7} P(X_1 + X_2 = i) = 0{,}46$
- $$\begin{aligned} P(X_1 > 2 \mid X_2 = 1) &= P(X_1 = 3 \mid X_2 = 1) + P(X_1 = 4 \mid X_2 = 1) \\ &= \frac{p_{X_1X_2}(3;1)}{p_{X2}(1)} + \frac{p_{X_1X_2}(4;1)}{p_{X2}(1)} \\ &= \frac{0{,}03}{0{,}3} + \frac{0{,}01}{0{,}3} \\ &= 0{,}1333 \end{aligned}$$
- Una vez resuelto el ítem b)
$$P(X_2 \leq 2) = 1 - p_{X_2}(3) = 1 - 0{,}23 = 0{,}77$$

(b) Las probabilidades marginales para la variable X_1 y X_2 pueden ser calculadas sumando las probabilidades conjuntas indicadas en la tabla por fila y por columna respectivamente:

$$p_{X_1}(k) = \sum_{l=0}^{3} p_{X_1X_2}(k;l) = \begin{cases} 0{,}19 & \text{si } k = 0 \\ 0{,}3 & \text{si } k = 1 \\ 0{,}25 & \text{si } k = 2 \\ 0{,}14 & \text{si } k = 3 \\ 0{,}12 & \text{si } k = 4 \\ 0 & \text{en otro caso.} \end{cases}$$

$$p_{X_2}(l) = \sum_{k=0}^{4} p_{X_1X_2}(k;l) = \begin{cases} 0{,}19 & \text{si } l = 0 \\ 0{,}3 & \text{si } l = 1 \\ 0{,}28 & \text{si } l = 2 \\ 0{,}23 & \text{si } l = 3 \\ 0 & \text{en otro caso.} \end{cases}$$

(c) Definimos la variable

$$\begin{aligned} T &= \text{“}\textit{Tiempo que demoran las cajas en atender a todos los clientes (en minutos)}.\text{”} \\ &= Max\{8 \cdot X_1, 3 \cdot X_2\} \end{aligned}$$

se tiene entonces que $R_T = \{0,\ 3,\ 6,\ 8,\ 9,\ 16,\ 24,\ 32\}$ y

$$p_T(t) = \begin{cases} p_{X_1X_2}(0;0) & \text{si } t = 0 \\ p_{X_1X_2}(0;1) & \text{si } t = 3 \\ p_{X_1X_2}(0;2) & \text{si } t = 6 \\ p_{X_1X_2}(1;0) + p_{X_1X_2}(1;1) + p_{X_1X_2}(1;2) & \text{si } t = 8 \\ p_{X_1X_2}(0;3) + p_{X_1X_2}(1;3) & \text{si } t = 9 \\ p_{X_1X_2}(2;0)p_{X_1X_2}(2;1) + p_{X_1X_2}(2;2) + p_{X_1X_2}(2;3) & \text{si } t = 16 \\ p_{X_1X_2}(3;0)p_{X_1X_2}(3;1) + p_{X_1X_2}(3;2) + p_{X_1X_2}(3;3) & \text{si } t = 24 \\ p_{X_1X_2}(4;0)p_{X_1X_2}(4;1) + p_{X_1X_2}(4;2) + p_{X_1X_2}(4;3) & \text{si } t = 32 \\ 0 & \text{en otro caso.} \end{cases}$$

$$= \begin{cases} 0{,}08 & \text{si } t = 0 \\ 0{,}07 & \text{si } t = 3 \\ 0{,}04 & \text{si } t = 6 \\ 0{,}26 & \text{si } t = 8 \\ 0{,}04 & \text{si } t = 9 \\ 0{,}25 & \text{si } t = 16 \\ 0{,}14 & \text{si } t = 24 \\ 0{,}12 & \text{si } t = 32 \\ 0 & \text{en otro caso.} \end{cases}$$

La esperanza queda

$$E(T) = 0{,}08 \cdot 0 + 0{,}07 \cdot 3 + 0{,}04 \cdot 6 + 0{,}26 \cdot 8 + 0{,}04 \cdot 9 + 0{,}25 \cdot 16 + 0{,}14 \cdot 24 + 0{,}12 \cdot 32 = 14{,}09$$

(d)

$$p_{X_1|X_2}(k \mid X_2 = 2) = p_{X_1|X_2=2}(k) = \frac{p_{X_1X_2}(k;2)}{P_{X_2}(2)} = \begin{cases} \frac{0{,}04}{0{,}28} & \text{si } k = 0 \\ \frac{0{,}05}{0{,}28} & \text{si } k = 1 \\ \frac{0{,}1}{0{,}28} & \text{si } k = 2 \\ \frac{0{,}04}{0{,}28} & \text{si } k = 3 \\ \frac{0{,}05}{0{,}28} & \text{si } k = 4 \\ 0 & \text{en otro caso.} \end{cases} = \begin{cases} 0{,}1429 & \text{si } k = 0 \\ 0{,}1786 & \text{si } k = 1 \\ 0{,}3571 & \text{si } k = 2 \\ 0{,}1429 & \text{si } k = 3 \\ 0{,}1785 & \text{si } k = 4 \\ 0 & \text{en otro caso.} \end{cases}$$

2. Un pequeño ferry transporta motos y autos desde Buenos Aires a Colonia. Suponé que el número de motos X y el número de autos Y transportados en un viaje son dos variables aleatorias con distribución conjunta

		Y		
		0	1	2
X	0	0.025	0.015	0.010
	1	0.050	0.030	0.020
	2	0.125	0.075	0.050
	3	0.150	0.090	0.060
	4	0.100	0.060	0.040
	5	0.050	0.030	0.020

Calculá la esperanza del ingreso de un viaje, si el precio del transporte de una moto es de 300 pesos y el de un auto es de 1000 pesos.

Solution: Sea la variable aleatoria $I = Ingreso\ dn\ un\ viaje\ (en\ pesos) = 300 \cdot X + 1000 \cdot Y$, por la linealidad de la esperanza se tiene que:

$$E(I) = E(300 \cdot X + 1000 \cdot Y) = 300 \cdot E(X) + 1000 \cdot E(Y),$$

luego para calcular $E(X)$ y $E(Y)$ calculamos primero las probabilidades marginales de X e Y sumando las probabilidades conjuntas de la tabla por fila y por columna respectivamente:

$$p_X(k) = \sum_{l=0}^{2} p_{XY}(k;l) = \begin{cases} 0{,}05 & \text{si } k = 0 \\ 0{,}1 & \text{si } k = 1 \\ 0{,}25 & \text{si } k = 2 \\ 0{,}3 & \text{si } k = 3 \\ 0{,}2 & \text{si } k = 4 \\ 0{,}1 & \text{si } k = 5 \\ 0 & \text{sino.} \end{cases}$$

$$p_Y(l) = \sum_{k=0}^{5} p_{XY}(k;l) = \begin{cases} 0{,}5 & \text{si } l = 0 \\ 0{,}3 & \text{si } l = 1 \\ 0{,}2 & \text{si } l = 2 \\ 0 & \text{sino.} \end{cases}$$

Las esperanzas dan como resultado:

$$E(X) = 0 \cdot 0{,}05 + 1 \cdot 0{,}1 + 2 \cdot 0{,}25 + 3 \cdot 0{,}3 + 4 \cdot 0{,}2 + 5 \cdot 0{,}1 = 2{,}8$$
$$E(Y) = 0 \cdot 0{,}5 + 1 \cdot 0{,}3 + 2 \cdot 0{,}2 = 0{,}7$$
$$E(I) = 300 \cdot E(X) + 1000 \cdot E(Y) = 300 \cdot 2{,}8 + 1000 \cdot 0{,}7 = 1540$$

3. El 60 % de los aspirantes a un cargo ejecutivo de mucha responsabilidad tienen el equilibrio emocional como para sobrellevar situaciones de mucho estrés. Clasificá como "adecuados" a aquellos con el adecuado equilibrio emocional. Como parte del proceso de selección el aspirante debe rendir un exámen que consta de 5 preguntas sobre cómo gestionar distintas situaciones de conflicto. Suponé que la probabilidad de que un aspirante adecuado conteste satisfactoriamente a una pregunta es 0,8 y la probabilidad de que un aspirante inadecuado conteste satisfactoriamente a una pregunta es 0.5. Suponé además que para una persona dada, las respuestas son independientes entre sí. Juan es un aspirante que rinde el examen. Definí las siguientes variables aleatorias. X : la variable Bernoulli que vale 1 si Juan es adecuado y 0 si no lo es, Y : número de respuestas satisfactorias.
   (a) Calulá la distribución conjunta de X e Y
   (b) Calculá la distribución marginal de Y
   (c) Calculá la distribución condicional de X dado $Y = y$ para todo $y \in \{0, 1, 2, 3, 4, 5\}$
   (d) Si Juan hubiese salido airoso de todas las restantes intancias del proceso de selección, y en el exámen hubiese respondido cuatro preguntas adecuadamente, ¿le ofrecerías el cargo? Y si hubiese respondido solo dos preguntas adecuadamente, ¿se lo ofrecerías? Justificá tu respuesta.

