Universidad Manuela Beltrán
Acreditada en Alta Calidad Multicampus

Conceptos básicos

Sesión I

19 de octubre de 2024

William Javier Rodríguez Cruz
phD. en física

Contenido

1. Definiciones básicas
2. Datos
3. Introducción R
   - Álgebra vectorial y matricial
   - Vectores, matrices y operaciones
   - Ecuaciones algebraicas
   - Funciones

Estadística

Figura: Diagramas estadísticos [1]

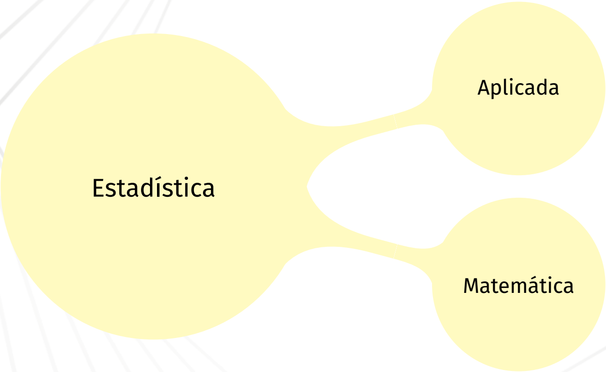
Estadística

Figura: Diagramas estadísticos [1]

Definición

Consiste en la recolección, cuantificación, sintesis, análisis e interpretación de un conjunto de datos [2].

Definiciones



Bioestadística

Figura: Estudios bioestadísticos [1].

Bioestadística

Figura: Estudios bioestadísticos [1].

Definición

Rama de la estadística que se enfoca en los problemas planteados dentro de los sistemas de las ciencias de la vida [2].

Bioestadística

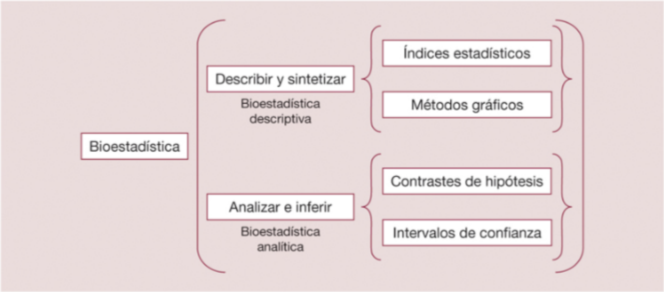


Figura: Aspectos estudiados por la bioestadística [2].

Muestras y poblaciones

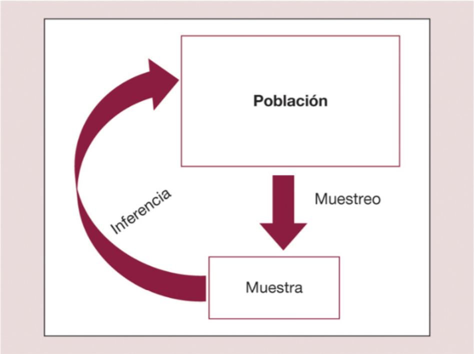


Figura: Muestras y poblaciones: proceso de muestreo e inferencia [2].

Investigación científica

Figura: Proceso iterativo de avance del conocimiento biomédico [2].

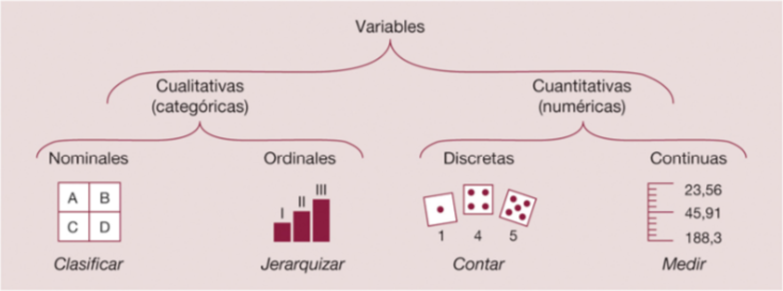


Figura: Tipos de variables [1].

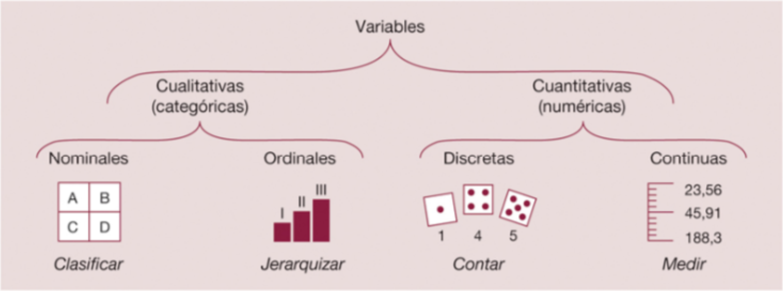


Figura: Tipos de variables [1].

