Hoja de Problemas de la asignatura

1. Tipo Examen

1.1. Probabilidad condicionada, total y Bayes

1. Se está investigando una nueva prueba de detección de cáncer. Si se realiza la prueba a una persona sana, la probabilidad de que la prueba experimental dé positivo es 0.05 y de que dé negativo, 0.95. Sabemos que 1 persona de cada 100.000 padece la enfermedad. Pero cuando se trata de una persona enferma, la prueba es infalible. Si una persona seleccionada al azar presenta una reacción positiva, ¿cuál es la probabilidad de que padezca el cáncer?

Solución

Definimos los siguientes eventos:

- C: Persona padece cáncer.
- S: Persona no padece cáncer.
- T^+: Prueba experimental da positivo.
- T^-: Prueba experimental da negativo.

Las probabilidades conocidas son:

- $P(C) = \frac{1}{100000} = 0.00001$: Probabilidad de que una persona padezca cáncer.
- $P(S) = 1 - P(C) = 1 - 0.00001 = 0.99999$: Probabilidad de que una persona no padezca cáncer.
- $P(T^+ \mid C) = 1$: Probabilidad de que la prueba dé positivo si la persona tiene cáncer.
- $P(T^+ \mid S) = 0.05$: Probabilidad de que la prueba dé positivo si la persona no tiene cáncer (falsos positivos).
- $P(T^- \mid S) = 0.95$: Probabilidad de que la prueba dé negativo si la persona no tiene cáncer (verdaderos negativos).

Queremos calcular la probabilidad de que una persona tenga cáncer dado que la prueba dio positivo, es decir, $P(C \mid T^+)$. Para esto, utilizamos el **Teorema de Bayes**:

$$P(C \mid T^+) = \frac{P(T^+ \mid C)P(C)}{P(T^+)}$$

Donde $P(T^+)$ se calcula usando la regla de la probabilidad total:

$$P(T^+) = P(T^+ \mid C)P(C) + P(T^+ \mid S)P(S)$$

Cálculo de $P(T^+)$:

$$P(T^+) = (1) \cdot (0.00001) + (0.05) \cdot (0.99999)$$

$$P(T^+) = 0.00001 + 0.05 \cdot 0.99999 \approx 0.00001 + 0.0499995 = 0.0500095$$

Cálculo de $P(C \mid T^+)$:

$$P(C \mid T^+) = \frac{(1) \cdot (0.00001)}{0.0500095} \approx \frac{0.00001}{0.0500095} \approx 0.0001998$$

Conclusión:

La probabilidad de que una persona padezca cáncer dado que la prueba experimental dio positivo es aproximadamente 0.0001998, o alrededor del 0.01998 %.

- . -

2. En un casino el 20 % de los dados están equilibrados, pero hay 4 partidas restantes (también al 20 %) de dados falsos. La probabilidad de obtener 6 es, respectivamente $\frac{2}{6}, \frac{3}{6}, \frac{4}{6}$, y $\frac{5}{6}$. Se lanza un dado una vez y sale 6. ¿Cuál es la probabilidad de que sea equilibrado? Y si repetimos el lanzamiento 10 veces y salen 7 seises. ¿Cuál es la probabilidad de que esté equilibrado?

3. En el lanzamiento de una nueva red social, se estima que el producto alcanzaría las siguientes cuotas de mercado con la probabilidad indicada en la tabla:

Cuota de mercado	Probabilidad
10 %	0.3
15 %	0.45
20 %	0.25

Para contrastar la hipótesis, se realiza una encuesta entre 5 usuarios *beta* resultando que 2 de ellos finalmente adquieren el producto. Posteriormente se realiza una segunda encuesta a 10 potenciales usuarios, resultando que tres de ellos adquieren nuestro producto. Calcula las probabilidades de que, tras los sucesivos resultados, podamos asignar a cada una de las cuotas de mercado.

4. A y B compiten en el siguiente juego: A lanza dos dados y gana cuando la suma es 4. Si no sale 4, lanza B y gana si la suma es 6. En caso de no obtener 6, se repite la jugada de A y B sucesivamente,hasta que el primero obtenga el valor ganador. ¿Cuál es la probabilidad de ganar cada jugador?

Solución

1. Probabilidad de ganar en un turno

Jugador A: A gana si la suma de los dos dados es igual a 4. Hay 3 combinaciones que suman 4: $(1, 3)$, $(2, 2)$, $(3, 1)$, de las 36 combinaciones posibles. Entonces, la probabilidad de que A gane es:

$$P(A \text{ gana}) = \frac{3}{36} = \frac{1}{12}$$

Jugador B: B gana si la suma de los dos dados es igual a 6. Hay 5 combinaciones que suman 6: $(1, 5)$, $(2, 4)$, $(3, 3)$, $(4, 2)$, $(5, 1)$. Entonces, la probabilidad de que B gane es:

$$P(B \text{ gana}) = \frac{5}{36}$$

2. Probabilidad de que ninguno gane en un turno

La probabilidad de que A no gane es:

$$P(A \text{ no gana}) = 1 - P(A \text{ gana}) = \frac{11}{12}$$

La probabilidad de que B no gane es:

$$P(B \text{ no gana}) = 1 - P(B \text{ gana}) = \frac{31}{36}$$

Entonces, la probabilidad de que ni A ni B ganen en un turno es:

$$P(\text{ninguno gana}) = \frac{11}{12} \cdot \frac{31}{36} = \frac{341}{432}$$

3. Probabilidad de que A gane

La probabilidad de que A gane en la primera ronda es $P(A \text{ gana}) = \frac{1}{12}$.
Si ninguno gana en la primera ronda, la probabilidad de que A gane en la segunda ronda es:

$$P(A \text{ gana en la segunda ronda}) = P(\text{ninguno gana}) \cdot P(A \text{ gana}) = \frac{341}{432} \cdot \frac{1}{12}$$

En general, la probabilidad de que A gane en la n-ésima ronda es:

$$P(A \text{ gana en la } n\text{-ésima ronda}) = \left(\frac{341}{432}\right)^{n-1} \cdot \frac{1}{12}$$

La probabilidad total de que A gane es la suma de esta serie geométrica:

$$P(A \text{ gana}) = \sum_{n=1}^{\infty} \left(\frac{341}{432}\right)^{n-1} \cdot \frac{1}{12}$$

Esto es una serie geométrica de la forma $\sum_{n=0}^{\infty} ar^n$, cuyo valor es $\frac{a}{1-r}$, donde $a = \frac{1}{12}$ y $r = \frac{341}{432}$:

$$P(A \text{ gana}) = \frac{\frac{1}{12}}{1 - \frac{341}{432}} = \frac{\frac{1}{12}}{\frac{91}{432}} = \frac{432}{12 \cdot 91} = \frac{36}{91}$$

4. Probabilidad de que B gane

La probabilidad de que B gane es:

$$P(B \text{ gana}) = 1 - P(A \text{ gana}) = 1 - \frac{36}{91} = \frac{55}{91}$$

Conclusión:

- La probabilidad de que el jugador A gane es $\frac{36}{91}$.
- La probabilidad de que el jugador B gane es $\frac{55}{91}$.

Alternativamente

La probabilidad de que gane A, una vez que se han jugado los dos primeros lanzamientos sin haber ganado nadie es la misma a partir de la siguiente jugada, es decir:

$$P = \frac{3}{36} + \frac{33}{36} \cdot \frac{31}{36} \cdot P$$

Despejando de esta ecuación, se obtiene $P = \frac{36}{91}$

- . -

5. En el CGPJ hay 3 componentes progresistas y 4 conservadores. Se decide crear una comisión secreta de tres miembros por sorteo. Se ha descubierto que un miembro es progresista. ¿Cuál es la probabilidad de que sea conservador el siguiente miembro que se considere de la comisión?

Solución

Se puede plantear en términos de urnas. Inicialmente tenemos una urna (el CGPJ) con 7 bolas, 3 **P**úrpura y 4 **C**eleste. Y decidimos construir una urna nueva (que sería la comisión) con tres bolas al azar de la anterior urna. Ahora consideramos los sucesos :

A_1, pasar tres bolas púrpura. A_2, pasar 2 bolas púrpura y una celeste.
A_3, pasar 1 bola púrpura y 2 celeste. A_4, pasar ninguna púrpura y 3 celeste.

Observa que:

$$P(A_1) = \frac{\binom{3}{3}}{\binom{7}{3}} = \frac{1}{35} \quad P(A_2) = \frac{\binom{3}{2}\binom{4}{1}}{\binom{7}{3}} = \frac{12}{35} \quad P(A_3) = \frac{\binom{3}{1}\binom{4}{2}}{\binom{7}{3}} = \frac{18}{35} \quad P(A_4) = \frac{\binom{4}{3}}{\binom{7}{3}} = \frac{4}{35}$$

Ahora planteamos el suceso B = obtener una bola Púrpura en la comisión (*segunda urna*):

$$P(B|A_1) = 1,\ P(B|A_2) = \frac{2}{3},\ P(B|A_3) = \frac{1}{3},\ P(B|A_4) = 0$$

Por tanto:

$$P(B) = \frac{1}{35} \cdot 1 + \frac{12}{35} \cdot \frac{2}{3} + \frac{18}{35} \cdot \frac{1}{3} + 0 = \frac{3}{7}$$

Utilizando ahora la fórmula de Bayes:

$$P(A_1|B) = \frac{\frac{1}{35} \cdot 1}{\frac{3}{7}} = \frac{1}{15} \quad P(A_2|B) = \frac{\frac{12}{35} \cdot \frac{2}{3}}{\frac{3}{7}} = \frac{8}{15} \quad P(A_3|B) = \frac{\frac{18}{35} \cdot \frac{1}{3}}{\frac{3}{7}} = \frac{6}{15} \quad P(A_4|B) = \frac{0}{\frac{3}{7}} = 0$$

Conclusión:

La probabilidad de que al sacar un nuevo miembro de la urna 2, éste sea conservador, sería

$$\frac{1}{2} \cdot \frac{8}{15} + 1 \cdot \frac{6}{15} = \frac{2}{3}$$

.

- . -

6. Una urna contiene tres bolas blancas y cuatro azules. Tres bolas son transferidas a una segunda urna. Una bola es seleccionada de la segunda urna y resulta ser blanca. ¿Cuál es la probabilidad de extraer una bola azul entre las dos restantes de la segunda urna?

Solución

Obsérvese que el problema es equivalente al ejercicio anterior.

- . -

7. En las próximas elecciones a Rector/a de la UA, el 30 % de las personas son seguidoras de la candidatura A, el 50 % de la candidatura B y el resto de la candidatura C. Una estimación muy fiable a concluido que van a ir a votar el 65 % de quienes siguen a A, el 82 % de quienes votan a B y el 50 % de los partidarios de C. El día de la votación, si seleccionamos al azar una persona en el bar (o sea que no ha votado), ¿Cuál es la probabilidad que sea seguidor de la candidatura A? ¿Cuál es la probabilidad de que una persona al azar sí haya votado?

Solución

Dadas las siguientes definiciones de los eventos, elegida una persona:

- A: La persona es seguidora de la candidatura A.
- B: La persona es seguidora de la candidatura B.
- C: La persona es seguidora de la candidatura C.
- V: La persona ha votado.
- $\bar{V}$: La persona no ha votado.

Probabilidades dadas

- Probabilidades de ser seguidor de una candidatura:

$$P(A) = 0.3, \quad P(B) = 0.5, \quad P(C) = 0.2$$

- Probabilidades de votar dado que son seguidores de una candidatura:

$$P(V \mid A) = 0.65, \quad P(V \mid B) = 0.82, \quad P(V \mid C) = 0.5$$

Paso 1: Probabilidad de que una persona sí haya votado $P(V)$

La probabilidad total de votar se puede calcular utilizando la fórmula de la probabilidad total:

$$P(V) = P(V \mid A) \cdot P(A) + P(V \mid B) \cdot P(B) + P(V \mid C) \cdot P(C)$$

Sustituyendo los valores:

$$P(V) = (0.65) \cdot (0.3) + (0.82) \cdot (0.5) + (0.5) \cdot (0.2)$$

$$P(V) = 0.195 + 0.41 + 0.1 = 0.705$$

Entonces, la probabilidad de que una persona al azar haya votado es $P(V) = 0.705$.

