Laboratorio de Probabilidad y Estadística

Valeria Naigé Rodríguez Martínez

16 de febrero de 2025

1- Tipos de datos y medidas de centralidad

Datos de 10 empleados en una empresa.

1- Clasifique las variables en cualitativas o cuantitativas
Nombre – Cualitativa, describe una cualidad
Edad – Cuantitativa, puede expresarse numéricamente
Área de trabajo – Cualitativa, se puede clasificar en categorías

2- Determine la media, mediana y moda de la variable "edad"
Media
Usamos la fórmula para calcular la media de los valores:

$$\bar{x} = \frac{\Sigma_{i=1}^{10} x_i}{10}$$

$$\bar{x} = \frac{25 + 30 + 40 + \cdots + 33 + 27}{10}$$

$$\bar{x} = 35.1$$

Mediana
Ordenamos los elementos

$$25, 30, 40, 35, 28, 50, 45, 38, 33, 27$$
$$\rightarrow$$
$$25, 27, 28, 30, 33, 35, 38, 40, 45, 50$$

Como se trata de un conjunto de número par, la mediana será el promedio de los dos valores centrales, en este caso, aquellos en la posición 5 y 6.

$$25, 27, 28, 30, 33, 35, 38, 40, 45, 50$$

$$\tilde{x} = \frac{33 + 35}{2}$$

$$\tilde{x} = 34$$

Moda
Como se observa, todos los valores de Edad se repiten exactamente una vez, por lo tanto, el conjunto de valores es **amodal**.

3- Interprete los resultados
Se puede concluir que, entre los 10 empleados, la edad promedio es de 35 años, la

edad media entre los valores será de 34 años y no hay una edad que se repita entre los empleados.

2- Medidas de dispersión

Conjunto de datos de calificaciones de 8 estudiantes en un examen.

$$X = \{70, 85, 90, 95, 88, 92, 75, 80\}$$

1. <u>Calcule la varianza y la desviación estándar de los datos</u>
   Se supondrá que se trata de una muestra de la población total.
   Se calculará la media de los datos:

$$\bar{x} = \frac{\Sigma_{i=1}^{n} x_i}{n}$$
$$\bar{x} = \frac{70 + 85 + 90 + \cdots + 75 + 80}{8}$$
$$\bar{x} = 84.375$$
$$\bar{x} = 84.4$$

   Varianza
   Se utilizará la fórmula de la varianza para muestras.

$$v(x) = S^2 = \frac{\Sigma_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}$$
$$v(x) = \frac{\Sigma_{i=1}^{8} (x_i - \bar{x})^2}{8 - 1}$$
$$v(x) = \frac{(70 - 84.4)^2 + (85 - 84.4)^2 + \cdots + (80 - 84.4)^2}{7}$$
$$v(x) = 75.69$$
$$v(x) = 75.7$$

   Desviación estándar

$$\sigma(x) = S = \sqrt{\frac{\Sigma_{i=1}^{n} (x_i - \bar{x})^2}{n - 1}}$$
$$\sigma(x) = \sqrt{v(x)}$$
$$\sigma(x) = \sqrt{75.7}$$
$$\sigma(x) = 8.7$$

2. <u>Interprete la dispersión de los datos</u>
   La varianza nos indica que los datos se encuentran, en promedio, a 75.7 unidades cuadradas de la media, si bien esto lo podríamos interpretar como que los datos no están ni muy alejados ni muy cercanos a la media, podemos tomar la desviación

estándar para un mejor análisis. La desviación nos indica que los datos se encuentran, en promedio, a 8.7 unidades de la media.

3- Probabilidades y Teorema de Bayes

60% de empleados programadores y 40% diseñadores. El 70% de los programadores y 30% de los diseñadores tienen conocimientos en IA.

Definimos los eventos

A − Es programador

B − Tiene conocimientos de IA

Definimos la probabilidad condicional de eventos dependientes

$$P\left({}^{A}/_{B}\right) = \frac{P(A \cap B)}{P(B)}$$

$$P\left({}^{A}/_{B}\right) = \frac{P\left({}^{B}/_{A}\right)P(A)}{P(B)}$$

Sustituimos los valores

$$P\left({}^{A}/_{B}\right) = \frac{(0.7)(0.6)}{((0.7 * 0.6) + (0.3 * 0.4))}$$

$$P\left({}^{A}/_{B}\right) = \frac{(0.7)(0.6)}{0.54}$$

$$P\left({}^{A}/_{B}\right) = 0.7778$$

∴La probabilidad de que un empleado elegido al azar con conocimientos de IA sea programador es de 77.78%

