Universidad de la República
Facultad de Ingeniería - IMERL

Probabilidad y Estadística
Segundo Semestre 2016

SOLUCIÓN PRIMER PARCIAL
MARTES 27 DE SETIEMBRE 2016.

Ejercicio 1. [**16 puntos**]

El número de clientes por día que compra diarios en un quiosco céntrico de la ciudad de Montevideo, se puede modelar según una variable aleatoria Poisson de párametro λ. Cada día el quiosco tiene N diarios a disposición.

1. Hallar $c(\lambda, k)$ constante que depende de λ y de k tal que $P(X = k + 1) = c(\lambda, k)P(X = k)$.
2. Para todo lo que sigue se asume ahora que $\lambda = 4$:

   (a) Utilizar la recursión anterior para hallar primer cuartil, mediana, moda (o modas) y tercer cuartil de la distribución de X. Puede ser de utilidad realizar una tabla como la que se muestra:

k	$P(X = k)$	$P(X \leq k)$

   (b) Hallar μ el valor esperado de X. ¿Cuál es la probabilidad de que el número de clientes en un día sea mayor que μ?

   (c) ¿Cuál es el mínimo número de diarios que debe disponer el quiosco para que la probabilidad de que la demanda supere a la oferta sea menor que 0.1?

3. Si consideramos que el quiosco está abierto de lunes a sábados y que el stock es $N = 7$:

   (a) ¿Cuál es la probabilidad de que ningún día de la semana la demanda supere el stock?

   (b) ¿Cuál es la probabilidad de que el sábado sea el primer día en que la demanda supere al stock?

4. Asumiendo que:
   - cada diario tiene un costo de 70 pesos y se vende a 100 pesos
   - el stock es $N = 7$

   ¿Cuál es la ganancia diaria esperada por venta de diarios? Tener en cuenta que lo recaudado puede no cubrir el costo del stock.

Ejercicio 1 - Solución

1. Por ser X Poisson de parámetro λ resulta que:

$$\frac{\mathbb{P}(X = k + 1)}{\mathbb{P}(X = k)} = \frac{e^{-\lambda}\lambda^{k+1}}{(k + 1)!} \times \frac{k!}{e^{-\lambda}\lambda^{k}} = \frac{\lambda}{k + 1}$$

   Por lo tanto $c(\lambda, k) = \frac{\lambda}{k+1}$.

   (a) Si $\lambda = 4$, entonces la tabla resulta:

k	$\mathbb{P}(X=k)$	$\mathbb{P}(X\leq k)$
0	0.0183	0.0183
1	0.073	0.0913
2	0.146	0.2373
3	0.195	0.4323
4	0.195	0.6273
5	0.156	0.7833
6	0.104	0.8873
7	0.059	0.9463

Luego, $Q_1 = 3$, $Q_2 = 4$ (mediana de X), $Q_3 = 5$ y las modas son 3 y 4.

(b) Como $X \sim \mathcal{P}(4)$, entonces $\mu = \mathbb{E}(X) = \lambda = 4$:

$$\mathbb{E}(X) = \sum_{k\geq 0} kP(X=k) = \sum_{k\geq 0} k\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k\geq 1}\frac{\lambda^k}{(k-1)!} = e^{-\lambda}\lambda\sum_{j\geq 0}\frac{\lambda^j}{j!} = \lambda.$$

Entonces:

$$\mathbb{P}(X > \mu) = 1 - \mathbb{P}(X \leq 4) = 1 - 0.6273 = 0.3727.$$

(c) Buscamos el primer $s = stock$ tal que:

$$\mathbb{P}(X > s) < 0.1 \Leftrightarrow 1 - \mathbb{P}(X \leq s) < 0.1 \Leftrightarrow \mathbb{P}(X \leq s) > 0.9.$$

De la tabla obtenemos entonces que $s = 7$.

2. (a) Sean X_1 la demanda de diarios del lunes (día 1), X_2 la demanda del martes (día 2), ..., X_6 la demanda del sábado (dá 6). Asumimos que las demandas de días diferentes son independientes, luego la probabilidad de que ningún día de la semana la demanda supere al stock es:

$$\mathbb{P}(X_1 \leq 7, ..., X_6 \leq 7) = 0.9463^6 = 0.7181$$

También es posible definir la variable aleatoria Y número de días en la semana en que la demanda supera al stock. Entonces se tiene que $Y \sim \text{Bin}(n,p)$ donde $n = 6$ y $p = P(\text{un día la demanda supere al stock}) = P(X > 7) = 0.0537$. Por lo tanto la probabilidad de que ningún día de la semana la demanda supere el stock es

$$P(Y = 0) = (1-p)^6 = 0.7181.$$

(b) En este caso:

$$\mathbb{P}(X_1 \leq 7; ...; X_5 \leq 7; X_6 > 7) = (1 - 0.0537)^5 \times 0.0537 = 0.0407$$

Como antes es posible definir una variable aleatoria Z primer día en que la demanda supere el stock, entonces $Z \sim \text{Geo}(p)$. Por lo tanto la probabilidad de que el primer día sea el sábado está dado por $P(Z = 6) = (1-p)^5 p = 0.0407$.