Solution:

(a) Tenemos que

$$X = \begin{cases} 1 & \text{si el aspirante es adecuado} \\ 0 & \text{si no es adecuado} \end{cases}$$

luego $X \sim Bernoulli(0{,}6)$. Además

$$Y \mid X = 0 \sim Bi(5; 0{,}5)$$
$$Y \mid X = 1 \sim Bi(5; 0{,}8)$$

Nos interesa calcular

$$p_{XY}(k,l) \quad \forall \quad 0 \le k \le 1; \ 0 \le l \le 5$$

Empecemos por el caso $k = 0$ y $l = 0$, usando la regla de multiplicación

$$p_{XY}(0,0) = P(Y = 0 \mid X = 0) \cdot P(X = 0) = \left[\binom{5}{0} \cdot 0{,}5^0 \cdot 0{,}5^5\right] \cdot 0{,}4 = 0,0125$$

Para el caso $k = 1$ y $l = 0$, repitiendo los pasos

$$p_{XY}(1,0) = P(Y = 0 \mid X = 1) \cdot P(X = 1) = \left[\binom{5}{0} \cdot 0{,}8^0 \cdot 0{,}2^5\right] \cdot 0{,}6 = 0,000192$$

Esto completaría la primer columna de la siguiente tabla (todas las demás se calculan de manera análoga):

		Y					
		0	1	2	3	4	5
X	0	0.0125	0.0625	0.125	0.125	0.0625	0.0125
	1	0.000192	0.00384	0.03072	0.12288	0.24576	0.196608

(b) Para calcular la distribución marginal de Y sumamos por columnas

$$p_Y(l) = \begin{cases} 0{,}012692 & \text{si } l = 0 \\ 0{,}06634 & \text{si } l = 1 \\ 0{,}15572 & \text{si } l = 2 \\ 0{,}24788 & \text{si } l = 3 \\ 0{,}30826 & \text{si } l = 4 \\ 0{,}209108 & \text{si } l = 5 \\ 0 & \text{en otro caso.} \end{cases}$$

(c) Calculemos el caso la función de probabilidad de masa de X cuando $Y = 0$:

$$p_{X|Y=0}(k) = \begin{cases} \frac{p_{XY}(0,0)}{p_Y(0)} & \text{si } k = 0 \\ \frac{p_{XY}(1,0)}{p_Y(0)} & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases} = \begin{cases} 0{,}98487236 & \text{si } k = 0 \\ 0{,}01512764 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

Repitiendo los pasos calculamos los casos para $1 \leq Y \leq 5$

$$p_{X|Y=1}(k) = \begin{cases} 0{,}94211637 & \text{si } k = 0 \\ 0{,}05788363 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

$$p_{X|Y=2}(k) = \begin{cases} 0{,}8027228 & \text{si } k = 0 \\ 0{,}1972772 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

$$p_{X|Y=3}(k) = \begin{cases} 0{,}5042763 & \text{si } k = 0 \\ 0{,}4957237 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

$$p_{X|Y=4}(k) = \begin{cases} 0{,}2027509 & \text{si } k = 0 \\ 0{,}7972491 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

$$p_{X|Y=5}(k) = \begin{cases} 0{,}05977772 & \text{si } k = 0 \\ 0{,}94022228 & \text{si } k = 1 \\ 0 & \text{en otro caso.} \end{cases}$$

(d) A partir del inciso anterior se puede afirmar que aproximadamente el 79.72 % de los aspirantes que responden correctamente 4 preguntas, resultan adecuados. Mientras que aproximadamente solo el 19.73 % de los que responden correctamente 2 preguntas lo serán.
En base a esto, sería sensato ofrecerle el cargo a Juan si fueran 4 sus preguntas contestadas correctamente. Pero si en cambio fueran solo 2, estaríamos enfrentando un riesgo alto al ofrecercelo.

4. En la ciudad X, la probabilidad de que un hombre y una mujer que conviven no tengan hijos es de 0.4. Los gerentes de un gran restaurante de la ciudad X, de las que las parejas son habitué, saben que los pedidos de cada uno de los comensales se distribuye de la siguiente manera:

Si tienen hijos:

		Hombre		
		Carne	Pastas	Vegetariano
Mujer	Carne	0.2	0.2	0.05
	Pastas	0.13	0.1	0.07
	Vegetariano	0.05	0.1	0.1

Si no tienen hijos:

		Hombre		
		Carne	Pastas	Vegetariano
Mujer	Carne	0.32	0.17	0.02
	Pastas	0.1	0.15	0.03
	Vegetariano	0.08	0.08	0.05

El menú de Carne representa un beneficio neto para el restaurante de 20 pesos, el menú de Pastas de 30 y el menú Vegetariano de 50. Por otro lado, si una pareja va con los hijos al restaurante, esto representa un beneficio adicional de 25 (asumimos que las parejas que tienen hijos van siempre con los hijos). Para una pareja al azar:

(a) Cuál es la probabilidad de que el hombre consuma el menú de Carne si la pareja no tiene hijos?

(b) Cuál es la probabilidad de que la mujer consuma el menú Vegetariano si la pareja tiene hijos?

(c) Cuál es la probabilidad de que la mujer consuma el menú de Pastas?

(d) Cuál es la probabilidad de que ambos consuman lo mismo dado que tienen hijos?

(e) Cuál es la probabilidad de que ambos consuman lo mismo?

(f) Cuál es la ganancia esperada para el restaurante?

Solution: Defino las variables

$$X = \begin{cases} 1 & \text{si el hombre consume el menú de carne} \\ 2 & \text{si el hombre consume el menú de pastas} \\ 3 & \text{si el hombre consume el menú vegetariano} \end{cases}$$

$$Y = \begin{cases} 1 & \text{si la mujer consume el menú de carne} \\ 2 & \text{si la mujer consume el menú de pastas} \\ 3 & \text{si la mujer consume el menú vegetariano} \end{cases}$$

$$S = \begin{cases} 1 & \text{si la pareja tiene hijos} \\ 0 & \text{si la pareja no tiene hijos} \end{cases}$$

(a)

$$\begin{aligned} P(X = 1 \mid S = 0) &= P(X = 1; Y = 1 \mid S = 0) + P(X = 1; Y = 2 \mid S = 0) + P(X = 1; Y = 3 \mid S = 0S = 0) \\ &= 0{,}32 + 0{,}1 + 0{,}08 = 0{,}5 \end{aligned}$$

(b)

$$\begin{aligned} P(Y = 3 \mid S = 1) &= P(Y = 3; X = 1 \mid S = 1) + P(Y = 3; X = 2 \mid S = 1) + P(Y = 3; X = 3 \mid S = 1) \\ &= 0{,}05 + 0{,}1 + 0{,}1 = 0{,}25 \end{aligned}$$

(c)

$$\begin{aligned} P(Y = 2) &= P(Y = 2 \mid S = 0) \cdot P(S = 0) + P(Y = 2 \mid S = 1) \cdot P(S = 1) \\ &= (0{,}1 + 0{,}15 + 0{,}03) \cdot 0{,}4 + (0{,}13 + 0{,}1 + 0{,}07) \cdot 0{,}6 = 0{,}292 \end{aligned}$$

(d)

$$\begin{aligned} P(X = Y \mid S = 1) &= P(X = 1; Y = 1 \mid S = 1) + P(X = 2; Y = 2 \mid S = 1) + P(X = 3; Y = 3 \mid S = 1) \\ &= 0{,}2 + 0{,}1 + 0{,}1 = 0{,}4 \end{aligned}$$

