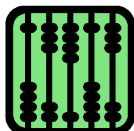




CURSO DE ESTADÍSTICA BÁSICA



ESQUEMA DEL CURSO



SESIÓN 1 DISEÑO DE ESTUDIOS EN INVESTIGACIÓN MÉDICA

- 1.1 Estudios observacionales
- 1.2 Estudios experimentales

SESIÓN 2 DESCRIPTIVA

- 2.1 Escalas de medida
- 2.2 Variables
- 2.3 Resumen de datos con números
- 2.4 Resumen de datos con gráficos

III

SESIÓN 3 ESTIMACIÓN

3.1 Concepto de estimación
3.2 Error estándar
3.3 Intervalo de confianza

IV

SESIÓN 4 INFERENCIA ESTADÍSTICA I

4.1 Contraste de hipótesis
4.2 Métodos Paramétricos
4.3 Transformaciones de datos
4.4 Métodos no Paramétricos

V

SESIÓN 5 INFERENCIA ESTADÍSTICA II

5.1 Tablas de contingencia
5.2 Contraste de hipótesis
5.3 Medidas de asociación

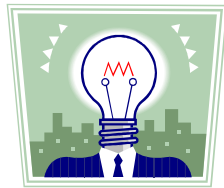
VI

SESIÓN 6 REGRESIÓN

6.1 Correlación
6.2 Regresión Lineal Simple

**“Si oigo algo lo olvido.
Si lo veo lo entiendo.
Si lo hago lo aprendo”.**

Confucio (551-478 A.C)



1.1 ESTUDIOS OBSERVACIONALES

DEFINICIÓN

Los estudios observacionales son aquellos en los que sólo se observa a los sujetos

Son de cuatro tipos:

1. Series de casos
2. Estudios de casos de control
3. Estudios de corte transversal
4. Estudios de cohorte

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

DEFINICIÓN

Simple recuento descriptivo de características observadas en un grupo de pacientes que han aparecido en un intervalo de tiempo

- En caso de nuevas enfermedades o fármacos contribuyen a la caracterización del perfil
- Incluyen pacientes vistos en un período breve
- No comprenden sujetos control, es decir, personas sin el problema que se describe
- Suelen ser precursores de otros estudios

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

VENTAJAS:

- Son fáciles de redactar
- Las observaciones son muy útiles

INCONVENIENTE:

- Son susceptibles a muchos sesgos relacionados con la selección del sujeto

1 2 3 4 5 6

Fundación para la Formación e Investigación Científica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

- Empiezan con ausencia o presencia de un resultado
- Los casos son individuos escogidos por alguna enfermedad
- Los controles son personas sin enfermedad
- Se analizan los antecedentes para detectar causas o posibles factores de riesgo

EJEMPLO → ¿INFLUYEN LAS ANTENAS DE TELEFONÍA MÓVIL EN LOS CASOS DE LEUCEMIA?

1 2 3 4 5 6

Fundación para la Formación e Investigación Científica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

VENTAJAS:

- Son útiles para estudiar eventos raros
- Permiten el estudio con tamaños muestrales relativamente pequeños

INCONVENIENTE:

- Selección de un grupo control adecuado

1 2 3 4 5 6

Fundación para la Formación e Investigación Científica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

- Analizan los datos obtenidos de un grupo de personas en un momento determinado más que en un período dado
- Determinan **¿qué está pasando?**

EJEMPLOS → ENCUESTAS ESTUDIOS ORIENTADOS AL CONSUMIDOR SOBRE SU SATISFACCIÓN

1 2 3 4 5 6

Fundación para la Formación e Investigación Científica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

VENTAJA:

- Son muy buenos para determinar la prevalencia de una enfermedad en una población determinada

EJEMPLO: prevalencia de VIH en una población determinada

INCONVENIENTE:

- Proporciona sólo una "imagen instantánea en el tiempo"

1 2 3 4 5 6 7 8 9 10 11 12

Fundación para la Formación e Investigación Biomédica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

DEFINICIÓN

Una cohorte es un grupo de personas que tiene algo en común y permanecen como parte de ese grupo por un largo período