4- Distribuciones de probabilidad

Distribución Poisson con media $\lambda = 3$ defectos por lote

1) $x \rightarrow número\ de\ defectos\ por\ lote$
2) $x \sim P\ con\ \lambda = 3$
3) $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}\ , x = 0,1,2, \ldots$
4) $f(x) = \frac{3^x e^{-3}}{x!}\ , x = 0,1,2, \ldots$

<u>1. Calcule la probabilidad de que un lote tenga exactamente 2 defectos.</u>

$$P(x = 2) = f(2) = \frac{3^2 e^{-3}}{2!}$$

$$P(x = 2) = 0.22$$

∴La probabilidad de que el lote tenga exactamente 2 elementos defectuosos es de 22%

2. Calcule la probabilidad de que un lote tenga al menos 1 defecto.

Como no se tiene el dato de la cantidad total de unidades por lote, se utilizará una de las reglas de la probabilidad, que establece que su valor máximo puede ser igual a 1.

Como se desea calcular la probabilidad de que el lote tenga 1 o más defectos, pero no se sabe el total de defectos posibles, se calculará la probabilidad como la probabilidad total menos la probabilidad de que existan 0 defectos.

$$P(x \geq 1) = 1 - P(x < 1) = 1 - P(x = 0)$$

Para ello, se calculará primero la probabilidad de que ocurran 0 defectos.

$$P(x = 0) = f(0) = \frac{3^0 e^{-3}}{0!}$$

$$P(x = 0) = 0.049$$

Ahora bien, se calculará la probabilidad inicial

$$P(x \geq 1) = 1 - P(x = 0)$$

$$P(x \geq 1) = 1 - (0.049)$$

$$P(x \geq 1) = 0.95$$

∴La probabilidad de que el lote tenga al menos un elemento defectuoso es de 95%.

5- Funciones de densidad y distribución acumulativa

$$x \sim N(50{,}10)$$

Se utilizará la estandarización para la resolución del problema donde

$$\mu = 50$$

$$\sigma = 10$$

Así bien, el nuevo valor de x será interpretado como z, en conjunto con la tabla de valores de la Normal.

1- Determine la probabilidad de que X tome un valor menor que 45.

$$P(x < 45)$$

Al estandarizar:

$$= P(\frac{x-\mu}{\sigma} < \frac{45-\mu}{\sigma})$$

$$= P(\frac{x-50}{10} < \frac{45-50}{10})$$

$$= P(z < -0.5)$$

$$= 0.3085$$

∴La probabilidad de que X tome un valor menor que 45 es de 30.85%

2- Determine la probabilidad de que X esté entre 40 y 60.

$$P(40 \leq x \leq 60)$$

Al estandarizar:

$$P\left(\frac{40-\mu}{\sigma} \leq \frac{x-\mu}{\sigma} \leq \frac{60-\mu}{\sigma}\right)$$

$$= P\left(\frac{40-50}{10} \leq z \leq \frac{60-50}{10}\right)$$

$$= P(-1 \leq z \leq 1)$$

$$= 1 - 0.1587 - 0.1587$$

$$= 0.6826$$

∴La probabilidad de que X esté entre 40 y 60 es de 68.26%

3- <u>Use la función de distribución acumulativa para verificar sus respuestas.</u>

Se utilizó la función de distribución acumulativa de la Distribución Normal. Los valores fueron tomados de una tabla de área bajo la curva de la Distribución Normal obtenida del curso de Probabilidad de la carrera.

Para verificar los resultados, se utilizará una calculadora virtual.

Primeramente, definiremos la función de probabilidad de una distribución normal:

$$f(x) = \frac{e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

Para aplicar la distribución acumulativa, integramos la función y simplificamos

$$\int \frac{e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} dx$$

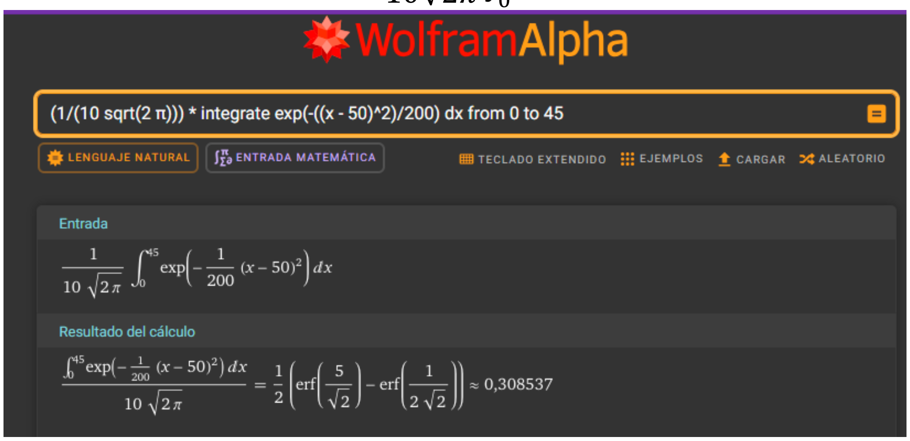
$$\frac{1}{\sigma\sqrt{2\pi}} \int e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx$$

