1	2	3	4	Calificación

Probabilidad y Estadística (C)

Recuperatorio del primer parcial - 14/07/2022

Complete esta hoja y entréguela con el resto del examen. ***Realizar cada ejercicio en hoja separada****. Escribir el nombre en cada una. Al retirarse debe firmar una hoja de asistencia.*

APELLIDO Y NOMBRES:.................................... Nº DE LIBRETA: ....

mail:..........................@.................... FIRMA:......................

Turno: | Tarde: 14 a 17 hs | Noche: 19 a 22 hs | Nº de hojas entregadas:

Criterio de aprobación: Para aprobar este examen es necesario sumar al menos 60 puntos y tener un ejercicio bien resuelto.

Recuerde definir con palabras los eventos y/o las variables aleatorias involucradas, nombres y parámetros de las distribuciones. ***Justifique claramente sus afirmaciones****.*

1. (25 puntos)

   (a) Para un estudio sobre trastornos en el sueño en una población se tiene la siguiente información:

   - El 18% de las mujeres y el 14% de los varones tienen trastornos en el sueño.
   - El 16,2 % de la población tiene trastornos en el sueño.

   i) (10 puntos) ¿Cuál es la probabilidad de que una persona con transtornos del sueño sea mujer?

   ii) (5 puntos) Calcular la probabilidad de que una persona elegida al azar sea mujer y no tenga trastornos en el sueño.

   (b) (10 puntos) En un edificio viven 11 personas, de las cuales 4 tienen trastornos en el sueño. Se eligen al azar 6 habitantes del edificio, ¿cuál es la probabilidad de que exactamente 3 tengan trastornos en el sueño?

2. (25 puntos) Sea X la cantidad de horas semanales que Juan estudia Probabilidad y Estadística (C). Su distribución puede modelarse con la siguiente función de densidad:

$$f_X(x) = \begin{cases} \dfrac{18}{x^3} & \text{si } x \geq 3 \\ 0 & \text{en cc} \end{cases}.$$

   (a) (9 puntos) Juan tiene 8 semanas (inclusive) para estudiar antes de rendir el primer parcial y se sentirá confiado para aprobar cuando llegue a la sexta semana en la que estudie más de 7 horas semanales. Hallar la probabilidad de que Juan llegue confiado al examen. Suponer que la cantidad de horas semanales que Juan estudia esta materia son independientes.

   (b) (8 puntos) Sea $x_{0.1}$ el 10-percentil de X. Calcular $P(X < 2\, x_{0.1} \mid X > x_{0.1})$.

   (c) (8 puntos) Juan mira Netflix $Y = e^{9-2X}$ horas semanales (con lo cual, depende del tiempo X que le dedica a estudiar esta materia). Hallar la función de densidad de la variable aleatoria Y.

3. (25 puntos) Se tira un dado equilibrado 5 veces seguidas. Sea X_i la variable aleatoria que cuenta la cantidad de veces que aparece el número i (para i entre 1 y 6).

   (a) (15 puntos) Calcular $P(X_1 < 2 | X_3 = 1)$.

   (b) (5 puntos) Calcular $P(X_1 + 2X_3 = 3)$

   (c) (5 puntos) ¿Son X_1 y X_3 independientes? Justificar

4. (25 puntos) El cortador de césped de un club de golf cobra \$600 por hectárea cortada y paga diariamente \$200 de combustible para la máquina cortadora. La cantidad de hectáreas que corta en un día está dada por una variable aleatoria X con función de densidad:

$$f_X(x) = \begin{cases} 1 - \frac{x}{4} & \text{si } 1 \leq x \leq 3 \\ 0 & \text{en otro caso} \end{cases} .$$

   Sea $G =$"ganancia (neta) diaria del cortador de césped".

   (a) (8 puntos) ¿Cuál es la probabilidad de que un día de trabajo gane más de \$850 ?

   (b) (8 puntos) Verificar que $E(G) = 900$ y $\text{Var}(G) = 110000$.

