Repaso Inicial de Estadística Aplicada

1. Introducción

La estadística aplicada es una disciplina que se encuentra en la intersección de la matemática y la ciencia de datos, proveyendo las herramientas necesarias para entender, interpretar y representar datos de una manera significativa. Se aplica en una amplia gama de campos incluyendo, pero no limitándose a, economía, psicologíá´ıa, ciencias naturales, y más. En este repaso, exploramos conceptos fundamentales tales como la probabilidad, distintas distribuciones y medidas centralizadas y de dispersión.

2. Probabilidad

2.1 Definición y propiedades

La **probabilidad** es una medida matemática que refleja la certeza o posibilidad de que ocurra un evento específico. Se representa en una escala de 0 a 1, donde 0 indica que el evento es imposible y 1 indica que el evento es seguro. La probabilidad se define como:

$$P(A) = \frac{\text{Número de casos favorables}}{\text{Número de casos posibles}}$$

2.2 Eventos dependientes e independientes

Es vital distinguir entre eventos dependientes e independientes. Los eventos independientes son aquellos donde el resultado de uno no afecta el resultado del otro, mientras que los eventos dependientes están correlacionados.

2.3 Problemas prácticos

Ejemplo 1: En el lanzamiento de un dado equilibrado, la probabilidad de que salga un número impar es:

$$P(\text{Número impar}) = \frac{3}{6} = 0.5$$

Ejemplo 2: En una bolsa con 5 bolas rojas y 3 bolas azules, la probabilidad de sacar una bola roja y luego una bola azul sin reemplazo es:

$$P(\text{Rojo y luego azul}) = \frac{5}{8} \cdot \frac{3}{7} = \frac{15}{56} \approx 0.2679$$

3. Distribuciones

3.1 Distribución normal

3.1.1 Definición y propiedades

La **distribución normal** es una distribución de probabilidad continua que está completamente caracterizada por su media (μ) y desviación estándar (σ). Es simétrica y tiene una forma de campana, y su función de densidad de probabilidad est´a dada por:

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

3.1.2 Regla 68-95-99.7

Esta regla, también conocida como regla empírica, establece que en una distribución normal:

- Aproximadamente el 68% de los datos caen dentro de una desviación estándar de la media.
- Aproximadamente el 95% de los datos están dentro de dos desviaciones estándar de la media.
- Aproximadamente el 99.7% de los datos están dentro de tres desviaciones estándar de la media.

3.1.3 Problemas prácticos

Ejemplo 1: Considerando una población con una distribución normal de alturas con μ = 170 cm y σ = 10 cm, podemos usar la regla 68-95-99.7 para determinar qué porcentaje de la población tiene una altura entre 160 cm y 180 cm (es decir, dentro de una desviación estándar de la media).

Ejemplo 2: En un set de datos de calificaciones de exámenes con una distribución normal, se puede utilizar la distribución normal para encontrar la probabilidad de que un estudiante seleccionado al azar tenga una calificación superior a un valor dado.

3.2 Distribución Uniforme

3.2.1 Definición y propiedades

La **distribución uniforme** describe un escenario donde cada evento tiene la misma probabilidad de ocurrir. La función de densidad de probabilidad para una distribución uniforme continua en el intervalo.

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } x \in [a,b] \\ 0 & \text{para } x \notin [a,b] \end{cases}$$

3.2.2 Problemas prácticos

Ejemplo 1: Si estamos esperando un autobús que pasa cada 15 minutos, pero no sabemos nada sobre el horario específico, entonces el tiempo de espera tiene una distribución uniforme entre 0 y 15 minutos.

Ejemplo 2: En una producción de piezas de maquinaria, si cada pieza tiene la misma probabilidad de tener un defecto, entonces la distribución de los defectos sigue una distribución uniforme.

4. Medidas de Tendencia Central y Dispersión

4.1 Media

4.1.1 Definición y propiedades

La **media**, también conocida como promedio, es una medida central que se calcula sumando todos los valores de un conjunto de datos y dividiendo por el número de observaciones:

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

4.1.2 Problemas prácticos

Ejemplo 1: Para encontrar la media de las calificaciones de un grupo de estudiantes, sumamos todas las calificaciones y dividimos por el número de estudiantes.

Ejemplo 2: En un estudio sobre los salarios de una población, la media ayuda a obtener una idea general del nivel de salarios en esa población.

4.2 Varianza

4.2.1 Definición y propiedades

La varianza es una medida de dispersión que indica cómo están dispersos los valores en un conjunto de datos respecto a la media. La varianza se calcula como:

$$\sigma^2$$

$$= 1\frac{}{n\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

4.2.2 Problemas prácticos

Ejemplo 1: En una empresa, la varianza de los salarios puede indicar la equidad en la distribución de los salarios.

Ejemplo 2: Al analizar los resultados de un experimento, la varianza ayuda a entender la variabilidad de los resultados.

5. Conclusión

En este repaso, hemos explorado los fundamentos de la estadística aplicada, brindando una comprensión profunda de conceptos críticos como la probabilidad, diferentes distribuciones, y medidas de tendencia central y dispersóon. La comprensión de estos conceptos es vital para aplicar técnicas estadísticas avanzadas en la práctica y para interpretar datos de manera efectiva en una amplia variedad de campos.