(e)

$$\begin{aligned} P(X = Y) &= P(X = Y \mid S = 0) \cdot P(S = 0) + P(X = Y \mid S = 1) \cdot P(S = 1) \\ &= (0{,}32 + 0{,}15 + 0{,}05) \cdot 0{,}4 + 0{,}4 \cdot 0{,}6 = 0{,}448 \end{aligned}$$

(f) Si G es la ganancia del restaurante, la esperanza la podríamos calcular conociendo la función de probabilidad de masa de G. Para lo cual primero sería necesario identificar su soporte. Por ejemplo, $65 \in R_G$, pues es la ganancia que se obtiene si tanto la mujer como el hombre piden el menú de carne y además tienen hijos, entonces:

$$P(G = 65) = P(X = 1; Y = 1; S = 1) = P(X = 1; Y = 1 \mid S = 1) \cdot P(S = 1) = 0{,}2 \cdot 0{,}6 = 0{,}12$$

Identificando el soporte de G y realizando los cálculos pertinentes se llega a que

Solution:

$$p_G(k) = \begin{cases} 0{,}128 & \text{si } k = 40 \\ 0{,}108 & \text{si } k = 50 \\ 0{,}06 & \text{si } k = 60 \\ 0{,}12 & \text{si } k = 65 \\ 0{,}04 & \text{si } k = 70 \\ 0{,}198 & \text{si } k = 75 \\ 0{,}044 & \text{si } k = 80 \\ 0{,}06 & \text{si } k = 85 \\ 0{,}06 & \text{si } k = 95 \\ 0{,}02 & \text{si } k = 100 \\ 0{,}102 & \text{si } k = 105 \\ 0{,}06 & \text{si } k = 125 \\ 0 & \text{en caso contrario.} \end{cases}$$

se concluye que

$$E(G) = 40 \cdot 0,128 + \ldots + 0,06 \cdot 125 = 74,1$$

5. Suponé que de los libros antiguos que hay en el mundo, el 10 % son raros. Suponé también que de los libros antiguos con tapa dorada, también el 10 % son raros. Suponé además que el 30 % de los libros antiguos tienen tapa dorada. Te enteraste que tengo un libro antiguo. Llama X a la variable aleatoria Bernoulli que vale 1 si mi libro es raro y llamá Y a la variable aleatoria Bernoulli que vale 1 si mi libro antiguo tiene tapa dorada. Calculá la distribución conjunta de X e Y y demostrá que X e Y son independientes. Suponé ahora que te cuento que vendí mi libro a un librería de libros antiguos que exhibe un libro en la vidriera si y solo si el libro es raro o tiene tapa dorada. Llamá Z a la variable aleatoria Bernoulli que vale 1 si mi libro será exhibido en la vidriera y vale 0 de otro modo. Calculá la probabilidad de éxito de Z. ¿Son X e Y independientes dado $Z = 1$? **Nota:** este problema ilustra el así llamado "*sesgo de Berkson*"que en algunas disciplinas también lo denominan "*sesgo de selección*". Este sesgo ocurre cuando dos variables que no tienen relación entre sí (en nuestro ejemplo X e Y) se vuelven dependientes si nos enteramos - es decir, al condicionar en - otra variable aleatoria (en nuestro ejemplo, Z).

Solution: Contamos con los siguientes datos

$$\begin{aligned} P(X = 1) &= 0{,}1 \\ P(X = 1 \mid Y = 1) &= 0{,}1 \\ P(Y = 1) &= 0{,}3 \end{aligned}$$

a partir de los cuales podemos deducir

- $P(X = 0) = 1 - P(X = 1) = 0{,}9$
- $P(X = 1; Y = 1) = P(X = 1 \mid Y = 1) \cdot P(Y = 1) = 0,03$
- $P(Y = 0) = 1 - P(Y = 1) = 0,7$

Volcando estos datos en una tabla

		Y		
		0	1	p_X
X	0	a	b	0,9
	1	c	0.03	0,1
	p_Y	0,7	0,3	

donde a, b y c deberán cumplir con la condición de sumar por filas y columnas los valores de las probabilidades marginales. Quedando entonces que

		Y		
		0	1	p_X
X	0	0,63	0,27	0,9
	1	0,07	0.03	0,1
	p_Y	0,7	0,3	

Vemos entonces que las variables son independientes porque se cumple

$$p_{XY}(k,l) = p_X(k) \cdot p_Y(l) \quad \forall \quad k \in \{0;1\},\ l \in \{0;1\}$$

Calculemos ahora la probabilidad de que el libro sea exhibido

$$P(Z = 1) = 1 - P(X = 0; Y = 0) = 1 - 0,63 = 0,37$$

Para que X e Y sean independientes dado $Z = 1$ debería ocurrir que

$$P(X = k; Y = l \mid Z = 1) = P(X = k \mid Z = 1) \cdot P(Y = l \mid Z = 1) \quad \forall \quad k \in \{0;1\},\ l \in \{0;1\}$$

sin embargo tomando $k = 0$ y $l = 0$ se tiene

$$0 \neq P(X = 0 \mid Z = 1) \cdot P(Y = 0 \mid Z = 1)$$

por lo que no hay independencia.

6. El número de emergencias médicas que sufre un alumno de una cierta universidad durante el período lectivo es una v.a. Pois(λ). Supongamos que hay dos tipos de alumnos, aquellos mas propensos a sufrir emergencias médicas, para los cuales $\lambda = 2$ y aquellos poco propensos a las emergencias médicas, para las cuales $\lambda = 0{,}2$. El 10 % de los alumnos son del tipo de los propensos a sufrir emergencias médicas. Si cada asistencia a una emergencia médica tiene un costo para la universidad de 2000 pesos, ¿cuál es el gasto esperado en asistencia medica para cada alumno por ciclo lectivo? ¿Cuál es el desvío estandard del gasto por alumno y período? ¿Cuál es la probabilidad de que un nuevo ingresante no ocasione ningún gasto en emergencias médicas durante el próximo ciclo lectivo? ¿Cuál es la probabilidad de que ocasione exactamente 4000 pesos en gastos por emergencias médicas? Para resolver este problema, defini claramente las variables aleatorias que usas.

Solution: Consideramos la variable $X \sim Bernoulli(0,1)$ definida como

$$X = \begin{cases} 1 & \text{si el alumno es propenso a sufrir emergencias médicas} \\ 0 & \text{si no es propenso} \end{cases}$$

Sabemos que la cantidad de emergencias médicas depende del tipo de alumno de la siguiente manera

$$Y \mid X = 0 \sim Poi(0{,}2)$$
$$Y \mid X = 1 \sim Poi(2)$$

y que el gasto en pesos durante un período lectivo por alumno será $G = 2000 \cdot Y$.
Nos interesa calcular $E(G) = 2000 \cdot E(Y)$ y para calcular $E(Y)$ usaremos que

$$E(Y) = E(E(Y \mid X)) = E(g(X))$$

con $g(x) = E(Y \mid X = x)$ para $x \in R_X$.
Luego como $Y \mid X = x$ es en ambos casos una Poisson

$$g(x) = \begin{cases} 0{,}2 & \text{si } x = 0 \\ 2 & \text{si } x = 1 \end{cases}$$

y la función de probabilidad de $g(X)$ es

$$\begin{aligned} p_{g(X)}(k) = P(g(X) = k) &= \begin{cases} P(X = 0) & \text{si } k = 0{,}2 \\ P(X = 1) & \text{si } k = 2 \end{cases} \\ &= \begin{cases} 0{,}9 & \text{si } k = 0{,}2 \\ 0{,}1 & \text{si } k = 2 \end{cases} \end{aligned}$$

se tiene entonces

$$E(Y) = E(g(X)) = 0{,}9 \cdot 0{,}2 + 0{,}1 \cdot 2 = 0{,}38$$

y por lo tanto

$$E(G) = 2000 \cdot 0{,}38 = 760$$

Para el cálculo del desvío nos ocupamos primero de la varianza

$$Var(G) = Var(2000 \cdot Y) = 2000^2 \cdot Var(Y)$$

para lo cual usaremos la fórmula

$$Var(Y) = Var(E(Y \mid X)) + E(Var(Y \mid X))$$

El primer sumando es

$$Var(g(X)) = E(g^2(X)) - E^2(g(X)) = (0{,}9 \cdot 0{,}2^2 + 0{,}1 \cdot 2^2) - 0{,}38^2 = 0{,}2916$$

y para el segundo sumando debemos tener en cuenta que $Var(Y \mid X) = h(X)$ con $h(x) = Var(Y \mid X = x)$ para $x \in R_X$. Pero como $Y \mid X = x$ es en ambos casos una Poisson resulta que entonces $h(x) = g(x)$ y por lo tanto

$$E(Var(Y \mid X)) = E(g(X)) = 0{,}38$$

Obtenemos entonces

$$Var(Y) = 0{,}2916 + 0{,}38 = 0{,}6716$$

y el desvío del gasto queda

$$\sigma_G = \sqrt{2000^2 \cdot Var(Y)} = 2000 \cdot \sqrt{0{,}6716} = 1639{,}024$$