Paso 2: Probabilidad de que una persona seleccionada en el bar sea seguidora de la candidatura A

Queremos calcular $P(A \mid \bar{V})$, es decir, la probabilidad de que una persona no haya votado y sea seguidora de A. Usamos el teorema de Bayes:

$$P(A \mid \bar{V}) = \frac{P(\bar{V} \mid A) \cdot P(A)}{P(\bar{V})}$$

Donde:

$$P(\bar{V} \mid A) = 1 - P(V \mid A) = 1 - 0.65 = 0.35$$

$$P(\bar{V}) = 1 - P(V) = 1 - 0.705 = 0.295$$

Sustituyendo:

$$P(A \mid \bar{V}) = \frac{0.35 \cdot 0.3}{0.295}$$

$$P(A \mid \bar{V}) = \frac{0.105}{0.295} \approx 0.356$$

Conclusiones

- La probabilidad de que una persona seleccionada al azar en el bar sea seguidora de la candidatura A es aproximadamente 0.356.
- La probabilidad de que una persona al azar haya votado es 0.705.

- . -

8. Un sensor de determinado tipo de polen lo detecta en un 90 % de las ocasiones que efectivamente hay presencia atmosférica de dicho polen. Si no hay presencia, detecta la ausencia con una probabilidad del 80 %. La AEMET ha informado que hoy hay una probabilidad del 20 % de tener presencia de dicho polen, calcula la probabilidad de que:

   a) Que efectivamente haya presencia al haber dado positivo el sensor.

   b) Que hay presencia al haber dado negativo el sensor.

   c) De que haya polen y el sensor dé positivo.

   d) Que no haya polen y el sensor dé negativo.

Solución

- $P(P)$: Probabilidad de que haya presencia de polen $= 0.20$.
- $P(\overline{P})$: Probabilidad de que no haya presencia de polen $= 0.80$.

- $P(\text{Pos} \mid P)$: Probabilidad de que el sensor dé positivo si hay polen $= 0.90$.
- $P(\text{-} \mid \overline{P})$: Probabilidad de que el sensor dé negativo si no hay polen $= 0.80$.
- $P(\text{Pos} \mid \overline{P})$: Probabilidad de que el sensor dé positivo si no hay polen $= 1 - P(\text{-} \mid \overline{P}) = 0.20$.

a) Probabilidad de que haya presencia al haber dado positivo el sensor

Utilizamos el teorema de Bayes para calcular $P(P \mid +)$:

$$P(P \mid +) = \frac{P(+ \mid P) \cdot P(P)}{P(+)}$$

donde:

$$P(+) = P(+ \mid P) \cdot P(P) + P(+ \mid \overline{P}) \cdot P(\overline{P}) = 0.34$$

Por tanto:

$$P(P \mid +) = \frac{P(+ \mid P) \cdot P(P)}{P(+)} = \frac{0.18}{0.34} \approx 0.5294117647058825$$

b) Probabilidad de que haya presencia al haber dado negativo el sensor

De nuevo, aplicamos el teorema de Bayes para calcular $P(P \mid \text{-})$:

$$P(P \mid \text{-}) = \frac{P(\text{-} \mid P) \cdot P(P)}{P(\text{-})}$$

donde:

$$P(\text{-}) = 1 - P(+) = 0.66$$

Por tanto :

$$P(P \mid \text{-}) = \frac{P(\text{-} \mid P) \cdot P(P)}{P(\text{-})} = \frac{0.02}{0.66} = 0.\widehat{03}$$

c) Probabilidad de que haya polen y el sensor dé positivo

Esta probabilidad conjunta se calcula como:

$$P(P \cap +) = P(+ \mid P) \cdot P(P) = 0.2 \cdot 0.9 = 0.18$$

d) Probabilidad de que no haya polen y el sensor dé negativo

La probabilidad conjunta se calcula como:

$$P(\overline{P} \cap \text{-}) = P(\text{-} \mid \overline{P}) \cdot P(\overline{P}) 0.8 \cdot 0.8 = 0.64$$

- . -

1.2. Variable Aleatoria

9. La calificación promedio del examen de Inferencia ha sido de 6.25 puntos con una desviación estándar de 1. Sospecho que el examen ha sido difícil, así que ajustaré las notas para que el promedio sea 7 y la desviación estándard 8. Su la nota la transformo como $Y = aX + b$, ¿Qué deberían valer a y b?

Solución

Dada una transformación lineal de la forma:

$$Y = aX + b$$

donde X tiene un promedio μ_X de 6.25 y una desviación estándar σ_X de 1. Se desea que Y tenga un promedio μ_Y de 7 y una desviación estándar σ_Y de 8.

Paso 1: Relación entre promedios

Sabemos que para el promedio de Y:

$$E[Y] = a \cdot E[X] + b$$

Sustituyendo los valores dados:

$$7 = a \cdot 6.25 + b$$

Paso 2: Relación entre desviaciones estándar

Para la desviación estándar de Y, se cumple:

$$Var[Y] = a^2 \cdot Var[X]$$

Sustituyendo los valores dados:

$$8 = |a| \cdot 1 \implies a = 8$$

Paso 3: Determinar b

Ahora que conocemos $a = 8$, sustituimos en la ecuación del promedio:

$$7 = 8 \cdot 6.25 + b$$

$$7 = 50 + b$$

$$b = 7 - 50 = -43$$

Finalmente tenemos que

$$a = 8, \quad b = -43$$

Por lo tanto, la transformación será:

$$Y = 8X - 43$$

- . -

10. La siguiente distribución de probabilidad corresponde a la variable aleatoria X: *Número de veces que un modelo generativo responde tarde a los prompts de un usuario en hora punta*

x	0	1	2	3	4
p(x)	$0.3 + k$	$2k$	0.2	$0.1 + 5k^2$	0.05

a) Halla el valor de k

b) Halla la probabilidd de que el número de veces que se produzca demora sea superior a 3.

c) Halla la probabilidad que que no se produzcan demoras.

d) Halla el número más probable de tardanzas. ¿Concincide con el valor esperado?

Solución

a) Observa que $\sum_{x=0}^{4} p(x) = 1$, por tanto, $k = \frac{1}{10}$, $k = \frac{-7}{10}$. La segunda solución se debe descartar porque daría probabilidades negativas ($P(X = 1)$, p.ej.). Por lo que nos quedamos con que $k = \frac{1}{10} = 0.1$.

b) Independientemente de lo que valga k, $P(X = 4) = 0.05$.

c) $P(X = 0) = 0.4$

d) El número más probable de tardanzas es $X = 0$ pues $P(X = 0)$ es mayor que todos los demás. En cambio, la esperanza viene dada por:

$$\sum_{x=0}^{4} x \cdot p(x) = 0 \cdot 0.4 + 1 \cdot 0.2 + 2 \cdot 0.2 + 3 \cdot 0.15 + 4 \cdot 0.05 = 1.25$$

- . -

11. En el proceso de fabricación de un robot un algoritmo de visión artificial detecta defectos. Cuando es así, la fabricación se detiene. Considera la variable aleatoria X: *Cantidad de veces que se detiene la máquina por día*. Su función de probabilidad viene dada por:

$$f(x) = \begin{cases} \frac{16}{31}(0.5)^x & x = 0, 1, 2, 3, 4 \\ 0 & \text{en otro caso} \end{cases}$$

a) ¿Cuál es la probabilidad de que se detenga la máquina un determinado día?

b) Las detenciones en días consecutivos son independientes. Halla la probabilidad de que el máximo de detenciones comparando dos días consecutivos sea exactamente 2.

c) Halla la función de probabilidad del número máximo de detenciones por día comparando dos días consecutivos.

Solución

a) $P(X \geq 1) = 1 - P(X = 0) = 1 - f(0) = \frac{15}{31}$

b) Que el primer día sean 0 ó 1 detenciones y el segundo 2 detenciones ó el primero 2 y el segundo sean 0, 1 ó 2, es decir:

$$\left(\frac{16}{31}+\frac{16}{2\cdot 31}\right)\cdot\left(\frac{16}{4\cdot 31}\right)+\frac{16}{4\cdot 31}\cdot\left(\frac{16}{31}+\frac{16}{2\cdot 31}+\frac{16}{4\cdot 31}\right)=\frac{208}{961}$$

c)

X	p(x)
0	$\frac{256}{961}$
1	$\frac{320}{961}$
2	$\frac{208}{961}$
3	$\frac{116}{961}$
4	$\frac{61}{961}$

- . -

12. Supongamos que la probabilidad de elegir una palabra de la frase *EL PROFESOR DE INFERENCIA TIENE POCO PELO EN LA CABEZA* es proporcional al número de letras que forman dicha palabra. Si X es el número de letras de la palabra seccionada, define:

    a) El espacio muestral
    b) La función de probabilidad asociada
    c) La variable aleatoria. Indica su tipo.
    d) $E[X]$ y $Var[X]$
    e) Si $X = 2$, ¿Cuál es la probabilidad de que la palabra seleccionada haya sido *EN*?

Solución

Dada la frase: *EL PROFESOR DE INFERENCIA TIENE POCO PELO EN LA CABEZA*, identificamos las palabras y sus respectivas longitudes:

- *EL* (2 letras)
- *PROFESOR* (8 letras)
- *DE* (2 letras)
- *INFERENCIA* (10 letras)
- *TIENE* (5 letras)
- *POCO* (4 letras)
- *PELO* (4 letras)
- *EN* (2 letras)
- *LA* (2 letras)
- *CABEZA* (6 letras)

Paso 1: Definición de la función de probabilidad

La probabilidad de elegir una palabra es proporcional al número de letras que tiene. Si X es el número de letras de una palabra, entonces la función de probabilidad es:

$$P(X=x)=\frac{x}{\text{Total de letras en la frase}}$$

Calculamos el total de letras en la frase:

$$2 + 8 + 2 + 10 + 5 + 4 + 4 + 2 + 2 + 6 = 45\text{letras}$$

Por lo tanto, la función de probabilidad es:

$$P(X = x) = \frac{x}{45}$$

Paso 2: Cálculo de la esperanza $\mathbb{E}(X)$
La esperanza $\mathbb{E}(X)$ se calcula como:

$$\mathbb{E}(X) = \sum_{i=1}^{10} P(X = x_i) \cdot x_i$$

Sustituyendo los valores de x_i:

$$\mathbb{E}(X) = \frac{2}{45} \cdot 2 + \frac{8}{45} \cdot 8 + \frac{2}{45} \cdot 2 + \frac{10}{45} \cdot 10 + \frac{5}{45} \cdot 5 + \frac{4}{45} \cdot 4 + \frac{4}{45} \cdot 4 + \frac{2}{45} \cdot 2 + \frac{2}{45} \cdot 2 + \frac{6}{45} \cdot 6$$

$$\mathbb{E}(X) = \frac{2 \cdot 2 + 8 \cdot 8 + 2 \cdot 2 + 10 \cdot 10 + 5 \cdot 5 + 4 \cdot 4 + 4 \cdot 4 + 2 \cdot 2 + 2 \cdot 2 + 6 \cdot 6}{45}$$

$$\mathbb{E}(X) = \frac{4 + 64 + 4 + 100 + 25 + 16 + 16 + 4 + 4 + 36}{45}$$

$$\mathbb{E}(X) = \frac{273}{45} \approx 6.07$$

Paso 3: Cálculo de la varianza $\text{Var}(X)$
La varianza se calcula como:

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Primero, calculemos $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \sum_{i=1}^{10} P(X = x_i) \cdot x_i^2$$

$$\mathbb{E}(X^2) = \frac{2}{45} \cdot 2^2 + \frac{8}{45} \cdot 8^2 + \frac{2}{45} \cdot 2^2 + \frac{10}{45} \cdot 10^2 + \frac{5}{45} \cdot 5^2 + \frac{4}{45} \cdot 4^2 + \frac{4}{45} \cdot 4^2 + \frac{2}{45} \cdot 2^2 + \frac{2}{45} \cdot 2^2 + \frac{6}{45} \cdot 6^2$$

$$\mathbb{E}(X^2) = \frac{2 \cdot 4 + 8 \cdot 64 + 2 \cdot 4 + 10 \cdot 100 + 5 \cdot 25 + 4 \cdot 16 + 4 \cdot 16 + 2 \cdot 4 + 2 \cdot 4 + 6 \cdot 36}{45}$$

$$\mathbb{E}(X^2) = \frac{8 + 512 + 8 + 1000 + 125 + 64 + 64 + 8 + 8 + 216}{45}$$

$$\mathbb{E}(X^2) = \frac{2013}{45} \approx 44.73$$

Finalmente, calculemos la varianza:

$$\text{Var}(X) = 44.73 - (6.07)^2$$

$$\text{Var}(X) = 44.73 - 36.84$$

$$\text{Var}(X) = 7.89$$

Por tanto:

- La esperanza de X es $\mathbb{E}(X) \approx 6.07$. - La varianza de X es $\text{Var}(X) \approx 7.89$.

