Soluciones del primer parcial

1. Para poder canjear un premio se deben coleccionar 3 figuritas diferentes, que vienen en paquetes de galletitas. Cada paquete trae una figurita al azar, es decir, la probabilidad de que una figurita particular aparezca en un paquete dado es la misma para cada una de las tres figuritas.

   a) (6p) ¿Cuál es la probabilidad de poder canjear al premio comprando exactamente 3 paquetes de galletitas?

   b) (14p) ¿Cuántos paquetes de galletitas hay que comprar, como mínimo, para que la probabilidad de poder canjear el premio sea mayor a 0,97?

   Solución: Supongamos que las figuritas son A, B, C. Si compramos n paquetes, nuestro espacio muestral será $\Omega = \{A, B, C\}^n$, podemos considerar $\mathcal{F} = \mathcal{P}(\Omega)$ y, como es equiprobable, $\mathbb{P}(\omega) = \dfrac{1}{\#\Omega} = \dfrac{1}{3^n} \forall \omega \in \Omega$.

   Para a), $n = 3$ y los casos favorables serán $3! = 3 \times 2 \times 1 = 6$, ya que deben salirnos todas las figuritas. Por lo tanto la probabilidad es $\dfrac{6}{3^3} = \dfrac{2}{9}$.

   Para b), consideremos los eventos $S_A = \{\text{entre las n salió la figurita A alguna vez}\}$, S_B y S_C, definidos análogamente.

   Queremos que $\mathbb{P}(S_A \cap S_B \cap S_C) > 0{,}97$, pero $\mathbb{P}(S_A \cap S_B \cap S_C) = 1 - \mathbb{P}((S_A \cap S_B \cap S_C)^c) = 1 - \mathbb{P}(S_A^c \cup S_B^c \cup S_C^c)$.

   Calculemos esta última probabilidad usando inclusión-exclusión:

   $$\mathbb{P}(S_A^c \cup S_B^c \cup S_C^c) = \mathbb{P}(S_A^c) + \mathbb{P}(S_B^c) + \mathbb{P}(S_C^c) - \mathbb{P}(S_A^c \cap S_B^c) - \mathbb{P}(S_A^c \cap S_C^c) - \mathbb{P}(S_B^c \cap S_C^c) + \mathbb{P}(S_A^c \cap S_B^c \cap S_C^c).$$

   Para calcular $\mathbb{P}(S_A^c)$, tengamos en cuenta que $S_A^c = \{\text{entre las n, la figurita A no salió nunca}\}$, por lo que para cada una de las n figuritas, tendremos 2 opciones. Es decir $\mathbb{P}(S_A^c) = \dfrac{2^n}{3^n}$. Cambiando A por B y C, tenemos los siguientes dos términos de la suma.

   Para calcular $\mathbb{P}(S_A^c \cap S_B^c)$, tengamos en cuenta que $S_A^c \cap S_B^c = \{\text{en las n, no salió A ni B}\}$, por lo que para cada una de las n figuritas, tendremos una sola opción (que sean todas C). Es decir $\mathbb{P}(S_A^c \cap S_B^c) = \dfrac{1}{3^n}$, y las otras intersecciones dobles son iguales.

   La intersección triple es vacía, por lo que su probabilidad es 0.

   Finalmente $\mathbb{P}(S_A \cap S_B \cap S_C) > 0{,}97 \iff 1 - 3\left(\dfrac{2}{3}\right)^n + 3\dfrac{1}{3^n} > 0{,}97$. Como el lado izquierdo es creciente, probando obtenemos que el mínimo n es $n = 12$.

2. Lucio juega al ajedrez. Cada día, juega reiteradas partidas hasta el momento en que pierde la primera; en ese momento descansa hasta el día siguiente. Supongamos que Lucio *gana* una partida con probabilidad p, *pierde* una partida con probabilidad q y hace *tablas* con probabilidad r, con $p + q + r = 1$; y suponemos que todas las partidas son independientes.
   Definimos las variables aleatorias

   $$J = \text{cantidad de partidas jugadas hoy}$$
   $$G = \text{cantidad de partidas ganadas hoy}$$

   a) (7p) Calcular $p_{(J,G)}$, la función de probabilidad puntual conjunta de J y G.

   b) (7p) Calcular las marginales p_J y p_G. ¿Son J y G independientes?

c) (6p) Si $p = 0{,}45$, $q = 0{,}5$ y $r = 0{,}05$, calcular $\mathbb{P}(J = 10 \mid G = 5)$.

