

TEMA 1. CONCEPTOS BÁSICOS Y ORGANIZACIÓN DE DATOS

ÍNDICE

- 1.1. La investigación en Psicología
- 1.2. Concepto y funciones de la estadística: Descripción e inferencia
- 1.3. Variables: medición y clasificación
- 1.4. Descripción de variables: distribución de frecuencias y representación gráfica
- 1.5. Tendencia central, variabilidad y forma de una variable: aproximación gráfica

TEMA 1. CONCEPTOS BÁSICOS Y ORGANIZACIÓN DE DATOS

1.1. La investigación en Psicología

En el siglo XVII apareció el método científico como fuente de conocimiento más utilizada.

Se caracteriza por ser **sistemático** y **replicable**.

- **SISTEMÁTICO**: Hace referencia a que se trata de un proceso que tiene unos pasos perfectamente establecidos.

- **REPLICABLE**: Hace referencia a que los datos obtenidos mediante su uso pueden ser replicados o refutados por cualquier investigador que lo desee.

Por lo tanto, el **método científico** nos proporciona una manera de afrontar una investigación siguiendo los siguientes pasos:

1. **PLANTEAMIENTO DEL PROBLEMA**: Cuestión sin responder que surge del conocimiento previo.
2. **FORMULACIÓN DE HIPÓTESIS**: Solución tentativa del problema de investigación.
3. **PROCEDIMIENTO PARA LA RECOGIDA DE DATOS**: Muestra, instrumentos, materiales, recogida de datos.
4. **ANÁLISIS DE DATOS**
5. **DISCUSIÓN DE LOS RESULTADOS**
6. **ELABORACIÓN DEL INFORME DE INVESTIGACIÓN**

1.2. Concepto y funciones de la estadística: descripción e inferencia

La estadística es la rama de las matemáticas que se encarga del estudio de determinadas características en una población, recogiendo datos, agrupándolos, organizándolos en tablas, representándolos gráficamente y analizándolos para sacar conclusiones de dicha población. Con esta definición, podemos establecer dos grandes áreas dentro de la estadística:

- **Estadística descriptiva:** Es la parte que se encarga de recoger, almacenar, ordenar, realizar tablas o gráficos y calcular parámetros básicos sobre el conjunto de datos.
- **Estadística inferencial:** Es la encargada de hacer deducciones, es decir, inferir propiedades, conclusiones y tendencias, a partir de una muestra del conjunto.

Hay una serie de conceptos que debemos tener claros antes de continuar:

- **POBLACIÓN:** conjunto de todos los elementos que cumplen una determinada característica objeto de estudio.
- **MUESTRA:** es un subconjunto cualquiera de una población.
- **PARÁMETRO:** es una propiedad descriptiva (una medida) de una población. Se denota con letras griegas.
- **ESTADÍSTICO:** es una propiedad descriptiva (una medida) de una muestra. Se denota con letras latinas.

1.3. Variables: medición y clasificación

Una variable es un conjunto de valores resultantes de medir una característica de interés sobre cada elemento individual de una población o muestra. Se representan con letras latinas en mayúsculas.

La medición es el proceso por el cual se asignan números a objetos o sucesos según determinadas reglas.

En Psicología se utilizan diferentes escalas de medidas que se definen a continuación:

ESCALA NOMINAL:

Consiste en la asignación arbitraria de números o símbolos a cada uno de los valores de la variable. La única relación que se tiene en cuenta es la de **igualdad o desigualdad**. Los valores de la variable se denominan categorías. A las variables que presentan un nivel de medida nominal se les denomina **variables cualitativas o categóricas** y pueden ser **dicotómicas** (dos categorías) o **politómicas** (varias categorías).

ESCALA ORDINAL:

Se asignan números a objetos para indicar la extensión relativa en que se posee una característica. Cuando se asignan números es sólo para indicar el orden de las posiciones de lo que se está clasificando.

Esta escala permite establecer relaciones del tipo “mayor que” o “menor que”. Las variables ordinales también se llaman **cuasicuantitativas**.

ESCALA DE INTERVALO:

Aquellas que ordenan los objetos según la magnitud del atributo que representan y proveen intervalos iguales entre las unidades de medida. La distancia entre los distintos valores consecutivos de la variable es la misma. Se caracteriza por tener una unidad de medición común y constante. El origen es arbitrario y no refleja en ningún momento ausencia de la magnitud que estamos midiendo.

