Descripción y Exploración de Datos en Psicología

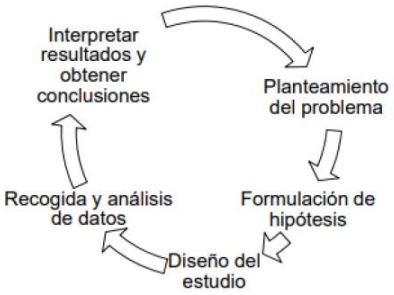
Grado en Psicología UGR

Apuntes

1. INTRODUCCIÓN

ROL DE LA ESTADÍSTICA EN PSICOLOGÍA

Estadística:



En **psicología**, se aplica a la Psicología clínica, al contexto escolar, al contexto servicio social, a departamentos de marketing online...

CONCEPTOS FUNDAMENTALES

Las **variables** son propiedades o atributos del objeto de estudio, que pueden asumir más de un valor.

Tipos de variables:

Interpersonales:

Intrapersonales:

Independientes:

Dependientes:

Extrañas o contaminadoras:

Cualitativas:

Cuasi-cuantitativas u ordinales:

Cuantitativas:

- **Discretas:**
- **Continuas:**

Dicotómica:

Politómica:

Las categorías de una variables deben ser **exhaustivas** y mutuamente **excluyentes**.

TIPOS DE ESTADÍSTICA

La **estadística descriptiva** se encarga de resumir y describir un conjunto de datos para su comprensión.

La **inferencia estadística** se encarga de extraer conclusiones sobre una población a partir del estudio de una muestra mediante técnicas probabilísticas.

POBLACIÓN Y MUESTRA

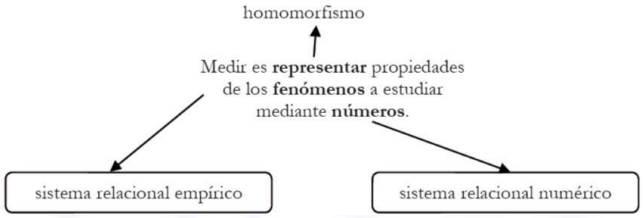
Una **población** es el conjunto de todos los elementos que cumplen una o varias características.

Un **parámetro** es una propiedad descriptiva de la población que generalmente es desconocida.

Una **muestra** es un subconjunto de los elementos de una población.

Un **estadístico** una propiedad descriptiva de la muestra.

QUÉ ES MEDIR



ESCALAS DE MEDIDA

Nominal:

Ordinal:

Intervalo:

Razón:

2. ORGANIZACIÓN Y REPRESENTACIÓN GRÁFICA DE LOS DATOS

DISTRIBUCIONES DE FRECUENCIAS

Frecuencia absoluta (n_i): número de veces que se repite un determinado valor. Puede ser individual o acumulada (N_i) (en les variables nominales no tiene sentido acumular).

Frecuencia relativa (f_i): proporción que representa las apariciones de ese valor respecto al total $f_i = \frac{n_i}{n}$

También puede ser individual o acumulada (F_i). Si se multiplica por 100 nos dará un porcentaje.

Amplitud del intervalo: $a_i = L_i - L_{i-1}$

Marca de clase: $c_i = \frac{L_i + L_{i-1}}{2}$

Altura del intervalo: $h_i = \frac{n_i}{a_i}$

Ejemplo numérico discreto, número de hijos por familia: 2, 1, 1, 0, 1, 2, 2, 0, 1, 1.

x_i	n_i	N_i	f_i	F_i	%	%a
0						
1						
2						

Ejemplo: numérico continuo, notas de un examen: 3,4; 5,3; 6,8; 9,1; 7,2; 8,3; 1,9; 5,1; 4,4; 3,2.

x_i	n_i	N_i	f_i	F_i	a_i	c_i	h_i
[0-2)							
[2-4)							
[4-6)							
[6-8)							
[8-10)							