Para que un estudiante no ocasione gastos debe ocurrir que $Y = 0$, usando probabilidad total

$$\begin{aligned} P(Y = 0) &= P(Y = 0 \mid X = 0) \cdot P(X = 0) + P(Y = 0 \mid X = 1) \cdot P(X = 1) \\ &= \frac{e^{-0{,}2} \cdot 0{,}2^0}{0!} \cdot 0{,}9 + \frac{e^{-2} \cdot 2^0}{0!} \cdot 0{,}1 \\ &= 0{,}7503912 \end{aligned}$$

Para que un estudiante genere gastos por 4000 pesos debe ocurrir que $Y = 2$, usando probabilidad total

$$\begin{aligned} P(Y = 2) &= P(Y = 2 \mid X = 0) \cdot P(X = 0) + P(Y = 2 \mid X = 1) \cdot P(X = 1) \\ &= \frac{e^{-0{,}2} \cdot 0{,}2^2}{2!} \cdot 0{,}9 + \frac{e^{-2} \cdot 2^2}{2!} \cdot 0{,}1 \\ &= 0{,}04180421 \end{aligned}$$

7. Suponé que el número de visitantes entre las 9 y 10 de la mañana a una pagina web de una aerolínea es una variable aleatoria Pois(50000) . Suponé que la probabilidad de que un visitante a la página efectúe una compra de pasaje es 0.2.Asumiendo que los visitantes a la página web no están relacionados entre sí, calculá la esperanza y el desvío estándar del número de compras de pasaje entre las 9 y 10 de la mañana. Para resolver este problema definí claramente las variables aleatorias que considerás e indicá que resultados teóricos usás en tu resolución.

Solution: Sean

$$N = \text{“}\mathit{Cantidad\ de\ visitantes\ en\ la\ página\ web}\text{”}$$

$$X_i = \begin{cases} 1 & \text{si el visitante i-ésimo compra} \\ 0 & \text{caso contrario} \end{cases}$$

se tiene entonces que $N \sim Poiss\,(50000)$ y $X_i \sim Bernoulli(0{,}2)$, todas independientes entre sí. Si llamamos C = “*Cantidad de compras*” tendremos que

$$C = \sum_{i=i}^{N} X_i$$

Para calcular $E(C)$ nos gustaría utilizar la linealidad de la esperanza, pero se nos presenta el problema de que la cantiadad de términos en la suma no es una cantidad fija, sino una cantidad aleatoria. Luego con el objetivo de obtener una suma con una cantidad fija de términos (para luego poder aplicar la propiedad de la linealidad de la esperanza) vamos a condicionar C de la siguiente forma:

$$C \mid (N = n) = \sum_{i=1}^{n} X_i,$$

y de esta manera:

$$E(C \mid N = n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} 0{,}2 = 0{,}2 \cdot n$$

Ahora bien, por la ley de esperanza total es sabido que $E(C) = E\big(E(C \mid N)\big)$, donde:

$$E(C \mid N) = g(N) \quad con \quad g(n) = E(Y \mid N = n) = 0{,}2 \cdot n.$$

Entonces ya habiendo calculado $E(C \mid N) = g(N) = 0{,}2 \cdot N$ se concluye que:

$$E(C) = E\big(E(C \mid N)\big) = E(0{,}2 \cdot N) = 2 \cdot E(N) = 0{,}2 \cdot 50000 = 10000.$$

Para el cálculo del desvío, calculamos la varianza utilizando la fórmula:

$$Var(C) = E\big(Var(C \mid N)\big) + Var\big(E(C \mid N)\big)$$

El segundo sumando de la fórmula es:

$$Var\big(E(C \mid N)\big) = Var\big(0{,}2 \cdot N\big) = 0{,}2^2 \cdot Var(N) = 0{,}04 \cdot 50000 = 2000.$$

Para el primer sumando necesitamos calcular $Var(C \mid N)$ que se define como:

$$Var(C \mid N) = h(N) \quad con \quad h(n) = Var(C \mid N = n),$$

luego por la independencia que podemos asumir dado que los visitantes de la página no están relacionados entre sí, se tiene que:

$$Var(C \mid N = n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} 0{,}2 \cdot (1 - 0{,}2) = 0{,}16 \cdot n,$$

entonces $Var(C \mid N) = h(N) = 0{,}16 \cdot N$ y el primer sumando queda:

$$E\big(Var(C \mid N)\big) = E(0{,}16 \cdot N) = 0{,}16 \cdot E(N) = 0{,}16 \cdot 50000 = 8000.$$

Finalmente concluímos que:

$$Var(C) = E\big(Var(C \mid N)\big) + Var\big(E(C \mid N)\big) = 8000 + 2000 = 10000,$$

y por lo tanto

$$\sigma_C = \sqrt{10000} = 100$$

8. El número de clientes que entran a un comercio de prendas de vestir en un día dado es una variable aleatoria $Pois\,(30)$. Suponé que la cantidad de prendas que compra un cliente en una visita al comercio es una variable aleatoria Poiss(2) y que las decisiones de comprar o no una prenda de cada cliente no está influida por las compras de otros clientes ni por cuán lleno está el comercio.
   (a) ¿Cuáles son la esperanza y el desvío estándar del número de prendas vendidas en un día por el comercio?.
   (b) Si la ganancia neta de una compra es una variable aleatoria con media 100 y desvío estanadard 20 y es independiente del número de compras y de visitas al comercio ¿cuáles son la media y el desvío estandard de la ganancia neta diaria del comercio?

Solution: Los datos que tenemos del enunciado son:

$$N = \text{“}\textit{Cantidad de clientes que entran al negocio}\text{”} \sim Poiss(30)$$
$$X_i = \text{“}\textit{Cantidad de prendas que compra el } i-\textit{ésimo cliente}\text{”} \sim Poiss(2)$$

con todas las variables independientes entre sí.

(a) Definimos

$$Y = \text{“}\textit{Número de prendas vendidas}\text{”} = \sum_{i=1}^{N} X_i$$

Para calcular su esperanza vamos a condicionar Y de la siguiente forma:

$$Y \mid (N = n) = \sum_{i=1}^{n} X_i,$$

y de esta manera:

$$E(Y \mid N = n) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} 2 = 2 \cdot n$$

Ahora bien, por la ley de esperanza total es sabido que $E(Y) = E\big(E(Y \mid N)\big)$, donde:

$$E(Y \mid N) = g(N) \quad con \quad g(n) = E(Y \mid N = n) = 2 \cdot n.$$

Entonces ya habiendo calculado $E(Y \mid N) = g(N) = 2 \cdot N$ se concluye que:

$$E(Y) = E\big(E(Y \mid N)\big) = E(2 \cdot N) = 2 \cdot E(N) = 2 \cdot 30 = 60.$$

Para el cálculo del desvío, calculamos la varianza utilizando la fórmula:

$$Var(Y) = E\big(Var(Y \mid N)\big) + Var\big(E(Y \mid N)\big)$$

El segundo sumando de la fórmula es:

$$Var\big(E(Y \mid N)\big) = Var\big(2 \cdot N\big) = 2^2 \cdot Var(N) = 4 \cdot 30 = 120.$$

