Estadística

Examen Parcial 2

21 de Abril de 2022 (Curso 2021-2022/2)
Resuelve el problema en la hoja del enunciado.
Anota tu nombre completo en mayúsculas, DNI, grupo e identificador (Id).

APELLIDOS:.. **NOMBRE:**........................

Puedes utilizar una calculadora no programable. **DNI:** .... **GRUPO:**........ **Id:** ....

Problema 1

1. [**30%**] Cierto tipo de componente está empaquetado en lotes de cuatro. Sea X el número de componentes que funcionan de modo adecuado en un lote elegido de manera aleatoria. Suponiendo que la probabilidad de que exactamente x componentes funcionen es proporcional a x; en otras palabras, suponiendo que la función de probabilidad de X venga dada por

$$p(x) = \begin{cases} c \cdot x, & x \in \{1, 2, 3, 4\} \\ \\ 0, & \text{de otro modo} \end{cases}$$

donde c es una constante. Se pide:

(a) Encuentra el valor de c para que $p(x)$ sea función de probabilidad y calcula la media del número de componentes que funcionan adecuadamente en un lote. [**5 puntos**]

(b) El precio de cada lote está relacionado con el número de componentes que funcionan adecuadamente en el mismo. Definimos la variable Y como el precio en euros de un lote. Si la relación entre el precio del lote y las componentes que funcionan adecuadamente es $Y = 15 \cdot X + 5$, calcula la desviación estándar del precio de los lotes. [**5 puntos**]

Solución

(a) Encuentra el valor de c para que $p(x)$ sea función de probabilidad y calcula la media del número de componentes que funcionan adecuadamente. Para que $p(x)$ sea función de probabilidad su suma debe valer 1:

$$\sum_{-\infty}^{\infty} p(x) = 1$$

$$\sum_{x=1}^{x=4} c \cdot x = 1$$

$$c(1 + 2 + 3 + 4) = 1$$

$$c = \frac{1}{10}$$

Calculamos el valor medio:

$$E(X) = \sum_{-\infty}^{\infty} x \cdot p(x) = \sum_{x=1}^{x=4} x \cdot c \cdot x = \sum_{x=1}^{x=4} c \cdot x^2 = \frac{1}{10}(1^2 + 2^2 + 3^2 + 4^2) = \frac{1}{10}30 = 3$$

(b) El precio de cada lote está relacionado con el número de componentes que funcionan adecuadamente en el mismo. Definimos la variable Y como el precio en euros de un lote. Si la relación entre el precio del lote y las componentes que funcionan adecuadamente es $Y = 15 \cdot X + 5$, calcula la desviación estándar del precio de los lotes.

La relación de las varianzas de dos variables aleatorias ligadas linealmente, $Y = a \cdot X + b$, es:

$$VAR(Y) = a^2 VAR(X)$$

Por su parte, la varianza de X la podemos calcular a partir de la expresión:

$$VAR(X) = E(X^2) - (E(X))^2$$

Del apartado anterior aprovechamos el valor de $E(X) = 3$ y ahora calculamos $E(X^2)$:

$$E(X^2) = \sum_{-\infty}^{\infty} x^2 \cdot p(x) = \sum_{x=1}^{x=4} x^2 \cdot c \cdot x = \sum_{x=1}^{x=4} c \cdot x^3 = \frac{1}{10}(1^3 + 2^3 + 3^3 + 4^3) = \frac{1}{10}100 = 10$$

En consecuencia, la varianza de X es $VAR(X) = E(X^2) - (E(X))^2 = 10 - 3^2 = 1$ y la varianza de Y vale $VAR(Y) = 15^2 VAR(X) = 225 \cdot 1 = 225$.

Así, la desviación estándar de Y es $\sigma_Y = \sqrt{VAR(Y)} = \sqrt{225} = 15$

Estadística

Examen Parcial 2

21 de Abril de 2022 (Curso 2021-2022/2)
Resuelve el problema en la hoja del enunciado.
Anota tu nombre completo en mayúsculas, DNI, grupo e identificador (Id).

APELLIDOS:.. **NOMBRE:**.......................