   (c) (9 puntos) Supongamos independencia entre los cortes de los distintos días. ¿Cuántos días como mínimo (aproximadamente) tiene que trabajar para que su ganancia promedio supere los \$850 con probabilidad 0.89 .

Recuperatorio primer parcial Proba C

Ejercicio 1

a) Para calcular la proba de que exactamente 3 de las 6 elegidas a la azar tengan transtorno podemos hacerlo modelando el espacio muestral como el conjunto de todas las formas de elegir 6 de 11. Los casos favorables son todas las formas de elegir 3 de los 4 que tienen transtornos y de los 7 que no, elegir otros 3. Es decir se tiene una distribución hipergeométrica de tamaño total de población $N = 11$, tamaño de muestra extraída $n = 6$ y cantidad de éxitos $e = 4$.

Si $X \sim H(N, n, e)$ entonces $P(X = k) = \frac{\binom{e}{k}\binom{N-e}{n-k}}{\binom{N}{n}}$.

$$P(\text{ exactamente 3 personas de las 6 elegidas tengan transtornos}) = \frac{\binom{4}{3}\binom{7}{3}}{\binom{11}{6}} = 0.3030303$$

b) Sean los siguientes eventos:

$$T = \text{ Tiene transtornos de sueño} \qquad M = \text{ ser mujer ,} \qquad H = \text{ ser hombre}$$

entonces la información que nos dan es:

$$P(T|M) = 0.18, \qquad P(T|H) = 0.14, \qquad P(T) = 0.162$$

i) Nos piden calcular $P(M|T)$. Para esto usamos Bayes:

$$P(M|T) = \frac{P(T|M)P(M)}{P(T)} = \frac{0.18P(M)}{0.162}$$

Para calcular $P(M)$ usamos probabilidad total donde particionamos de acuerdo a los dos eventos M y H:

$$P(T) = 0.162 = P(T|M)P(M) + P(T|H)(1 - P(M))$$

despejando $P(M) = \frac{0.162-0.14}{0.04} = 0.55$ Entonces $P(M|T) = \frac{0.18P(M)}{0.162} = 0.6111111$

ii) Nos piden $P(M \cap T^c) = P(T^c|M)P(M) = (1 - P(T|M))P(M) = 0.451$

Ejercicio 2

$X =$ cantidad de horas semanales tiene densidad

$$f_X(x) = \frac{18}{x^3} I_{[3,+\infty]}(x)$$

a) Calculemos primero el valor de $x_{0.1}$:

$$0.1 = P(X < x_{0.1}) = \int_3^{x_{0.1}} \frac{18}{x^3} dx = -\left[\frac{9}{x_{0.1}^2} - \frac{9}{3^2}\right]$$

despejando $x_{0.1} = \sqrt{10}$.

Para calcular la proba pedida:

$$P(X < 2x_{0.1}|X > x_{0.1}) = \frac{P(\{X > 2x_{0.1}\} \cap \{X > x_{0.1}\})}{P(X > x_{0.1})} = \frac{P(\sqrt{10} < X < 2\sqrt{10})}{P(X > \sqrt{10})}$$

Pero $P(X > \sqrt{10}) = P(X > x_{0.1}) = 1 - 0.1 = 0.9$ y

$$P(\sqrt{10} < X < 2\sqrt{10}) = \int_{\sqrt{10}}^{2\sqrt{10}} \frac{18}{x^3} dx = -\left[\frac{9}{(2\sqrt{10})^2} - \frac{9}{\sqrt{10}^2}\right] = \frac{27}{40}$$

luego

$$P(X < 2x_{0.1} | X > x_{0.1}) = \frac{27/40}{0.9} = \frac{27 * 10}{40 * 9} = \frac{3}{4}$$

b) Sea $p = P(\text{Juan estudie más de 7 horas en una semana}) = P(X > 7) = \int_7^{+\infty} \frac{18}{x^3} dx = \frac{9}{49}$.