Para el primer sumando necesitamos calcular $Var(Y \mid N)$ que se define como:

$$Var(Y \mid N) = h(N) \quad con \quad h(n) = Var(Y \mid N = n),$$

luego por la independencia entre la cantidad de prendas compradas por cliente, se tiene que:

$$Var(Y \mid N = n) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} 2 = 2 \cdot n,$$

entonces $Var(Y \mid N) = h(N) = 2 \cdot N$ y el primer sumando queda:

$$E\big(Var(Y \mid N)\big) = E(2 \cdot N) = 2 \cdot E(N) = 2 \cdot 30 = 60.$$

Finalmente concluimos que:

$$Var(Y) = E\big(Var(Y \mid N)\big) + Var\big(E(Y \mid N)\big) = 60 + 120 = 180,$$

y por lo tanto

$$\sigma_Y = \sqrt{180} = 13{,}41641$$

(b) Sea Z = "*Ganancia neta por compra*", definimos

$$G = \text{“}Ganancia\ neta\ diaria\text{”} = \sum_{i=1}^{N} Z_i$$

luego

$$(G \mid N = n) = \sum_{i=1}^{n} Z_i,$$

por linealidad e independencia:

$$\begin{aligned} g(n) &= E(G \mid N = n) = n \cdot E(Z_i) = 100 \cdot n \\ h(n) &= Var(G \mid N = n) = n \cdot Var(Z_i) = 20^2 \cdot n = 400 \cdot n \end{aligned}$$

Luego la esperanza será

$$E(G) = E(g(N)) = E(100 \cdot N) = 100 \cdot E(N) = 100 \cdot 30 = 3000$$

Y la varianza

$$\begin{aligned} Var(G) &= E\left(Var(G \mid N)\right) + Var\left(E(G \mid N)\right) \\ &= E(h(N)) + Var(g(N)) \\ &= E(400 \cdot N) + Var(100 \cdot N) \\ &= 400 \cdot E(N) + 100^2 \cdot Var(N) \\ &= 400 \cdot 30 + 100^2 \cdot 30 = 312000 \end{aligned}$$

y finalmente

$$\sigma_G = \sqrt{312000} = 558{,}5696$$

9. La aseguradora más importante de una ciudad comercializa dos tipos de seguro: para hogares y para automóviles. Las pólizas se cobran mensualmente, con cuotas \$160 y \$95, respectivamente, y reembolsan un monto de \$500 y \$150 si ocurre un siniestro con el hogar o el vehículo. Es decir, el asegurado recibe \$500 por cada siniestro que ocurra con el hogar y \$150 por cada siniestro que ocurra con el vehículo.
Basados en estudios previos, los gerentes de la aseguradora saben que la cantidad de siniestros en el hogar de un potencial cliente sigue una distribución Poisson con tasa de 2 por año, mientras que la cantidad de accidentes de auto sigue una distribución Poisson con tasa de 5 por año. Además, saben que la ocurrencia de ambos tipos de siniestros son independientes entre sí.
Para un potencial cliente al azar:

(a) Cuántos siniestros en el hogar esperaría la aseguradora que tuviera en el año? ¿Cuántos en un mes?

(b) Cuántos siniestros con su vehículo esperaría la aseguradora que tuviera en el año? Cuántos en un mes?

(c) Cuál es la ganancia esperada en un año para la aseguradora?

(d) Si el individuo en cuestión no posee vehículo pero sí casa, cuál es la ganancia esperada en un año para la aseguradora?

Supongamos ahora que el contrato de seguros especifica que si el individuo sufre 10 siniestros en un mismo año el contrato se cancela, lo que representa para la aseguradora un costo neto de 100. Asumamos además que todos los clientes tienen ambos tipos de seguros.

(e) Cuál es la probabilidad de que un cliente vea cancelada su póliza?

Solution: Consideramos las variables

$$H = \text{“}\textit{Cantidad de siniestros anuales en el hogar}\text{”} \sim Poiss(2)$$
$$A = \text{“}\textit{Cantidad de siniestros anuales automovilisticos}\text{”} \sim Poiss(5)$$

(a) $E(X) = 2$. Si la tasa de siniestros se distribuye uniformemente durante los meses del año y no existe relación alguna entre lo que pasa en un mes y en otro, lo esperado sería que tuviera $\frac{2}{12}$ siniestros mensuales.

(b) $E(X) = 5$. Si la tasa de siniestros se distribuye uniformemente durante los meses del año y no existe relación alguna entre lo que pasa en un mes y en otro, lo esperado sería que tuviera $\frac{5}{12}$ siniestros mensuales.

(c) La ganancia anual por cliente es

$$G = 12 \cdot (160 + 95) - 500 \cdot H - 150 \cdot A$$

Por linealidad de la esperanza

$$E(G) = 3060 - 500 \cdot E(H) - 150 \cdot E(A) = 3060 - 500 \cdot 2 - 150 \cdot 5 = 1310$$

(d) La ganancia anual por cliente que posee casa pero no vehículo es

$$G_H = 12 \cdot 160 - 500 \cdot H$$

Por linealidad de la esperanza

$$E(G_H) = 1920 - 500 \cdot E(H) = 1920 - 500 \cdot 2 = 920$$

(e) La probabilidad de que un cliente vea cancelada la póliza es $P(H+A \geq 10)$. Para hacer calcular esta probabilidad usaremos que $H + A \sim Poiss(2+5)$ pues ambas son Poisson independientes de parametros 2 y 5 respectivamente:

$$P(H+A \geq 10) = 1 - P(H+A \leq 9) = 1 - \left(\frac{e^{-7} \cdot 7^0}{0!} + \ldots + \frac{e^{-7} \cdot 7^9}{9!} \right) = 1 - 0{,}8304959 = 0{,}1695041$$

10. El siguiente juego es una variación del juego que se conoce en economía como el juego de las billeteras. Supongamos que hay n personas no relacionadas entre sí, la persona i tiene en su billetera una cantidad X_i de pesos, la cual es desconocida para el resto del grupo. Se subasta la suma $V = X_1 + ... + X_n$ del dinero en las n billeteras, cada una de las n personas hará una oferta y ganará la subasta aquel cuya oferta sea la más alta. Lo único que es sabido por todas las personas es que cada billetera puede a lo sumo tener 100 pesos.
    (a) Si la persona i NO conoce cuanto dinero hay en su billetera, cuál es la esperanza V?
    (b) Si la persona i sólo conoce la cantidad de dinero que hay en su billetera, cuál es la esperanza de V ahora?
    (c) Suponé que cada persona i oferta el valor esperado en el punto (b). Suponé que la persona $i = 1$ gana la subasta. Condicional en conocer el valor de X_1 y en haber ganado la subasta, cuál es la esperanza de V? Explicá intuitivamente por qué esta esperanza es menor que la que encontraste en el punto (b).

Solution:

(a) Como no sabemos nada sobre cada X_i, las asumimos $U(0;100)$, luego

$$E(V) = n \cdot E(X_1) = 50 \cdot n.$$

(b) Ahora sabemos que $X_i = x_i$

$$\begin{aligned} E(V \mid X_i = x_i) &= E(X_1 + \ldots + X_i + \ldots + X_n \mid X_i = x_i) \\ &= E(X_1 + \ldots + x_i + \ldots + X_n \mid X_i = x_i) \\ &= x_i + \sum_{j \neq i} E(X_j \mid X_i = x_i) \end{aligned}$$

como las X_i son independientes se tiene que $E(X_j \mid X_i = x_i) = E(X_j)$ si $j \neq i$, entonces

$$E(V \mid X_i = x_i) = x_i + \sum_{j \neq i} E(X_j) = x_i + (n-1) \cdot 50$$

(c) Nos interesa calcular

$$E(V \mid X_1 = x_1, \max(X_2, ..., X_n) \leq x_1) = \sum_{j=1}^{n} E(X_j \mid X_1 = x_1, \max(X_2, ..., X_n) \leq x_1)$$

$$= E(X_1 \mid X_1 = x_1, \max(X_2, ..., X_n) \leq x_1) + \sum_{j=2}^{n} E(X_j \mid X_1 = x_1, \max(X_2, ..., X_n) \leq x_1)$$

por independencia entre las X_i

$$\begin{aligned} E(V \mid X_1 = x_1, \max(X_2, ..., X_n) \leq x_1) &= E(X_1 \mid X_1 = x_1) + \sum_{j=2}^{n} E(X_j \mid X_j \leq x_1) \\ &= E(x_1 \mid X_1 = x_1) + (n-1) \cdot E(X_2 \mid X_2 \leq x_1) \end{aligned}$$

por el ejercicio 16 del TP 8 sabemos que $X_2 \mid X_2 \leq x_1 \sim U(0; x_i)$, entonces

$$E\left(V \mid X_1 = x_1, \max\left(X_2, ..., X_n\right) \leq x_1\right) = x_1 + (n-1) \cdot \frac{x_1}{2}$$

Es razonable que la esperanza sea menor que la calculada en (b) porque en en este caso estamos condicionando a información que acota el valor de cada una de las X_j por x_1 mientras que en (b), para $j \neq 1$, solamente sabemos que $X_j < 100$.