Sugerencia: $\sum_{n=k}^{\infty} \binom{n}{k} t^n = \dfrac{t^k}{(1-t)^{k+1}}$, *si* $0 \leq t < 1$.

Solución: a) Observemos que $R_J = \mathbb{N}$ y que $R_G = \mathbb{N}_0$. Si $(j, g) \in R_J \times R_G$, calculemos $p_{(J,G)}(j, g)$:

Notemos que no puede ganar más partidos de los que juega, ni tampoco ganar la misma cantidad que juega (ya que el último que juega siempre lo pierde), por lo que $p_{(J,G)}(j, g) = 0$ si $j \leq g$.

Si $j > g$, pasemos la intersección al condicional:

$$p_{(J,G)}(j, g) = \mathbb{P}(J = j, G = g) = \mathbb{P}(G = g \mid J = j)\mathbb{P}(J = j)$$

Sabemos que J es una v.a. geométrica de parámetro q, por lo que $\mathbb{P}(J = j) = (1 - q)^{j-1}q$.

Ahora pensemos $\mathbb{P}(G = g \mid J = j)$: como jugó j, tendremos una tira de j lugares independientes (donde en cada lugar G=ganó, P=perdió ó T=tablas). Sabemos que el último lugar es una P (con proba 1, porque **lo sabemos**) y los primeros $j - 1$ lugares tendrán G o T (ya que si hubiera una P, habría parado de jugar antes). Por lo tanto, la probabilidad de que haya una G en algún lugar es $\hat{p} = \frac{p}{p+r}$. Entonces, si lo pensamos como una binomial de parámetros $j - 1$ y $\hat{p}$,

$$\mathbb{P}(G = g \mid J = j) = \tbinom{j-1}{g} \left(\frac{p}{p+r}\right)^g (1 - \frac{p}{p+r})^{j-1-g}$$

Finalmente,

$$\begin{aligned} p_{(J,G)}(j, g) &= \binom{j-1}{g} \left(\frac{p}{p+r}\right)^g (1 - \frac{p}{p+r})^{j-1-g}(1-q)^{j-1}q \\ &= \binom{j-1}{g} p^g r^{j-1-g} q \end{aligned}$$

si $j > g$ y 0 si no.

Obs: la última igualdad sale notando que $p + r = 1 - q$ y simplificando las cosas.

b) Si $j \in R_J$,

$$\begin{aligned} p_J(j) &= \sum_{g \in R_G} p_{(J,G)}(j, g) \\ &= \sum_{0 \leq g < j} \binom{j-1}{g} p^g r^{j-1-g} q \\ &= r^{j-1} q \sum_{0 \leq g < j} \binom{j-1}{g} \left(\frac{p}{r}\right)^g \\ &= r^{j-1} q \left(1 + \frac{p}{r}\right)^{j-1} \\ &= q(1-q)^{j-1} \end{aligned}$$

ya que esa suma es un binomio de Newton y reagrupando.

Si $g \in R_G$,

$$\begin{aligned}
p_G(g) &= \sum_{j \in R_J} p_{(J,G)}(j,g) \\
&= \sum_{j \geq g} \binom{j-1}{g} p^g r^{j-1-g} q \\
&= \frac{p^g q}{r^g} \cdot \sum_{j > g} \binom{j-1}{g} r^{j-1} \\
&= \frac{p^g q}{(1-r)^{g+1}}
\end{aligned}$$

ya que en esa última suma usamos la Sugerencia del problema con $n = j-1, k = g, r = t$

Veamos que no son independientes: Si $(j,g) \in R_J \times R_G$ y $j \leq g$, la probabilidad conjunta es 0 mientras que las dos marginales son distintas de cero, por lo tanto no valdría que $\mathbb{P}(J = j, G = g) = \mathbb{P}(J = j)\mathbb{P}(G = g)$, y entonces no son independientes.

c) Usando la definición de probabilidad condicional,

$$\begin{aligned}
\mathbb{P}(J = 10 \mid G = 5) &= \frac{\mathbb{P}(G = 5, J = 10)}{\mathbb{P}(G = 5)} \\
&= \frac{\binom{10-1}{5}(0{,}45)^5(0{,}05)^{10-1-5}(0{,}5)}{\frac{(0{,}45)^5(0{,}5)}{(1-0{,}05)^{5+1}}} \\
&= \binom{9}{5}(0{,}05)^4(1-0{,}05)^6 \\
&= 0{,}00057
\end{aligned}$$