Actividad 1.

1. ¿Las variables cualitativas o categóricas pueden representarse con números?
2. Consulte los dos tipos de variables y sus sub-clasificaciones.

Definiciones

El lenguaje de programación en R es orientado a objetos que almacenan datos de cualquier clase por medio de arreglos (vectores, matrices o data frame).

Vectores: Todos los `datos` son de la misma clase

$$\mathbf{a} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \end{pmatrix} \tag{1}$$

$$\mathbf{b} = \begin{pmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{pmatrix}. \tag{2}$$

Definiciones

Matrices: Todos los `datos` son de la misma clase

$$\mathbf{W} = \begin{pmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{pmatrix} \tag{3}$$

Ejemplo: Matriz de características

$$\mathbf{C} = \begin{pmatrix} M & \text{Grande} & \text{San bernardo} \\ H & \text{mediano} & \text{Beagle} \\ H & \text{pequeño} & \text{Pincher} \end{pmatrix}. \tag{4}$$

Cree un ejemplo de una matriz.

Definiciones

Data frame: Admite cualquier clase de datos

$$\mathbf{d} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{pmatrix} \tag{5}$$

Cree un data frame y luego escribalo en R

Definiciones

Data frame: Admite cualquier clase de datos

$$\mathbf{d} = \begin{pmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{pmatrix} \tag{5}$$

Cree un data frame y luego escribalo en R ¿Qué es R?

Suma y resta

- Si se suma un ESCALAR a un vector cada uno de los componentes del vector es modificado.
- Para la suma de vectores es necesario que la dimensión (longitud) de los vectores sea idéntica.
- Una matriz corresponde a un arreglo conformado por vectores de la misma clase.
- Para sumar vectores y matrices no es necesario que estos concuerden en longitud y dimension.
- En caso de que la longitud de la matriz (filas por columnas) no sea múltiplo de la longitud del vector, la operación se realizará, pero habrá un mensaje de error.

Producto y división (elemento por elemento)

- Para la multiplicación se usa el símbolo (*) y para la división el símbolo (/).
- En términos matemáticos tenemos:

$$c\mathbf{v} = c(v_1 . \dots, v_n) = (cv_1, \dots, cv_n) \tag{6}$$

$$c/\mathbf{v} = c/(v_1 . \dots, v_n) = (c/v_1, \dots, c/v_n) \tag{7}$$

- Para la multiplicación matricial y vectorial es necesario que la dimensión sea idéntica.
- También es posible la multiplicación entre un vector fila y una matriz, cuyas dimensiones coincidan con la longitud del vector.

Ejemplo de aplicación

- Cree un vector de los primeros cuatro número pares, es decir,

$$r = (2, 4, 6, 8) \tag{8}$$

- Obtenga la mitad de cada uno de los cuatro elementos que compone el vector r.
- Empleando la operación de división, obtenga el siguiente resultado:

$$r_1 = (2; 1; 0{,}66; 0{,}5) \tag{9}$$

Producto escalar

El producto escalar se define como

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \tag{10}$$

Producto escalar

El producto escalar se define como

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \tag{10}$$

Un ejemplo de papel y lápiz

$$\mathbf{a} = (1, 2, 3) \tag{11}$$

$$\mathbf{b} = (2, 5, 8) \tag{12}$$

Producto escalar

El producto escalar se define como

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \ldots + a_n b_n \tag{10}$$

Un ejemplo de papel y lápiz

$$\mathbf{a} = (1, 2, 3) \tag{11}$$

$$\mathbf{b} = (2, 5, 8) \tag{12}$$

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} =& (1, 2, 3) \cdot (2, 5, 8) \\ =& (1 \times 2) + (2 \times 5) + (3 \times 8) \\ =& 2 + 10 + 24 = 36 \end{aligned} \tag{13}$$

Producto escalar

Si dos vectores son ortogonales el producto es nulo

Multiplique los siguientes vectores

$$\mathbf{a} = (1, 1) \quad \mathbf{b} = (-1, 1) \tag{14}$$

Para una mejor comprensión grafique cada uno de los vectores en la recta numérica ¿Qué concluye?

Norma de un vector

Esta operación nos arroja la magnitud o longitud del vector.

$$|v| = \sqrt{v_1^2 + \ldots + v_n^2} \tag{15}$$

El lenguaje de programación R no proporciona una función o librería que desarrolle esta operación.

