

# Descripción de datos con medidas numéricas

## Estadística

Gamaliel Moreno Chávez

Centro de Crecimiento Humanista

Enero  
2021

# Generalidades

Las gráficas pueden ayudar a describir la forma básica de una distribución de datos pero hay limitaciones

- No siempre es posible presentar gráficas.
- Son poco precisas.
- No son comparables

Una forma de superar estos problemas es usar medidas numéricas, que se pueden calcular para una muestra o una población de mediciones.

# Generalidades

## Aplicaciones

# Generalidades

## Aplicaciones

# Generalidades

## Aplicaciones

# Generalidades

## Aplicaciones

# Introducción

La Estadística resuelve problemas de

## Análisis de muestras

Se elige una muestra de una población para hacer inferencias respecto a esa población a partir de lo observado en la muestra (sondeos de opinión, control de calidad, etc).

# Introducción

La Estadística resuelve problemas de

## Descripción de datos

Procedimientos para resumir la información contenida en un conjunto (amplio) de datos.



# Introducción

La Estadística resuelve problemas de

## Contraste de hipótesis

Metodología estadística para diseñar experimentos que garanticen que las conclusiones que se extraigan sean válidas. Sirve para comparar las predicciones resultantes de las hipótesis con los datos observados (medicina eficaz, diferencias entre poblaciones, etc).

# Introducción

La Estadística resuelve problemas de

## Medición de relaciones

Medición de relaciones entre variables estadísticas (contenido de gas hidrógeno neutro en galaxias y la tasa de formación de estrellas, etc)

## Predicción

Prever la evolución de una variable estudiando su historia y/o relación con otras variables.

# Introducción

## Estadística Descriptiva

La cual es la que se utiliza en la descripción y análisis de conjuntos de datos o población.

## Inferencia Estadística

La cual hace posible la estimación de una característica de una población, o la toma de una decisión con respecto a una población, con base únicamente en resultados muestrales.

# Universo

Población o Universo: Conjunto completo de individuos, objetos, o medidas los cuales poseen una característica común observable y que serán considerados en un estudio.

Con información obtenida de una muestra representativa de esa población, tratamos de describir o pronosticar el comportamiento de la población.

# Parámetros

Caracteres cuantitativos o cualitativos. El objeto de nuestra medida pueden ser caracteres de tipos muy diversos. De ahí que normalmente se clasifiquen en:

- **Cuantitativos:** aquellos que toman valores numéricos
- **Cualitativos:** también llamados atributos, son aquellos que no podemos representar numéricamente y describen cualidades.

# Variable Aleatoria

Una **variable** es una característica que cambia o varía con el tiempo y/o para diferentes personas u objetos bajo consideración.

- **Variables Cualitativas:** son aquellas que pueden expresarse sólo en forma de atributo.
- **Variables Cuantitativas:** son aquellas variables que pueden expresarse en forma numérica. Se dividen en discretas y continuas.

# Variable Aleatoria

## Ejemplos de Variables cualitativas

# Variable Aleatoria

**Variables Cuantitativas Discretas**, son respuestas numéricas que surgen de un proceso de conteo, siendo siempre un número entero.

- 1) Número de asignaturas inscritas en el semestre.
- 2) Número de integrantes del grupo familiar.
- 3) Número de salas de clases.



# Variable Aleatoria

**Variables Cuantitativas Continuas**, son respuestas numéricas que surgen de un proceso de medición, las cuales pueden tomar valores entre dos números enteros.

- 1) Estatura
- 2) Temperatura
- 3) Peso

# Variable Aleatoria

# Gráficas para datos

Una vez recolectados los datos, éstos pueden consolidarse y resumirse para mostrar la siguiente información:

- La **frecuencia** o número de mediciones en cada categoría
- La **frecuencia relativa** o proporción de mediciones en cada categoría
- El **porcentaje** de mediciones en cada categoría

# Gráficas para datos

Por ejemplo, si con  $n$  representamos el número total de mediciones en el conjunto, se puede hallar la frecuencia relativa y porcentaje usando estas relaciones:

$$\text{Frecuencia relativa} = \frac{\text{Frecuencia}}{n} \quad (1)$$

$$\text{porcentaje} = 100 \times \text{Frecuencia relativa} \quad (2)$$

# Gráficas para datos

Ejercicio 1. En una encuesta respecto a la educación pública, a 400 administradores de escuelas se les pidió calificar la calidad de la educación. Sus respuestas están resumidas en la tabla. Construya una gráfica de pastel y una de barras a partir de este conjunto de datos.

# Gráficas para datos

Ejercicio 2. Una bolsa de tamaño botana de dulces de cacahuete M&M's contiene 21 dulces con los colores que se indican en la tabla. La variable "color" es cualitativa. Encuentre la frecuencia, frecuencia relativa y porcentaje. Realice una grafica de barras ordenadas de mayor a menor (gráfica de Pareto).

# Gráficas de líneas

Cuando una variable cuantitativa se registra en el tiempo a intervalos igualmente espaciados (por ejemplo diario, semanal, mensual, trimestral o anual), el conjunto de datos forma una serie de tiempo. Los datos de una serie de tiempo se presentan con más efectividad en una gráfica de líneas con el tiempo como eje horizontal.

# Gráficas de líneas



# Gráficas de puntos

La gráfica más sencilla para datos cuantitativos es la gráfica de puntos. Para un conjunto pequeño de mediciones.

# Gráficas de tallo y hoja

Otra forma sencilla de exhibir la distribución de un conjunto de datos cuantitativos es la gráfica de tallo y hoja. Esta gráfica presenta una exhibición gráfica de los datos usando los valores numéricos reales de cada punto de datos.

# Gráficas de tallo y hoja

Para crear el tallo y hoja, se puede dividir cada observación entre las unidades y las decenas. El número a la izquierda es el tallo; el de la derecha es la hoja.

# Gráficas de tallo y hoja

A veces las opciones de tallo disponibles resultan en una gráfica que contiene muy pocos tallos y un gran número de hojas dentro de cada tallo. En esta situación, se pueden prolongar los tallos al dividir cada uno en varias líneas, dependiendo de los valores de hojas que se les asignen.

# Gráficas de tallo y hoja

# Histograma

Un histograma de frecuencia relativa, para un conjunto de datos cuantitativo es una gráfica de barras en la que la altura de la barra muestra “con qué frecuencia” (medida como proporción o frecuencia relativa) las mediciones caen en una clase o subintervalo particular. Las clases o subintervalos se grafican a lo largo del eje horizontal.

# Histograma

# Histograma



# Histograma

# Ejercicios

Preescolar A continuación se dan las edades (en meses) a los que se inscribieron por primera vez 50 niños en una escuela preescolar.

# Histograma

- a. Construya una gráfica de tallo y hoja para los datos.
- b. Construya un histograma de frecuencia relativa para estos datos. Empiece en la frontera inferior de la primera clase en 30 y use un ancho de clase de 5 meses.
- c. Compare las gráficas de los incisos a) y b). ¿Hay alguna diferencia importante que le haría escoger una como el mejor método para exhibir los datos?

# Histograma

- d. ¿Qué proporción de los niños tenían 35 meses (2 años, 11 meses) o más, pero menos de 45 meses (3 años, 9 meses) de edad cuando se inscribieron por primera vez en preescolar?
- e. Si un niño fuera seleccionado al azar de este grupo de niños, ¿cuál es la probabilidad de que tuviera menos de 50 meses de edad (4 años, 2 meses) cuando se inscribió por primera vez en preescolar?