3. La medida, en centímetros, de la longitud de la cintura de los hombres en Buenos Aires sigue una distribución normal de parámetros $\mu = 75$ y $\sigma = 16$. Se sabe que todos los hombres de menos de 71 centímetros de cintura usan cinturón talle 1, los de cintura entre 71 y 83 centímetros usan talle 2, y los restantes talle 3.

   a) (4p) Hallar la distribución del talle de cinturón de los hombres de Buenos Aires.

   b) (5p) ¿Cuál debería ser la longitud máxima de cintura del talle 1 si se quiere que el 30 % de los hombres de Buenos Aires usen talle 1?

   c) (6p) En una tienda, un cliente acaba de comprar un cinturón de talle 2, ¿cuál es la probabilidad de que su cintura mida más de 75 centímetros?

   d) (5p) Si en una tienda entran hombres a comprar un cinturón al azar, ¿cuál es la probabilidad de que el cuarto cinturón vendido sea el primero de talle 1?

Solución: Llamemos X a la v.a. "longitud de cintura" y T a la v.a. "talle de cinturón". Sabemos que $X \sim N(75, 16)$ y que $R_T = \{1, 2, 3\}$, por lo que T es una v.a. discreta.

Definimos la v.a. $Z = \dfrac{X - 75}{16}$ y por lo tanto $Z \sim N(0,1)$. Además, sabemos que:

(1) $\mathbb{P}(Z < \alpha) = \mathbb{P}(Z > -\alpha) \forall \alpha$, ya que Z es simétrica respecto de 0.
(2) En la tabla tenemos los valores de $\mathbb{P}(Z \geq \alpha)$ para $\alpha \in [0, 4)$ (aprox).

a) Para calcular la distribución de T basta con calcular $\mathbb{P}(T = 1), \mathbb{P}(T = 2)$ y $\mathbb{P}(T = 3)$.

- $\mathbb{P}(T=1) = \mathbb{P}(X<71) = \mathbb{P}\left(\frac{X-75}{16} < \frac{71-75}{16}\right) = \mathbb{P}(Z<-0{,}25) =_{(1)} \mathbb{P}(Z>0{,}25) =_{(2)}$ 0,4013
- $\mathbb{P}(T=2) = \mathbb{P}(71 \leq X < 83) = \mathbb{P}\left(\frac{71-75}{16} \leq \frac{X-75}{16} < \frac{83-75}{16}\right) = \mathbb{P}(-0{,}25 \leq Z < 0{,}5) = \mathbb{P}(Z<0{,}5) - \mathbb{P}(Z<-0{,}25) =_{(1)} 1 - \mathbb{P}(Z>0{,}5) - \mathbb{P}(Z>0{,}25) =_{(2)} 1 - 0{,}3085 - 0{,}4013 = 0{,}2902$
- $\mathbb{P}(T=3) = \mathbb{P}(X \geq 83) = \mathbb{P}\left(\frac{X-75}{16} \geq \frac{83-75}{16}\right) = \mathbb{P}(Z \geq 0{,}5) =_{(2)} 0{,}3085$

b) Queremos buscar un x tal que $\mathbb{P}(X<x) = 0{,}3$. Estandarizando, $\mathbb{P}\left(Z < \frac{x-75}{16}\right) = 0{,}3$ y, por (1), $\mathbb{P}\left(Z > -\frac{x-75}{16}\right) = 0{,}3$. Buscamos en la tabla, y el valor más cercano es $z = 0{,}52$, por lo tanto $-\frac{x-75}{16} = 0{,}52 \iff x = 66{,}68$.

c) Queremos calcular

$$\begin{aligned}\mathbb{P}(X>75 \mid T=2) &= \mathbb{P}(X>75 \mid 71 \leq X < 83)\\ &= \frac{\mathbb{P}(\{X>75\} \cap \{71 \leq X < 83\})}{\mathbb{P}(71 \leq X < 83)}\\ &= \frac{\mathbb{P}(75 \leq X < 83)}{0{,}2902}\end{aligned}$$

