

APUNTES, MANUALES, PRESENTACIONES



Universidad de Jaén

Facultad de Humanidades y Ciencias de la Educación

Tema1 (ADPI): La medición en psicología

Antonio Ibáñez Molina
Yasmina Crespo Cobo

08/01/2025

Análisis de datos en psicología I



CREA

Tema 1: La medición en Psicología y el Análisis de Datos

En esta lección vamos a hacer una breve introducción a la Estadística y el Análisis de datos. Trataremos de dar un enfoque que nos sitúe en las áreas de conocimiento relacionadas con la Psicología.

Profesor: Antonio Ibáñez Molina y Yasmina Crespo Cobo

El índice de esta lección es el siguiente:

1. El análisis de datos y el método científico

2. La estadística y la psicología

3. ¿Qué es y para qué sirve el Análisis de Datos?

3.1 Definiciones: Población, muestra, parámetro y estadístico.

4. La medición en Psicología

4.1 Las escalas de medición

4.2 Consideraciones en la aplicación de escalas de medida

5. Definición y clasificación de variables

6. Notación estadística.

1. El análisis de datos y el método científico

Los científicos estudian el mundo natural a través de la observación y la experimentación. Un proceso utilizado para hacer esto es el **método científico**. El método científico comprende hacer una observación, formular una pregunta, crear una hipótesis, experimentar, analizar y concluir el experimento. Si bien el método científico es fluido, cada paso depende del anterior.

Durante la experimentación, se recopilan datos que deberán analizarse para llegar a una conclusión. El análisis se refiere a inspeccionar y desglosar la información recopilada que podría ser cualitativa o cuantitativa para comprender mejor el mundo natural.

Los **datos cualitativos** son datos que observan las cualidades de un experimento y no son numéricos:

- Cambio de color
- Textura
- Notas de observación

Los **datos cuantitativos** son datos que pueden cuantificarse o medirse numéricamente:

- 35 grados centígrados
- 54 segundos
- 57 años

2. La estadística y la psicología

La Psicología y la Estadística comparten una relación estrecha. La Estadística es una herramienta importante para los psicólogos ya que les ayuda a entender mejor la información que recopilan. Esta información sirve como base para la investigación, la toma de decisiones y el diseño de estrategias. Por ejemplo, al analizar los resultados de una encuesta, los psicólogos pueden utilizar la estadística para comprender mejor los resultados y tomar decisiones en función de lo que han descubierto. La Estadística también se utiliza para predecir el comportamiento humano, lo que es esencial para una variedad de campos de la Psicología, como la psicología clínica, la psicología educativa y la psicología del trabajo. La Estadística se utiliza ampliamente en la Psicología para ayudar a los psicólogos a entender mejor la información que recopilan y tomar decisiones basadas en sus hallazgos.

3. ¿Qué es y para qué sirve el Análisis de Datos?

Cuando hablamos de **análisis de datos** en cualquier contexto de investigación de carácter empírico, nos referimos a la **aplicación de la estadística a un conjunto de datos para describirlos e interpretarlos**. Estos dos procesos se corresponden con las dos ramas en las que clásicamente se divide la disciplina de la Estadística. Por un lado, está la Estadística Descriptiva que se ocupa de procesar y caracterizar los datos resumiéndolos en unos pocos números, que llamamos estadísticos descriptivos, o en un conjunto de gráficas. Por otro lado, se encuentra la Estadística Inferencial que se ocupa de analizar con el objetivo de interpretar los resultados de la muestra con respecto a la población total. En este curso, nos centraremos casi exclusivamente en la Estadística Descriptiva y aparcaremos la parte inferencial para cursos más avanzados.

3.1 Definición de Población, muestra, parámetro y estadístico.

La Estadística Descriptiva caracteriza los datos de una muestra tomada de una población. Con población nos referimos a un conjunto de objetos o sujetos que verifican una definición operativa. En Psicología serían las personas que, por sus características, participan en nuestro estudio. Como estas personas son normalmente muchísimas, tomamos una muestra que es un subconjunto de la población más manejable. Existen dos conceptos importantes relacionados con población y muestra:

El primero es el Parámetro: Operación o función definida sobre los valores numéricos de la población. Los parámetros se estiman mediante la Estadística inferencial. Ejemplo: media aritmética del peso de TODOS los universitarios españoles.