ESCALA DE RAZÓN:

Los números asignados a los objetos admiten como válidas las relaciones de igualdad-desigualdad, orden, suma, resta, multiplicación y división.

Las variables medidas en escala de intervalo y de razón son **variables cuantitativas** y éstas, a su vez, pueden ser **continuas o discretas**. Una variable continua es aquella para la que, dados dos valores siempre se puede encontrar un tercer valor que esté incluido entre los dos anteriores. Una variable discreta es aquella que adopta valores aislados.

TABLA RESUMEN

| Tipo de variable | Escala de medida | Características básicas | Relaciones válidas | Ejemplos |
|---|------------------|---|---|--|
| CUALITATIVA: - Dicotómica - Politémica | Nominal | Los números identifican y clasifican objetos | Relaciones del tipo "igual que" o "distinto que" | Sexo, estado civil, raza, diagnóstico clínico |
| CUASICUANTITATIVA | Ordinal | Además, los números indican las posiciones relativas de los objetos | Además, relaciones del tipo "mayor que" o "menor que" | Dureza, posición en el ranking ATP, grado de satisfacción. |
| CUANTITATIVA: - Discreta - Continua | Intervalo | Además, hay una unidad de medición común | Además, igualdad o desigualdad de diferencias | Temperatura en grados centígrados, inteligencia |
| | Razón | Además, el punto cero es absoluto | Además, igualdad o desigualdad de razones | Longitud, peso, altura, tiempo de reacción |

1.4. Descripción de variable: distribución de frecuencias y representación gráfica

Una vez que se tenga la base de datos se utilizará para extraer la información relevante. Si tenemos pocos datos es posible que la simple inspección visual de los mismos sea suficiente, pero esto no es frecuente, y por ello se debe organizar la información a través de una distribución de frecuencias.

Para ello debemos tener claro los siguientes conceptos:

- **Frecuencia absoluta (n_i):** Es el número de veces que se repite cada uno de los valores de las variables. La suma de todas las frecuencias absolutas es el total de la muestra (n).
- **Proporción o frecuencia relativa (p_i):** Es el cociente entre la frecuencia absoluta de cada valor y el número de observaciones. $p_i = n_i/n$
- **Porcentaje (P_i):** Es el valor de la frecuencia relativa multiplicado por 100.
- **Frecuencia absoluta acumulada (n_a):** Es el número de veces que se repiten cada uno de los valores o cualquiera de los valores inferiores.
- **Proporción acumulada o frecuencia relativa acumulada (p_a):** Es el cociente entre la frecuencia absoluta acumulada y el total de observaciones. $p_a = n_a/n$
- **Porcentaje acumulado (P_a):** Es el valor de la frecuencia relativa acumulada multiplicado por 100.

Si la variable con la que estamos trabajando se puede ordenar, podremos calcular todos los conceptos anteriores, en cambio, si es una variable que no se puede ordenar, solamente se calcularán los tres primeros.

Veamos algunos ejemplos.

VARIABLE DE CARÁCTER CUALITATIVO

En esta tabla tenemos una distribución de frecuencias de carácter cualitativo, en este caso la zona de procedencia de inmigrantes. En ella se resumen de manera organizada los datos correspondientes a 150 sujetos entrevistados.

| X | n_i | p_i | P_i |
|-----------------|-------|-------|-------|
| Sudamérica | 24 | 0,16 | 16 |
| Norte África | 60 | 0,40 | 40 |
| Centro África | 12 | 0,08 | 8 |
| Europa del este | 18 | 0,12 | 12 |
| China | 21 | 0,14 | 14 |
| Sur Asia | 15 | 0,10 | 10 |
| | n=150 | 1 | 100 |

VARIABLE ORDINAL

En esta tabla nos encontramos con una distribución de frecuencias de una variable ordinal, sobre la preocupación por las cuestiones políticas. En este añadimos tres columnas más, las correspondientes a los valores acumulados.

| X | n_i | p_i | P_i | n_a | p_a | P_a |
|----------|-------|-------|-------|-------|-------|-------|
| Máxima | 12 | 0,10 | 10 | 12 | 0,10 | 10 |
| Moderada | 48 | 0,40 | 40 | 60 | 0,50 | 50 |
| Mínima | 42 | 0,35 | 35 | 102 | 0,85 | 85 |
| Nula | 18 | 0,15 | 15 | 120 | 1,00 | 100 |
| | n=120 | 1 | 100 | | | |

