

# Introducción a la Inferencia Estadística

ESCUELA POLITÉCNICA  
SUPERIOR DE CÓRDOBA

Universidad de Córdoba

DEPARTAMENTO DE ESTADÍSTICA





## Nociones previas

Población o colectivo: es un conjunto de elementos que pueden ser agrupados dentro de una determinada clase, ya sean existentes como posibles, es decir el conjunto de elementos o individuos que se trata de analizar.

Individuo o unidad estadística: cada uno de los elementos que forman la población.

Censo: recuento de todos los elementos que conforman toda la población.

Dificultad en realización de censos: tiempo, costos, destrucción del ente, entes relativos, etc.

Su uso más habitual está asociado al censo poblacional o censo de población, donde se contabilizan los habitantes de un pueblo o de un país y se recogen diversos datos con fines estadísticos.

Muestra: es el subconjunto de la población que se observa, y que se usa para inducir o extrapolar información sobre toda la población.

## Nociones previas

Muestra genérica:  $\vec{X} = (X_1, X_2, \dots, X_n)$  v.a. n-dimensional.

Para cada valor de  $\vec{X} \Rightarrow$  muestra distinta: realización de la muestra.

Conjunto de todas las posibles realizaciones: espacio muestral de  $\vec{X}$ .

$x_1$	$x'_1$	$x''_1$	...	$X_1$
$x_2$	$x'_2$	$x''_2$	...	$X_2$
$x_3$	$x'_3$	$x''_3$	...	$X_3$
...	...	...	...	...
$x_n$	$x'_n$	$x''_n$	...	$X_n$
<hr/>				
$\bar{x}$	$\bar{x}'$	$\bar{x}''$	...	$\bar{X}$
$S_x$	$S_{x'}$	$S_{x''}$	...	$S_X$



## Tipos de muestreo

Según el diseño:

Muestreo aleatorio simple: las distintas observaciones se obtienen con igual probabilidad y de forma independiente unas de otras.

Muestreo estratificado: la población se divide en varios subconjuntos o estratos y en cada uno de éstos se toma una muestra; los estratos se seleccionan de forma que sean lo más homogéneos posibles internamente y heterogéneos entre sí.

Muestreo por conglomerados: la población se divide en varios subconjuntos o conglomerados donde se toman las correspondientes muestras; los conglomerados deben tomarse homogéneos entre sí para que cada uno de ellos sea representativo de toda la población.

Muestreo sistemático: se basa en tomar muestras de una manera directa y ordenada a partir de una regla determinística, también llamada sistemática.

Muestreo polimetálico: los datos se seleccionan en varios instantes de tiempo o etapas.



## Tipos de muestreo

Según la forma en que se toman las observaciones:

Muestreo independiente o aleatorio simple: consiste en elegir al azar las unidades experimentales de tal forma que cualquier elemento de la población tenga la misma probabilidad de ser elegido y que las elecciones sucesivas sean independientes entre sí.

Muestreo dependiente: se realiza una observación, después una intervención y posteriormente una nueva observación.

Según el tamaño de la muestra:

Muestreo en poblaciones finitas

Con reemplazamiento: seleccionada una observación para la muestra, puede volver a ser seleccionada.

Sin reemplazamiento: al tomar una observación, esa unidad queda excluida al tomar la observación siguiente.

Muestreo en poblaciones infinitas.

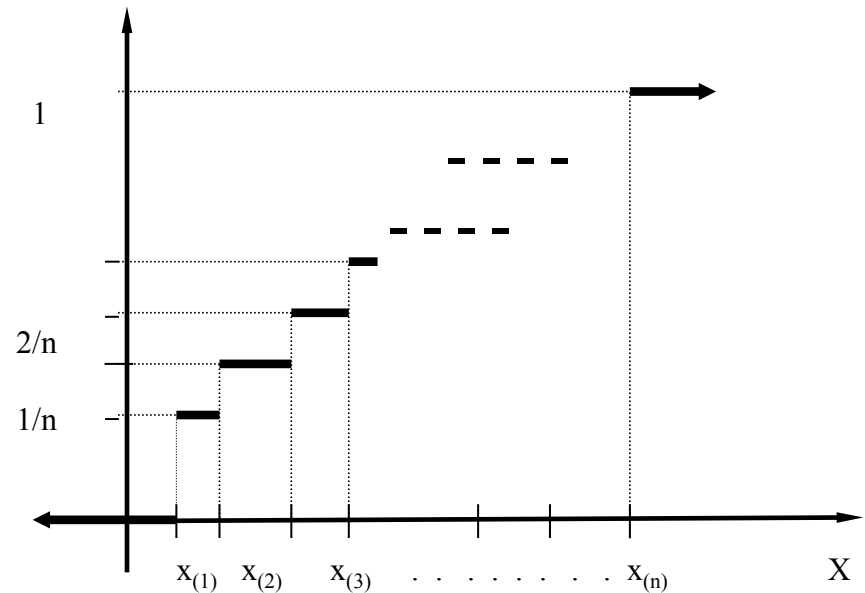
**Función de densidad** de  $\vec{X}$  (supuesto el muestreo aleatorio simple):

También llamada función de verosimilitud.