El segundo es el Estadístico: Operación o función definida sobre valores numéricos de la muestra. Los estadísticos se calculan directamente sobre la muestra. Ejemplo: media aritmética del peso de un subconjunto de universitarios españoles.

La Estadística Descriptiva, nos proporciona un retrato cuantitativo de nuestro conjunto de datos. Imagínate que hemos recogido la puntuación en nivel de ansiedad rasgo en una muestra de 300 personas. El significado de todos esos datos puestos en una tabla sería difícilmente accesible; podríamos ir explorándolos uno a uno para encontrar datos anómalos, o podríamos considerar si los valores son semejantes entre sí o no... pero en realidad, no estableceríamos un retrato o descripción objetiva, pues si esto mismo lo hiciera otra persona, podría llegar a conclusiones muy distintas. La Estadística Descriptiva nos proporciona las herramientas necesarias para realizar ese retrato objetivo del conjunto de datos. ¿Cómo lleva a cabo esta tarea? Simplemente usa procedimientos matemáticos que nos proporcionan estadísticos de tendencia central o posición, dispersión y forma. Estos tres tipos de estadísticos nos indican respectivamente cuál es el valor aproximado en el que se sitúan los datos; cómo de dispersos o

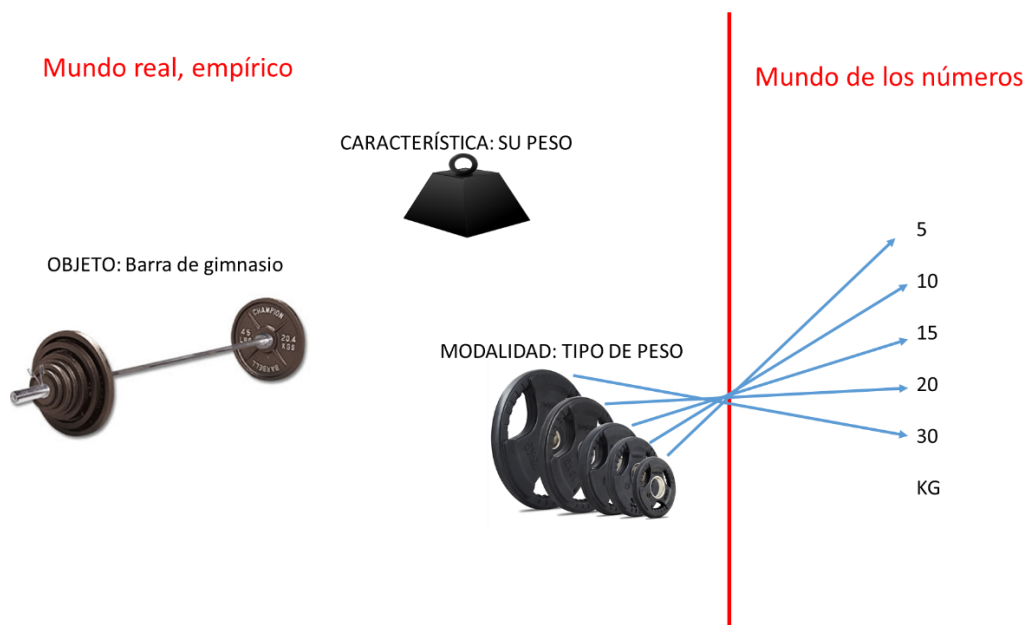
separados se encuentran entre sí; y cómo es el perfil de frecuencias de los valores obtenidos (frecuencia o número de veces que registramos cada posible dato). Estos tres conceptos serán abordados con detalle en los siguientes capítulos. Por ahora bastará con saber que obteniendo estos indicadores y/o su representación gráfica, tendremos un buen resumen del conjunto de datos.