Los estudios de cohorte hacen la pregunta *¿qué pasará?*

EJEMPLO → FUMAR Y CÁNCER DE PULMÓN

1 2 3 4 5 6 7 8 9 10 11 12

Fundación para la Formación e Investigación Biomédica de la Región de Murcia

1.1 ESTUDIOS OBSERVACIONALES

SERIES DE CASOS ESTUDIOS DE CASOS DE CONTROL ESTUDIOS DE CORTE TRANSVERSAL ESTUDIOS DE COHORTE

VENTAJA:

- Son el diseño de elección para estudiar las causas de un problema, la evolución de una enfermedad o sus factores de riesgo

INCONVENIENTE:

- Son vulnerables a problemas relacionados con el seguimiento de los pacientes

1 2 3 4 5 6 7 8 9 10 11 12

Fundación para la Formación e Investigación Biomédica de la Región de Murcia

1.2 ESTUDIOS EXPERIMENTALES

DEFINICIÓN

Los estudios experimentales son aquellos en los que se efectúa alguna intervención

Los estudios experimentales en los que participan humanos se llaman **ENSAYOS CLÍNICOS**, y se agrupan en dos categorías:

- Los que utilizan controles
- Los que no utilizan controles

2

Fundación para la Formación e Investigación Biomédica de la Región de Murcia

SESIÓN 2 DESCRIPTIVA

- Escalas de medida
- Variables
- Resumen de datos con números
- Resumen de datos con gráficos

2.1 ESCALAS DE MEDIDA

1. ESCALAS NOMINALES 2. ESCALAS ORDINALES 3. ESCALAS NUMÉRICAS

- Se usan para el nivel más simple de medición cuando los datos se agrupan en categorías.
- Los datos nominales se describen en términos de porcentajes y proporciones.
- No se miden, se cuentan sus observaciones.
- Se utilizan tablas de contingencia y gráficas de barras para mostrar este tipo de información.
- Pueden ser dicotómicas o multinomiales.

Ejemplos: Sexo, Tratamiento, etc.

2.1 ESCALAS DE MEDIDA

1. ESCALAS NOMINALES 2. ESCALAS ORDINALES 3. ESCALAS NUMÉRICAS

- Las observaciones se clasifican en categorías como las nominales pero algunas observaciones son más ó mayores que otras.
- Al igual que con las escalas nominales se usan porcentajes y proporciones. A veces los datos medidos en una escala ordinal se resumen en un valor medio.
- Los mismos cuadros y gráficas que se usan para escalas nominales pueden usarse para escalas ordinales.

Ejemplos: Cantidad de riesgo de un paciente, grado de desarrollo de un tumor, etc.

2.1 ESCALAS DE MEDIDA

1. ESCALAS NOMINALES 2. ESCALAS ORDINALES 3. ESCALAS NUMÉRICAS

- A las observaciones que miden cantidades se les llama observaciones cuantitativas.
- Existen dos tipos de escala numérica: continuas y discreta.

Ejemplos: peso, talla, tiempo de supervivencia, etc.

2.2 VARIABLES

DEFINICIÓN

Característica observable y medible que varía entre los diferentes individuos de una población

Características de los individuos que se pueden resumir en variables:

- Edad
- Sexo
- Nivel de glucosa en sangre
- Religión
- Número de hijos
- Grado de satisfacción
- Etc...

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

Si sus valores no se pueden asociar naturalmente a un número (no se pueden hacer operaciones con ellos)

Pueden ser de dos tipos:

- Nominales
- Ordinales

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

NOMINALES

Si sus valores no se pueden ordenar

EJEMPLOS

SEXO

GRUPO SANGUÍNEO

RELIGIÓN

Pueden ser a su vez de dos tipos:

- Dicotómicas
 - Sano/Enfermo
 - Expuesto/No expuesto
- Policotómicas
 - Grupo Sanguíneo (A/B/0/AB)
 - Tratamiento

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

ORDINALES

Si sus valores se pueden ordenar

EJEMPLOS

MEJORA DE UN TRATAMIENTO

GRADO DE SATISFACCIÓN

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

Si sus valores son numéricos (se pueden hacer operaciones con ellos)

Pueden ser de dos tipos:

- Discretas
- Continuas

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

DISCRETAS

Si toma valores enteros, aislados



Permiten operaciones de igualdad-desigualdad y orden, pero también operaciones algebraicas

2.2 VARIABLES

TIPOS DE VARIABLES

1. CUALITATIVAS

2. CUANTITATIVAS

CONTÍNUAS

Toman valores numéricos y entre dos de ellos hay infinitos valores intermedios



Permiten todo tipo de operaciones algebraicas

2.2 VARIABLES

CATEGORIZACIÓN DE UNA VARIABLE CUANTITATIVA

Puede realizarse una transformación de una variable cuantitativa y pasarla a una escala ordinal. A esto se le llama **CATEGORIZACIÓN**.



