Probabilidad y Estadística

El propósito de este laboratorio es familiarizarse con conceptos fundamentales de probabilidad y estadística, a través de problemas prácticos relacionados con los subtemas estudiados.

1. Tipos de datos y Medidas de tendencia central

En una empresa se han recolectado los siguientes datos de 10 empleados:

Nombre	Edad (años)	Área de trabajo
Ana	25	Ventas
Luis	30	Administración
Marta	40	Producción
Carlos	35	Ventas
Elena	28	Recursos Humanos
Juan	50	Producción
Sofía	45	Administración
Pedro	38	Ventas
Daniel	33	Producción
Laura	27	Recursos Humanos

1. Clasifique las variables en cualitativas o cuantitativas.

   Variables cualitativas:
   Nombre y Área de trabajo

   Variables cuantitativas:
   Edad

2. Determine la media, mediana y moda de la variable "Edad".

   La media es $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

   $\bar{x} = 35.1$

   Mediana = X25, X27, X28, X30, <u>33</u>, <u>35</u>, X38, X40, X45, X50
   Mediana = 34 (promedio de 33 y 35)

   Moda: Como ningún dato se repite, en conjunto de datos es amodal.

3. Interprete los resultados obtenidos.

   En el conjunto de datos estudiado, se puede decir que el grupo de edades tiene un distribución relativamente equilibrada, pero con una ligera tendencia hacia los valores más altos, dado que la media (35.1) es un poco mayor a la mediana (34). Sobre la moda, se puede decir que no hay una edad predominante y que las edades están distribuidas de manera dispersa.

2. Medidas de dispersión

Dado el siguiente conjunto de datos correspondiente a las calificaciones de 8 estudiantes en un examen:

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

1. Calcule la varianza y la desviación estándar de los datos.

   La varianza se calcula con la siguiente fórmula:

   $$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

   $\sigma^2 = 66.23$

   La desviación estándar se calcula con la siguiente fórmula:

   $$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

   $\sigma = 8.13$

2. Interprete la dispersión de los datos.

   Con el uso de la desviación estándar, podemos inferir que tan alejados en promedio estan los datos del promedio de la muestra ya sea hacia la derecha o la izquieda, por lo cual podemos decir que una gran parte de los datos estarana en el area creada por 84.37 ± 8.13

3. Probabilidades y Teorema de Bayes

Una empresa de tecnología ha identificado que el 60% de sus empleados son programadores, y el 40% son diseñadores. Se sabe que el 70% de los programadores tienen conocimientos de inteligencia artificial (IA), mientras que solo el 30% de los diseñadores tienen estos conocimientos.

Si se elige un empleado al azar y se sabe que tiene conocimientos de IA, ¿cuál es la probabilidad de que sea programador?

Tenemos que la probabilidad de que un empleado sepa sobre IA es del 70% y la probabilidad de que al escoger a un programador de entre los empleados sea del 60%, por lo cual podemos obtener:

$$P(P|IA) = \frac{P(IA|P)P(P)}{P(IA)}$$

Para obtener $P(IA)$ hacemos la siguiente operación:

$$P(IA) = P(IA|P)P(P) + P(IA|D)P(D)$$

Que es:

$$P(IA) = (0.7)(0.6) + (0.3)(0.4) = 0.54$$

Posterior a esto, calculamos $P(IA|P)P(P) = (0.7)(0.6)$, lo cual es igual a 0.42, y realizamos la división:

$$P(P|IA) = \frac{0.42}{0.54} = 0.77778$$

Por lo cual, tenemos que la probabilidad de que la persona escogida sea un programador es del 77.778%.

4. Distribuciones de probabilidad

Suponga que el número de defectos en un lote de producción sigue una distribución de Poisson con media $\lambda = 3$ defectos por lote.

1. Calcule la probabilidad de que un lote tenga exactamente 2 defectos.

   Tenemos que $\lambda = 3$ y que $k = 2$ buscamos este valor en la tabla de Probabilidades de la distribución Poisson y encontramos
   $P(X = 2, \lambda) = 0.2240$
   por lo cual podemos decir que la probabilidad de que se encuentren 2 defectos es de 22.4%

2. Calcule la probabilidad de que un lote tenga al menos 1 defecto.
   Tenemos que $\lambda = 3$ y que $k \geq 1$ buscamos el valor de la probabilidad que no deseamos en la tabla de Probabilidades de la distribución Poisson y encontramos
   $P(X = 0, \lambda) = 0.0498$
   Posterioir a esto restamos este valor al total de casosy obtenemos:
   $P(X \geq 1) = 1 - 0.0498 = 0.9502$
   por lo cual podemos decir que la probabilidad de que se encuentren al menos 1 defectos es de 95.02%

5. Funciones de densidad y distribución acumulativa

Sea X una variable aleatoria con distribución normal de media $\mu = 50$ y desviación estándar $\sigma = 10$.

1. Determine la probabilidad de que X tome un valor menor que 45.
   Con los datos obtenidos procedemos a estandarizar la muestra dada tal que:

   $$Z = \frac{x - \mu}{\sigma} = \frac{45 - 50}{10} = \frac{-5}{10} = -.5$$

   Posterior a esto buscamos el valor de $-.5$ en la tabla de distribucion normal y obtenemos:

   $$P(X \leq 45) = .3085$$

   Por lo cual podemos decir que la probabilidad de tomar un valor ≤ 45 es de 30.85%

2. Determine la probabilidad de que X esté entre 40 y 60.
   Con los datos obtenidos procedemos a estandarizar la muestra dada tal que:

   $$Z = \frac{x - \mu}{\sigma} = \frac{40 - 50}{10} = \frac{-10}{10} = -1$$

   $$Z = \frac{x - \mu}{\sigma} = \frac{60 - 50}{10} = \frac{10}{10} = 1$$

   Posterior a esto buscamos los valores de -1 y 1 en la tabla de distribución normal y obtenemos:

   $$P(40 \leq X \leq 60) = .8413 - .1587 = .6826$$

   Por lo cual podemos decir que la probabilidad de tomar un valor entre 40 y 60 es de 68.26%

3. Use la función de distribución acumulativa para verificar sus respuestas.
   Use la función de distribución acumulativa para realizar el procedimientos, usando las tablas de distribución normal para encontrar la probabilidad de que los resultados estuvieran dentro de los parámetros establecidos y de esta manera sacar los porcentajes.