4. La medición en Psicología: Teoría representacional de la medición

Si queremos cumplir con el rigor de una ciencia experimental, el proceso de obtención y análisis de datos en Psicología puede ser muy complejo. Esto es porque requiere la cuantificación de características psicológicas difícilmente mensurables objetivamente. Así, tratar de medir el miedo de un niño a las arañas, o el nivel de ansiedad de una persona durante una entrevista de trabajo, podría parecernos una tarea casi imposible. Y es que por otros cursos sabemos, que la medición requiere una buena definición del fenómeno a medir (constructo psicológico); y también sabemos que necesitamos alguna prueba o aparato para recoger la medida. Además, esa prueba debe incluir algún tipo de regla o sistema de medir para obtener el número que finalmente representa nuestro fenómeno. Por ejemplo, para obtener el nivel de ansiedad del que hablábamos anteriormente, debemos elegir un instrumento mediante el cual podamos asignar un número que caracterice la ansiedad de cada persona (e.g., STAI, inventarios psicológicos). Por ejemplo, la prueba podría indicarnos el nivel bajo-medio-alto de ansiedad mediante los números 1-2-3. Otra alternativa podría ser que la prueba asigne un número del 1 al 10 desde más bajo a más alto nivel de ansiedad, etc. Esta asignación numérica será importante para las posibles conclusiones que saquemos a partir de nuestros estadísticos.

Entonces, ¿qué es medir y cómo podemos hacerlo con las suficientes garantías que exige el rigor de una ciencia empírica? Para contestar a esta pregunta tan compleja vamos a tomar la perspectiva teórica de la **Teoría Representacional de la Medición**. Esta teoría fue impulsada por Stanley Smith Stevens a mediados del siglo XX (1946). Parte del trabajo de Stevens fue el estudio de la psicoacústica, en la que es de suma importancia tratar de medir la sensación subjetiva de un sonido, tratando de usar escalas similares a las usadas con la longitud o el peso. La aplicación de estas escalas sobre sensaciones recibió muchas críticas y Stevens consideró que, en realidad, el debate debería situarse en el significado mismo de medir. Probablemente por eso presentó su trabajo sobre las escalas de medida. Stevens comienza sus razonamientos con una definición simple de Campbell: **medir es atribuir números a objetos según reglas**. Evidentemente, esta definición está incompleta puesto que el número no se asigna al objeto sino a una de sus características, por ejemplo, su peso. Además, cada característica del objeto, tiene muchas posibilidades o modalidades. Por ejemplo, un objeto puede ser de muchos colores y entonces, la característica COLOR tendría muchas modalidades diferentes (rojo, amarillo, verde, etc). Pues bien, **medir significa asignar un número a una característica del objeto siguiendo**

una regla que preserve la lógica de las distintas modalidades. Por ejemplo, para medir la característica peso de un objeto, esta característica puede tomar muchas modalidades, desde pesos más ligeros a más pesados. Medir el peso sería asignar números siguiendo una regla que permita cuantificar las diferencias entre los distintos pesos posibles de los objetos. Si tengo 2 objetos, y usando algún tipo de balanza sé que B pesa como tres veces A, entonces podríamos poner $A=1$ y $B=3$. La medición, entonces, pone en relación las modalidades y los números, y la regla a aplicar debe ser tal que los números resultantes de la medición preserven las relaciones originales en las modalidades. Este es el sistema relacional empírico-numérico. En la siguiente figura resumimos esta idea: por un lado, tenemos un objeto, una barra de gimnasio; una de sus características es su peso, y si quisiéramos medirlo tendríamos que considerar que existen muchos posibles pesos (modalidades) y que cada uno de ellos debe recibir un número que represente el peso real.



4.1 Las escalas de medición

A la relación entre lo empírico y lo numérico es a lo que llamamos escala, y según Stevens hay sólo 4 escalas posibles:

- Nominal:** Sirve para establecer relaciones de igualdad o desigualdad entre los valores atribuidos a las modalidades. Los números (o símbolos) que atribuimos a las modalidades son solo etiquetas. Ejemplo: Profesión de una persona, nacionalidad, afiliación política.
- Ordinal:** Además de la relación anterior, se establecen relaciones de orden entre los valores. Pero los valores no determinan la distancia

entre dos puntos. Ejemplo: Nivel académico de una persona, puesto ocupado en una empresa.

- **De intervalo:** Además de las relaciones anteriores, permite establecer diferencias de magnitud entre los valores. El valor cero en esta escala es arbitrario. Ejemplo: la temperatura en grados Celsius o Fahrenheit, puntuaciones de CI.