¿EN QUÉ CONSISTE LA CATEGORIZACIÓN?

En pasar a considerar una variable que inicialmente poseía características de cuantitativa (porque estaba expresado en números), como cualitativa ordinal, de manera que los valores se reagrupan formando unos pocos grupos

2.2 VARIABLES

CODIFICACIÓN DE LAS VARIABLES

Para facilitar sus operaciones con el ordenador, las variables se deben **CODIFICAR** como números

Los posibles valores de una variable suelen denominarse **MODALIDADES**. Las modalidades se agrupan en intervalos llamados "CLASES"

Para estar bien diseñadas las clases deben ser...

- Exhaustivas: no olvidar ningún posible valor de la variable
- Excluyente: nadie puede presentar dos posibles valores de la misma variable

¡MAL!
No incluye a los que tienen 3 hijos

¡MAL!
Los que tienen 4 hijos están en ambas categorías

2.3 RESUMEN DE DATOS CON NÚMEROS

▪ Cuando un investigador obtiene muchas observaciones, los números que resumen los datos pueden proporcionar mucha información

▪ Las medidas más utilizadas son:

1. Medidas de tendencia central
2. Medidas de dispersión

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

MEDIA

- Poca variabilidad de una muestra a otra
- No debe emplearse con datos ordinales
- Es sensible a valores extremos

9,06

MEDIANA

- Es el punto a partir del cual la mitad de los datos son menores y la otra mitad mayores
- Es menos sensible a valores extremos
- Puede usarse con observaciones ordinales

9

MODA

- Es el valor que se presenta con mayor frecuencia
- Útil con grandes cantidades de datos
- Cuando un grupo de datos tiene dos modas se llama bimodal

8

Ejemplo: 6, 6, 7, 8, 8, 8, 8, 9, 9, 9, 10, 11, 11, 12, 14

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

¿CÓMO SE CALCULAN?

MEDIA

Se calcula sumando todas las observaciones y dividiendo entre el número total

FORMULA

$$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$$

MEDIANA

Se disponen los datos de menor a mayor y se cuentan hasta encontrar el valor medio

- Para un número impar de observaciones es el valor medio
- Para un número par es la media de los dos valores medios

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

Media, Mediana, Moda... ¿CÚAL ES LA MÁS ADECUADA PARA UN GRUPO PARTICULAR DE DATOS?

Debe tenerse en cuenta:

1. La escala de medición: ordinal o numérica
2. La forma de distribución de los resultados: si se carga a la derecha o a la izquierda de la media, si son o no simétricos, etc.

GUÍA

MEDIA: para datos numéricos y distribuciones simétricas

MEDIANA: para datos ordinales o numéricos si la distribución está cargada (es decir, cuando los resultados extremos se inclinan en una sola dirección)

MODA: para datos cualitativos o distribuciones bimodales

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL 2. MEDIDAS DE DISPERSIÓN 3. TABLAS DE FRECUENCIAS

PERCENTILES, DECILES Y CUARTILES

La mediana divide a la muestra en 2 partes iguales. Como generalización de ella se definen los **percentiles o cuantiles**

Aquellos que dividen la muestra en 100 partes iguales

- Se denotan por p_i
- En general, el percentil i es aquel valor que deja a su izquierda el $i\%$ de los valores de la muestra ordenada
- Los percentiles son muy utilizados para describir los casos "raros" de las poblaciones