Ahora, sustituimos los valores

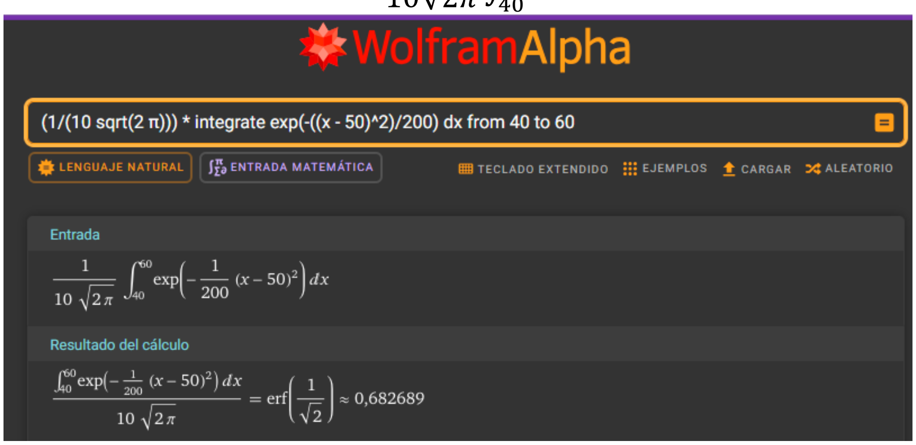
$$\frac{1}{10\sqrt{2\pi}} \int e^{\frac{-(x-50)^2}{200}} dx$$

Calculamos para ambos casos, utilizando una calculadora matemática.

$$\frac{1}{10\sqrt{2\pi}}\int_{0}^{45} e^{\frac{-(x-50)^2}{200}}\,dx$$



$$\frac{1}{10\sqrt{2\pi}}\int_{40}^{60} e^{\frac{-(x-50)^2}{200}}\,dx$$



Como se observa, se llegó a los resultados esperados.

6- Probabilidad condicional

Un dado justo de seis caras se lanza dos veces.

1- <u>¿Cuál es la probabilidad de obtener un número par en el segundo lanzamiento, dado que en el primero salió impar?</u>

Definimos los eventos como

$A - Obtener\ par\ en\ el\ segundo\ lanzamiento$

$B - Obtener\ impar\ en\ el\ segundo\ lanzamiento$

Se define la probabilidad condicional como

$$P\left({}^{A}/_{B}\right) = \frac{P(A \cap B)}{P(B)}$$

En este caso, se trata de dos eventos independientes, dado a que el resultado del lanzamiento del segundo dado no se ve afectado por el primer lanzamiento. De esta forma, tenemos:

$$P\left({}^{A}/_{B}\right) = \frac{P(A) * P(B)}{P(B)}$$

Ahora bien, calculamos las probabilidades de A y B

- En el caso del evento A tenemos

$$M = \{1,2,3,4,5,6\}$$
$$E = \{2,4,6\}$$
$$P(A) = \frac{3}{6} = 0.5$$

- En el caso del evento B tenemos

$$M = \{1,2,3,4,5,6\}$$
$$E = \{1,3,5\}$$
$$P(B) = \frac{3}{6} = 0.5$$

Sustituimos los valores en la probabilidad original

$$P\left({}^{A}/_{B}\right) = \frac{(0.5) * (0.5)}{(0.5)}$$

$$P\left({}^{A}/_{B}\right) = 0.5$$

∴La probabilidad de que en el segundo lanzamiento se obtenga un número par dado que en el primero se obtuvo uno impar de 50%

2- <u>Interprete los resultados obtenidos.</u>

Los datos demuestran que, independientemente del valor obtenido en el primer lanzamiento, la probabilidad de obtener un valor par en el segundo lanzamiento será del 50% ya que los casos de éxito (E) para este caso representan la mitad del espacio muestral (M).

7- Distribución binomial

Examen con 5 preguntas, cada una con 4 posibles respuestas. Un estudiante responde al azar.

1) $x \rightarrow aciertos\ del\ estudiante$
2) $x \sim B\ con\ n = 5\ y\ \theta = 0.25$
3) $f(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, x = 0,1,2,3,4,5$
4) $f(x) = \binom{5}{x} 0.25^x (0.75)^{5-x}, x = 0,1,2,3,4,5$

1- ¿Cuál es la probabilidad de que el estudiante acierte exactamente 3 respuestas?

$$P(x = 3) = f(3)$$
$$= f(3) = \binom{5}{3} 0.25^3 (0.75)^{5-3}$$
$$= f(3) = 0.0878$$

∴La probabilidad de acertar exactamente 3 respuestas es de 8.78%.