VARIABLE CUANTITATIVA CON POCOS DATOS

En este caso tenemos una distribución de frecuencias de una variable cuantitativa con pocos datos, nos da información sobre la calificación obtenida en el examen de Introducción al Análisis de datos de una muestra de 40 alumnos.

| X | n_i | p_i | P_i | n_a | p_a | P_a |
|----|-------|-------|-------|-------|-------|-------|
| 1 | 1 | 0,025 | 2,5 | 1 | 0,025 | 2,5 |
| 2 | 1 | 0,025 | 2,5 | 2 | 0,050 | 5,0 |
| 3 | 1 | 0,025 | 2,5 | 3 | 0,075 | 7,5 |
| 4 | 8 | 0,200 | 20 | 11 | 0,275 | 27,5 |
| 5 | 5 | 0,125 | 12,5 | 16 | 0,400 | 40,0 |
| 6 | 7 | 0,175 | 17,5 | 23 | 0,575 | 57,5 |
| 7 | 8 | 0,200 | 20 | 31 | 0,775 | 77,5 |
| 8 | 7 | 0,175 | 17,5 | 38 | 0,950 | 95,0 |
| 9 | 1 | 0,025 | 2,5 | 39 | 0,975 | 97,5 |
| 10 | 1 | 0,025 | 2,5 | 40 | 1,000 | 100 |
| | n=40 | 1 | 100 | | | |

VARIABLE CUANTITATIVA CON GRAN CANTIDAD DE DATOS

Dentro de las variables cuantitativas podemos encontrarnos con algún caso en el que el número de valores que tome la variable sea muy amplio y será necesario agrupar las variables en intervalos, grupos de valores consecutivos de la variable.

Primero se debe decidir el número de intervalos que tendrá la distribución de frecuencias. Pueden ser pocos, pero muy amplios o podemos dividirlo en más intervalos con una amplitud más pequeña. En cualquier caso, se perderá información, por lo tanto queda a la elección del investigador.

Los intervalos tienen un valor inferior y un valor superior. Estos son los llamados **límites aparentes**. Tienen la misma unidad de medida que los valores de la variable.

El problema que tenemos con estos límites es que existe una pequeña discontinuidad, ya que el límite superior de un intervalo no coincide con el inferior del siguiente. Por ello, calcularemos los **límites exactos**.

Además de estos límites, también debemos calcular el **punto medio** del intervalo, que lo llamaremos **marca de clase**. Y con ello se harán los posteriores cálculos.

Y por último, otro de los conceptos a tener en cuenta es la **amplitud del intervalo**, que es la diferencia entre el límite superior exacto y el límite inferior exacto.