Finalmente, para responder a la última pregunta, podríamos utilizar el Teorema de Bayes pero acaba siendo más ágil el plantear que si $X = 2$, solo han podido ser cuatro palabras, además con la misma probabilidad:

- *EL*
- *DE*
- *EN*
- *LA*

Por tanto, calcular la probabilidad condicional viene dada por

$$P(\text{EN} \mid X = 2) = \frac{1}{\text{Número de palabras de 2 letras}} = \frac{1}{4} = 0.25$$

- . -

13. La duración de determinados sensores, T medido en años, sigue una función de densidad de probabilidad:

$$f(t) = \begin{cases} 1 & si \quad 0 \leq t < 0.5 \\ 0.5e^{-(t-0.5)} & t > 0.5 \\ 0 & si \quad \text{otros casos} \end{cases}$$

   a) Halla la función de distribución acumulada F

   b) Los sensores se venden como *regulares* si duran menos de tres meses, *buenos* si duran entre tres meses y tres años y *muy buenos* si duran más de tres años. ¿Qué porcentajes de sensores tendremos en el proceso de fabricación.

   c) Los empaquetamos en cajas de 20. Si se encuentran uno o más artículos regulares en una caja, la fábrica proporciona una caja nueva gratuitamente. ¿Cuál es la probabilidad de obtener una de regalo? Y si somos tres amigos los que hemos comprado una caja cada uno, ¿Cuál es la probabilidad de que entre los tres tengamos, al menos, una caja de regalo?

Solución

a) La función de distribución acumulada $F(t)$ se define como:

$$F(t) = \int_{-\infty}^{t} f(x)\,dx$$

Por tanto, para calcular $F(t)$ con $t \leq 0$
Si $t \leq 0$, la función de densidad $f(x) = 0$, por lo que:

$$F(t) = 0 \quad \text{para } t \leq 0$$

Posteriormente, para calcular $F(t)$ para $0 < t < 0.5$:
En este intervalo, $f(x) = 1$. Por lo tanto, la integral de $f(x)$ desde 0 hasta t es:

$$F(t) = \int_{0}^{t} 1\,dx = t \quad \text{para } 0 < t < 0.5$$

A continuación, calculamos $F(t)$ para $t \geq 0.5$:
En este intervalo, $f(x) = 0.5 \cdot \exp(-(x - 0.5))$. La función de distribución acumulada se calcula como la integral desde 0 hasta 0.5 (donde la densidad es constante) más la integral desde 0.5 hasta t:

$$F(t) = \int_{0}^{0.5} 1\,dx + \int_{0.5}^{t} 0.5\exp(-(x - 0.5))\,dx$$

$$F(t) = 0.5 + 0.5\int_{0.5}^{t} \exp(-(x - 0.5))\,dx$$

Calculamos la segunda integral. Sea $u = x - 0.5$, entonces $du = dx$, y los límites de integración cambian de $x = 0.5$ a $x = t$, lo cual implica que u varía de 0 a $t - 0.5$. La integral se convierte en:

$$\int_{0.5}^{t} \exp(-(x - 0.5))\,dx = \int_{0}^{t-0.5} \exp(-u)\,du$$

La integral de $\exp(-u)$ es $-\exp(-u)$, por lo que:

$$\int_{0}^{t-0.5} \exp(-u)\,du = -\exp(-u)\Big|_{0}^{t-0.5} = -\exp(-(t - 0.5)) + 1$$

Sustituyendo en la expresión de $F(t)$:

$$F(t) = 0.5 + 0.5\,[1 - \exp(-(t - 0.5))]$$

$$F(t) = 0.5 + 0.5 - 0.5\exp(-(t - 0.5)) = 1 - 0.5\exp(-(t - 0.5))$$

Ya solo queda definir $F(t)$
Juntando todos los intervalos, la función de distribución acumulada $F(t)$ es:

$$F(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ t & \text{si } 0 < t < 0.5 \\ 1 - 0.5\exp(-(t - 0.5)) & \text{si } t \geq 0.5 \end{cases}$$

b) Para determinar el porcentaje de sensores *regulares*, tendremos que determinar la probabilidad de encontrar uno de ellos. Esta probabilidad definirá la proporción buscada. Así, solo tendremos que calcular $P(T \leq 0.25) = F(0.25) = 0.25$. (tres meses es la cuarta parte de un año, o sea, $T = 0.25$)

Para los *buenos*, deberemos calcular $P(0.25 < T \leq 3) = F(3) - F(0.25)$.

Y finalmente los *muy buenos* se calculan como $P(T > 3) = 1 - F(3)$.

c) La probabilidad de obtener una caja de regalo, viene dada por la probabilidad $P(X \geq 1)$, donde X es la v.a. que mide número de sensores regulares en una caja de 20 sensores. Como $X \sim B(20.0.25)$, solo tendremos que calcular dicha probabilidad: $P(X \geq 1) = 1 - P(X = 0) = 1 - (0.75)^{20} \approx 0.996828788061066$

Por último, sin consideramos la variable W que mide el número de cajas *NO regulares*, tendremos que $W \sim B(3, 0.003171211938933993)$, y nos estarían preguntando $1 - P(W = 0) = 1 - (0.003171211938933993)^3 = 0.999999968108437$

- . -

14. La función de distribución de la demanda de energía eléctrica (en MW) a la hora punta de las 12:00 en un determinado DataCenter viene dada por :

$$F(x) = \begin{cases} 0 & si \quad x < 0 \\ bx^2 & si \quad 0 \leq x < 1 \\ b(2(x-1)+1) & si \quad 1 \leq x < 3 \\ 1 - b(x-4)^2 & si \quad 3 \leq x < 4 \\ 1 & si \quad x \geq 4 \end{cases}$$

a) Halla la función de densidad de la demanda de energía eléctrica.

b) Halla la demanda superada sólo por el 20 % de los días a la hora punta.

Solución

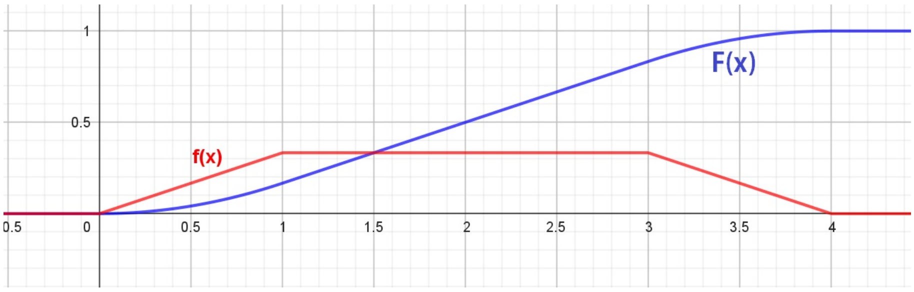
a) Sabemos que la f.d.p $f(x)$ es la derivada de la función de distribución $F(x)$, es decir:

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ 2bx & si \quad 0 \leq x < 1 \\ 2b & si \quad 1 \leq x < 3 \\ -2b(x-4) & si \quad 3 \leq x < 4 \\ 0 & si \quad x \geq 4 \end{cases}$$

Solo deberíamos determinar el valor de b para que estuviera bien definida la función de densidad:

1) $\int_0^1 2bxdx = b$
2) $\int_1^3 2bdx = 2bx]_1^3 = 4b$
3) $\int_3^4 -2b(x-4)dx = 2bx]_1^3 = 1 - b(x-4)^2]_3^4 = b$

Por tanto, $b + 4b + b = 1$, por lo que $b = \frac{1}{6}$

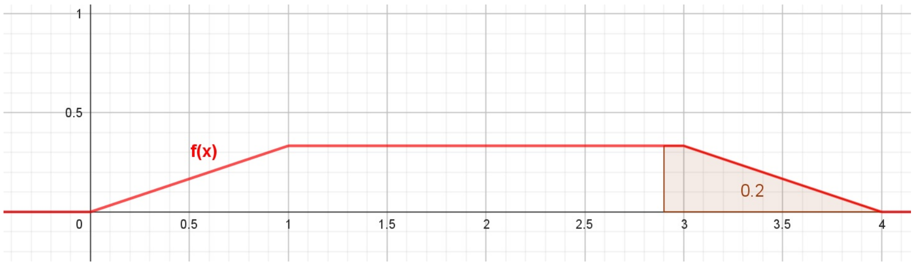


b) Para la segunda cuestión planteada, nos están preguntando por el valor de x para que se cumpla que $P(X \geq x) = 0.20$

Observa que el área entre 3 y 4, definida por un triángulo, es $\frac{1 \cdot \frac{2}{6}}{2} = \frac{1}{6}$

Por lo que se trata de calcular x tal que:

$$(3 - x) \cdot \frac{2}{6} + \frac{1}{6} = 0.2 \rightarrow x = \frac{29}{10} = 2.9$$



- . -

1.3. Binomial

15. Supón que la proporción de mujeres entre los usuarios de *NetPlix* es del 51 % y 49 % para el resto. Calcula la probabilidad de que entre 6 usuarios se tenga:

   a) Al menos una mujer
   b) Como máximo, una mujer.
   c) Exactamente 3 mujeres.
   d) Un número par de mujeres.

Solución

Sea X la variable aleatoria que representa el número de mujeres entre 6 usuarios de *NetPlix*. Entonces, X sigue una distribución binomial con parámetros:

- $n = 6$ (número de usuarios)
- $p = 0.51$ (probabilidad de que un usuario sea mujer)

La probabilidad de que haya exactamente k mujeres en n intentos se calcula como:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

donde $\binom{n}{k}$ es el coeficiente binomial que se calcula como:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

a) Al menos una mujer

Para calcular la probabilidad de tener al menos una mujer, calculamos:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{6}{0}(0.51)^0(0.49)^6 \approx 1 - 0.013841287201 \approx 0.986158712799$$

b) Como máximo, una mujer

Para calcular la probabilidad de tener como máximo una mujer, sumamos las probabilidades de tener 0 o 1 mujer:

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

Ya hemos calculado $P(X = 0)$. Ahora calculamos $P(X = 1)$:

$$P(X = 1) = \binom{6}{1}(0.51)^1(0.49)^5 = 6 \cdot (0.51) \cdot (0.49)^5 \approx 0.086437426194$$

$$P(X \leq 1) = P(X = 0) + P(X = 1) \approx 0.100278713395$$

c) Exactamente 3 mujeres

Para calcular la probabilidad de tener exactamente 3 mujeres:

$$P(X = 3) = \binom{6}{3}(0.51)^3(0.49)^3$$

Calculamos $\binom{6}{3} = 20$:

$$P(X = 3) = 20 \cdot (0.51)^3 \cdot (0.49)^3 \approx 0.31212514998$$

d) Un número par de mujeres

Un número par de mujeres puede ser 0, 2, 4 o 6 (Podríamos descartar el 0) . Por lo tanto, calculamos:

$$P(X \text{ es par}) = P(X = 0) + P(X = 2) + P(X = 4) + P(X = 6)$$

Ya hemos calculado $P(X = 0)$. Ahora calculamos los demás:

1. Para $P(X = 2)$:

$$P(X = 2) = \binom{6}{2}(0.51)^2(0.49)^4 \approx 0.224913711015$$

2. Para $P(X = 4)$:

$$P(X = 4) = \binom{6}{4}(0.51)^4(0.49)^2 =\approx 0.243648714015$$

3. Para $P(X = 6)$:

$$P(X = 6) = \binom{6}{6}(0.51)^6(0.49)^0 =\approx 0.017596287801$$

Ahora sumamos:

$$P(X \text{ es par}) \approx P(X = 0) + P(X = 2) + P(X = 4) + P(X = 6) \approx 0.500000000032$$

- . -

16. La tasa de baja en un año entre los clientes de TicTac es del 10 %. De entre el grupo de 10 amigos que somos, cuál es la probabilidad en el próximo año:

    a) Nos demos todos de baja.

    b) La mitad se den de baja.

    c) Al menos tres se den de baja.

    d) Solo se den de baja tres.