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL 2. MEDIDAS DE DISPERSIÓN 3. TABLAS DE FRECUENCIAS

PERCENTILES, DECILES Y CUARTILES

EJEMPLO

El percentil 10 del peso de los niños varones recién nacidos es 2700 gr

Sólo un 10% de ellos tienen un peso inferior a 2700 gr

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL 2. MEDIDAS DE DISPERSIÓN 3. TABLAS DE FRECUENCIAS

PERCENTILES, DECILES Y CUARTILES

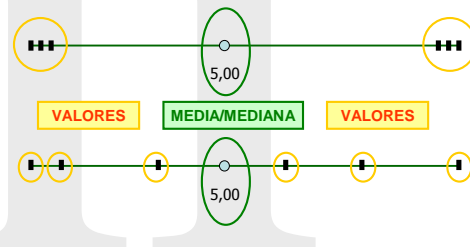
Algunos percentiles reciben nombres especiales

p_{25}	→	Primer cuartil	} Dejan a su izquierda la cuarta parte, la mitad y las tres cuartas partes respectivamente de la muestra ordenada. Es decir, dividen la muestra en cuatro partes iguales.
p_{50}	→	Segundo cuartil	
p_{75}	→	Tercer cuartil	

De igual modo a los percentiles $p_{10}, p_{20}, \dots, p_{90}$ se les llama **deciles**

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL 2. MEDIDAS DE DISPERSIÓN 3. TABLAS DE FRECUENCIAS



2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

Las medidas de tendencia, si bien dan información acerca de cuáles son los valores de la muestra, no indican nada sobre cómo de diferentes son los mismos

Las medidas que indican cómo de agrupados ó dispersos están los datos se conocen con el nombre de **MEDIDAS DE DISPERSIÓN**

Toda medida de dispersión toma el valor cero cuando los datos no presentan variabilidad, es decir, todos los datos de la muestra son iguales (p. ej.: 2,2,2,2)
En caso contrario toman un valor positivo, tanto mayor cuanto más grande sea la dispersión de los datos

Las medidas de dispersión más útiles son **el recorrido, la desviación típica y la varianza**

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

RECORRIDO

- Es la más simple y de mayor valor intuitivo
- Se calcula como la diferencia entre el valor más grande y el más pequeño de la muestra

$$R = X_{\max} - X_{\min}$$



VENTAJAS:

- Se expresa en las mismas unidades de la muestra
- Es fácil de calcular



INCONVENIENTES:

- Utiliza sólo dos elementos de la muestra
- Aumenta con el tamaño de la muestra

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

VARIANZA Y DESVIACIÓN TÍPICA

Una medida de dispersión adecuada utiliza la distancia entre cada dato con respecto a la media

Para evitar valores negativos se eleva esta diferencia al cuadrado

VARIANZA

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

Para que tenga las mismas unidades de la muestra se calcula su raíz cuadrada

DESVIACIÓN TÍPICA

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

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

En las revistas científicas a menudo se presenta información de las distribuciones de frecuencias

DEFINICIONES

Frecuencia absoluta

Número de individuos que presentan una modalidad

Frecuencia relativa

Proporción de individuos referidos al total que presentan una modalidad

Frecuencia acumulada

Número de observaciones menores o iguales que la modalidad que estamos estudiando

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

EJEMPLO

Número de visitas al especialista

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
1	19	8,8	8,8	8,8
2	3	1,4	1,4	10,1
3	195	89,9	89,9	100,0
TOTAL	217	100,0	100,0	

V
A
L
I
D
O
S

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

¿CÓMO CONSTRUIR UNA TABLA DE FRECUENCIAS?

Se dividen las observaciones en clases y luego se cuentan los datos de cada clase

PASOS

1. Identificar los resultados mayor y menor
2. Restar el resultado menor del mayor para obtener la variación de los datos
3. Determinar el número de clases (entre 6 y 14)
4. Hacer un recuento del número de datos de cada clase

CONSEJO

Para obtener la amplitud de las clases, dividir la variación de los datos entre el número de clases

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

EJEMPLO

Hacer una tabla de frecuencia con los siguientes datos correspondientes a la altura de una serie de pacientes de fisioterapia:

N=33

160, 168, 175, 183, 170, 164, 170, 184, 171, 168, 187, 161, 183, 175, 185, 186, 187, 164, 165, 175, 162, 192, 169, 163, 166, 172, 173, 167, 174, 176, 178, 179, 177.