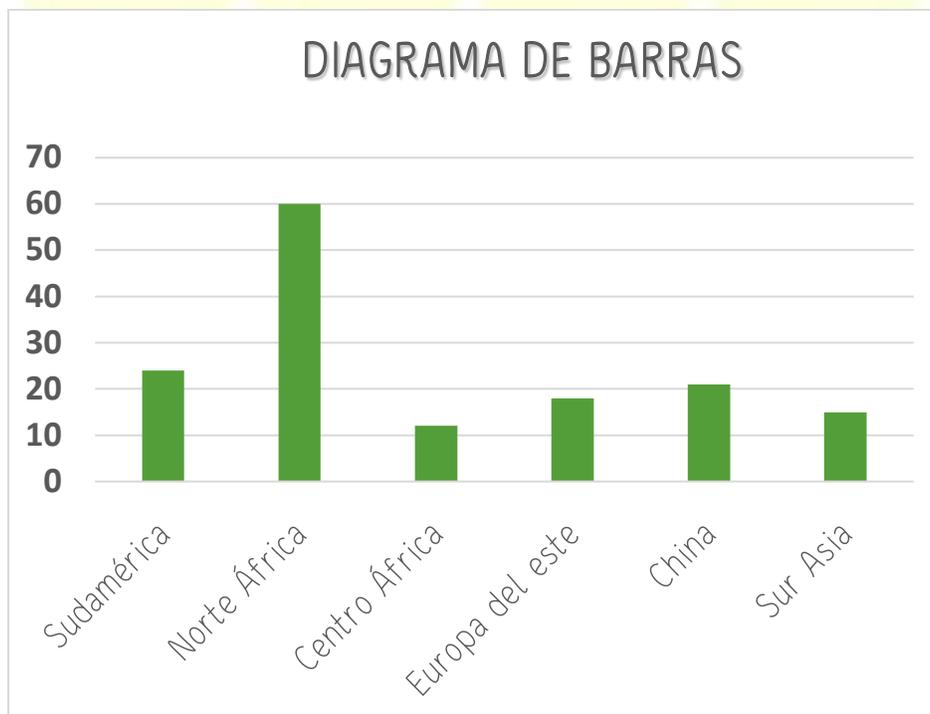
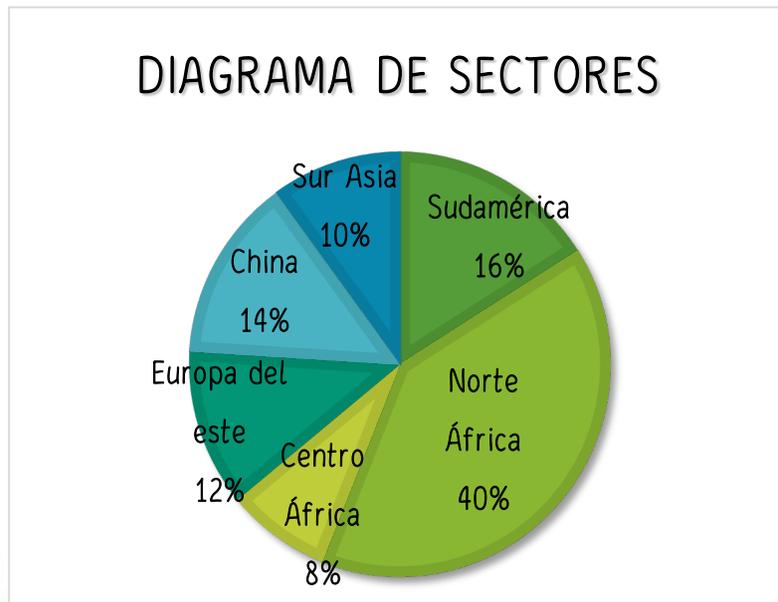
En la siguiente tabla tenemos una distribución de frecuencias con los datos agrupados en intervalos de la variable CI para un grupo de niños.