    (Considera que somos amigos muy independientes y que lo que hagamos no depende en absoluto de lo que hagan los demás).

Solución

Dado que somos amigos independientes y lo que hagamos no depende de los demás, podemos modelar la situación usando una **distribución binomial**.

Definiciones Sean:

- $n = 10$ (número de amigos).
- $p = 0.10$ (probabilidad de que un amigo se dé de baja en un año).

La variable aleatoria X representa el número de amigos que se dan de baja, y sigue una distribución binomial $B(n, p)$.

a) Probabilidad de que todos se den de baja

$$P(X = 10) = \binom{10}{10}(0.10)^{10}(1 - 0.10)^0 = 10^{10} \times 0.10^{10} \approx 1 \times 10^{-10}$$

b) Probabilidad de que la mitad se den de baja

$$P(X = 5) = \binom{10}{5}(0.10)^5(1 - 0.10)^5 \approx 0.0014880348$$

c) Probabilidad de que al menos tres se den de baja Queremos calcular:

$$P(X \geq 3) = 1 - P(X < 3) = 1 - [P(X = 0) + P(X = 1) + P(X = 2)] \approx 0.07019082639999996$$

d) Probabilidad de que solo tres se den de baja

$$P(X = 3) = \binom{10}{3}(0.10)^3(1 - 0.10)^7 \approx 0.057395628$$

- . -

17. El 60 % de usuarios de InstaKilo quiere eliminar el contenido violento en los contenidos de la red social. La dueña de la Red, no estando segura, decide elegir una muestra de 20 personas y preguntar directamente. Cuál es la probabilidad de que:

    a) Entre 10 y 13 sean partidarios de la eliminación .

b) Que no haya nadie partidario de dicha eliminación.

c) Que haya mayoría de no filtrar el contenido violento.

Solución

Modelamos esta cuestión usando una **distribución binomial**.

Definiciones Sean:

- $n = 20$ (tamaño de la muestra).
- $p = 0.60$ (probabilidad de que una persona sea partidaria de la eliminación del contenido violento).

La variable aleatoria X representa el número de personas en la muestra que son partidarias de la eliminación, y sigue una distribución binomial $B(n, p)$.

a) Probabilidad de que entre 10 y 13 personas sean partidarias Queremos calcular $P(10 \leq X \leq 13)$:

$$P(10 \leq X \leq 13) = P(X = 10)+P(X = 11)+P(X = 12)+P(X = 13) = F(13)-F(9) \approx 0.6224680819149585$$

b) Probabilidad de que no haya nadie partidario de la eliminación Queremos calcular $P(X = 0)$:

$$P(X = 0) = \binom{20}{0}(0.60)^0(1 - 0.60)^{20} \approx 1.099511627776008 \times 10^{-8}$$

c) Probabilidad de que haya mayoría de no filtrar el contenido violento Queremos calcular $P(X < 10)$, ya que una mayoría se define como más de la mitad en contra de la eliminación:

$$P(X < 10) = \sum_{k=0}^{9} P(X = k) = F(9) \approx 0.1275212461472175$$

- . -

18. La ruleta francesa consta de 37 casillas numeradas del 1 al 36, numeradas alternativamente como roja y negra más el cero que suele ser verde (o blanco). Si lanzo 10 veces, ¿Cuál es el número esperado de rojos que puede salir? ¿Cuál es la probabilidad de que no salga 0? ¿Cuál es la probabilidad de que salgan más rojos que del resto? ¿Cuál es la probabilidad de que el primer rojo salga en la décima tirada? Han salido 7 verdes, ¿Cuál es la probabilidad de que haya más rojos que negros?

Solución

Podemos modelar las tiradas como eventos de Bernoulli con ciertas probabilidades, y trabajar con distribuciones binomiales o geométricas dependiendo del apartado.

1. Número esperado de rojos en 10 lanzamientos

Sea $p_R = \frac{18}{37}$ la probabilidad de que salga un número rojo. En 10 lanzamientos, el número esperado de rojos se calcula a partir de una binomial $X \sim B(10, \frac{18}{37})$ como:

$$E(X) = n \cdot p_R = 10 \cdot \frac{18}{37} \approx 4.86$$

2. Probabilidad de que no salga un 0 en 10 lanzamientos

La probabilidad de que en un lanzamiento no salga un 0 es $P = 1 - \frac{1}{37}$. La probabilidad de que no salga un 0 en 10 lanzamientos consecutivos es:

$$P(\text{no sale 0 en 10 lanzamientos}) = \left(1 - \frac{1}{37}\right)^{10} \approx 0.7603398748776844$$

3. Probabilidad de que salgan más rojos que del resto

Queremos calcular $P(\text{más rojos que del resto})$. Esto se puede aproximar como:

$$P(X > 5) \approx 0.3441783291991802$$

Ten en cuenta que se considera **el resto** como el conjunto de bolas de cualquier color que no sea rojo.

4. Probabilidad de que el primer rojo salga en la décima tirada

Sea $p_R = \frac{18}{37}$ la probabilidad de éxito. La probabilidad de que el primer rojo salga en la décima tirada se modela con una distribución geométrica:

$$P(\text{primer rojo en la décima tirada}) = (1 - p_R)^9 \cdot p_R \approx 0.001207918611834908$$

5. Probabilidad de que haya más rojos que negros si han salido 7 verdes Solo hay una casilla verde, es imposible que salgan 7 verdes

- . -

19. La probabilidad de que la bolsa suba un entero en un día es $\frac{1}{3}$ y de que baje un entero $\frac{2}{3}$. Las alzas y bajas son independientes de un día a otro. Considera X la variable que da la posición de la bolsa al cabo de 4 días, sabiendo que el día 0, $X = 100$. Se pide:

    a) función de probabilidad de X

    b) Esperanza de X

    c) Probabilidad de que haya subido subido la bolsa al cabo de 4 días.

Solución

a) Función de probabilidad de X

Sea Y el número de días en los que la bolsa sube. Como hay 4 días y las alzas son independientes, Y sigue una distribución binomial:

$$Y \sim \text{Binomial}(4, \frac{1}{3}).$$

La posición de la bolsa después de 4 días se puede expresar como:

$$X = 100 + Y - (4 - Y) = 100 + 2Y - 4 = 96 + 2Y.$$

Dado que Y sigue una distribución binomial, podemos calcular las probabilidades para cada valor de Y:

$$P(Y=k)=\binom{4}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{4-k},\quad k=0,1,2,3,4.$$

Los valores posibles de X y sus probabilidades correspondientes son:

$$\begin{aligned}
P(X=96)=P(Y=0)&=\binom{4}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^4=\left(\frac{2}{3}\right)^4=\frac{16}{81},\\
P(X=98)=P(Y=1)&=\binom{4}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3=4\cdot\frac{1}{3}\cdot\left(\frac{2}{3}\right)^3=\frac{32}{81},\\
P(X=100)=P(Y=2)&=\binom{4}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2=6\cdot\left(\frac{1}{3}\right)^2\cdot\left(\frac{2}{3}\right)^2=\frac{24}{81},\\
P(X=102)=P(Y=3)&=\binom{4}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1=4\cdot\left(\frac{1}{3}\right)^3\cdot\frac{2}{3}=\frac{8}{81},\\
P(X=104)=P(Y=4)&=\binom{4}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^0=\left(\frac{1}{3}\right)^4=\frac{1}{81}.
\end{aligned}$$

b) Esperanza de X Dado que $X=96+2Y$, podemos calcular la esperanza de X como:

$$\mathbb{E}(X)=96+2\mathbb{E}(Y).$$

Dado que Y sigue una distribución binomial $\text{Bin}(4,\frac{1}{3})$, su esperanza es:

$$\mathbb{E}(Y)=4\cdot\frac{1}{3}=\frac{4}{3}.$$

Por lo tanto, la esperanza de X es:

$$\mathbb{E}(X)=96+2\cdot\frac{4}{3}=96+\frac{8}{3}=98.\overline{6}.$$

c) Probabilidad de que haya subido la bolsa al cabo de 4 días La bolsa habrá subido si $X>100$, es decir, si $Y\geq 3$. La probabilidad de que $Y\geq 3$ es:

$$P(Y\geq 3)=P(Y=3)+P(Y=4).$$

Calculamos:

$$P(Y=3)=\frac{8}{81},\quad P(Y=4)=\frac{1}{81}.$$

Por lo tanto:

$$P(Y\geq 3)=\frac{8}{81}+\frac{1}{81}=\frac{9}{81}=\frac{1}{9}.$$

- . -

20. En una sala con diez servidores se ha detectado un ruido excesivo que denota mal funcionamiento. Supongamos que 4 de estos 10 servidores tienen ese mal funcionamiento mientras que los otros 6 entran dentro de lo aceptable.Si se examinan al azar 5 de estos 10 servidores y se define la variable aleatoria X: *número de servidores defectuosos en la muestra*, indicar

    a) La distribución de la variable X

b) La probabilidad de que no todos sean defectuosos.

c) La probabilidad de que, a lo sumo, 4 sean defectuosos.

Solución

a) La distribución de la variable X es hipergeométrica. Ten en cuenta que como estamos sacando la muestra sin reemplazamiento, no podríamos enmarcarla dentro de la binomial. Los parámetros serían: N = Tamaño de la población = 10, K = Número de éxitos = 4, n = Tamaño de la muestra = 5

b) Podríamos calcular $1 - P(X = 0)$.

$$1 - P(X = 0) = 1 - \frac{\binom{4}{0}\binom{6}{5}}{\binom{10}{5}} \approx 1 - 0.0238095238 \approx 0.976190$$

c) $P(X \leq 4) = 1 - P(X = 5)$. Si consideramos que no se puede dar que $X = 5$, tendríamos que la probabilidad solicitada es $P(X \leq 4) = 1$

- . -

21. La compañía de vuelos JetProb sabe que el 4 % de los pasajeros que reserva un viaje Alicante-Mallorca no se presenta al vuelo. Por eso, en un avión con 70 asientos decide vender 72 billetes. ¿Cuál es la probabilidad de que puedan viajar todos los pasajeros que se presentan a embarcar? Y ¿Cuál la probabilidad de que tenga que afrontar quejas? Si esta política se sigue durante una semana, ¿Cuál es la probabilidad de que, al menos un día, deba afrontar quejas?

Solución

Veamos como modelar el problema

Denotamos por X el número de pasajeros que se presentan. $X \sim B(72, 0.96)$.

Vamos a resolver el ejercicio aproximando la distribución binomial mediante una distribución normal, dado que n está por encima de las tablas usuales y que para valores relativamente *grandes*, la normal aproxima satisfactoriamente dicho valor. Usamos:

$$\mu = n \cdot p = 72 \cdot 0.96 = 69.12, \quad \sigma^2 = n \cdot p \cdot (1 - p) = 72 \cdot 0.96 \cdot 0.04 = 2.7648.$$

Por lo tanto, $\sigma \approx 1.662$. La aproximación normal es $X \sim N(69.12, 1.662^2)$.

