

Resumen de estadística y probabilidad

J. I. Corradi

Tercer semestre

Índice general

Estadística	3
Conceptos básicos	3
Población	3
Población	3
Variable	3
Variables estadísticas	3
Cualitativas o categóricas	3
Nominal	3
Ordinal	3
Cuantitativas o numéricas	3
Continuas	4
Discretas	4
Medidas de localización o posición	4
Media	4
Mediana	4
Cuartil	4
Primer cuartil	4
Tercer cuartil	5
Percentil	5
Segundo cuartil	5
Segundo cuartil	5
Moda	5
Medidas de variabilidad o dispersión	6
Varianza muestral	6
Desviación estándar	6
Frecuencia	6
Probabilidad	6
Espacio muestral Ω	7
Evento	7
Teoría de conjuntos	7
Complemento	7
Unión	7
Intersección	7

Evento nulo	8
Probabilidad	8

Estadística

Conceptos básicos

Población

Una población estadística N es el total de individuos o conjunto de ellos que presentan o podrían presentar el rasgo característico que se desea estudiar.

Muestra

Una muestra n es un subconjunto de casos o individuos de una población. Interesa que una muestra sea representativa, y para ello debe escogerse una técnica de muestra adecuada que produzca una muestra aleatoria adecuada.

Variable

Una variable es cualquier característica cuyo valor puede cambiar de un objeto a otro en la población

Variables estadísticas

Cualitativas o categóricas

Son el tipo de variables que expresan distintas cualidades, características o atributo.

Nominal

Es aquella variable cualitativa cuya categoría no sigue ningún orden, se agrupa sin ninguna jerarquía entre sí. Ej.: estado civil, grupo sanguíneo, sexo.

Ordinal

Son aquellas variables categóricas con orden secuencial o progresión natural esperable o jerarquía. Ej.: nivel de educación.

Cuantitativas o numéricas

Describen una característica respecto a un valor numérico o cantidad.

Continuas

Son aquellas características que se miden dentro de un rango continuo infinito de valores numéricos y se registran con números reales. Pueden presentar cualquier valor dentro de cierto intervalo. Ej.: estatura, peso, ingresos.

Discretas

Este tipo de variables numéricas están asociadas a conteos o enumeraciones, razón por la cual solo pueden registrarse con números enteros. Ej.: número de hermanos, edad.

Medidas de localización o posición

Las medidas de localización están diseñadas para brindar al analista algunos valores cuantitativos de la ubicación central o de otro tipo de los datos en una muestra.

Media

La media es simplemente un promedio numérico.

$$\bar{x} = \sum_{i=1}^n \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_i}{n}$$

Mediana

El propósito de la mediana de la muestra es reflejar la tendencia central de la muestra de manera que no sea influida por los valores extremos. Los valores deben estar ordenados.

Para n impar:

$$M_e = \frac{2(n+1)}{4}$$

Para n par:

$$M_e = \frac{2n}{4}$$

Cuartil

Los cuartiles son tres valores que dividen una muestra en cuatro partes iguales. Los valores deben estar ordenados.

Primer cuartil

Para n impar:

$$Q_1 = \frac{n+1}{4}$$

Para n par:

$$Q_1 = \frac{n}{4}$$

Tercer cuartil

Para n impar:

$$Q_3 = \frac{3(n+1)}{4}$$

Para n par:

$$Q_3 = \frac{3n}{4}$$

Percentil

El percentil es una medida que indica el valor de la variable por debajo del cual se encuentra un porcentaje dado de observaciones en un grupo. Los valores deben estar ordenados.

Percentil 5

Para n impar:

$$P_5 = \frac{5(n+1)}{100}$$

Para n par:

$$P_5 = \frac{5n}{100}$$

Percentil 95

Para n impar:

$$P_{95} = \frac{95(n+1)}{100}$$

Para n par:

$$P_{95} = \frac{95n}{100}$$

Moda

La moda es el valor más repetido del conjunto de datos, es decir, el valor cuya frecuencia relativa es mayor. En un conjunto puede haber más de una moda (bimodal).

Medidas de variabilidad o dispersión

Varianza muestral

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

Desviación estándar

Es una medida del grado de dispersión de los datos con respecto al valor promedio.

$$s = \sqrt{s^2}$$

Frecuencia

La frecuencia es una medida que sirve para comparar la aparición de un elemento x_i en un conjunto de elementos (x_1, x_2, \dots, x_n) .

aquí falta colocar todos los tipos de frecuencias.

Probabilidad

Espacio muestral (Ω)

Al conjunto de todos los resultados posibles de un experimento estadístico se le llama espacio muestral Ω y se representa con el símbolo S .

Ej. 1: lanzar un dado.

$$\Omega_1 = \{1, 2, 3, 4, 5, 6\}$$

Ej. 2: par o impar.

$$\Omega_2 = \{par, impar\}$$

Evento

Un evento es el conjunto de uno o más resultados contenidos en el espacio muestral Ω . Un evento es simple si consiste en exactamente un resultado y compuesto si consiste en más de un resultado.

Teoría de conjuntos

Complemento

El complemento de un evento A , denotado por A' , es el conjunto de todos los resultados en Ω que no están contenidos en A .

Unión

La unión de dos eventos A y B , denotados por $A \cup B$ y leídos " A o B ", es el evento que consiste en todos los resultados que están en A o en B o en ambos eventos, es decir, todos los resultados en por lo menos uno de los eventos. El evento $A \cup B$ ocurre si por lo menos ocurre A , B o ambos.

Intersección

La intersección de dos eventos A y B , denotada por $A \cap B$ y leída " A y B ", es el evento que consiste en todos los resultados que están tanto en A como en B . El evento $A \cap B$ ocurre si A y B ocurren a la vez.

Evento nulo

Se denomina \emptyset a un evento sin resultados. Cuando $A \cap B = \emptyset$, se dice que A y B son eventos mutuamente excluyentes o disjuntos.

Probabilidad

$$P(A) = \frac{n_i}{N}$$