Sea Y la variable que da cuenta de la cantidad de semanas en las que Juan estudia más de 7 horas, entonces $Y \sim Bin(8, p)$, con $p = 9/49$ calculada arriba.

Juan se siente confiado si estudia 6 o mas semanas, más de 7 horas por dia. La probabilidad de que Juan llegue confiado al examen es

$$P(Y \geq 6) = P(Y = 6) + P(Y = 7) + P(Y = 8) = 0.0007638$$

c) $Y = \exp(9 - 2X)$. Para calcular f_Y, calculamos su acumulada $F_Y(y)$ y derivamos con respecto de y. si $y > 0$:

$$F_Y(y) = P(Y \leq y) = P(\exp(9 - 2X) \leq y) = P\left(X \geq \frac{9 - \ln(y)}{2}\right) = \int_{(9-\ln(y))/2}^{+\infty} \frac{18}{x^3} dx = \frac{9}{\left(\frac{9-\ln(y)}{2}\right)^2} = \frac{36}{(9 - \ln(y))^2}$$

La ultima igualdad vale si $\frac{9-\ln(y)}{2} > 3$ es decir si $y < e^3$, de lo contrario la integral es 0 y $F_Y(y) = 0$. Entonces, derivando con respecto de y

$$f_Y(y) = \frac{72}{y(9 - \ln(y))^3} I_{(0,e^3)}(y)$$

Ejercicio 3

Se tira un dado equilibrado 5 veces. Sea X_i = cantidad de veces que sale el número i con $i = 1, 2, 3, 4, 5, 6$.

Sea Y = sale el 2, el 4, el 5 o el 6. Consideremos el vector aleatorio $(X_1, X_3, Y) \sim Multinom(5, p_1, p_3, p_y)$ donde $p_1 = p_3 = 1/6$ son las probabilidades de que salga el 1 y el 3 respectivamente. p_y es la probabilidad de que salga el 2 o el 4 o el 5 o el 6, luego $p_y = 1/6 + 1/6 + 1/6 + 1/6 = \frac{4}{6}$.

a)

$$\begin{aligned} P(X_1 < 2 | X_3 = 1) &= P(X_1 = 0 | X_3 = 1) + P(X_1 = 1 | X_3 = 1) \\ &= \frac{P(X_1 = 0, X_3 = 1, Y = 4)}{P(X_3 = 1)} + \frac{P(X_1 = 1, X_3 = 1, Y = 3)}{P(X_3 = 1)} \end{aligned}$$

Pero

$$P(X_1 = 0, X_3 = 1, Y = 4) = \binom{5}{1}\binom{4}{4}(1/6)^1(4/6)^4$$

esto sale de: de las 5 tiradas elijo una donde cae el num 3 y de las 4 restantes elijo 4 donde caen los otros numeros que no son ni el 1 ni el 3.

$$P(X_1 = 1, X_3 = 1, Y = 3) = \binom{5}{1}\binom{4}{1}\binom{3}{3}(1/6)^1(1/6)^1(4/6)^3$$

esto pues: de las 5 tiradas elijo una donde cae el num 1, de las 4 restantes elijo 1 donde el num 3 y de las 3 restantes elijo 3 donde caen los otros numeros que no son ni el 1 ni el 3.

Además $P(X_3 = 1) = \binom{5}{1}(1/6)^1(5/6)^4$

luego

$$
\begin{aligned}
P(X_1 < 2|X_3 = 1) &= \frac{\binom{5}{1}\binom{4}{4}(1/6)^1(4/6)^4 + \binom{5}{1}\binom{4}{1}\binom{3}{3}(1/6)^1(1/6)^1(4/6)^3}{\binom{5}{1}(1/6)^1(5/6)^4} \\
&= \frac{5 * (4^4/6^5) + 5 * 4 * 4^3/6^5}{5 * 5^4/6^5} = 0.8192
\end{aligned}
$$