- **De razón:** Además de las relaciones anteriores, esta escala permite establecer razones (a/b) o proporciones entre los valores, es decir, entre dos números atribuidos a las modalidades también son válidas las operaciones de multiplicación y división. El valor cero en esta escala es absoluto, es insensible a cualquier transformación. Ejemplo: Ingresos mensuales, tiempo de reacción ante un estímulo, velocidad, grados Kelvin.

4.2 Consideraciones en la aplicación de escalas de medida

En este contexto, de aplicación de escalas de medida, se considera que existen tres problemáticas que deben tenerse en cuenta:

Representación: El sistema relacional numérico debe tener una estructura similar al relacional empírico. Es decir, las relaciones entre las modalidades deben preservarse entre los números.

Unicidad: La forma de representar NO es única. Las características pueden describirse con muchos tipos de conjuntos de datos. Esto quiere decir que existen muchas maneras equivalentes de asignar los números, y por tanto son susceptibles de ser transformados de múltiples maneras.

Significación: Se refiere a las afirmaciones válidas que se pueden hacer sobre una medida concreta (interpretaciones que se pueden hacer de los datos). Qué estadísticos debemos usar, qué transformaciones de los datos son admisibles

En cada tipo de escala se pueden aplicar transformaciones a los valores numéricos, y estas transformaciones serán admisibles en los siguientes casos:

- Escala nominal: transformación que preserve la igualdad-desigualdad
- Escala ordinal: transformación que preserve igualdad-desigualdad y el orden creciente/decreciente
- Escala de intervalo: Transformaciones de tipo $g(x)=a + bx$; siempre que $b>0$

Ejemplo: transformación de grados centígrados en Fahrenheit mediante $F = 32 + (9/5)C$. O lo que es lo mismo: $g(x) = 32 + (9/5)x$

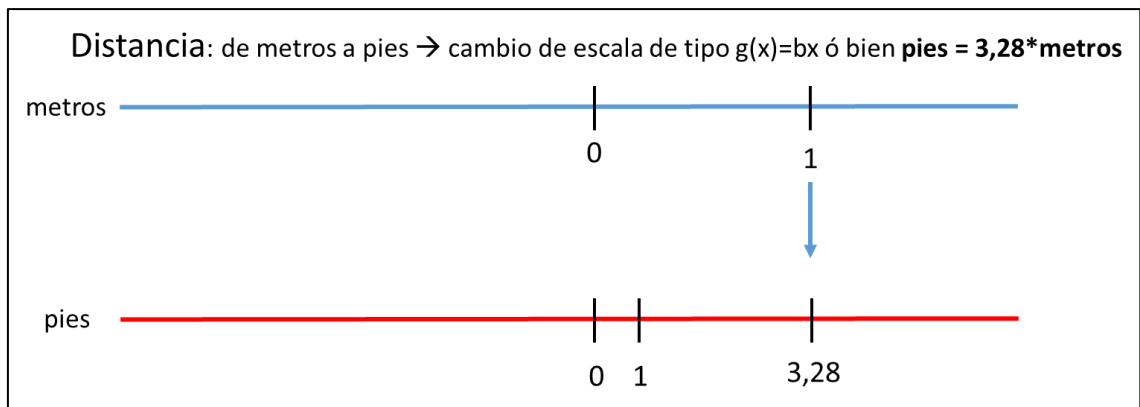
- Escala de razón: Transformaciones de tipo $g(x) = bx$; siempre que $b > 0$