11. Un grupo de 3 amigos pasan juntos un fin de semana (Sáb. y Dom.) en una playa. Acuerdan que cada día harán un sorteo equitativo para decidir quién paga la cena. Sea la v.a. X: nro. de veces que salió sorteado Juan y sea Y la v.a. que vale 1 si Juan pagó la cena del Domingo y 0 en otro caso.
    (a) Obtener una tabla para la distribución conjunta del par (X, Y). ¿Son X e Y independientes? Justifique su respuesta.
    (b) Sea $Z = X + Y$. Hallar $E(Z)$, $\mathrm{Var}(Z)$.
    (c) Si se sabe que Juan pagó la cena del Domingo, cuál es la probabilidad de que también haya pagado la del Sábado?

Solution: Notemos que las variables aleatorias:

$$X = \textit{cantidad de noches que Juan pagó la cena}$$

$$Y = \begin{cases} 0 & \text{si Juan no pagó el Domingo} \\ 1 & \text{si Juan pagó el Domingo} \end{cases}$$

son tal que $X \sim Bi(2; 1/3)$ y $Y \sim Bernoulli(1/3)$:

$$P_X(k) = \binom{2}{k} \cdot \left(\frac{1}{3}\right)^k \cdot \left(1 - \frac{1}{3}\right)^{2-k} = \begin{cases} 4/9 & \text{si } k = 0 \\ 4/9 & \text{si } k = 1 \\ 1/9 & \text{si } k = 2 \\ 0 & \text{sino.} \end{cases}$$

$$P_Y(l) = \begin{cases} 2/3 & \text{si } l = 0 \\ 1/3 & \text{si } l = 1 \end{cases}$$

(a) Para completar la tabla con los valores de la probabilidad conjunta empecemos mencionando que $p_{XY}(0;1) = 0$ y $p_{XY}(2;0) = 0$, ya que no puede ocurrir que Juan no haya pagado alguna cena y haya salido sorteado el Domingo como tampoco puede ocurrir que haya pagado las 2 cenas y no haya salido sorteado el Domingo.
Esto nos permite completar 2 posiciones de la tabla:

		Y		
		0	1	P_X
	0	-	0	4/9
X	1	-	-	4/9
	2	0	-	1/9
	P_Y	2/3	1/3	

Y ahora usando que al sumar por filas o columnas se obtienen las probabilidades marginales (valores previamente calculados) podemos completar el resto de las posiciones de la tabla:

		Y		
		0	1	P_X
	0	4/9	0	4/9
X	1	2/9	2/9	4/9
	2	0	1/9	1/9
	P_Y	2/3	1/3	

Las variables no son independientes pues $p_{XY}(0;1) = 0$ pero $P_X(0) \cdot P_Y(1) = \frac{4}{9} \cdot \frac{1}{3} \neq 0$.

(b) Calculamos la varianza aplicando la propiedad de linealidad:

$$E(Z) = E(X+Y) = E(X) + E(Y) = \left(0 \cdot \frac{4}{9} + 1 \cdot \frac{4}{9} + 2 \cdot \frac{1}{9}\right) + \left(0 \cdot \frac{2}{3} + 1 \cdot \frac{1}{3}\right) = \frac{2}{3} + \frac{1}{3} = 1$$

Para la varianza usamos la formula de la varianza de la suma:

$$\begin{aligned} Var(Z) &= Var(X+Y) = Var(X) + Var(Y) + 2 \cdot Cov(X;Y) \\ &= \left(E(X^2) - E^2(X)\right) + \left(E(Y^2) - E^2(Y)\right) + 2 \cdot \left(E(X \cdot Y) - E(X) \cdot E(Y)\right) \end{aligned}$$

Calculamos los términos faltantes:

- $E(X^2) = 0^2 \cdot \frac{4}{9} + 1^2 \cdot \frac{4}{9} + 2^2 \cdot \frac{1}{9} = \frac{8}{9}$
- $E(Y^2) = 0^2 \cdot \frac{2}{3} + 1^2 \cdot \frac{1}{3} = \frac{1}{3}$
- $E(X \cdot Y) = 1 \cdot 1 \cdot \frac{2}{9} + 2 \cdot 1 \cdot \frac{1}{9} = \frac{4}{9}$

Reemplazando todo:

$$Var(X+Y) = \left[\frac{8}{9} - \left(\frac{2}{3}\right)^2\right] + \left[\frac{1}{3} - \left(\frac{1}{3}\right)^2\right] + 2 \cdot \left(\frac{4}{9} - \frac{2}{3} \cdot \frac{1}{3}\right) = \frac{4}{9} + \frac{2}{9} + \frac{4}{9} = \frac{10}{9} = 1{,}1111$$

(c) $P(Juan\ pague\ el\ sábado \mid Y = 1) = \frac{P(Juan\ pague\ el\ sábado\ y\ Y=1)}{P(Y=1)} = \frac{P(X=2)}{P(Y=1)} = \frac{1/9}{1/3} = 0{,}3333$

12. Vas a invertir 10000 $ que tenés la opción de distribuidos en en tres portafolios. Llamá X_1, X_2 y X_3 a las tasas de retorno al cabo de un año de los tres portafolios expresadas en proporciones -NO en porcentajes-, de modo que por ejemplo, si invertís 3000$ en el portafolio 1, al cabo de un año tu retorno será de $3000X_1$$.

Suponé que las medias y desvíos estandard de X_1, X_2 y X_3 son

	X_1	X_2	X_3
Media	0.08	0.12	0.1
Desvío	0.02	0.05	0.03

Suponé que las correlaciones son

	X_1	X_2	X_3
X_1	1	-0.4	-0.2
X_2	-0.4	1	0.6
X_3	-0.2	0.6	1

(a) Supone que invertís a_j en el portafolio $j = 1,2,3$. Como tenés 10000$, $a_1 + a_2 + a_3 = 10000$. Entonces, tu retorno será $Z = a_1X_1 + a_2X_2 + a_3X_3$.
Encontrá la función $g(a_1, a_2, a_3)$ tal que $g(a_1, a_2, a_3) = E(Z)$

(b) Encontrá la función $h(a_1, a_2, a_3)$ tal que $h(a_1, a_2, a_3) = Var(Z)$

(c) Suponé que sos averso al riesgo, y que querés distribuir tu inversión de modo tal de que el desvío de tu inversión nunca supere el desvío del portafolio con menor desvío, es decir del portafolio 1. ¿Es posible encontrar una estrategia de inversión para la cual el desvío sea menor o igual a $10000 \times 0{,}02 = 200$ (el desvio del retorno si invirtieras 10000$ en el portafolio 1) pero cuya esperanza de tasa de retorno sea mayor que $10000 \times 0{,}08 = 800$ (la esperanza del retorno si invirtieras todo tu dinero en el portafolio 1)?

(d) Que problema de maximización deberías plantear formalmente para encontrar la mejor estrategia, en el sentido de maximizar la esperanza del retorno sujeto a que el desvío del retorno no supere a 200$?

