Clase de repaso

Estadística I - División Moneta Pizarro

Jonás Alvarado   Diana Choulet   Rafael Hofmann
Rosario Martos Vocos   Nicolas Laffitte   Lara Schellhorn

Índice

Enunciados

Ejercicio 1

Un estudio realizado en los bancos de la ciudad de Córdoba sobre el atraso en el pago de cuotas de créditos personales, sugiere que las demoras en ingresar la cuota por los deudores siguen una distribución normal con media de 45 días y desviación estándar de 16 días. Dicha demora se mide en cantidad de días transcurridos desde el vencimiento de la cuota hasta la fecha de pago de la misma. Sobre la base de esta información:

1. Calcular la probabilidad que un crédito al azar tenga una demora superior a 57 días.
2. Calcular la probabilidad que un crédito al azar tenga una demora inferior a 30 días.
3. Calcular entre qué cantidad de días de demora se encuentra el 90 de los deudores, considerando que estos valores sean equidistantes respecto de la media. Ayuda extra:

$$F(1,282) = 0,90$$

$$F(1,645) = 0,95$$

$$F(1,96) = 0,975$$

4. Suponer ahora que se desconoce la distribución de probabilidad de esta variable y encontrar la cota de probabilidad para que la demora de un crédito difiera de la media en más de 20 días.

Ejercicio 2

Un estudio de las filas de clientes en las cajas de un importante banco de la ciudad reveló que el número promedio de clientes atendidos por hora fue igual 4. Determinar la probabilidad de que:

1. Ningún cliente haya sido atendido durante una hora.
2. Más de 3 clientes hayan sido atendidos en un lapso de media hora.
3. Comenzando la observación en cualquier momento de tiempo, calcular la probabilidad de que transcurran como máximo 20 minutos hasta que llegue el primer cliente.

Ejercicio 3

$$f(x) = \begin{cases} \frac{x}{16} & x \in (2,6) \\ 0 & \text{c.c} \end{cases}$$

1. Obtener la función de distribución.
2. Calcular la probabilidad de que x sea mayor a 3.
3. Obtener el Percentil 90.

Resoluciones

Ejercicio 1 (Nicolás y Rafael)

1. **Calcular la probabilidad que un crédito al azar tenga una demora superior a 57 días.**

Se busca la siguiente probabilidad:

$$P(X > 57) = 1 - P(X \leq 57)$$

Dadas las salidas que se disponen como ayuda, será conveniente trabajar con una variable aleatoria $X \sim N(0, 1)$ por lo que será necesario estandarizar:

$$Z = \frac{X - \mu}{\sigma}$$

$$Z = \frac{57 - 45}{16} = 0,75$$

$$P(X > 57) = 1 - P(X \leq 57) = 1 - P(Z \leq 0,75) = 1 - 0,77337 = \boxed{0,22663}$$

2. **Calcular la probabilidad que un crédito al azar tenga una demora inferior a 30 días.**

Se busca la probabilidad:

$$P(X < 30) = P(X \leq 30)$$

Trabajando nuevamente con una distribución normal estándar:

$$Z = \frac{30 - 45}{16} = -0,9375$$

$$P(X < 30) = P(X \leq 30) = P(Z \leq -0,9375) = 1 - P(Z > -0,9375) = 1 - 0,82639 = \boxed{0,17361}$$

3. **Calcular entre qué cantidad de días de demora se encuentra el 90% de los deudores, considerando que estos valores sean equidistantes respecto de la media.**

Lo que se busca es:

$$P(|X| \leq x) = P(-x < X < x) = 0,9$$

En este caso, al querer buscar la cantidad de días de demora donde se encuentra el 90% de los deudores **al rededor de la media**, tendremos un 10% de deudores restantes que estarán en ambos extremos de la distribución, repartidos de igual forma, o sea, 5% del lado izquierdo de la cola y 5% del lado derecho. De esta forma, lo conveniente es es calcular la densidad acumulada hasta el 95% de la distribución **menos** la densidad acumulada hasta el 5% de la distribución, por lo que la ayuda brindada por el ejercicio que nos será de utilidad es $F(1,645) = 0,95$

Entonces, trabajando con una distribución normal estándar:

$$P(-1,645 < Z < 1,645) = P(Z < 1,645) - P(Z < -1,645) = 0,9$$

Al tener los valores de Z, utilizando la formula para estandarizar:

$$-1,645 = \frac{X - 45}{16} \implies X = \boxed{18,68}$$

$$1,645 = \frac{X - 45}{16} \implies X = \boxed{71,32}$$

Por lo que el 90% de los deudores se encuentra demorado entre 19 y 71 días.