Aplicación en R

- Desarrolle todas las operaciones que se presentaron en las transparencias 11 y12.
- Calcule la norma del vector

$$\mathbf{s} = (4, 3) \tag{16}$$

- Grafique el vector **s** y calcule la longitud de la hipotenusa ¿Qué concluye?

Producto matricial

- El producto matricial no es conmutativo $A_{n\times m}B_{m\times p} \neq B_{m\times p}A_{n\times m}$.
- El número de columnas de la matriz A debe ser igual al númeo de filas de la matriz B.
- La matriz resultante tendrá el mismo número de filas de la matriz A y el mismo número decolumnas de la matriz B.
- Aplicación en R Cree dos matrices y obtenga el producto matricial.

Polinomios

- Una ecuación es un objeto abstracto que aplicado a casos reales arroja soluciones codificadas en la variable x.

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0 \tag{17}$$

- En R una forma de solucionar las ecuaciones consite en crear vectores con lps coeficientes del polinomio en forma creciente $(a_0, a_1, \ldots, a_n)$, luego restarlos y buscar la raíz medianbte la función `polyroot`, que es la solución.

Ejemplo: Resolver la ecuación

$$\begin{aligned} 2x^2 + 5x + 5 &= x^2 + 3x + 4 \\ x^2 + 2x + 1 &= 0 \end{aligned} \tag{18}$$

Sistemas de ecuaciones lineales

Un sistema de ecuaciones líneales tiene la forma,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \ldots \quad + \ldots \quad + \ldots \quad + \ldots &= \quad \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \tag{19}$$

Aquí x_j corresponde a las incognitas, a_{ij} a los coficientes y b_i etiqueta los términos independientes.

Tipos de soluciones

- Incompatible: Si no tiene solución
- Compatible: Si tiene solución. Puede ser *determinado* si el sistema admite una unica solución e *indeterminado* si admite un conjunto infinito de soluciones.

Sol. Método matricial

$$A \cdot x = b \tag{20}$$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{21}$$

Existen diferentes métodos para solucionar sistemas de ecuaciones líneales teanto análiticamente como por métodos numéricos. En R la función `solve` resuelve este tipo de ecuaciones.

Solve R

Resolver el siguiente sistema de ecuaciones:

$$\begin{aligned} 3x + y + 2z &= 10 \\ 4x + 3y + 4z &= 21 \\ 2x + y + 2z &= 9 \end{aligned} \tag{22}$$

Solve R

Resolver el siguiente sistema de ecuaciones:

$$\begin{aligned} 3x + y + 2z &= 10 \\ 4x + 3y + 4z &= 21 \\ 2x + y + 2z &= 9 \end{aligned} \tag{22}$$

Paso 1: Sistema mediante matrices

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 4 & 3 & 4 \\ 2 & 1 & 2 \end{pmatrix}; \quad b = \begin{pmatrix} 10 \\ 21 \\ 9 \end{pmatrix} \tag{23}$$

Solve R

- Cree tres vectores en R que contengan las tres filas de la matriz A.
- Cree un vector para el vector b.
- Use la operación `solve` para resolver el sistema de ecuaciones líneales.
- Compruebe la solución obtenida.

Funciones

Tabla 1. Clases básicas de funciones

Clase	Forma funcional
Lineal	$mx + b$
Potencia	x^n
Raíz	$\sqrt[n]{x}$
Racional	$P(x)/Q(x)$
Exponencial	a^x
Logarítmica	$log_a(x)$
Trigonométrica	$sin(x)$
	$cos(x)$
	$tan(x)$

Figura: Imágen tomada de [?]

Gráficos de una función

Tabla 2. Opciones gráficas elementales

Descripción	Instrucción	Opciones
Título	`main`	`"texto"`
Nombre del eje x	`xlab`	`"texto"`
Nombre del eje y	`ylab`	`"texto"`
Dominio	`xlim`	`c(a,b)`
Rango	`ylim`	`c(a,b)`
Tipo de gráfico	`type`	Líneas: `"l"` Puntos: `"p"` Ambos: "o"
Ancho de línea	`lwd`	`1,2,...`
Color de línea	`col`	"red","blue",...

Figura: Imágen tomada de [?]

Bibliografía

- https://towardsdatascience.com/40-statistics-interview-problems-and-answers-for-data-s
  https://www.masterspublichealth.net/faq/what-is-biostatistics/.
- Martínez. González. M. A, Sánchez-Villegas. A, Toledo. Atucha. E. A & Faulin. Fajardo. J. *Bioestadístca amigable* (2014). España: Elsevier.