Solution:

(a) Usando linealidad de la esperanza

$$E(Z) = a_1 \cdot E(X_1) + a_2 \cdot E(X_2) + a_3 \cdot E(X_3) = 0{,}08 \cdot a_1 + 0{,}12 \cdot a_2 + 0{,}1 \cdot a_3$$

(b) Se puede ver que la varianza de una combinación lineal de 3 variables aleatorias es

$$\begin{aligned} Var(Z) = {} & {a_1}^2 \cdot Var(X_1) + {a_2}^2 \cdot Var(X_2) + {a_3}^2 \cdot Var(X_3) \\ & + 2 \cdot a_1 \cdot a_2 \cdot Cov(X_1; X_2) + 2 \cdot a_1 \cdot a_3 \cdot Cov(X_1; X_3) + 2 \cdot a_2 \cdot a_3 \cdot Cov(X_2; X_3) \end{aligned}$$

Reemplazando, teniendo en cuenta que $\rho_{X_i X_j} = \frac{Cov(X_i;X_j)}{\sigma_{X_i} \cdot \sigma_{X_j}}$ queda que

$$\begin{aligned} Var(Z) = {} & 0{,}0004 \cdot {a_1}^2 + 0{,}0025 \cdot {a_2}^2 + 0{,}009 \cdot {a_3}^2 \\ & - 0{,}0008 \cdot a_1 \cdot a_2 - 0{,}00024 \cdot a_1 \cdot a_3 + 0{,}0018 \cdot a_2 \cdot a_3 \end{aligned}$$

(c) Sí es posible, por ejemplo diversificando

$$a_1 = 9000$$
$$a_2 = 1000$$
$$a_3 = 0$$

(d) Habría que resolver el problema de extremos restringidos que consiste en maximizar la función $g(a_1, a_2, a_3) = 0{,}08 \cdot a_1 + 0{,}12 \cdot a_2 + 0{,}1 \cdot a_3$ sujeta a las restricciones

$$\begin{cases} a_1 + a_2 + a_3 & \leq 10000 \\ h(a_1, a_2, a_3) & \leq 200^2 \\ 0 & \leq a_1 \\ 0 & \leq a_2 \\ 0 & \leq a_3 \end{cases}$$

Se puede demostrar que la región en cuestión es acotada y sumado a que la función g es continua, entonces necesariamente alcanzará un valor máximo dentro de la región.

13. Para el ejercicio 1,

(a) ¿Son X_1 y X_2 variables independientes?

(b) Hallá la distribución de la variable aleatoria $g(X_2) = E(X_1 | X_2)$ y calculá su esperanza.

(c) Calculá la esperanza de X_1 y corroborá que coincide con la esperanza de $g(X_2)$ que calculaste en el item anterior. ¿Es una casualidad que coincida o es la consecuencia de alguna ley general?

(d) Calcular la varianza de $X_1 + 2X_2$.

Solution:

(a) Sabemos que una condición necesaria y suficiente que tiene que verificarse para que las variables sean independientes es:

$$p_{X_1X_2}(k;l) = p_{X_1}(k) \cdot p_{X_2}(l) \quad \forall \quad k \in R_{X_1} \ y \ l \in R_{X_2},$$

pero tomando los valores $k = 3$ y $l = 0$ se tiene que $p_{X_1X_2}(3;0) = 0$ y sin embargo:

$$p_{X_1}(3) \cdot p_{X_2}(2) = (0 + 0{,}03 + 0{,}04 + 0{,}07) \cdot 0{,}28 = 0{,}0392 \neq 0,$$

por lo que no se verifica la condición mencionada y entonces concluímos que no hay independencia entre las variables.

(b) Para trabajar con la variable aleatoria $g(X_2) = E(X_1 \mid X_2)$ primero vamos a averiguar la fórmula de la función $g(l) = E(X_1 \mid X_2 = l)$ con $l \in R_{X_2}$.
Notemos que $g(2) = E(X_1 \mid X_2 = 2)$ es la esperanza de la variable aleatoria cuya función de probabilidad calculamos en el ítem anterior:

$$E(X_1 \mid X_2 = 2) = 0 \cdot 0{,}1429 + \ldots + 4 \cdot 0{,}1785 = 2{,}0357.$$

Con esto tendríamos que:

$$g(l) = E(X_1 \mid X_2 = l) = \begin{cases} E(X_1 \mid X_2 = l) & \text{si } l = 0 \\ E(X_1 \mid X_2 = l) & \text{si } l = 1 \\ 2{,}0357 & \text{si } l = 2 \\ E(X_1 \mid X_2 = l) & \text{si } l = 3 \end{cases}$$

Repitiendo estos pasos para calcular las funciones de probabilidad restantes y sus respectivas esperanzas se puede ver que:

$$g(l) = E(X_1 \mid X_2 = l) = \begin{cases} 0{,}8421 & \text{si } l = 0 \\ 1{,}2 & \text{si } l = 1 \\ 2{,}0357 & \text{si } l = 2 \\ 2{,}6522 & \text{si } l = 3 \end{cases}$$

Entonces la esperanza condicional queda definida como:

$$g(X_2) = \begin{cases} 0{,}8421 & \text{si } X_2 = 0 \\ 1{,}2 & \text{si } X_2 = 1 \\ 2{,}0357 & \text{si } X_2 = 2 \\ 2{,}6522 & \text{si } X_2 = 3 \end{cases}$$

y su función de probabilidad es:

$$p_{g(X_2)}(k) = \begin{cases} P(X_2 = 0) & \text{si } g(X_2) = 0{,}8421 \\ P(X_2 = 1) & \text{si } g(X_2) = 1{,}2 \\ P(X_2 = 2) & \text{si } g(X_2) = 2{,}0357 \\ P(X_2 = 3) & \text{si } g(X_2) = 2{,}6522 \\ 0 & \text{sino.} \end{cases} = \begin{cases} 0{,}19 & \text{si } g(X_2) = 0{,}8421 \\ 0{,}3 & \text{si } g(X_2) = 1{,}2 \\ 0{,}28 & \text{si } g(X_2) = 2{,}0357 \\ 0{,}23 & \text{si } g(X_2) = 2{,}6522 \\ 0 & \text{sino.} \end{cases}$$

Por último calculamos su esperanza:

$$E\big(g(X_2)\big) = 0{,}8421 \cdot 0{,}19 + 1{,}2 \cdot 0{,}3 + 2{,}0357 \cdot 0{,}28 + 2{,}6522 \cdot 0{,}23 = 1{,}7$$

(c) Para calcular la $E(X_1)$ usamos su función de probabilidad hallada en el primer ejercicio

$$E(X_1) = 0 \cdot 0{,}19 + 1 \cdot 0{,}3 + 2 \cdot 0{,}25 + 3 \cdot 0{,}14 + 4 \cdot 0{,}12 = 1{,}7$$

Vemos que coincide con $E\left(E\left(X_1|X_2\right)\right)$. Este hecho no es causalidad, es consecuencia de la ley de esperanza total.

(d) Por propiedad, sabemos que $Var(X_1+2\cdot X_2) = 1^2\cdot Var(X_1)+2^2\cdot Var(X_2)+2\cdot 1\cdot 2\cdot Cov(X_1;X_2)$. Utilizando las distribuciones marginales se puede ver que:

$$\begin{aligned} Var(X_1) &= E({X_1}^2) - E^2(X_1) = 1{,}59 \\ Var(X_2) &= E({X_2}^2) - E^2(X_2) = 1{,}0875 \\ Cov(X_1;X_2) &= E(X_1 \cdot X_2) - E(X_1)\cdot E(X_2) = 3{,}33 - 1{,}7 \cdot 1{,}55 = 0{,}695 \end{aligned}$$

Finalmente se tiene

$$Var(X_1 + 2 \cdot X_2) = 1{,}59 + 4 \cdot 1{,}0875 + 4 \cdot 0{,}695 = 8{,}72$$

14. Para el ejercicio 2, calcular $Cov(X;Y)$

Solution:

		Y		
		0	1	2
X	0	0.025	0.015	0.010
	1	0.050	0.030	0.020
	2	0.125	0.075	0.050
	3	0.150	0.090	0.060
	4	0.100	0.060	0.040
	5	0.050	0.030	0.020

$$Cov(X;Y) = E(X \cdot Y) - E(X)\cdot E(Y),$$

por el ejercicio 2 sabemos que $E(X) = 2{,}8$ y $E(Y) = 0{,}7$. Y además se tiene que

$$\begin{aligned} E(X\cdot Y) &= \sum_{k\in R_X}\sum_{l\in R_Y} k \cdot l \cdot p_{XY}(k;l) \\ &= 0\cdot 0\cdot 0{,}025 + \ldots 5 \cdot 0 \cdot 0{,}05 \\ &\quad + 1\cdot 1\cdot 0{,}03 \\ &\quad + 1 \cdot 2 \cdot 0{,}02 + 2 \cdot 1 \cdot 0{,}075 \\ &\quad \vdots \\ &\quad + 5\cdot 2\cdot 0{,}02 \\ &= 1{,}96 \end{aligned}$$

por lo tanto

$$Cov(X;Y) = 1{,}96 - 2{,}8 \cdot 0{,}7 = 0$$

(es posible demostrar que las variables son independientes, entonces la covarianza nula resulta ser coherente con la independencia de las variables)

15. Para el ejercicio 7, calcular la correlación entre la cantidad de visitantes del sitio web y la cantidad de pasajes vendidos.