2.3 RESUMEN DE DATOS CON NÚMEROS

1. MEDIDAS DE TENDENCIA CENTRAL

2. MEDIDAS DE DISPERSIÓN

3. TABLAS DE FRECUENCIAS

EJEMPLO

160, 168, 175, 183, 170, 164, 170, 184, 171, 168, 187, 161, 183, 175, 185, 186, 187, 164, 165, 175, 162, 192, 169, 163, 166, 172, 173, 167, 174, 176, 178, 179, 177.

PASOS

1. Identificar los resultados mayor y menor
Valor máximo: 192 Valor mínimo: 160
2. Restar el resultado menor del mayor para obtener la variación de los datos
 $V_{\text{máx}} - V_{\text{mín}} = 32$
3. Determinar el número de clases (entre 6 y 14)
Como hay pocos datos elegimos pocas clases $\rightarrow 6$
4. Hacer un recuento del número de datos de cada clase
Con este recuento obtenemos la tabla de frecuencias

2.4 RESUMEN DE DATOS CON GRÁFICOS

- Los gráficos se utilizan mucho en Medicina
- Sirven para **complementar** un análisis, no para **realizarlo**
- Los más útiles son: los gráficos de sectores, los histogramas, los gráficos de barras y los polígonos de frecuencias

2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES

GRÁFICOS DE BARRAS

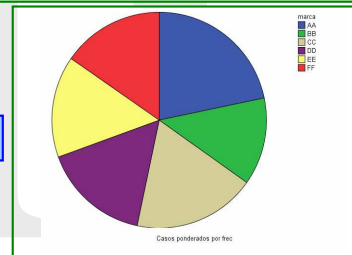
HISTOGRAMAS

POLÍGONOS DE FRECUENCIAS

Es un gráfico muy sencillo. Se consigue la información que proporcionan los datos con un simple golpe de vista

EJEMPLO

¡BIEN APLICADO!



2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES

GRÁFICOS DE BARRAS

HISTOGRAMAS

POLÍGONOS DE FRECUENCIAS

Es un gráfico muy sencillo. Se consigue la información que proporcionan los datos con un simple golpe de vista

¡MAL APLICADO!

Se pierde el orden de las categorías



2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES

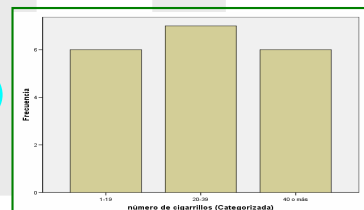
GRÁFICOS DE BARRAS

HISTOGRAMAS

POLÍGONOS DE FRECUENCIAS

Presentan las categorías en el eje X y el nº o porcentaje de observaciones en el eje Y. Las bases de las columnas deben de ser iguales y la altura representa la frecuencia. Se deben ordenar las representaciones. Pensado para variables cualitativas.

EJEMPLO

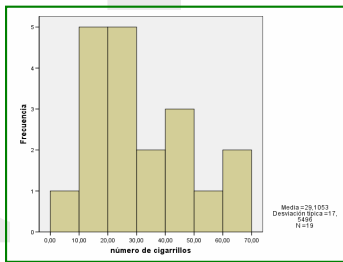


2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES GRÁFICOS DE BARRAS **HISTOGRAMAS** POLÍGONOS DE FRECUENCIAS

Están pensados para variables cuantitativas que siguen una escala numérica

EJEMPLO

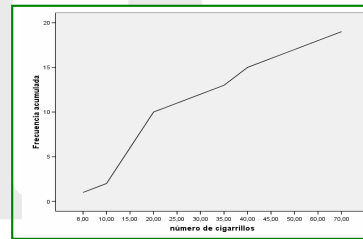


2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES GRÁFICOS DE BARRAS HISTOGRAMAS **POLÍGONOS DE FRECUENCIAS**

Gráficas lineales similares a los histogramas, muy útiles para comparar dos distribuciones en la misma gráfica. Un caso particular es el de frecuencias acumuladas.

EJEMPLO



2.4 RESUMEN DE DATOS CON GRÁFICOS

GRÁFICOS DE SECTORES GRÁFICOS DE BARRAS HISTOGRAMAS **POLÍGONOS DE FRECUENCIAS**

Gráficas lineales similares a los histogramas, muy útiles para comparar dos distribuciones en la misma gráfica. Un caso particular es el de frecuencias acumuladas.

¡MAL APLICADO!