Estandarizando, $\mathbb{P}(75 \leq X < 83) = \mathbb{P}(0 < Z < 0{,}5) = \mathbb{P}(Z<0{,}5) - \mathbb{P}(Z<0) = 1 - \mathbb{P}(Z>0{,}5) - \mathbb{P}(Z>0) = 1 - 0{,}3085 - 0{,}5 = 0{,}1915$. Por lo tanto, lo pedido es $\frac{0{,}1915}{0{,}2902} = 0{,}6598$.

d) Definimos la v.a. $Y =$"número del cliente que compra el primer cinturón de talle 1", e $Y \sim \mathcal{G}(p)$, con $p = \mathbb{P}(T=1) = 0{,}4013$. Por lo tanto, queremos calcular $\mathbb{P}(Y=4) = (1-p)^3 p = (0{,}5987)^3(0{,}4013) = 0{,}0861$.

4. Para diseñar estructuras en una ciudad andina, es necesario entender cómo se distribuye la magnitud de los terremotos que pueden ocurrir. Para ello, se sabe que la magnitud de un terremoto sigue una distribución exponencial de parámetro $\lambda = 0{,}5$, y que las magnitudes de distintos terremotos son independientes.

   a) (5p) Si en un año hubo 20 terremotos, ¿cúal es la probabilidad de que todos ellos hayan tenido magnitud mayor a 1?

   b) (7p) Si sabemos que la cantidad de terremotos que ocurren en un año se distribuye como una variable Poisson de parámetro $\lambda = 30$, ¿cúal es la probabilidad de que a lo largo de un año no haya ningún terremoto de magnitud mayor a 7?

   c) (8p) Si sabemos que la cantidad de terremotos que ocurren en t años se distribuye como una variable Poisson de parámetro $\lambda = 30t$, ¿cuántos años deben pasar, como mínimo, para que la probabilidad de observar al menos un terremoto de magnitud mayor a 7 sea mayor a 0,99?.

Solución: a) Sean $T_i \sim exp(0{,}5)$ con $i = 1,..,20$ las magnitudes de los terremotos. Queremos calcular

$$\begin{aligned}\mathbb{P}(T_1 > 1, ..., T_{20} > 1) &\overset{indep}{=} \prod_{i=1}^{20} \mathbb{P}(T_i > 1)\\ &\overset{id.dist}{=} (\mathbb{P}(T_1 > 1))^{20}\\ &= (1 - \mathbb{P}(T_1 \leq 1))^{20}\\ &= (1 - (1 - e^{-0,5\cdot 1}))^{20}\\ &= e^{-10}\end{aligned}$$

b) Si no tiene que haber ningún terremoto mayor a 7, quiere decir que todos son ≤ 7 y, equivalentemente, el máximo es ≤ 7. Definamos la v.a. $M =$"máxima magnitud observada en los terremotos que ocurrieron".

Queremos calcular $\mathbb{P}(M \leq 7)$ y para eso podemos hacer probabilidad total con todos los valores que puede tomar la variable $N_1 \sim \mathcal{P}(30)$. Es decir:

$$\mathbb{P}(M \leq 7) = \sum_{k=0}^{+\infty} \mathbb{P}(M \leq 7 \mid N_1 = k)\mathbb{P}(N_1 = k)$$

Ahora, si $N_1 = k$, tendremos k magnitudes de terremotos $T_1, ..., T_k$ que queremos que sean cada una ≤ 7.

$$\begin{aligned}\mathbb{P}(M \leq 7) &= \sum_{k=0}^{+\infty} \mathbb{P}(M \leq 7 \mid N_1 = k)\mathbb{P}(N_1 = k)\\ &= \sum_{k=0}^{+\infty} \mathbb{P}(T_1 \leq 7, ..., T_k \leq 7)\mathbb{P}(N_1 = k)\\ &\overset{iid}{=} \sum_{k=0}^{+\infty} (\mathbb{P}(T_1 \leq 7))^k \mathbb{P}(N_1 = k)\end{aligned}$$

Sabemos que la puntual de N_1 es $\mathbb{P}(N_1 = k) = \dfrac{e^{-30}30^k}{k!}$ y que $\mathbb{P}(T_1 \leq 7) = 1 - e^{-0,5\cdot 7} := p \sim 0{,}9698$.