1. Probabilidad de que puedan viajar todos los pasajeros Queremos calcular $P(X \leq 70)$. Estandarizando:

$$Z = \frac{70 - 69.12}{1.662} \approx 0.53.$$

De tablas de la normal:

$$P(Z \leq 0.53) \approx 0.701.$$

Por lo tanto, la probabilidad de que puedan viajar todos los pasajeros es $P(X \leq 70) \approx 0.701$.

2. Probabilidad de que tenga que afrontar quejas

$$P(X > 70) = 1 - P(X \leq 70) = 1 - 0.701 = 0.299.$$

3. Probabilidad de que haya quejas al menos un día en una semana

(Se supone que hay un viaje por día, los siete días de la semana)

$$P(\text{al menos un día}) = 1 - (1 - 0.299)^7 \approx 1 - (0.701)^7 \approx 1 - 0.082 = 0.918.$$

- . -

22. En un proyecto de Learning Analytics el 30 % del alumnado está en trabajos *de calidad* y el 50 % en trabajos regulares. Sabemos que el 70 % del alumnado está en, al menos, uno de los proyectos. Si consideramos 5 alumnos, calcula:

 a) Probabilidad de que al menos dos del alumnado seleccionado estén exactamente en uno de esos proyectos.

 b) Probabilidad de que a lo sumo, tres estén en ambos proyectos proyectos.

 c) Probabilidad de que todos los seleccionados estén en alguno de estos proyectos.

Solución

Definimos los siguientes eventos:

- Q: Un alumno está en un trabajo de calidad.
- R: Un alumno está en un trabajo regular.

Según la información proporcionada:

$$P(Q) = 0.30, \quad P(R) = 0.50, \quad P(Q \cup R) = 0.70.$$

La probabilidad de que un alumno esté en ambos proyectos se denota como $P(Q \cap R)$. De acuerdo con la fórmula de la probabilidad de la unión de dos eventos:

$$P(Q \cup R) = P(Q) + P(R) - P(Q \cap R).$$

Despejando $P(Q \cap R)$:

$$P(Q \cap R) = P(Q) + P(R) - P(Q \cup R).$$

Sustituyendo los valores proporcionados:

$$P(Q \cap R) = 0.30 + 0.50 - 0.70 = 0.10.$$

Ahora podremos calcular la probabilidad de que un alumno esté exactamente en un proyecto:

$$P(\text{exactamente uno}) = P(\text{Al menos uno}) - P(\text{exactamente uno}) = 0.70 - (0.10) = 0.60.$$

a) Probabilidad de que al menos dos del alumnado seleccionado estén exactamente en uno de los proyectos

Si denotamos como X el número de alumnos que están en exactamente un proyecto. $X \sim B(5, 0.6)$. La probabilidad de que al menos 2 alumnos estén en exactamente un proyecto es:

$$P(X \geq 2) = 1 - P(X \leq 1).$$

Donde $P(X = k)$ se calcula con la fórmula de la distribución binomial:

$$P(X = k) = \binom{5}{k}(0.60)^k(0.40)^{5-k}.$$

Calculamos:

$$P(X = 0) = \binom{5}{0}(0.60)^0(0.40)^5 = 0.01024.$$

$$P(X = 1) = \binom{5}{1}(0.60)^1(0.40)^4 = 0.0768.$$

Entonces:

$$P(X \geq 2) = 1 - 0.01024 - 0.0768 = 0.91296.$$

Por lo tanto, la probabilidad de que al menos dos alumnos estén en exactamente uno de los proyectos es aproximadamente 0.913.

b) Probabilidad de que a lo sumo tres estén en ambos proyectos

Denotamos el número de alumnos en ambos proyectos como Y. $Y \sim B(5, 0.1)$. La probabilidad de que como máximo 3 alumnos estén en ambos proyectos es:

$$P(Y \leq 3) = \approx 0.99954$$

c) Probabilidad de que todos los seleccionados estén en alguno de estos proyectos

La probabilidad de que un alumno esté en al menos uno de los proyectos es $P(Q \cup R) = 0.70$. Si consideramos la variable W que mide número de alumnos (de los cinco) que están en al menos, un proyecto, $W \sim B(5, 0.7)$. Por lo tanto, la probabilidad de que todos los alumnos estén en al menos un proyecto es:

$$P(\text{todos en alguno}) = (P(Q \cup R))^5 = 0.70^5 = 0.16807.$$

- . -

1.4. Poisson

23. En el sistema de venta de entradas online de una cantante de éxito, el número de compradores en una hora sigue una Poisson de media 1.500. Calcula la probabilidad de que:

 a) Durante un día se vendan más de 30.000 entradas.

 b) Las dos primeras horas se vendan menos de 2.000 entradas.

Solución

Dado que el número de compradores por hora sigue una distribución de Poisson con media $\lambda = 1500$, utilizaremos las propiedades de la distribución de Poisson para resolver las dos partes del problema.

En una distribución de Poisson $X \sim \text{Poisson}(\lambda)$, la probabilidad de que ocurra un determinado número de eventos en un intervalo de tiempo viene dada por:

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

a) Probabilidad de que durante un día se vendan más de 30.000 entradas

Dado que hay 24 horas en un día y la media de compradores por hora es 1.500, la media diaria será:

$$\lambda_{\text{día}} = 1500 \times 24 = 36.000$$

Queremos calcular $P(X > 30.000)$, donde X es el número de entradas vendidas durante el día y $X \sim \text{Poisson}(36.000)$.

Si utilizamos Geogebra para hacer el cálculo, obtenemos: $P(X > 30.000) \approx 1$

Pero por otro lado, dado que λ es grande, podemos aproximar la distribución de Poisson con una distribución normal:

$$X \sim \mathcal{N}(\mu, \sigma^2), \quad \mu = 36.000, \quad \sigma^2 = 36.000$$

Por lo tanto, la desviación estándar es $\sigma = \sqrt{36,000} \approx 189'74$. La probabilidad que buscamos es:

$$P(X > 30.000) \approx P\left(Z > \frac{30.000 - 36.000}{189'74}\right) = P(Z > -31'62)$$

Dado que el valor de $-31'62$ está muy lejos del promedio en una distribución normal estándar, podemos asumir que esta probabilidad es aproximadamente 1. Por lo tanto, $P(X > 30'000)$ es prácticamente 1 (o 100 %).

b) Probabilidad de vender menos de 2,000 entradas en las dos primeras horas

Para las dos primeras horas, la media será:

$$\lambda_{2\ \text{horas}} = 1500 \times 2 = 3000$$

Queremos calcular $P(Y < 2,000)$, donde Y es el número de entradas vendidas en las dos primeras horas y $Y \sim \text{Poisson}(3000)$. De nuevo, dado que λ es grande, podemos aproximar Y con una distribución normal:

$$Y \sim \mathcal{N}(3000, 3000)$$

La desviación estándar es $\sigma = \sqrt{3000} \approx 54.77$. La probabilidad que buscamos es:

$$P(Y < 2,000) \approx P\left(Z < \frac{2,000 - 3,000}{54.77}\right) = P\left(Z < -18.27\right)$$

Dado que el valor de -18.27 está muy lejos a la izquierda de la media de una distribución normal estándar, esta probabilidad es prácticamente 0. Por lo tanto, $P(Y < 2,000)$ es casi 0.

Resumen de resultados:

- $P(\text{Más de 30.000 entradas en un día}) \approx 1$
- $P(\text{Menos de 2,000 entradas en dos horas}) \approx 0$

- . -

24. Se estima que el número de usuarios que entran en ChatGbm en una hora sigue una distribución Poisson con $\lambda = 20$. El coste del acceso es de 2 €.

a) ¿Cuál es la probabilidad de que la empresa ingrese más de 15 €la primera hora.

b) ¿Cuál es la probabilidad de que la empresa ingrese más de 500 €el primer día.

c) Cuando ha ingresado 100 €¿Cuánto tiempo transcurrirá hasta que se produzca un aumento en el ingreso realizado?

25. Demuestra que si $X \sim Poi(\lambda)$ entonces se cumple que $P(X = k+1) = \frac{\lambda}{k+1} P(X = k)$

Solución

Sea X una variable aleatoria con distribución de Poisson con parámetro λ. Queremos demostrar que:

$$P(X = k+1) = \frac{\lambda}{k+1} \cdot P(X = k).$$

Demostración

Dado que $X \sim \text{Poisson}(\lambda)$, la función de probabilidad de X está dada por:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}.$$

Queremos demostrar que:

$$P(X = k+1) = \frac{\lambda}{k+1} \cdot P(X = k).$$

Paso 1: Expresión de $P(X = k+1)$

Sabemos que:

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

Como queríamos demostrar.

- . -

26. El número de automóviles que acuden a una ITV cada hora sigue una distribución Poisson de parámetro 10. Sabemos que el 30 % de vehículos los vehículos que acuden no la pasan (Se les detectan anomalías graves).

a) ¿Cuál es la probabilidad de que ninguno pase la ITV en una hora?

b) ¿Cuál es la probabilidad de que pasen la ITV más vehículos que los que no la pasan?

c) ¿Cuál es la función de probabilidad del número de vehículos que no pasan la ITV en una hora?

Solución

Partimos de que

- Los automóviles que acuden a la ITV siguen una distribución de Poisson con parámetro $\lambda = 10$.

- Sabemos que el 30 % de los vehículos que acuden no pasan la ITV

a) Probabilidad de que ninguno pase la ITV en una hora

Primero, veamos qué significa que *ninguno pase la ITV*. Como el 30 % de los vehículos no pasa, entonces el 70 % sí pasa la ITV.

Denotemos:

- N: Número de vehículos que acuden a la ITV en una hora.
- $P(\text{pasan}) = 0.70$.
- $P(\text{no pasan}) = 0.30$.

La probabilidad de que ninguno pase la ITV significa que todos los vehículos que llegan en una hora no pasan. En términos de probabilidades, esto es:

$$P(\text{ninguno pasa}) = P(\text{todos los vehículos no pasan})$$

Dado que N sigue una distribución de Poisson, podemos calcular la probabilidad de que el número de vehículos que pasan la ITV sea cero. El número de vehículos que pasan se puede modelar también como una distribución de Poisson, con un parámetro de λ ajustado al porcentaje de vehículos que pasan:

$$\lambda_{\text{pasan}} = 0.70 \times 10 = 7$$

Por lo tanto, la probabilidad de que ninguno pase es:

$$P(X = 0) = \frac{e^{-7} \cdot 7^0}{0!} = e^{-7}$$

Calculamos:

$$P(\text{ninguno pasa}) = e^{-7} \approx 0.00091$$

b) Probabilidad de que pasen la ITV más vehículos que los que no la pasan

Sea X el número de vehículos que pasan la ITV y Y el número de vehículos que no la pasan. Sabemos que $X \sim \text{Poisson}(7)$ y $Y \sim \text{Poisson}(3)$ (ya que el 30 % de 10 es 3).

Para resolver el problema deberíamos abordarlo con EXCEL. Calcularíamos la probabilidad de que lleguen n vehículos y depués consideraríamos la variable aleatoria Y_n de vehículos que sí pasan, que es una binomial, y calcularíamos la probabilidad de que $Y_n > \frac{n}{2}$, para sucesivos valores de $n = 1, 2, 3, 4, 5, ...$ hasta obtener que la probabilidad acumulada supera determinado número de decimales prefijado. Dicho valor sería: $P = 0.927730735510322$ **c) Función de probabilidad del número de vehículos que no pasan la ITV**

El número de vehículos que no pasan la ITV sigue una distribución de Poisson con parámetro $\lambda_{\text{no pasan}} = 0.30 \times 10 = 3$.

La función de probabilidad de una distribución de Poisson con parámetro λ está dada por:

$$P(Y = k) = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

En este caso:

$$P(Y = k) = \frac{e^{-3} \cdot 3^k}{k!}, \quad k = 0, 1, 2, \ldots$$

Resumen de resultados

- a) $P(\text{ninguno pasa}) \approx 0.00091$.
- b) $P(\text{pasan más vehículos que los que no pasan})$ requiere un enfoque numérico.
- c) $P(Y = k) = \frac{e^{-3} \cdot 3^k}{k!}$.