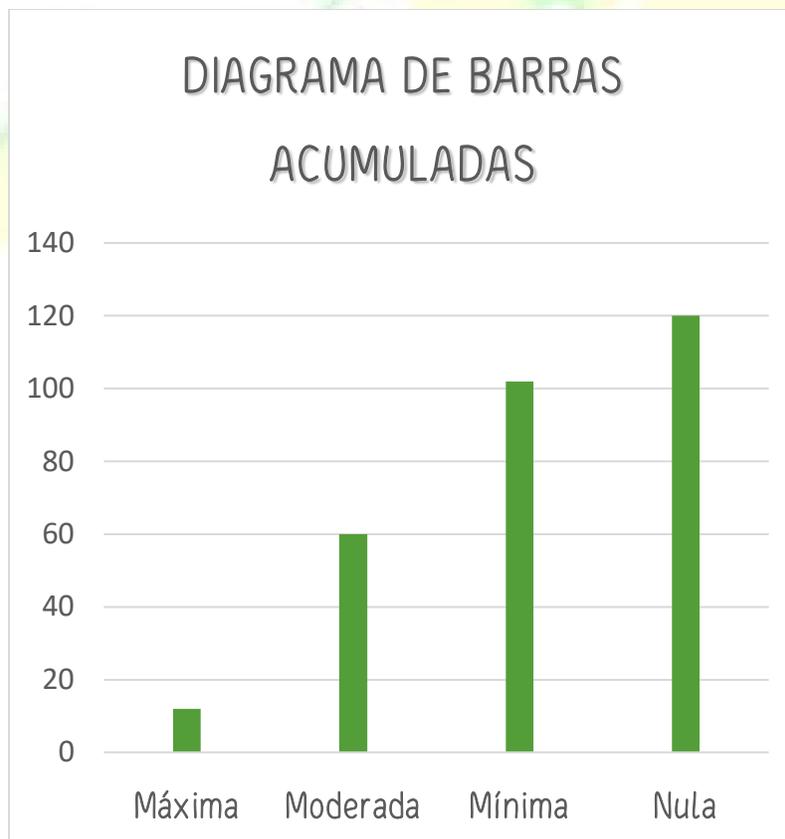
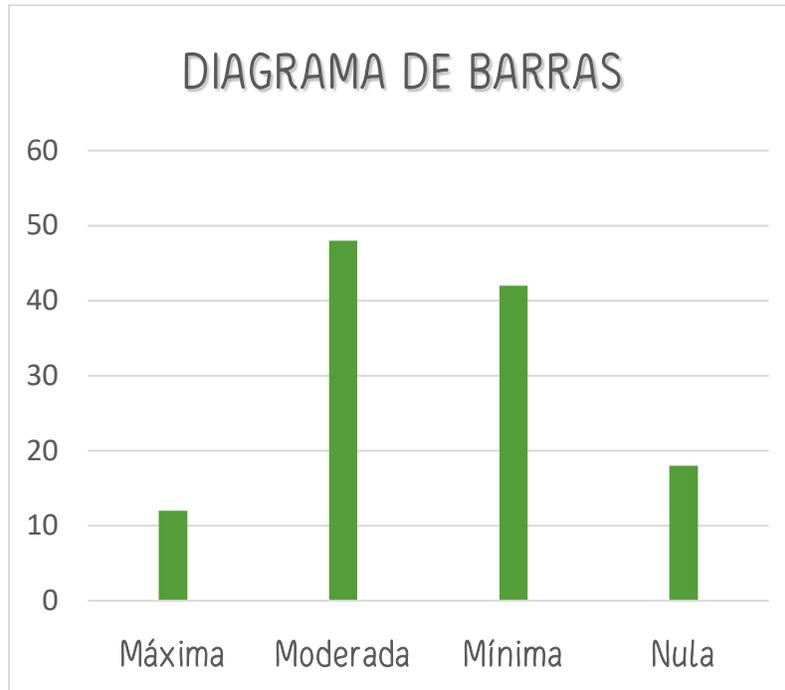
| X Límites aparentes | X Límites exactos | X Punto medio | n_i | p_i | P_i | n_a | p_a | P_a |
|---------------------------|-------------------------|---------------------|---------|--------|-------|-------|--------|-------|
| 60-69 | 59,5-69,5 | 64,5 | 4 | 0,0071 | 0,71 | 4 | 0,0071 | 0,71 |
| 70-79 | 69,5-79,5 | 74,5 | 20 | 0,0355 | 3,55 | 24 | 0,0426 | 4,26 |
| 80-89 | 79,5-89,5 | 84,5 | 40 | 0,0709 | 7,09 | 64 | 0,1135 | 11,35 |
| 90-99 | 89,5-99,5 | 94,5 | 100 | 0,1773 | 17,73 | 164 | 0,2908 | 29,08 |
| 100-109 | 99,5-109,5 | 104,5 | 250 | 0,4433 | 44,33 | 414 | 0,7341 | 73,41 |
| 110-119 | 109,5-119,5 | 114,5 | 115 | 0,2039 | 20,39 | 529 | 0,9380 | 93,80 |
| 120-129 | 119,5-129,5 | 124,5 | 25 | 0,0443 | 4,43 | 554 | 0,9823 | 98,23 |
| 130-139 | 129,5-139,5 | 134,5 | 10 | 0,0177 | 1,77 | 564 | 1,0000 | 100 |
| | | | $n=564$ | 1 | 100 | | | |