3. Sea V la variable aleatoria que cuenta cuántos diarios se venden cada día. El recorrido es $\mathcal{R}_V = \{0, 1, 2, ..., 6, 7\}$ y la función de probabilidad puntual está dada por $\mathbb{P}(V = k) = \mathbb{P}(X = k)$ (la de la tabla anterior) si $k = 0, 1, ..., 6$. Finalmente, $V = 7$ cuando o bien vinieron exactamente 7 clientes a comprar diarios (*i.e.* $\{X = 7\}$), o bien vinieron más de

7 clientes a comprar diarios por lo que a partir del octavo cliente todos se fueron con las manos vacías (*i.e.* cuando $\{X > 7\}$). Es decir que:

$$\{V = 7\} = \{X \geq 7\}.$$

Luego la ganancia según el número de diarios vendidos es:

$$G(V) = \text{ ingresos - costo fijo } = 100 \times V - 70 \times 7$$

y

$$\mathbb{E}(G(V)) = 100 \times \mathbb{E}(V) - 70 \times 7.$$

Como

$$\begin{aligned}\mathbb{E}(V) &= \sum_{k=0}^{k=7} k\mathbb{P}(V = k) = \sum_{k=0}^{k=6} k\mathbb{P}(X = k) + 7\mathbb{P}(X \geq 7) \\ &= 3.134 + 7 \times (1 - 0.8873) = 3.923.\end{aligned}$$

entonces

$$\mathbb{E}(G(V)) = 100 \times 3.923 - 70 \times 7 = -97.65.$$

Ejercicio 2. [**16 puntos**]

1. Se supone que tenemos tres letras: A, B y C. Para formar una palabra de 4 letras se elige cada letra al azar de manera independiente:
   (a) ¿Cuál es la probabilidad de formar la palabra "ABBA"?
   (b) ¿Cuál es la probabilidad de obtener una palabra que empiece en A y termine en A?
2. Supongamos ahora que las sucesivas letras no son independientes y que cada letra depende únicamente de la letra anterior. Sea $p_i(j) = P(j|i)$ la probabilidad de transición de i a j , es decir la probabilidad de que luego de la letra i aparezca la letra j. Además $p(i)$ es la probabilidad de aparición de la letra i y $p(i, j)$ es la probabilidad de aparición del par (i, j), es decir la letra i seguida de la letra j.

   Supongamos que tenemos las siguientes probabilidades de aparición:

   Table 1: Probabilidades de aparición $p(i, j)$

		j A	B	C
	A	0	1/5	1/5
i	B	1/15	1/15	1/15
	C	1/5	1/5	0

   De la tabla tenemos por ejemplo que $P(A, B) = 1/5$, lo que significa que las letras A y B aparecen juntas y en ese orden con probabilidad 1/5. Además la combinación AA tiene probabilidad cero de aparecer.

   (a) Hallar $P(A)$, $P(B)$ y $P(C)$.
   (b) Completar la siguiente tabla:
   (c) Se forma una palabra de 4 letras eligiendo la primer letra al azar, y eligiendo las siguientes letras de acuerdo a las probabilidades de transición halladas en la parte anterior:
      i. ¿Cuál es la probabilidad de formar la palabra ABBA?

Table 2: Probabilidades de transición $p_i(j)$

		j A	B	C
	A			
i	B			
	C			

ii. ¿Cuál es la probabilidad de obtener una palabra que empiece en A y termine en A?

Ejercicio 2 - Solución

1. (a) Como las letras se eligen de manera independiente se tiene que:

$$P(\text{“ABBA”}) = \left(\frac{1}{3}\right)^4 = \frac{1}{81}$$

   (b) Sumando todas las probabilidades de palabras que comienzan en A y terminan en A tenemos que:

$$P(A\cdot\cdot A) = 9\left(\frac{1}{3}\right)^4 = \frac{1}{9}$$

2. (a)

$$P(A) = P(A,A) + P(A,B) + P(A,C) = 0 + \frac{1}{5} + \frac{1}{5} = \frac{2}{5}.$$

   De igual modo para B y C se tiene que:

$$P(B) = \frac{1}{15} + \frac{1}{15} + \frac{1}{15} = \frac{1}{5}$$
$$P(C) = \frac{1}{5} + \frac{1}{5} + 0 = \frac{2}{5}$$