- . -

1.5. Uniforme

27. En una v.a. uniforme se sabe que su valor esperado es 15 y que $P(13 \leq x \leq 13.5) = 0.55$, halla su varianza y $P(X \leq 15)$

Solución

Dada una variable aleatoria X con distribución uniforme en el intervalo $[a, b]$, se nos proporciona la siguiente información:

- Valor esperado: $\mathbb{E}[X] = 15$.
- $P(13 < X < 13.5) = 0.55$.

Paso 1: Encontrar los límites a y b

El valor esperado para una distribución uniforme en $[a, b]$ es:

$$\mathbb{E}[X] = \frac{a + b}{2} = 15 \implies a + b = 30 \quad (1)$$

Por otro lado, la probabilidad de que X se encuentre en el intervalo $(13, 13.5)$ se calcula como:

$$P(13 < X < 13.5) = \frac{13.5 - 13}{b - a}$$

Dado que se nos proporciona $P(13 < X < 13.5) = 0.55$, tenemos:

$$0.55 = \frac{0.5}{b - a} \implies b - a = \frac{0.5}{0.55} \approx 0.909$$

Ahora tenemos dos ecuaciones:

$$\begin{aligned} a + b &= 30 \quad (1) \\ b - a &= 0.909 \quad (2) \end{aligned}$$

Sumando las dos ecuaciones:

$$2b = 30 + 0.909 \implies b = \frac{30.909}{2} \approx 15.455$$

Sustituyendo b en la ecuación (1):

$$a = 30 - 15.455 \approx 14.545$$

Paso 2: Calcular la varianza

La varianza de una distribución uniforme en $[a, b]$ está dada por:

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Sustituyendo los valores de a y b:

$$\text{Var}(X) = \frac{(15.455 - 14.545)^2}{12} = \frac{0.909^2}{12} \approx \frac{0.826}{12} \approx 0.0688$$

Paso 3: Calcular $P(X < 15)$

Dado que X es uniforme en $[a, b]$, la probabilidad de que X sea menor que 15 es:

$$P(X < 15) = \frac{15-a}{b-a} = \frac{15 - 14.545}{0.909} \approx \frac{0.455}{0.909} = 0.5$$

Resumen de resultados

- Varianza: $\text{Var}(X) \approx 0.0688$.
- $P(X < 15) = 0.5$.

- . -

28. Un repartidor de *GloGlo* debe entregar un paquete a las 10 de la mañana. A consecuencia del tráfico, el tiempo que tarda en recorrer el trayecto oscila entre 35 y 45 minutos. ¿A qué hora debe partir para entregar el paquete puntualmente con probabilidad del 0.85?

Solución

Dado que el tiempo que tarda en recorrer el trayecto oscila entre 35 y 45 minutos, podemos modelar este tiempo con una distribución uniforme X en el intervalo $[35, 45]$. Queremos que el repartidor llegue puntualmente con una probabilidad del 0.85.

Paso 1: Plantear la distribución uniforme

La variable aleatoria X tiene una distribución uniforme en el intervalo $[35, 45]$. La función de densidad para una distribución uniforme está dada por:

$$f_X(x) = \begin{cases} \frac{1}{45-35} = \frac{1}{10}, & \text{si } 35 \le x \le 45, \\ 0, & \text{en otro caso.} \end{cases}$$

Paso 2: Encontrar el cuantil que corresponde al 0.85

Queremos que el repartidor llegue con una probabilidad del 0.85, es decir, encontrar el valor x tal que:

$$P(X \leq x) = 0.85$$

La función de distribución acumulativa de X es:

$$F_X(x) = \frac{x - 35}{10}, \quad \text{para } 35 \leq x \leq 45.$$

Por lo tanto, para hallar x:

$$\frac{x - 35}{10} = 0.85 \implies x - 35 = 10 \cdot 0.85 \implies x = 35 + 8.5 = 43.5.$$

Paso 3: Calcular la hora de partida

El repartidor debe llegar a las 10:00 (es decir, 600 minutos después de medianoche). Si queremos que llegue puntualmente con una probabilidad del 0.85, entonces el tiempo de partida debe ser al menos 43.5 minutos antes de las 10:00.

$$\text{Hora de partida} = 10:00 - 43.5 \text{ minutos}.$$

Restando 43.5 minutos a las 10:00, obtenemos:

$$10:00 - 43.5 \text{ minutos} = 9:16:30.$$

Respuesta final

El repartidor debe partir **a las 9:16:30** para entregar el paquete puntualmente con una probabilidad del 0.85.

- . -

1.6. Exponencial

29. La duración del altavoz del asistente personal de *MiAlexia* y de *PearPod* se modela con una v.a. exponencial. Si el 10 % del mercado lo tiene *MiAlexia* con una duración media de 20.000 horas y el resto del sector lo tiene *PearPod* con una media de 50.000 horas, halla:

    a) La probabilidad que sea de *MiAlexia* sabiendo que su altavoz ha superado las 35.000 horas.

    b) Proporción de altavoces que fallan antes de las 60.000 horas.

    c) Probabilidad de que un altavoz que ya ha superado las 20.000 horas, supere finalmente las 40.000 horas.

Solución

Problema La duración del altavoz del asistente personal de MiAlexia y de PearPod se modela con una variable aleatoria exponencial. Se sabe que:

- El 10 % del mercado lo tiene MiAlexia, con una duración media de 20.000 horas.
- El resto del mercado lo tiene PearPod, con una duración media de 50.000 horas.

Se desea hallar:

a) La probabilidad de que el altavoz sea de MiAlexia sabiendo que ha superado las 35,000 horas.

b) La proporción de altavoces que fallan antes de las 60.000 horas.

c) La probabilidad de que un altavoz que ya ha superado las 20.000 horas, supere finalmente las 40.000 horas.

Solución

a) Probabilidad de que el altavoz sea de MiAlexia sabiendo que ha superado las 35.000 horas

Definimos las variables relevantes:

- X_M representa la duración del altavoz de MiAlexia, y $X_M \sim \text{Exponencial}(\lambda_M)$.
- X_P representa la duración del altavoz de PearPod, y $X_P \sim \text{Exponencial}(\lambda_P)$.

Dado que la duración media de MiAlexia es de 20.000 horas, tenemos:

$$\lambda_M = \frac{1}{20.000}.$$

Análogamente, dado que la duración media de PearPod es de 50.000 horas, tenemos:

$$\lambda_P = \frac{1}{50.000}.$$

Denotamos los sucesos

M = Haber elegido el asistente *MiAlexia*
$P = Ppod$ = Haber elegido el asistente *PearPod*

Queremos hallar $P(M \mid X > 35.000)$. Usamos el teorema de Bayes:

$$P(M \mid X > 35.000) = \frac{P(X > 35.000 \mid M)P(M)}{P(X > 35.000)}$$

Calculamos cada término:

- $P(X > 35.000 \mid M) = e^{-\lambda_M \cdot 35.000} = e^{-\frac{35.000}{20.000}} = e^{-1.75}$
- $P(X > 35.000 \mid Ppod) = e^{-\lambda_P \cdot 35.000} = e^{-\frac{35.000}{50.000}} = e^{-0.7}$
- $P(X > 35.000) = P(X > 35.000 \mid M)P(M) + P(X > 35.000 \mid Ppod)P(Ppod)$

Sustituyendo:

$$P(X > 35.000) = (e^{-1.75})(0.1) + (e^{-0.7})(0.9)$$

Finalmente, la probabilidad buscada es:

$$P(M \mid X > 35.000) = \frac{e^{-1.75} \cdot 0.1}{P(X > 35.000)}$$

b) Proporción de altavoces que fallan antes de las 60.000 horas

Queremos calcular $P(X < 60.000)$, es decir:

$$P(X < 60.000) = 1 - P(X > 60.000).$$

Donde:

$$P(X > 60.000) = P(X > 60.000 \mid M)P(M) + P(X > 60.000 \mid Ppod)P(Ppod)$$

Calculamos:

$$P(X > 60.000 \mid M) = e^{-\lambda_M \cdot 60.000} = e^{-3}, \quad P(X > 60.000 \mid Ppod) = e^{-\lambda_P \cdot 60.000} = e^{-1.2}.$$

Sustituyendo:

$$P(X > 60.000) = (e^{-3})(0.1) + (e^{-1.2})(0.9).$$

Finalmente:

$$P(X < 60.000) = 1 - P(X > 60.000).$$

c) Probabilidad de que un altavoz que ya ha superado las 20.000 horas, supere finalmente las 40.000 horas

Queremos hallar $P(X > 40.000 \mid X > 20.000)$. Dada la propiedad de falta de memoria de la distribución exponencial:

$$P(X > 40.000 \mid X > 20.000) = P(X > 20.000).$$

Entonces:

$$P(X > 20.000) = P(X > 20.000 \mid M)P(M) + P(X > 20.000 \mid P)P(P),$$

donde:

$$P(X > 20.000 \mid M) = e^{-\lambda_M \cdot 20.000} = e^{-1}, \quad P(X > 20.000 \mid P) = e^{-\lambda_P \cdot 20.000} = e^{-0.4}.$$

Por lo tanto:

$$P(X > 40.000 \mid X > 20.000) = e^{-1}(0.1) + e^{-0.4}(0.9).$$

- . -

30. Sabiendo que $X \sim Exp(\lambda)$, halla la distribución de $Y = CX$.

Solución

Sea X una variable aleatoria que sigue una distribución exponencial con parámetro λ, es decir:

$$X \sim \text{Exponencial}(\lambda),$$

lo que implica que la función de densidad de probabilidad (pdf) de X es:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0, \\ 0 & \text{si } x < 0. \end{cases}$$

Supongamos que multiplicamos X por una constante positiva C, y queremos encontrar la distribución de la variable aleatoria $Y = C \cdot X$.

Transformación de variable: $Y = C \cdot X$

Dado que C es una constante positiva, definimos el cambio de variable:

$$Y = C \cdot X \implies X = \frac{Y}{C}.$$

Además, dado que C es positiva, esta transformación es una función monótona creciente, lo cual simplifica :

$$F_Y(y) = P(CX \le y) = P(X \le \frac{y}{C}) = F_X(\frac{y}{C})$$

Aplicando derivadas, la obtención de la nueva f.d.p. mediante la regla de la cadena:

$$(F_Y(y))' = f_Y(y) = (F_X(\frac{y}{C}))' = f_X\left(\frac{y}{C}\right)\left|\left(\frac{y}{C}\right)'\right| = f_X\left(\frac{y}{C}\right)\frac{1}{C} \quad \text{para } y \ge 0.$$

Por lo tanto, la densidad de Y es:

$$f_Y(y) = \lambda e^{-\lambda \frac{y}{C}} \cdot \frac{1}{C} = \frac{\lambda}{C} e^{-\frac{\lambda}{C} y}, \quad \text{para } y \ge 0.$$

Esta es la forma de la función de densidad de una variable exponencial con parámetro $\frac{\lambda}{C}$. Por lo tanto, Y sigue una distribución exponencial con parámetro $\frac{\lambda}{C}$:

$$Y \sim \text{Exponencial}\left(\frac{\lambda}{C}\right).$$

- . -

1.7. Normal

31. Sea $X \sim N(5, \sigma = 10)$. Calcula

    a) $P(X < 0), P(X > 10), P(X \ge 15)$
    b) $P(-5 < X < 30), P(-5 \le X \le 30), P(-20 < X \le 15)$
    c) El valor de x que cumple que $P(X > x) = 0.05$

Solución

Dado que X es una variable aleatoria con distribución normal $N(5, 10)$, esto implica que X tiene una media de $\mu = 5$ y una desviación estándar de $\sigma = 10$. Para resolver los ejercicios, plantearemos una estandarizar las probabilidades utilizando la distribución normal estándar $Z \sim N(0, 1)$ y las tablas de la distribución normal.