REPRESENTACIÓN GRÀFICA

Diagrama de sectores

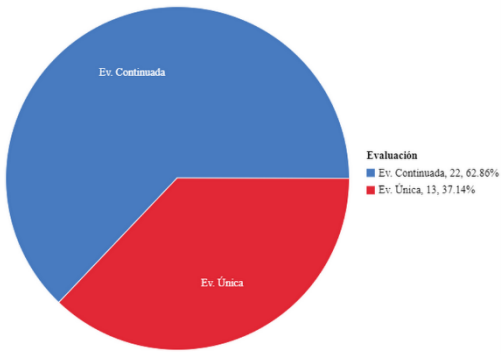


Diagrama de barras

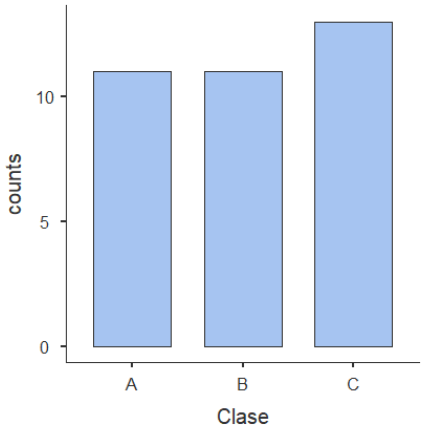
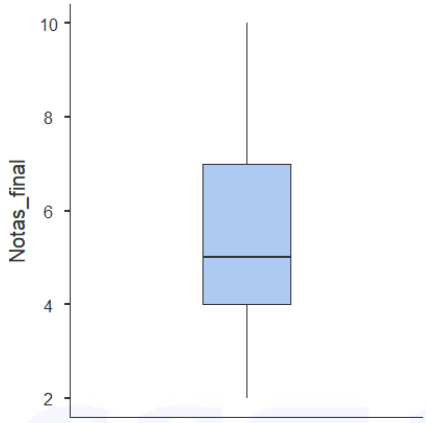
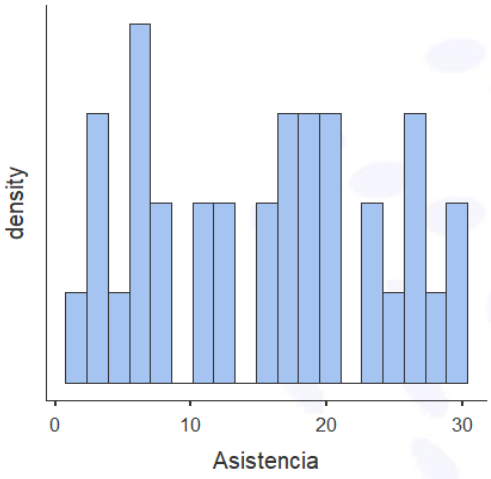


Diagrama de la caja / Boxplot



Histograma



Densidad

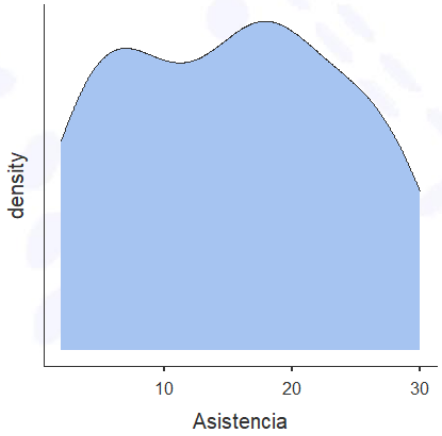


Diagrama de tallo y hojas

```
1 | 9999999999
2 | 00000000001111112334457 9
3 | 5
```

3. EXPLORACIÓN Y DESCRIPCIÓN UNIVARIADA

MEDIDAS DE POSICIÓN CENTRAL

Media aritmética: es el centro de gravedad de los datos.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{\sum x_i \cdot n_i}{n}$$

Mediana: es el valor de la distribución, tal que una vez ordenados todos los valores observados de menor a mayor (en sentido creciente), ocupa el lugar central.

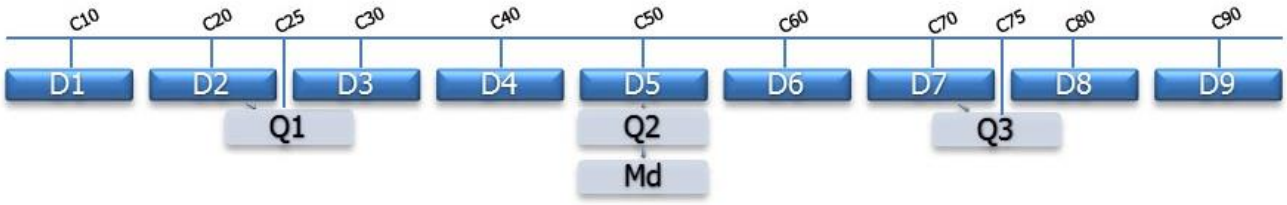
$$Pos_{Md} = \frac{n+1}{2} = 0{,}5 \cdot (n+1)$$

Moda: es el valor que más se repite.