b) Calcular $P(X_1 + 2X_3 = 3)$. Para que ocurra que $X_1 + 2X_3 = 3$ o bien $X_1 = 1$ y $X_3 = 1$ o bien $X_1 = 3$ y $X_3 = 0$, ambos eventos disjuntos, luego:

$$
\begin{aligned}
P(X_1 + 2X_3 = 3) &= P(X_1 = 3, X_3 = 0, Y = 2) + P(X_1 = 1, X_3 = 1, Y = 3) \\
&= \binom{5}{3}\binom{2}{2}(1/6)^3(4/6)^2 + \binom{5}{1}\binom{4}{1}\binom{3}{3}(1/6)^1(1/6)^1(4/6)^3 = 0.1851852
\end{aligned}
$$

¿Son independientes X_1 y X_3?

$$
\begin{aligned}
P(X_1 = 1, X_3 = 0) &= \binom{5}{1}\binom{4}{4}(1/6)^1(4/6)^4 = 0.1646091 \\
P(X_1 = 1)P(X_3 = 0) &= \binom{5}{1}(1/6)^1(5/6)^4\binom{5}{0}(1/6)^0 * (5/6)^5 = 0.1615056
\end{aligned}
$$

Luego no son independientes!

Ejercicio 4

La cantidad de hectáreas que corta en un dia es $X \sim f_X(x) = (1 - \frac{x}{4})I_{[1,3]}(x)$. Cobra \$600 por ha cortada y paga \$200 de combustible. La ganacia neta diaria es $G = 600X - 200$.

a)

$$
\begin{aligned}
P(G \geq 850) = P(600X - 200 \geq 850) = P(X \geq 7/4) &= \int_{7/4}^{3} (1 - x/4)dx \\
&= (3 - 9/8) - (7/4 - (7/4)^2/8) = 0.5078125
\end{aligned}
$$

b) Calulemos $E(G)$ y $Var(G)$:

$$
E(G) = E(600X - 200) = 600E(X) - 200 = 600 * \int_{1}^{3} x(1 - x/4)dx - 200 = 600 * \frac{11}{6} - 200 = 900
$$

$$
\begin{aligned}
Var(G) &= Var(600X - 200) = 600^2 Var(X) = 600^2(E(X^2) - E(X)^2) \\
&= 600^2\left[\int_{1}^{3} x^2(1 - x/4)dx - (11/6)^2\right] = 600^2(11/3 - 11^2/6^2) = 110000
\end{aligned}
$$

c) Sea $\bar{G}_n = \frac{\sum_{i=1}^{n} G_i}{n}$ la ganancia promedio en n dias. Queremos hallar n tal que $P(\bar{G}_n \geq 850) \geq 0.89$.

Por TCL sabemos que $\sqrt{n}\frac{(\bar{G}_n - E(G_1))}{\sqrt{Var(G_1)}}$ sigue aproximadamente una distribución $\mathcal{N}(0,1)$, luego

$$P(\bar{G}_n \geq 850) = P\left(\sqrt{n}\frac{(\bar{G}_n - E(G_1))}{\sqrt{Var(G_1)}} \geq \sqrt{n}\frac{850 - E(G_1)}{\sqrt{Var(G_1)}}\right) \approx P\left(Z \geq \frac{-50\sqrt{n}}{\sqrt{110000}}\right) = 1 - \Phi\left(\frac{-50\sqrt{n}}{\sqrt{110000}}\right) \geq 0.89$$

donde $Z \sim \mathcal{N}(0,1)$.

Despejando n:

$$\sqrt{n} \geq -\frac{\Phi^{-1}(0.11)\sqrt{110000}}{50} = 8.1358671$$

Notar que $\Phi^{-1}(0.11) < 0$ luego $\sqrt{n}$ es mayor que una cantidad positiva, elevando al cuadrado en ambos lados: $n \geq 66.1923341$, es decir, para que la ganacia promedio supere \$850 debe trabajar al menos 67 días.