Parte a)

Queremos calcular:

a) $P(X < 0)$
b) $P(X > 10)$

c) $P(X \leq 15)$

La fórmula de estandarización para convertir X a la distribución normal estándar Z es:

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

1. $P(X < 0)$

$$P(X < 0) = P\left(\frac{X - 5}{10} < \frac{0 - 5}{10}\right) = P\left(Z < -0.5\right).$$

De las tablas de la distribución normal, $P(Z < -0.5)$ se puede obtener como:

$$P(Z < -0.5) \approx 0.3085(0.3085375387259869)$$

2. $P(X > 10)$

$$P(X > 10) = P\left(\frac{X - 5}{10} > \frac{10 - 5}{10}\right) = P\left(Z > 0.5\right).$$

Dado que $P(Z > 0.5) = 1 - P(Z \leq 0.5)$, y de las tablas de la normal $P(Z \leq 0.5) \approx 0.6915(0.6914624612740131)$: Aunque por simetría ya podríamos haber planteado que

$$P(Z > 0.5) = P(>< -0.5) \approx 1 - 0.6915 = 0.3085.$$

3. $P(X \leq 15)$

$$P(X \leq 15) = P\left(\frac{X - 5}{10} \leq \frac{15 - 5}{10}\right) = P\left(Z \leq 1\right).$$

De las tablas de la normal, $P(Z \leq 1) \approx 0.8413(0.8413447460685429)$.

Parte b)

Queremos calcular:

a) $P(-5 < X < 30)$

b) $P(-5 \leq X \leq 30)$

c) $P(-20 < X \leq 15)$

1. $P(-5 < X < 30)$

$$P(-5 < X < 30) = P\left(\frac{-5 - 5}{10} < Z < \frac{30 - 5}{10}\right) = P(-1 < Z < 2.5).$$

Calculamos cada término:

$$P(-1 < Z < 2.5) = P(Z < 2.5) - P(Z < -1).$$

De las tablas, sabemos que:

$$P(Z < 2.5) \approx 0.9938, \quad P(Z < -1) \approx 0.1587$$

$$(P(Z < 2.5) \approx 0.9937903346742238, \quad P(Z < -1) \approx 0.1586552539314571)$$

Entonces:

$$P(-1 < Z < 2.5) \approx 0.9938 - 0.1587 \approx 0.8351(0.8351350807427668)$$

2. $P(-5 \leq X \leq 30)$
Dado que $P(-5 \leq X \leq 30)$ incluye los límites y para una distribución continua esto no cambia el resultado:

$$P(-5 \leq X \leq 30) = P(-5 < X < 30) = 0.8351.$$

3. $P(-20 < X \leq 15)$

$$P(-20 < X \leq 15) = P\left(\frac{-20-5}{10} < Z \leq \frac{15-5}{10}\right) = P(-2.5 < Z \leq 1).$$

Calculamos cada término:

$$P(-2.5 < Z \leq 1) = P(Z \leq 1) - P(Z < -2.5).$$

De las tablas, sabemos que:

$$P(Z \leq 1) \approx 0.8413, \quad P(Z < -2.5) \approx 0.0062.$$

Entonces:

$$P(-2.5 < Z \leq 1) = 0.8413 - 0.0062 = 0.8351.$$

Parte c)
Queremos encontrar el valor de x tal que $P(X > x) = 0.05$. Esto equivale a encontrar x tal que:

$$P(X \leq x) = 0.95.$$

Para una distribución normal estándar, buscamos $z_{0.95}$ tal que $P(Z \leq z_{0.95}) = 0.95$. De las tablas, sabemos que:

$$z_{0.95} \approx 1.645(1.6448536269514722)$$

Usamos la fórmula de estandarización para encontrar x:

$$\frac{x-5}{10} = 1.645 \implies x = 5 + 10 \cdot 1.645 = 21.45(21.448536269514723)$$

- . -

32. La distribución de pesos entre el alumnado de la UA sigue una distribución normal con $\mu = 75$ y $\sigma = 7$.

   a) Halla la probabilidad de un alumno elegido al azar pese más de 95 kilos.
   b) Estimar el número de alumnos de entre los 30.000 de la UA , con peso entre 80 y 95 alumnos.
   c) ¿Qué peso no superan el 10 % de los alumnos?
   d) Si 10 alumnos entran en el ascensor del edificio de la Politécnica I, ¿Qué probabilidad hay de que se superen los 800 kilos que soporta el ascensor? ¿Y si son 8 alumnos?

Solución

De nuevo, para plantear los correspondientes cálculos probabilísticos utilizaremos una estandarización de la Normal a efectos de poder utilizar la tabla.

Parte a) Probabilidad de un alumno que pese más de 95 kg

Queremos calcular $P(X > 95)$. La fórmula de estandarización es:

$$Z = \frac{X - \mu}{\sigma} = \frac{95 - 75}{7} = \frac{20}{7} \approx 2.86(2.857142857142857)$$

Buscamos la probabilidad $P(Z > 2.86)$, que se puede calcular usando las tablas de la normal o una calculadora:

$$P(Z > 2.86) \approx 1 - P(Z \leq 2.86) \approx 1 - 0.9979 = 0.0021(0.002137366980086264)$$

Por lo tanto, la probabilidad de que un alumno pese más de 95 kg es aproximadamente 0.0021.

Parte b) Número de alumnos entre 80 y 95 kg

Queremos calcular $P(80 < X < 95)$. Primero, estandarizamos los extremos:

- Para $X = 80$:

$$Z = \frac{80 - 75}{7} = \frac{5}{7} \approx 0.714$$

- Para $X = 95$:

$$Z = \frac{95 - 75}{7} = \frac{20}{7} \approx 2.86$$

Queremos calcular $P(0.714 < Z < 2.86)$, que se obtiene restando las probabilidades:

$$P(0.714 < Z < 2.86) = P(Z < 2.86) - P(Z < 0.714)$$

De las tablas de la normal, sabemos que:

- $P(Z < 2.86) \approx 0.9979$, - $P(Z < 0.714) \approx 0.7621$

Entonces:

$$P(0.714 < Z < 2.86) \approx 0.9979 - 0.7621 = 0.2358(0.2353878950468903)$$

Ahora, multiplicamos esta probabilidad por el número total de alumnos (30,000):

$$\text{Número de alumnos} = 0.2358 \times 30,000 \approx 7074$$

Parte c) Peso que no supera el 10 % de los alumnos

Queremos encontrar el peso x tal que $P(X \leq x) = 0.10$. Usando las tablas de la normal estándar, sabemos que el percentil 10 corresponde aproximadamente a $Z \approx -1.28(1.281551565544601)$

Estandarizamos para encontrar x:

$$\frac{x - 75}{7} = -1.28 \implies x = 75 + 7 \cdot (-1.28) = 75 - 8.96 \approx 66.04(66.02913904118779)$$

Por lo tanto, el peso que no superan el 10 % de los alumnos es aproximadamente 66.04 kg.

Parte d) Probabilidad de exceder los 800 kg en el ascensor

Si 10 alumnos entran al ascensor, la suma de sus pesos sigue una distribución normal debido a la propiedad de aditividad de la distribución normal. Sea S_{10} la suma de los pesos de los 10 alumnos. La media de S_{10} es:

$$\mu_{S_{10}} = 10 \times 75 = 750.$$

Y la desviación estándar de S_{10} es:

$$\sigma_{S_{10}} = \sqrt{10} \times 7 \approx 22.14.$$

Queremos calcular $P(S_{10} > 800)$. Estandarizamos:

$$Z = \frac{800 - 750}{22.14} \approx 2.26.$$

Buscamos la probabilidad $P(Z > 2.26)$:

$$P(Z > 2.26) \approx 1 - P(Z \leq 2.26) \approx 1 - 0.9881 = 0.0119.$$

Por lo tanto, la probabilidad de que 10 alumnos superen los 800 kg es aproximadamente 0.0119(0.01194885533975509)

Si son 8 alumnos

En este caso, la suma de los pesos sigue también una distribución normal. Sea S_8 la suma de los pesos de los 8 alumnos. La media de S_8 es:

$$\mu_{S_8} = 8 \times 75 = 600.$$

Y la desviación estándar de S_8 es:

$$\sigma_{S_8} = \sqrt{8} \times 7 \approx 19.80.$$

Queremos calcular $P(S_8 > 800)$. Estandarizamos:

$$Z = \frac{800 - 600}{19.80} \approx 10.10.$$

Dado que $Z = 10.10$ es un valor extremadamente grande, la probabilidad $P(Z > 10.10)$ es prácticamente 0. Por lo tanto, la probabilidad de que 8 alumnos superen los 800 kg es esencialmente nula.

- . -

33. El trayecto de la Playa de San Juan a la UA suele tardar un promedio de 24 minutos con una desviación de 3.8 minutos. Sabemos que la duración se ajusta a una distribución normal. Calcula:

    a) Probabilidad de que hoy tarde al menos media hora en llegar a la UA desde la Playa.

    b) Si la clase empieza a las 15:00 y salgo a las 14:45, ¿Qué probabilidad hay de llegar tarde?

    c) Si sales de la playa a las 14:35 y ese día hay tarta entre las 14:45 y las 15:00, ¿Qué probabilidad hay de quedarte sin tarta?

    d) Calcula la probabilidad de que 2 de los siguientes tres traslados duren menos de media hora.

Solución

En este ejercicio, en lugar de estandarizar, procedemos directamente al cálculo de probabilidades con excel:

	x	Valor Probabilidad	Formula
a)	30	0.057174065	`=1-DISTR.NORM.N(C8;$D$5;$E$5;VERDADERO)`
b)	15	0,991067904	`=1-DISTR.NORM.N(C9;$D$5;$E$5;VERDADERO)`
c)	25	0,396214441	`=1-DISTR.NORM.N(C10;$D$5;$E$5;VERDADERO)`
d)	30	0,942825935 0,152469637	`=1-DISTR.NORM.N(C11;$D$5;$E$5;VERDADERO)` `=DISTR.BINOM.N(2;3;D11;FALSO)`

- . -

34. En el afán de clasificar LLMs, se ha establecido un coeficiente de inteligencia para dichos modelos. Se ha constatado que los LLMs americanos tienen un coeficiente de media 100 y desviación 10 mientras que los europeos tienen media 105 y desviación 12. Todos tienen un comportamiento *normal*. Elegido un LLM al azar, se ha constatado que su coeficiente es superior a 120. Bajo ese supuesto de *normalidad* cuál es la probabilidad de que el LLM sea europeo? Elegido al azar un LLM americano y otro europeo ¿Cuál es la probabilidad de que el coeficiente de inteligencia del americano sea superior al del europeo?

Solución

Paso 1: Definir las variables

Sean:

- X: Coeficiente de inteligencia de un LLM americano, con $X \sim N(100, 10^2)$.
- Y: Coeficiente de inteligencia de un LLM europeo, con $Y \sim N(105, 12^2)$.

Paso 2: Calcular las probabilidades $P(X > 120)$ y $P(Y > 120)$

Para los LLMs americanos:

$$Z_X = \frac{120 - 100}{10} = 2.$$

La probabilidad $P(X > 120)$ se obtiene de la tabla de la distribución normal estándar:

$$P(X > 120) = 1 - P(Z_X \leq 2) \approx 1 - 0.9772 = 0.0228.$$

Para los LLMs europeos:

$$Z_Y = \frac{120 - 105}{12} = 1.25.$$

La probabilidad $P(Y > 120)$ es:

$$P(Y > 120) = 1 - P(Z_Y \leq 1.25) \approx 1 - 0.8944 = 0.1056.$$

Paso 3: Calcular la probabilidad condicional de que el LLM sea europeo dado que su coeficiente es superior a 120

Utilizamos el Teorema de Bayes. Sea E el evento de que el LLM sea europeo y A el evento de que el coeficiente sea superior a 120. Queremos calcular $P(E \mid A)$:

$$P(E \mid A) = \frac{P(A \mid E)P(E)}{P(A)},$$

donde:

- $P(A \mid E) = P(Y > 120)$.
- $P(E) = P(\overline{E}) = 0.5$.