MEDIDAS DE POSICIÓN NO CENTRAL

Cuantiles

$$Pos_{IP} = \frac{k}{100} \cdot (n+1)$$



	Cualitativo nominal	Cuasi-cuantitativo	Cuantitativo
Moda	Si	Si	Si
Media			Si
Mediana y Cuantiles		Si	Si

Ejemplo: calcula los indicadores anteriores con el ejemplo del número de hijos 2, 1, 1, 0, 1, 2, 2, 0, 1, 1

MEDIDAS DE DISPERSIÓN

Amplitud total: es la diferencia entre el máximo y el mínimo.

$$A_t = máx - mín$$

Amplitud intercuartil: es la diferencia entre el tercer cuartil y el primero.

$$A_{iq} = Q_3 - Q_1$$

Se considera un dato extremo si se supera en 1,5 veces la A_{iq} por encima del Q3 o por debajo del Q1

Amplitud semi-intercuartil: es la amplitud promedio entre Q1 y Q2 y entre Q2 y Q3.

$$ASI = \frac{Q_3 - Q_1}{2}$$

Desviación media: es la media de la diferencia de los valores a la propia media.

$$DM = \frac{\sum|x_i - \bar{x}|}{n} = \frac{\sum|x_i - \bar{x}| \cdot n_i}{n}$$

Varianza: es la media aritmética de los cuadrados de las diferencias de los valores de la variable con respecto a su media aritmética y coincide con el momento de segundo orden respecto a la media.

$$S_x{}^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{\sum(x_i - \bar{x})^2 \cdot n_i}{n} \qquad = \frac{\sum x_i{}^2}{n} - \bar{x}^2 = \frac{\sum x_i{}^2 \cdot n_i}{n} - \bar{x}^2$$

La desviación estándar/típica: es la raíz de la varianza.

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

Coeficiente de variación de Pearson: es el más empleado de los índices de dispersión relativa y se define como el cociente entre la desviación típica y la media aritmética. A mayor coeficiente de variación, menor representatividad de la media.

$$CV = \frac{S_x}{\bar{x}}$$

	Cualitativo nominal	Cuasi-cuantitativo	Cuantitativo
Amplitud			Si
AIQ y ASIQ		Si	Si
Desviación típica			Si
CV			SI

Ejemplo: calcula los indicadores anteriores con el ejemplo del número de hijos 2, 1, 1, 0, 1, 2, 2, 0, 1, 1

MEDIDAS DE FORMA

Asimetría:

As>0 asimetría positiva
As=0 simetría
As<0 asimetría negativa

Curtosis:

C>0 leptocúrtica
C=0 mesocúrtica
C<0 platicúrtica

Puntuaciones típicas y diferenciales:

$$d_i = x_i - \bar{x} \qquad z = \frac{x_i - \bar{x}}{S}$$

Ejemplo: un alumno del G1 ha sacado un 7 en Estadística, y otro del G2 ha sacado un 6. Sabiendo que la nota media del G1 ha sido de 6,5 con una desviación de 2 y la nota media del G2 ha sido 5 con una desviación de 3. ¿Quién tiene una mejor posición relativa?

Ejemplo:

Descriptives

	Notas_final
N	35
Mean	5.54
Median	5
Mode	4.00
Variance	6.26
Range	8
Minimum	2
Maximum	10
Skewness	0.233
Std. error skewness	0.398
Kurtosis	−1.04
Std. error kurtosis	0.778
25th percentile	4.00
50th percentile	5.00
75th percentile	7.00

Este dosier está hecho para seguir la clase de prueba.

Si te apuntas al curso te enviaremos por correo el dosier entero con todos los temas que faltan, ejercicios y exámenes de años anteriores

Más información en:

www.estadistix.com

Y si tienes cualquier consulta,
escríbenos un whatsapp al 644310902