Puedes utilizar una calculadora no programable. **DNI:** .... **GRUPO:**........ **Id:** ....

Problema 2

2. [**30%**] Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} x\alpha^2 e^{-(0.5(x\alpha)^2)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

donde α es un parámetro. Se pide:

(a) Calcula la función de distribución de la variable aleatoria continua X (es decir $p(X \leq x)$). Sugerencia: cambio de variable $t = 0.5(x\alpha)^2$. [**5 puntos**]

(b) Encuentra los cuartiles de X en función del parámetro $\alpha > 0$. [**5 puntos**]

Solución

(a) Comprueba que f es una función de densidad.

La función $f(x)$, por las funciones básicas que la componen no es nunca negativa, por lo que únicamente nos falta probar que su área es 1:

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

$$\int_{-\infty}^{\infty} f(x) \cdot dx = \int_{-\infty}^{0} f(x) \cdot dx + \int_{0}^{\infty} f(x) \cdot dx = \int_{-\infty}^{0} 0 \cdot dx + \int_{0}^{\infty} x \cdot \alpha^2 \cdot e^{-0.5 \cdot (x \cdot \alpha)^2} \cdot dx$$

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 0 + \int_{0}^{\infty} x \cdot \alpha^2 \cdot e^{-0.5 \cdot (x \cdot \alpha)^2} \cdot dx$$

Cambio de variable: $t = 0.5 \cdot \alpha^2 \cdot x^2$, lo que implica que $dt = \alpha^2 \cdot x \cdot dx$

$$\int_{-\infty}^{\infty} f(x) \cdot dx = \int_{0}^{\infty} x \cdot \alpha^2 \cdot e^{-0.5 \cdot (x \cdot \alpha)^2} \cdot dx = \int_{0}^{\infty} e^{-t} \cdot dt = \left[-e^{-t}\right]_0^{\infty} = 0 - (-1)) = 1$$

(b) Calcula la función de distribución de la variable aleatoria continua X (es decir $p(X \leq x)$).

Usaremos los cálculos del apartado anterior para deducir la función de distribución de X:

$$F(x) = P(X \leq x)$$

Para $x < 0$ la fución de distribución es 0.
Para $x > 0$:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y) \cdot dy = \int_{-\infty}^{0} f(y) \cdot dy + \int_{0}^{x} f(y) \cdot dy = \int_{-\infty}^{0} 0 \cdot dy + \int_{0}^{x} y \cdot \alpha^2 \cdot e^{-0.5 \cdot (y \cdot \alpha)^2} \cdot dy$$

$$F(x) = P(X \leq x) = \int_{0}^{0.5(x \cdot \alpha)^2} e^{-t} dt = \left[-e^{-t}\right]_0^{0.5(x \cdot \alpha)^2} = -e^{-x} - (-1) = 1 - e^{-0.5(x \cdot \alpha)^2}$$

(c) Encuentra los cuartiles de X en función del parámetro $\alpha > 0$.

Los cuartiles son los valores que reparten la variable aleatoria en sectores de un 25% de probabilidad cada uno.
Q_1 es el valor que acumula una probabilidad del 25%, es decir: $F(Q_1) = 0.25$
Q_2 es el valor que acumula una probabilidad del 50%, es decir: $F(Q_2) = 0.5$
Q_3 es el valor que acumula una probabilidad del 75%, es decir: $F(Q_3) = 0.75$

Encontramos cada uno de ellos a partir de la expresión de $F(x)$:

$$1 - e^{-0.5\cdot(Q_1\cdot\alpha)^2} = 0.25$$

$$e^{-0.5\cdot(Q_1\cdot\alpha)^2} = 1 - 0.25$$

Aplicando la función logaritmo a la expresión anterior:

$$\ln\left(e^{-0.5\cdot(Q_1\cdot\alpha)^2}\right) = \ln(0.75)$$

$$-0.5\cdot(Q_1\cdot\alpha)^2 = \ln(0.75)$$

$$(Q_1\cdot\alpha)^2 = -2\cdot\ln(0.75)$$

Usamos la raiz cuadrada para aislar Q_1:

$$\sqrt{(Q_1\cdot\alpha)^2} = \sqrt{-2\cdot\ln(0.75)}$$

$$Q_1\cdot\alpha = \sqrt{-2\cdot\ln(0.75)}$$

$$Q_1 = \frac{\sqrt{-2\cdot\ln(0.75)}}{\alpha}$$

Semejantemente hallaremos los valores de Q_2 y Q_3:

$$Q_2 = \frac{\sqrt{-2\cdot\ln(0.50)}}{\alpha}$$

$$Q_3 = \frac{\sqrt{-2\cdot\ln(0.25)}}{\alpha}$$

Estadística

Examen Parcial 2

21 de Abril de 2022 (Curso 2021-2022/2)
Resuelve el problema en la hoja del enunciado.
Anota tu nombre completo en mayúsculas, DNI, grupo e identificador (Id).

APELLIDOS:.. **NOMBRE:**.......................

Puedes utilizar una calculadora no programable. **DNI:** .... **GRUPO:**........ **Id:** ....

Problema 3

3. [**40%**] Estamos ayudando al departamento de matemáticas a planear un taller de programación de R avanzado. La universidad exige tener **un total de** 5 chicas en el taller. Teniendo en cuenta que 3 de cada 10 participantes de anteriores talleres son chicas, se pide:

(a) ¿Cuál es la probabilidad de abrir el taller con un máximo de 7 plazas si las inscripciones se aceptaran por orden de llegada? [**5 puntos**]

(b) En promedio un participante se apunta cada 2 horas a los talleres ofrecidos por el departamento. ¿Cuántas horas deberíamos abrir el periodo de inscripción para tener una probabilidad del 90% de que se apunten **un total de** 5 chicas? [**5 puntos**]

Solución

(a) Sea X la variable aleatoria que cuenta el número total de chicos inscritos después de observar 5 chicas. X es

por lo tanto una binomial negativa que cuenta el número de ensayos hasta observar $r = 5$ eventos de probabilidad $p = 3/10 = 0.3$ (según el enunciado).

El problema nos pide calcular $P(X \leq 2)$

$P(X \leq 2) =$ $\mathtt{F}(x) =$0.0287955

en R: `pnbinom(2,5,0.3)`

Respuesta: Hay una probabilidad del 2.88% de abrir el curso si las incripciones se hacen por orden de llegada y sólo podemos ofrecer hasta 7 plazas.

(b) Sea X el tiempo hasta el evento que consiste en completar 5 inscripciones de chicas. Si t_i es el periodo de incripción, el problema nos pide calcular su valor tal que: $P(X < t_i) = 0.90$.

Para esto hay que considerar que

1. El problema nos da el valor esperado de observar una inscripcion cada 2 horas. Como cada 3 de 10 incripciones son chicas debemos esperar la incripción de una chica cada $2*10/3 = 20/3$ horas en promedio. Sin embargo, estamos interesados en completar 5 inscripciones de chicas, por lo que tendremos que esperar en promedio 5 veces más, o sea $20/3 * 5 = 100/3$ horas, o 33.3 horas para completar las 5 inscripciones.

2. Por lo tanto, se esperan $\lambda_{5chicas} = 3/100 = 0.03$ eventos de poder completar 5 inscripciones de chicas en una sola hora.

Definiendo la variable aleatoria K como el evento binario de completar 5 incripciones de chicas cada hora, este tiene un valor esperado de $E(K) = \lambda_{5chicas} = 0.03$. Como K es una variable de Poisson, entonces X (que mide el tiempo hasta el primer evento de completar 5 inscripciones de chicas) es una variable exponencial con $\lambda_{5chicas} = 3/100$. Para una densidad de probabilidad exponencial, su función de distribución de probabilidad es:

$P(X < t_i) = F(t_i) = \int_{\infty}^{t_i} \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t_i}$

Por lo tanto

$P(X < t_i) = F(t_i) = 1 - e^{-\lambda t_i} = 0.9$ despejando t_i tenemos

$t_i = -\frac{\ln(0.1)}{\lambda} = 76.75.$

Respuesta: Debemos abrir las incripciones por 76.75 horas.