4. **Suponer ahora que se desconoce la distribución de probabilidad de esta variable y encontrar la cota de probabilidad para que la demora de un crédito difiera de la media en más de 20 días.**

Cuando se desconoce la distribución de probabilidad de una variable, se puede utilizar la Desigualdad de Tchebycheff. En nuestro caso, que se pide la probabilidad para que la demora de un crédito difiera de la media en **mas** de 20 días, se debe utilizar:

$$P\{|X - \mu| > d\} < \frac{V(x)}{d^2}$$

Para nuestro caso:

$$P\{|X - 45| > 20\} < \frac{16^2}{20^2} = \boxed{0,64}$$

Ejercicio 2 (Lara y Rosario)

1. **Probabilidad de que ningún cliente haya sido atendido durante una hora**

Recordando la forma funcional de una variable aleatoria $x \sim Po(\lambda)$:

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!}$$

En nuestro caso particular, con $\lambda = 4$ y $x = 0$:

$$P(x) = \frac{e^{-4}4^0}{0!} = e^{-4} = \boxed{0,01832}$$

2. **Probabilidad de que más de 3 clientes hayan sido atendidos en un lapso de media hora**

Cómo en promedio se atienden 4 clientes en una hora, podríamos decir que en media hora se atienden, en promedio, 2 clientes ($\lambda = 2$). De esta forma, debemos encontrar:

$$P(X > 3) = P(X \geq 4) = 1 - P(X \leq 3)$$

Donde la probabiliada acumulada hasta 3 inclusive vendrá dada por:

$$P(X \leq 3) = F(3) = \sum_{x=0}^{3} \frac{e^{-2}2^x}{x!} = \frac{e^{-2}2^0}{0!} + \frac{e^{-2}2^1}{1!} + \frac{e^{-2}2^2}{2!} + \frac{e^{-2}2^3}{3!}$$

$$P(X \leq 3) = F(3) = 0,1353 + 0,2706 + 0,2706 + 0,1804 = 0,8571$$

Por lo que:

$$P(X > 3) = P(X \geq 4) = 1 - P(X \leq 3) = 1 - 0,8571 = \boxed{0,1428}$$

3. **Comenzando la observación en cualquier momento de tiempo, calcular la probabilidad de que transcurran como máximo 20 minutos hasta que llegue el primer cliente.**

Se trabaja con una $X \sim Exp(\lambda = \frac{4}{16} = 0,666...)$.

A diferencia de los incisos anteriores, donde se buscaba un número de ocurrencias (clientes atendidos), siendo esta variable contable (discreta), al trabajar con una variable como lo es el tiempo, estamos frente a una variable aleatoria continua, por lo que no podremos trabajar con la distribución Poisson.

En este caso se busca:

$$F(20) = P(T \leq 20) = 1 - P(T > 20)$$

En este caso ya se dispone de $P(T > 20) = 0,2636$ por lo que:

$$F(20) = P(T \leq 20) = 1 - P(T > 20) = 1 - 0,2636 = \boxed{0,7364}$$

Otra forma es utilizar la forma funcional de la función de distribución de una variable aleatoria $X \sim Exp(\lambda)$

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}$$

Para nuestro caso particular:

$$F(20) = P(T \leq 20) = 1 - e^{-\frac{4}{60}20} = 1 - 0,2636 = \boxed{0,7364}$$

Ejercicio 3 (Diana y Jonás)

1. **Obtener la función de distribución**

La función de distribución estará definida como:

$$F(x) = \begin{cases} 0 & x \leq 2 \\ \int_2^x \frac{u}{16} du & 2 < x < 6 \\ 1 & x \geq 6 \end{cases}$$

Por lo que deberemos buscar:

$$F(x) = \int_2^x \frac{u}{16} du$$

$$F(x) = \frac{1}{32} u^2 |_2^x = \frac{1}{32}(x^2 - 2^2) = \frac{x^2}{32} - \frac{1}{8}$$

Quedando así:

$$\boxed{F(x) = \begin{cases} 0 & x \leq 2 \\ \frac{x^2}{32} - \frac{1}{8} & 2 < x < 6 \\ 1 & x \geq 6 \end{cases}}$$

2. **Calcular la probabilidad de que x sea mayor a 3**

Debemos buscar:

$$P(X > 3)$$

Que es equivalente a escribir:

$$1 - P(X \leq 3) = 1 - F(3)$$

Por lo que:

$$1 - F(3) = 1 - \frac{3^2}{32} + \frac{1}{8} = 1 - \frac{9}{32} + \frac{1}{8} = 1 - \frac{5}{32} = 1 - 0,15625 = \boxed{0,84375}$$

3. **Obtener el percentil 90**

Se busca obtener el x_0 tal que:

$$F(x_0) = 0,9$$

Entonces, reemplazando por la función de distribución:

$$\frac{x_0^2}{32} - \frac{1}{8} = 0,9$$

$$x_0^2 = (0,9 + 0,125).32$$

$$x_0^2 = (1,025).32$$

$$x_0^2 = 32,8$$

De esta forma:

$$x_0 = 5,7271$$

$$x_0 = -5,7271$$

Pero como $x \in (2,6)$, $F(-5,7271) = 0$, por lo que el percentil 90 será:

$$\boxed{x_0 = 5,7271}$$