6. Probabilidad condicional

Un dado justo de seis caras se lanza dos veces.

1. ¿Cuál es la probabilidad de obtener un número par en el segundo lanzamiento, dado que en el primero salió un número impar?
   Teniendo un espacio maestral de

   $$S = 1, 2, 3, 4, 5, 6$$

   tenemos que la probabilidad de tener un numero par es de

   $$P(X = (2, 4, 6)) = \frac{3}{4} = \frac{1}{2}$$

   y de tener un numero impar es de

   $$P(X = (1, 3, 5)) = \frac{3}{4} = \frac{1}{2}$$

   Independientemente de los resultados anteriores por lo cual la probabilidad de obtener un numero par es del 50%

2. Interprete los resultados obtenidos.
   Al ser eventos independientes el uno del otro y ninguno depender del resultado anterior (no haber cambios de casos favorables, en contra o totales) las probabilidades son iguales en todos los eventos con las mismas condiciones

7. Distribución binomial

Un examen de opción múltiple tiene 5 preguntas, cada una con 4 posibles respuestas, de las cuales solo una es correcta. Un estudiante responde al azar.

1. ¿Cuál es la probabilidad de que el estudiante acierte exactamente 3 respuestas?
   Tenemos que la probabilidad de acertar una respuesta es de $\frac{1}{4}$ por lo cual si queremos saber la probabilidad acertar 3 preguntas realizamos la siguiente operación:

   $$P(X=3)=\binom{5}{3}*(\frac{1}{4})^3*(\frac{3}{4})^2=.08789$$

   Teniendo una probabilidad de 8.789%

2. ¿Cuál es la probabilidad de que acierte al menos una respuesta?
   Tenemos que la probabilidad de acertar una respuesta es de $\frac{1}{4}$ por lo cual si queremos saber la probabilidad acertar 3 preguntas realizamos las siguientes operaciónes:

   $$P(X=0)=\binom{5}{0}*(\frac{1}{4})^0*(\frac{3}{4})^5=.2373$$

   esa siendo la probabilidad de no acertar ninguna respuesta

   $$P(X\geq 1)=1-.2373=.7627$$

   Teniendo una probabilidad de 76.27%

8. Regla de Laplace

Una urna contiene 5 bolas rojas y 7 bolas azules. Se extrae una bola al azar.

1. Determine la probabilidad de que la bola extraída sea roja.
   Tenemos que el total de casos es de $7+5=12$ por lo cual la probabilidad de extraer una bola del color rojo es de $P(R)=\frac{5}{12}=.41666$ entonces hay una probabilidad de 41.66%

2. Si se extraen dos bolas sin reemplazo, ¿cuál es la probabilidad de que ambas sean azules?
   Al ser eventos (la extracción de la bola) dependientes el primeo del segundo multiplicamos los resultados de la probabilidad de uno contra el otro teniendo

   $$(P(A_1)=\frac{7}{12})(P(A_2)=\frac{6}{11})=.5833*.5454=.3181$$

   Teniendo una probabilidad del 31.81% de sacar dos bolas azules sin remplazo

9. Esperanza matemática

Suponga que una persona juega una lotería donde el premio es de 1000 dólares con una probabilidad de 0.01, y el costo del boleto es de 10 dólares.

1. Calcule la esperanza matemática de la ganancia del jugador. Tenemos la siguiente información

   $$P(Ganar)=0.01$$

   $$P(Perder)=0.99$$

   Para lo cual tenemos que las ganancias sera de Premio - Costo_Boleto lo cual es $1000-10=900$ dólares y si no ganamos estaríamos perdiendo 10 dólares (-10) y tenemos que la esperanza de ganar es de

   $$E(X)=\sum P_i*X_i=(0.01*990)+(0.99*(-10))=0$$

   Por lo cual podemos decir que tenemos una esperanza de ganancia es de 0 dólares

2. Interprete el resultado obtenido. La esperanza de tener algún tipo de ganancia en la rifa es de 0 dlls por lo cual es bastante probable que no se tenga ganancia en la rifa

10. Ley de los grandes números

Un experimento consiste en lanzar una moneda justa 1000 veces y calcular la frecuencia relativa de obtener cara.

1. ¿Cuál es el valor esperado de la frecuencia relativa de obtener cara?

   $$FrecuenciaRelativa=\frac{\text{Veces que salio Cara}}{Lanzamientos}$$

   con una probabilidad de obtener cara del .5 podemos decir que de 1000 lanzamientos obtendremos muy posiblemente 500 caras en los 1000 lanzamientos de tal manera que tendremos

   $$FrecuenciaRelativa=\frac{500}{1000}=.5$$

   Por lo cual tendremos un 50% de caras en los lanzamientos

2. ¿Cómo se relaciona esto con la Ley de los Grandes Números?
   La probabilidad de tener cara es de 50% pero no significa que de cada 2 lanzamientos 1 sera cara y otro cruz por lo cual no siempre tendremos una distribución en la que tendremos 50% caras y 50% cruces, no obstante entre más lanzamientos realicemos lo mas probable es que nos acerquemos a este resultado viendo de esta manera los porcentajes de 50% y 50%