Calculamos $P(A)$:

$$P(A) = P(Y > 120) \cdot P(E) + P(X > 120) \cdot P(\overline{E}) = 0.1056 \cdot 0.5 + 0.0228 \cdot 0.5 = 0.0642.$$

Finalmente, la probabilidad condicional es:

$$P(E \mid A) = \frac{0.1056 \cdot 0.5}{0.0642} \approx 0.823.$$

Parte 2: Probabilidad de que el coeficiente de inteligencia del LLM americano sea superior al del europeo

Queremos calcular $P(X > Y)$. Definimos una nueva variable $Z = X - Y$, que seguirá una distribución normal:

$$Z \sim N(\mu_Z, \sigma_Z^2),$$

donde:

$$\mu_Z = \mu_X - \mu_Y = 100 - 105 = -5,$$

y

$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 = 10^2 + 12^2 = 100 + 144 = 244.$$

Entonces, $Z \sim N(-5, 244)$. Queremos calcular $P(Z > 0)$:

$$Z = \frac{0 - (-5)}{\sqrt{244}} \approx \frac{5}{15.62} \approx 0.32.$$

Buscamos la probabilidad $P(Z > 0.32)$:

$$P(Z > 0.32) = 1 - P(Z \leq 0.32) \approx 1 - 0.6255 = 0.3745.$$

- . -

35. Una empresa antivirus está desarrollando el antivirus a tres tipos nuevos de virus que han aparecido en el mercado (A,B y C). Acabamos de recibir una muestra de 3,2 y 5 ejecutables respectivamente, pero no los sabemos distinguir. La probabilidad de que el virus A corrompa el servidor de pruebas es $P(|X| < 4)$, donde $X \sim N(3, 5)$. La probabilidad de que B corrompa el servidor es $P(Y \leq 3)$ donde $Y \sim B(5, 0.7)$. Por último la probabilidad de que el virus C lo corrompa es $P(W \leq 5)$ donde $W \sim Poi(4)$. Elegimos un virus al azar y lo testeamos en el servidor y finalmente se corrompe. ¿Cuál es el tipus más probable que hemos podido utilizar?

Solución

Paso 1: Calcular las probabilidades de corrupción para cada virus

Virus A

$$X \sim N(3, 5).$$

Queremos calcular $P(|X| < 4)$:

$$P(-4 < X < 4) = P(X < 4) - P(X < -4).$$

Primero, estandarizamos:

$$Z_1 = \frac{4 - 3}{\sqrt{5}} \quad \text{y} \quad Z_2 = \frac{-4 - 3}{\sqrt{5}}.$$

Calculamos Z_1 y Z_2:

$$Z_1 = \frac{1}{\sqrt{5}} \approx 0.447 \quad \text{y} \quad Z_2 = \frac{-7}{\sqrt{5}} \approx -3.130.$$

Buscamos en la tabla de la normal:

$$P(X < 4) \approx P(Z < 0.447) \approx 0.6736,$$
$$P(X < -4) \approx P(Z < -3.130) \approx 0.0009.$$

Por lo tanto:

$$P(-4 < X < 4) = P(X < 4) - P(X < -4) \approx 0.6736 - 0.0009 = 0.6727.$$

Virus B

$$Y \sim B(5, 0.7).$$

Queremos calcular $P(Y \leq 3)$:

$$P(Y \leq 3) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3).$$

Usamos la fórmula de la binomial:

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Calculamos cada término:

$$\begin{aligned}
P(Y = 0) &= \tbinom{5}{0}(0.7)^0(0.3)^5 \approx 0.00243, \\
P(Y = 1) &= \tbinom{5}{1}(0.7)^1(0.3)^4 \approx 0.02835, \\
P(Y = 2) &= \tbinom{5}{2}(0.7)^2(0.3)^3 \approx 0.1323, \\
P(Y = 3) &= \tbinom{5}{3}(0.7)^3(0.3)^2 \approx 0.30253
\end{aligned}$$

Entonces:

$$P(Y \leq 3) \approx 0.00243 + 0.02835 + 0.1323 + 0.30253 \approx 0.46561$$

Virus C

$$W \sim Poi(4).$$

Queremos calcular $P(W \leq 5)$:

$$P(W \leq 5) = \sum_{k=0}^{5} \frac{e^{-4} 4^k}{k!}.$$

Calculamos cada término:

$$\begin{aligned}
P(W = 0) &= \tfrac{e^{-4}4^0}{0!} \approx 0.0183, \\
P(W = 1) &= \tfrac{e^{-4}4^1}{1!} \approx 0.0733, \\
P(W = 2) &= \tfrac{e^{-4}4^2}{2!} \approx 0.1465, \\
P(W = 3) &= \tfrac{e^{-4}4^3}{3!} \approx 0.1953, \\
P(W = 4) &= \tfrac{e^{-4}4^4}{4!} \approx 0.1953, \\
P(W = 5) &= \tfrac{e^{-4}4^5}{5!} \approx 0.1562
\end{aligned}$$

Entonces:

$$P(W \leq 5) \approx 0.0183 + 0.0733 + 0.1465 + 0.1953 + 0.1953 + 0.1562 \approx 0.5849.$$

Paso 2: Usar el teorema de Bayes para calcular la probabilidad posterior

Dado que elegimos un virus al azar y finalmente se corrompe, denotemos C como el evento de que el servidor se corrompe y A, B, y C como los eventos de que el virus es A, B o C, respectivamente.

Queremos encontrar:

$$P(A|C), \quad P(B|C), \quad P(C|C).$$

Por el teorema de Bayes, tenemos:

$$P(A|C) = \frac{P(C|A)P(A)}{P(C)},$$
$$P(B|C) = \frac{P(C|B)P(B)}{P(C)},$$
$$P(C|C) = \frac{P(C|C)P(C)}{P(C)}.$$

Asumiendo que hay una probabilidad igual de elegir cada virus:

$$P(A) = P(B) = P(C) = \frac{1}{3}.$$

La probabilidad de que el servidor se corrompa dado el virus A, B o C es simplemente $P(C|A) = P(-4 < X < 4)$, $P(C|B) = P(Y \leq 3)$, y $P(C|C) = P(W \leq 5)$ calculadas anteriormente.

Calculamos $P(C)$:

$$P(C) = P(C|A)P(A) + P(C|B)P(B) + P(C|C)P(C) = \frac{1}{3}(P(A|C) + P(B|C) + P(C|C)).$$

Sustituyendo los valores:

$$P(C) = \frac{1}{3}(0.6727 + 0.46561 + 0.5849) \approx \frac{1}{3}(1.72321) \approx 0.5744.$$

Finalmente:

$$P(A|C) = \frac{0.6727 \cdot \frac{1}{3}}{0.5744} \approx 0.3910,$$
$$P(B|C) = \frac{0.46561 \cdot \frac{1}{3}}{0.5744} \approx 0.2705,$$
$$P(C|C) = \frac{0.5849 \cdot \frac{1}{3}}{0.5744} \approx 0.3385.$$

Paso 3: Comparar las probabilidades

- $P(A|C) \approx 0.3910$
- $P(B|C) \approx 0.2705$
- $P(C|C) \approx 0.3385$

Resultado

La probabilidad más alta corresponde al virus A.

- . -

1.8. Para nota

36. Un sensor de temperatura debe tener longitud $500mm$. En el proceso de fabricación, y una vez calibrada la máquina con la longitud L (no necesariamente igual a 500) para que fabrique el sensor , se produce un error obteniendo longitud $L + X$, donde $X \sim N(0, \sigma = 3mm)$. En caso de obtener un sensor de longitud inferior a 500 mm, la pieza se descarta y se pierde. Si es mayor, la pieza se acepta, pero se pule hasta obtener los 500 mm, perdiéndose por tanto el exceso obtenido. Calcula la longitud L a la que debe calibrarse la máquina que los fabrica para minimizar el valor esperado de pérdidas de material.

Solución

Definiciones y Supuestos

- L: longitud calibrada de la máquina. - $X \sim N(0, \sigma^2)$ con $\sigma = 3\,\text{mm}$. - Longitud objetivo: $500\,\text{mm}$.

Cálculo de la Pérdida Esperada

La pérdida de material se desglosa en dos casos:

Caso 1: $L + X < 500$

En este caso, el sensor se descarta. Si consideramos la variable aleatoria $P_1 = L + X$, podremos considerar que como se descarta entera, $E[P_1] = E(L + X]$, es decir:

$$E[P_1] = \int_{-\infty}^{500-L} (L + x) \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} dx$$

Caso 2: $L + X \geq 500$

En este caso, la pérdida viene dada por $P_2 = X + L - 500$, por lo que el valor esperado será:

$$E[P_2] = \int_{500-L}^{\infty} (L + x - 500) \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} dx$$

Por tanto, la función g que mide el valor esperado de las pérdidas en función del calibre L de la máquina viene dada por:

$$g(L) = E[P_1] + E[P_2]$$

Por tanto el valor óptimo L_0 al que debe calibrarse la máquina para minimizar $g(L)$ se obtendrá derivando $g(L)$, igualando a cero y despejando dicho valor óptimo.

Recuerda que si pretendes derivar la función

$F(L) = \int_{a(L)}^{b(L)} f(x,L)dx$, se cumple que

$$\frac{dF}{dL} = F'(L) = \int_{a(L)}^{b(L)} \frac{df(x,L)}{dL}dx + f(b(L),L)\cdot b'(L) + f(a(L),L)\cdot a'(L)$$

Procedemos a plantear

$$\frac{dg}{dL} = 0$$

donde

$$g(L) = \int_{-\infty}^{500-L} (L+x)\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}dx + \int_{500-L}^{\infty} (L+x-500)\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}dx$$

$$g'(L) = \overbrace{\int_{-\infty}^{500-L} \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}dx + \int_{500-L}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-x^2}{2\sigma^2}}dx}^{1} + \ldots$$

$$\ldots\left[(L+500-L)\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(500-L)^2}{2\sigma^2}}(-1)+0\right] + \left[0-(L+500-L-500)\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(500-L)^2}{2\sigma^2}}\right]$$

Quedando:

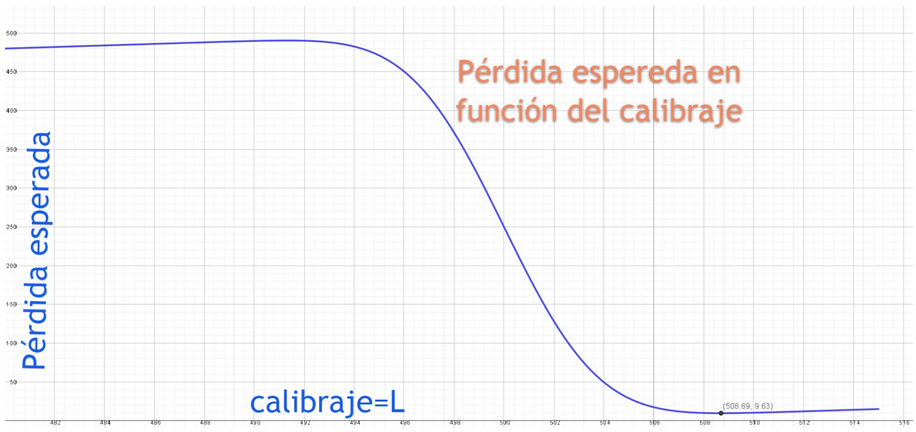
$$1-\frac{500}{\sigma\sqrt{2\pi}}e^{\frac{-(500-L)^2}{2\sigma^2}} = 0 \to e^{\frac{-(500-L)^2}{2\sigma^2}} = \frac{\sigma\sqrt{2\pi}}{500} \to -\frac{(500-L)^2}{2\sigma^2} = ln(\frac{\sigma\sqrt{2\pi}}{500})$$

Al despejar L, obtenemos:

$$L = 500 \pm \sqrt{2\sigma^2 ln\left(\frac{500}{\sigma\sqrt{2\pi}}\right)}$$

Que al sustituir los valores de σ obtenemos: $L_1 = 491.3082$, $L_2 = 508,6917$

Comprobándose con la derivada segunda que en L_2 es donde está el mínimo. La siguiente gráfica explica en función del calibraje elegido L cuál es la pérdida esperada.



- . -