Solution: Las variable involucradas son

$$N \sim Poiss(50000)$$
$$C = \sum_{i=1}^{N} X_i \quad con \quad X_i \sim Bernoulli(0{,}2)$$

se nos pide calcular

$$\begin{aligned}\rho_{NC} &= \frac{Cov(N;C)}{\sigma_N \cdot \sigma_C}\\ &= \frac{E(N\cdot C) - E(N)\cdot E(C)}{\sigma_N \cdot \sigma_C}\end{aligned}$$

dado que $N \sim Poiss(50000)$ sabemos que $E(N) = 50000$ y $\sigma_N = \sqrt{50000}$. Además por el ejercicio 7 se tiene que $E(C) = 10000$ y $\sigma_C = 100$. Luego el ejercicio se reduce a calcular $E(N\cdot C)$, para lo cual usaremos ley de esperanza total:

$$E(N\cdot C) = E(E(N\cdot C \mid N)) = E(g(N)),$$

siendo g la función definida como

$$\begin{aligned}g(n) &= E(N\cdot C \mid N = n)\\ &= E\left(N\cdot \sum_{i=1}^{N} X_i \mid N = n\right)\\ &= E\left(n\cdot \sum_{i=1}^{n} X_i \mid N = n\right) \quad \to usando\ linealidad\\ &= n\cdot \sum_{i=1}^{n} E(X_i \mid N = n) \quad \to por\ independencia\\ &= n\cdot \sum_{i=1}^{n} E(X_i) = n\cdot(0{,}2\cdot n) = 0{,}2\cdot n^2\end{aligned}$$

se tendrá entonces que

$$\begin{aligned}E(N\cdot C) &= E\left(0{,}2\cdot N^2\right)\\ &= 0{,}2\cdot E\left(N^2\right)\\ &= 0{,}2\cdot\left(Var(N) + E^2(N)\right)\\ &= 0{,}2\cdot(50000 + 50000^2)\\ &= 500010000\end{aligned}$$

y finalmente

$$\rho_{NC} = \frac{500010000 - 50000\cdot 10000}{\sqrt{5000}\cdot 100} = 0{,}447213595$$

16. Para el ejercicio 8, calcular la correlación entre la cantidad de clientes que entran al comercio y la cantidad de prendas vendidas durante el día. A partir de la correlación calculada, es posible determinar si dichas cantidades aleatorias son independientes?

Solution: Las variable involucradas son

$$N \sim Poiss(30)$$
$$Y = \sum_{i=1}^{N} X_i \quad con \quad X_i \sim Bernoulli(2)$$

se nos pide calcular

$$\begin{aligned} \rho_{NY} &= \frac{Cov(N;Y)}{\sigma_N \cdot \sigma_Y} \\ &= \frac{E(N \cdot Y) - E(N) \cdot E(Y)}{\sigma_N \cdot \sigma_Y} \end{aligned}$$

dado que $N \sim Poiss(30)$ sabemos que $E(N) = 30$ y $\sigma_N = \sqrt{30}$. Además por el ejercicio 8 se tiene que $E(Y) = 60$ y $\sigma_Y = \sqrt{180}$. Luego el ejercicio se reduce a calcular $E(N \cdot Y)$, para lo cual usaremos ley de esperanza total:

$$E(N \cdot Y) = E(E(N \cdot Y \mid N)) = E(g(N)),$$

siendo g la función definida como

$$\begin{aligned} g(n) &= E(N \cdot Y \mid N = n) \\ &= E\left(N \cdot \sum_{i=1}^{N} X_i \mid N = n\right) \\ &= E\left(n \cdot \sum_{i=1}^{n} X_i \mid N = n\right) \quad \rightarrow usando\ linealidad \\ &= n \cdot \sum_{i=1}^{n} E(X_i \mid N = n) \quad \rightarrow por\ independencia \\ &= n \cdot \sum_{i=1}^{n} E(X_i) = n \cdot (2 \cdot n) = 2 \cdot n^2 \end{aligned}$$

se tendrá entonces que

$$\begin{aligned} E(N \cdot Y) &= E\left(2 \cdot N^2\right) \\ &= 2 \cdot E\left(N^2\right) \\ &= 2 \cdot \left(Var(N) + E^2(N)\right) \\ &= 2 \cdot (30 + 30^2) \\ &= 1860 \end{aligned}$$

y finalmente

$$\rho_{NY} = \frac{1860 - 30 \cdot 60}{\sqrt{30} \cdot \sqrt{180}} = 0{,}816496581$$

Como las variables están correlacionadas (el coeficiente resultó ser no nulo), podemos afirmar que las variables no son independientes.

17. Para el ejercicio 11, calcular el coeficiente de correlación entre X e Y

Solution: En la resolución presentada para el ejercicio 11 se obtiene que

$$\begin{aligned} Var(X) &= \frac{8}{9} - \left(\frac{2}{3}\right)^2 = \frac{4}{9} \\ Var(Y) &= \frac{1}{3} - \left(\frac{1}{3}\right)^2 = \frac{2}{9} \\ Cov(X;Y) &= \frac{4}{9} - \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9} \end{aligned}$$

luego

$$\rho_{XY} = \frac{2/9}{\sqrt{4/9} \cdot \sqrt{2/9}} = 0{,}707106781$$

18. Considere a X una variable aleatoria que puede tomar únicamente los valores -1, 0 y 1 todos con igual probabilidad. Además considere la variable aleatoria Y definida como:

$$Y = \begin{cases} 1 & \text{si } X = 0 \\ 0 & \text{caso contrario.} \end{cases}$$

(a) Calcular la correlación entre X e Y

(b) Son X e Y variables aleatorias independientes?

Solution: Dado que

$$\begin{aligned} X = 0 &\Rightarrow Y = 1 \\ X = -1 &\Rightarrow Y = 0 \\ X = 1 &\Rightarrow Y = 0 \end{aligned}$$

entonces

$$P(Y = 0 \mid X = 0) = P(Y = 1 \mid X = -1) = P(Y = 1 \mid X = 1) = 0$$

y consecuentemente

$$\begin{aligned} P(Y = 0 \mid X = 0) &\Rightarrow p_{XY}(0;0) = 0 \\ P(Y = 1 \mid X = -1) &\Rightarrow p_{XY}(-1;1) = 0 \\ P(Y = 1 \mid X = 1) &\Rightarrow p_{XY}(1;1) = 0 \end{aligned}$$

Esta información, junto con que X es una variable uniforme, nos dice que

		Y		
		0	1	P_X
	-1	-	0	1/3
X	0	0	-	1/3
	1	-	0	1/3
	P_Y	-	-	

y luego el resto de las posiciones son fácilmente calculables

		Y		
		0	1	P_X
	-1	$1/3$	0	$1/3$
X	0	0	$1/3$	$1/3$
	1	$1/3$	0	$1/3$
	P_Y	$2/3$	$1/3$	

(a)

$$E(X) = (-1)\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3} = 0$$
$$E(Y) = 0\cdot\frac{1}{3} + 1\cdot\frac{2}{3} = \frac{2}{3}$$
$$E(X\cdot Y) = \sum_{k\in R_X}\sum_{l\in R_Y} k\cdot l\cdot p_{XY}(k;l) = 0$$

(notar que $E(X\cdot Y) = 0$ pues cada sumando de la doble sumatoria es nulo, ya que en la tabla se evidencia que para cualquier posición siempre ocurre que alguna de las variables o la conjunta toman valor cero). Finalmente

$$\rho_{XY} = \frac{Cov(X;Y)}{\sigma_X\cdot\sigma_Y} = \frac{E(X\cdot Y) - E(X)\cdot E(Y)}{\sigma_X\cdot\sigma_Y} = \frac{0 - 0\cdot 2/3}{\sigma_X\cdot\sigma_Y} = 0$$

(b) No hay independencia, pues:

$$p_{XY}(0,0) \neq p_X(0)\cdot p_Y(0)$$

ya que $0 \neq 1/3\cdot 2/3$.
Es importante reconocer que se trata de variables aleatorias no correlacionadas pero que sin embargo resultan ser dependientes. Lo cual muestra que, si bien independencia implica correlación nula (propiedad vista en clase), la vuelta es falsa y por lo tanto no se puede hablar de una equivalencia entre dichas características.