2- ¿Cuál es la probabilidad de que acierte al menos una respuesta?

$$P(x \geq 1) = f(1) + f(2) + f(3) + f(4) + f(5) = 1 - f(0)$$

Calculamos la probabilidad de acertar exactamente 0 respuestas

$$f(0) = \binom{5}{0} 0.25^0 (0.75)^{5-0} = 0.2373$$

Por lo tanto

$$P(x \geq 1) = 1 - 0.2373$$
$$P(x \geq 1) = 0.7627$$

∴La probabilidad de obtener al menos una respuesta correcta es de 76.27%

8- Regla de Laplace

5 bolas rojas y 7 bolas azules.

1- Si se extrae una bola al azar, determine la probabilidad de que la bola extraída sea roja.

Definimos el evento

$A - obtener\ una\ bola\ roja$

Así como el espacio muestral

$M = \{5\ rojas, 7\ azules\}$

Y los casos de éxito

$E = \{5\ rojas\}$

Calculamos la probabilidad:

$$P(A) = \frac{E}{M} = \frac{5}{12} = 0.4166$$

∴La probabilidad de obtener una bola roja extraída al azar de la urna es de 41.66%

2- Si se extraen dos bolas sin reemplazo, ¿cuál es la probabilidad de que ambas sean azules?
Definimos el evento 1
$A - obtener\ una\ bola\ azul$
Así como el espacio muestral
$M = \{5\ rojas, 7\ azules\}$
Y los casos de éxito
$E = \{7\ azules\}$

Calculamos la probabilidad:

$$P(A) = \frac{E}{M} = \frac{7}{12}$$

Así bien, definimos el evento 2, dado que ya sucedió el evento 1
$B - obtener\ una\ segunda\ bola\ azul$
Así como el espacio muestral
$M = \{5\ rojas, 6\ azules\}$
Y los casos de éxito
$E = \{6\ azules\}$

$$P(B) = \frac{E}{M} = \frac{6}{11}$$

Este caso será una probabilidad por etapas, por lo que la probabilidad total estará dada por

$$P(A) * P(B)$$

$$\frac{7}{12} * \frac{6}{11} = 0.3181$$

∴La probabilidad de obtener dos bolas azules, una después de otra, sin reemplazo es de 31.81%

9- Esperanza matemática

Premio 1000 dólares, probabilidad de 0.01 y costo de boleto de 10 dólares.

1- Calcule la esperanza matemática de la ganancia del jugador.

En total, se tienen 100 boletos, cada uno con un costo de 10 dólares, este costo se puede interpretar como una ganancia negativa de -10.
Así bien, solo habrá un boleto ganador, por lo que la persona que tenga este boleto tendrá una ganancia de -10+1000 = 990 dólares.
De esta manera, se tendrá a 99 personas con una ganancia de -10 y a una persona con una ganancia de 990.
Definimos la esperanza matemática con una X discreta

$$E(x) = \sum_{\forall x} xf(x)$$

Se puede sustituir de la forma

$$E(x) = 1(990) + \ 99(-10)$$

De esta forma se obtiene:

$$E(x) = 990 - 990$$
$$E(x) = 0$$

∴La esperanza matemática de la ganancia de un jugador será de 0.

2- Interprete el resultado
En promedio, se esperará que un jugador de la lotería no obtenga una ganancia significativa del juego.

10- Ley de los grandes números

Se lanza una moneda justa 1000 veces y se calcula la frecuencia relativa de obtener cara.

1- ¿Cuál es el valor esperado de la frecuencia relativa de obtener cara?
Definimos el evento como
$A - Lanzar\ una\ moneda\ justa$
Definimos el espacio muestral como
$M = \{C, A\}$ refiriéndose a C como "cara" y A como "águila" o "cruz"
Definimos los casos de éxito como
$E = \{C\}$

Calculamos la esperanza matemática de la frecuencia relativa como

$$E(A) = \frac{E * (total\ de\ éxitos)}{M * 1000} \approx \frac{1}{2} \approx 0.5$$

Al tener un alto número de eventos, el valor esperado de la frecuencia relativa tenderá a ser el valor esperado teórico del evento.
∴El valor esperado de la frecuencia relativa de obtener cara será de 0.5.

2- ¿Cómo se relaciona esto con la Ley de los Grandes Números?
La Ley de Grandes Números establece que, entre más grande sea la muestra de eventos al azar sobre la cual se somete el estudio, más cercano estará de representar

la media de la población a estudio. Es decir, entre más eventos tengamos en la muestra, más cercanos estarán los resultados a la realidad.
En este caso, al tener una muestra grande de un evento simple, el valor esperado de la frecuencia relativa tenderá a ser el valor teórico de la esperanza matemática en el estudio del caso simple.