Por lo tanto,

$$\begin{aligned}\mathbb{P}(M \leq 7) &= \sum_{k=0}^{+\infty} p^k \frac{e^{-30}30^k}{k!}\\ &= e^{-30} \sum_{k=0}^{+\infty} \frac{(30p)^k}{k!}\\ &= e^{-30}e^{30p}\\ &= e^{30(p-1)} \sim 0{,}4041\end{aligned}$$

c) Con la notación del ítem anterior, queremos hallar el mínimo t para el cual $\mathbb{P}(M > 7) > 0{,}99$, que es equivalente a $\mathbb{P}(M \leq 7) < 0{,}01$.

Si definimos $N_t \sim \mathcal{P}(30t)$ y hacemos la misma cuenta que en b) pero poniendo "$30t$" donde dice "30", llegamos a que $\mathbb{P}(M \leq 7) = 0{,}4041^t$. Finalmente, $0{,}4041^t < 0{,}01 \iff t > 5$. Luego, tienen que pasar $t = 6$ años como mínimo.

5. Sean X_1, X_2 variables aleatorias i.i.d. con distribución exponencial de parámetro λ. Sea

$$Y = \frac{X_1}{X_1 + X_2}.$$

   a) (8p) Calcular la distribución de Y.

   b) Si $\lambda = 1$, calcular

      1) (7p) $\mathbb{P}(Y > X_1 + X_2)$

      2) (5p) $\mathbb{P}(Y < 3 \mid X_1 + X_2 > 3)$.

Solución: Teniendo en cuenta todos los ítems, nos conviene calcular la distribución de Y junto con la de $Z = X_1 + X_2$. Por lo tanto calculemos la densidad conjunta del vector (Y, Z), utilizando el TCV.

Notemos que $(Y, Z) = g(X_1, X_2)$ para $g(x_1, x_2) = \left(\frac{x_1}{x_1 + x_2}, x_1 + x_2\right)$, y que $g^{-1}(y, z) = (yz, z - yz)$.

Por lo tanto, $J_{g^1}(y, z) = \det \begin{pmatrix} z & y \\ -z & 1-y \end{pmatrix} = z(1-y) + yz = z.$

Al ser, iid, sabemos que $f_{(X_1,X_2)}(x_1, x_2) = \lambda^2 e^{-\lambda(x_1+x_2)} \mathbf{1}_{(0,+\infty)}(x_1)\mathbf{1}_{(0,+\infty)}(x_2)$.

Entonces, vía TCV,

$$f_{(Y,Z)}(y, z) = f_{(X_1,X_2)}(g^{-1}(y, z))\, |z| = \lambda^2 e^{-\lambda z}\, |z|\, \mathbf{1}_{(0,+\infty)}(yz)\mathbf{1}_{(0,+\infty)}(z - yz)$$

Veamos las indicadoras: $\mathbf{1}_{(0,+\infty)}(yz)$ nos dice que ambas tienen el mismo signo y $\mathbf{1}_{(0,+\infty)}(z-yz)$ nos dice que $0 < z(1-y) < +\infty$. Si fuesen negativas ambas, no se cumpliría esto último. Por lo tanto son ambas positivas, pero entonces $y < 1$. Es decir tenemos la siguiente equivalencia:

$$\begin{cases} 0 < yz \\ 0 < z(1-y) \end{cases} \iff \begin{cases} 0 < z \\ 0 < y < 1 \end{cases}$$

Por lo tanto, $f_{(Y,Z)}(y, z) = \lambda^2 z e^{-\lambda z} \mathbf{1}_{(0,+\infty)}(z)\mathbf{1}_{(0,1)}(y)$. Notar que pudimos sacarle el módulo a z, ya que es positiva.

Además, resultan independientes (pues la densidad conjunta se factoriza como el producto de dos funciones, cada una dependiendo de una variable), y es fácil notar que $Y \sim \mathcal{U}(0, 1)$ y $Z \sim \Gamma(2, \lambda)$.

Para b),

$$\mathbb{P}(Y > Z) = \mathbb{P}((Y, Z) \in B) = \iint_B f_{YZ}\,dydz$$

con $B = \{(y, z) \in \mathbb{R}^2 : y > z\}$. Por lo tanto

$$\mathbb{P}(Y > Z) = \int_0^1 \int_z^1 z e^{-z}\,dydz = \frac{3}{e} - 1$$

Para c), notemos que al ser independientes esa probabilidad sería igual a $\mathbb{P}(Y < 3)$, pero al ser Y uniforme $(0, 1)$, esa probabilidad da 1.