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \cdots f(x_n)$$

Función de distribución empírica de la muestra concreta:

$$F^*(x) = \begin{cases} 0 & si \quad x < x_{(1)} \\ 1/n & si \quad x_{(1)} \leq x < x_{(2)} \\ 2/n & si \quad x_{(2)} \leq x < x_{(3)} \\ \dots & \dots \dots\dots\dots\dots\dots \\ (n-1)/n & si \quad x_{(n-1)} \leq x < x_{(n)} \\ 1 & si \quad x_{(n)} \geq x \end{cases}$$

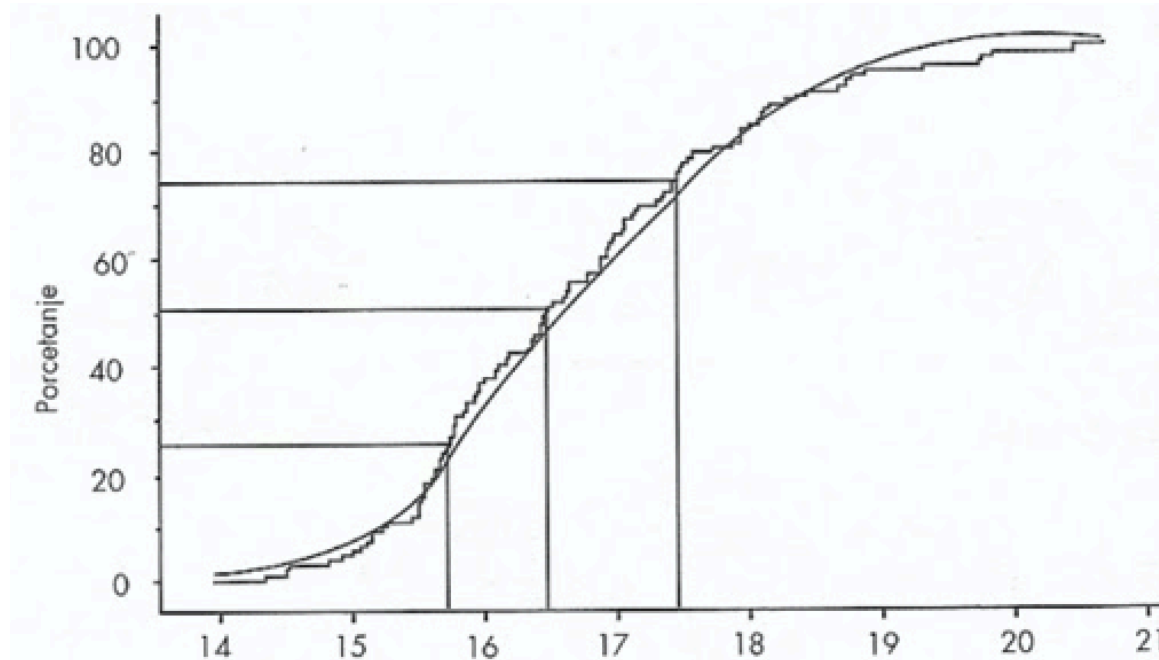


## Teorema Central de la Estadística o de Glivenko-Cantelli.

“La distribución empírica de la muestra converge en probabilidad a la distribución de la población cuando el tamaño muestral tiende a ser grande”.

Es decir que:  $F^*(x) \xrightarrow{P} F(x) \quad \forall x \in \mathcal{R}, n \rightarrow \infty$

$$P\left(\sup |F^*(x) - F(x)| \rightarrow 0\right) \rightarrow 1$$



## Estadístico

Función de la v.a.  $\vec{X}$   $t : (X_1, X_2, \dots, X_n) \rightarrow \mathcal{R}$   
 $(x_1, x_2, \dots, x_n) \rightarrow t(x_1, x_2, \dots, x_n)$

Para cada muestra concreta, el estadístico toma un valor real concreto.

Es una variable aleatoria cuya distribución (distribución muestral) depende de la distribución de la población.

Un **Estimador** es un estadístico que se usa para estimar parámetros poblacionales desconocidos.

Estimador  $\Rightarrow$  Estadístico  
 $\nLeftarrow$



## Tipos de Inferencia

Paramétrica: usadas principalmente para estimar parámetros poblacionales desconocidos. Asumen distribuciones estadísticas subyacentes a los datos.

Estimación: Por punto y por intervalo.

Contrastes de hipótesis.

No paramétrica: los datos no deben ajustarse a ninguna distribución. También pueden ser usadas para estimar parámetros poblacionales o incluso la forma de la distribución.

La ventaja de usar una prueba paramétrica en lugar de una no paramétrica consiste en que la primera tiene más **potencia** estadística que la segunda.

Además, las técnicas no paramétricas son mas complejas y exigen un volumen de cálculo mucho mayor.

# Introducción a la Inferencia Estadística

**ESCUELA POLITÉCNICA  
SUPERIOR DE CÓRDOBA**

Universidad de Córdoba

**DEPARTAMENTO DE ESTADÍSTICA**