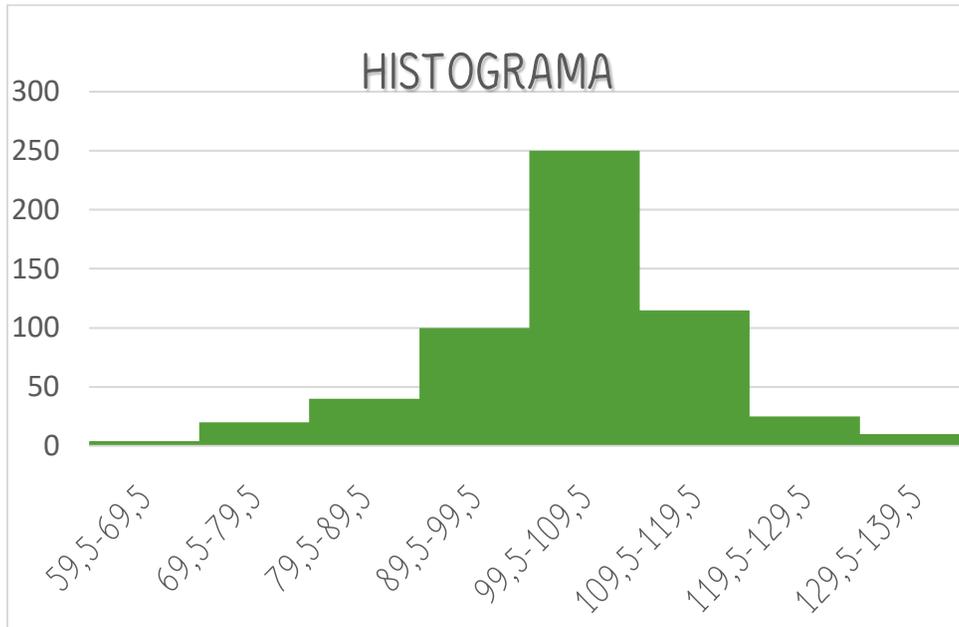
REPRESENTACIÓN GRÁFICA DE LOS DATOS

Los gráficos más habituales en la descripción de variables cualitativas o variables ordinales son los gráficos de barras y los de sectores.

En el caso de las variables cuantitativas discretas, los gráficos utilizados son el diagrama de barras y el de líneas. En cambio, cuando tenemos variables continuas, utilizaremos el histograma.





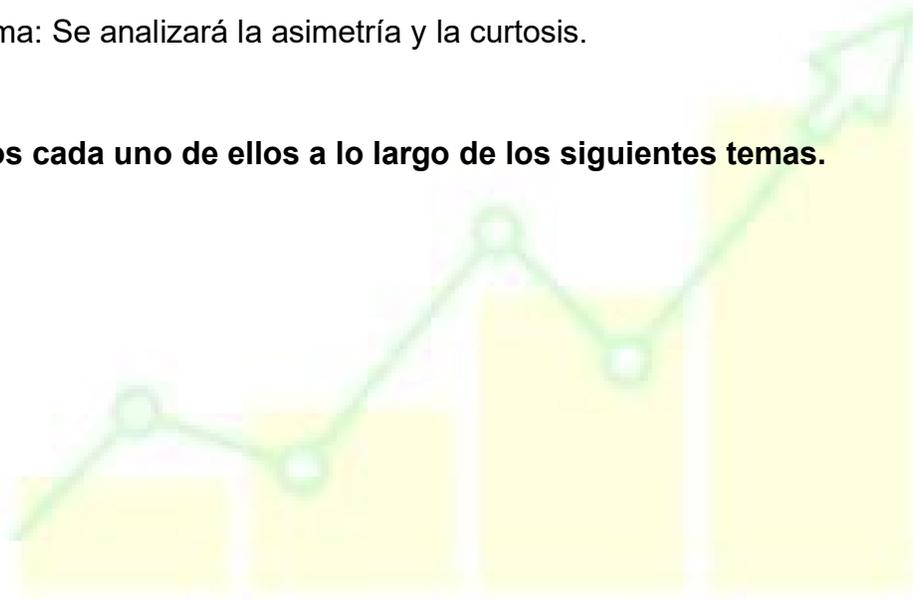


1.5. Tendencia central, variabilidad y forma de una variable: aproximación gráfica

En los temas siguientes se calcularán una serie de índices estadísticos que los utilizaremos para medir:

- La tendencia central: lugar donde se centra una distribución particular en la escala de valores.
- La variabilidad: Se refiere al grado de concentración de los valores entre sí o con respecto al valor central. Será homogénea si los valores están cercanos al promedio y será heterogénea si los valores se dispersan mucho.
- La forma: Se analizará la asimetría y la curtosis.

Veremos cada uno de ellos a lo largo de los siguientes temas.





 INFO@ADEFACIL.COM |



WWW.TUACADEMIAFACIL.COM