   (b) Completar la siguiente tabla:

Table 3: Probabilidades de transición $p_i(j)$

		j A	B	C
	A	0	1/2	1/2
i	B	1/3	1/3	1/3
	C	1/2	1/2	0

   Donde para calcular cada entrada se utiliza que $p_i(j) = \frac{p(i,j)}{P(j)}$. Por ejemplo,

$$P(B|A) = P(A,B)P(A) = \frac{1/5}{2/5} = \frac{1}{2}$$

   También es posible pensar del siguiente modo: la combinación AA tiene probabilidad 0, mientras que las combinaciones AB y AC tienen igual probabilidad de aparición (1/5), por lo tanto dado que la letra es A, tenemos igual probabilidad de que la siguiente letra sea B o C y probabilidad 0 de que sea A. Por lo tanto $P(B|A) = P(B|C) = 1/2$ y $P(A|A) = 0$.

(c) i. Usando las probabilidades halladas en la parte anterior se tiene que:

$$P(ABBA) = P(A)P(B|A)P(B|B)P(A|B) = \frac{2}{5} \times \frac{1}{2} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{45}$$

ii. Nuevamente usando las partes anteriores tenemos:

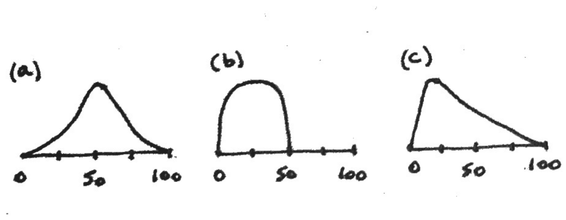
$$P(A \cdot \cdot A) = \sum_{k_1, k_2 \in \{A,B,C\}} P(Ak_1k_2A)$$

Debido a que los pares AA y CC no aparecen juntos dicha suma se puede desarrollar explícitamente como sigue:

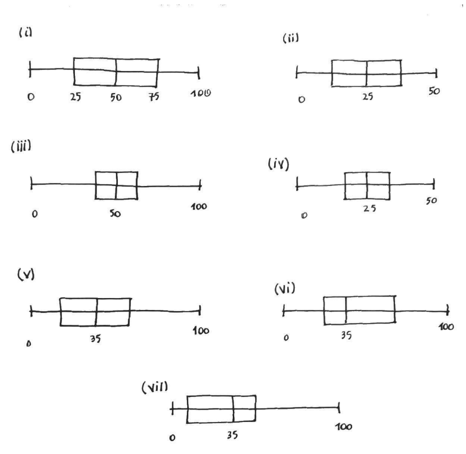
$$P(ABBA) + P(ABCA) + P(ACBA) = \frac{2}{5}\left(\frac{1}{3}\frac{1}{3}\frac{1}{2} + \frac{1}{2}\frac{1}{3}\frac{1}{2} + \frac{1}{3}\frac{1}{2}\frac{1}{2}\right) = \frac{4}{45}.$$

Ejercicio 3. [**8 puntos**]

Abajo se bosquejan histogramas de tres tablas de datos.



1. Indicar aproximadamente el promedio de cada uno de los datos.
2. En cada caso indicar si la mediana es menor, igual o mayor que el promedio.
3. Indicar cuál de los siguientes boxplots se corresponden con los histogramas de las parte 1:



Ejercicio 3 - Solución

1. Por simetría del histograma, en la primer gráfica el promedio es aproximadamente 50. De modo similar en la segunda gráfica el promedio es aproximadamente 25.

   En la tercera gráfica, que es ligeramente asimétrica hacia la izquierda, el promedio es aproximadamente 40.

2. En las primeras dos gráficas la mediana es igual al promedio, por ser estas simétricas. En la tercera gráfica la mediana es menor al promedio ya que el histograma es asimétrico y concetra su área sobre el sector izquierdo.

3. El primer histograma es simétrico y su mediana es 50. Los boxplot posibles para ello son el (i) y el (iii). Observando la dispersión de los datos respecto de la media se concluye que el boxplot más adecuado es el (i).

   El segundo histograma también es simétrico pero con mediana 25 con lo cual los boxplot posibles seían preliminarmente (ii) y (iv). Nuevamente observando la dispersíon vemos que el que mejor se ajusta es el (ii). Sin embargo, si la justificación lo amerita el boxplot (iv) también podría considerarse correcto.

   El tercer histograma es asimétrico y es razonable suponer que su mediana es 35 lo cual deja disponibles los boxplot (v), (vi) y (vii). Como los datos son más dispersos a derecha y menos dispersos a izquierda de la mediana, el boxplot coherente es el (vi).