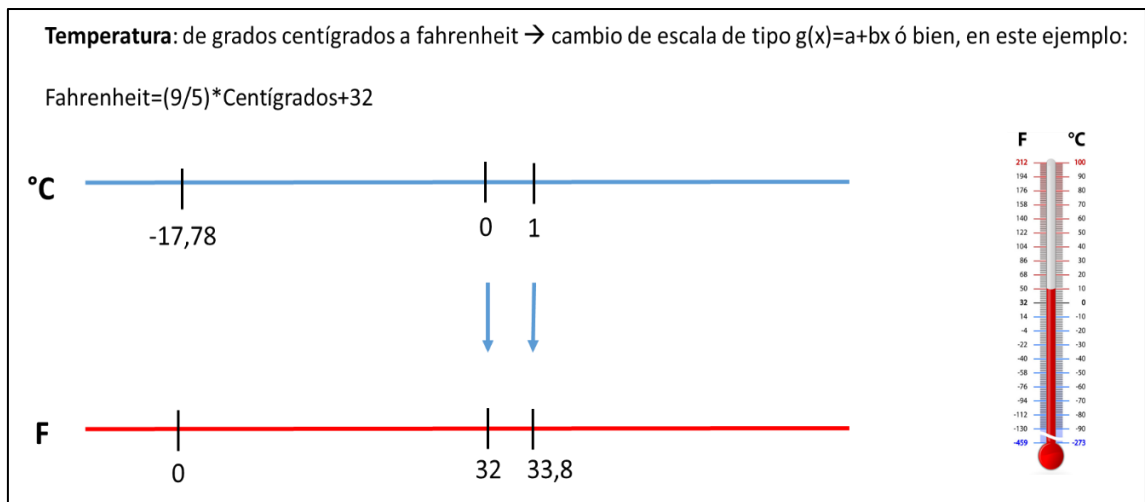
Debemos tener en cuenta que las transformaciones no son más que la sustitución de un conjunto de medidas por otras equivalentes. Esa equivalencia sólo puede obtenerse si se aplican correctamente las relaciones anteriormente expuestas.

La significación para cada escala no se alcanza usando los mismos estadísticos. Según la publicación original de Stevens los estadísticos más frecuentes admisibles en cada escala son:

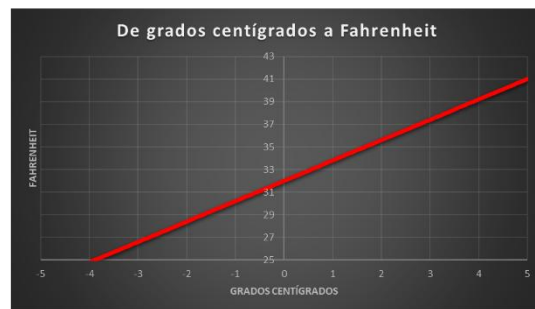
- Nominal: frecuencia y moda, coeficiente de contingencia C
- Ordinal: mediana y percentiles, correlación de Spearman
- De intervalo: media, desviación típica, correlación de Pearson
- De razón: media, desviación típica, correlación de Pearson

En las figuras siguientes incluimos dos ejemplos de transformación de escala. El primero con escala de razón y el segundo con escala de intervalo. Observa como en la primera la posición del cero no cambia mientras que en la segunda sí lo hace. Además, hemos incluido dos representaciones gráficas de estas transformaciones. Fíjate que la relación entre la escala de partida (eje X) y escala final (eje Y) es una línea recta que en el caso de la escala de razón pasa por el origen (0,0) y en la escala de intervalo corta el eje Y en el valor de "a", de la regla de transformación.





- Gráficamente: aquí se ve claro que los cambios de escala de medición son transformaciones lineales



Tipo de escala	Representación	Trasformaciones admisibles	Estadísticos admisibles
Nominal	Igualdad/desigualdad	Transformaciones inyectivas e isomorficas.	Frecuencia y moda, coeficiente de contingencia C
Ordinal	Igualdad/desigualdad Orden	Funciones monótonas crecientes	Mediana y percentiles, correlación de Spearman
Intervalo	Igualdad/desigualdad Orden Igualdad de Intervalos	Transformación lineal positiva: $A+BX$ $B>0$	Media, desviación típica, correlación de Pearson
Razón	Igualdad/desigualdad Orden Igualdad de intervalos Igualdad de razones	Transformación multiplicativa: BX $B>0$	Media, desviación típica, correlación de Pearson Coeficiente de variación.

Tabla 1. Resumen características escalas de medida

5. Definición y clasificación de variables

En este apartado trataremos brevemente el concepto de variable. Es un concepto amplio, ya tratado en otras ocasiones en otras áreas de conocimiento.

- Las **variables**, estadísticamente, son características que se expresan mediante números en dos o más modalidades. Evidentemente, cuando hay una sola modalidad en la característica estamos ante una **constante**

El significado más utilizado de variable en psicología está en relación al **papel que juega en una investigación**. Desde este punto de vista, la variable se refiere al tipo de manipulación de un estudio, o bien, al conjunto de datos recogidos. En el primer caso, se denomina variable independiente, y en el segundo, variable dependiente. También debemos incluir las variables extrañas que son aquellas que pueden afectar a la investigación, y que, si se conocen o detectan, deben ser controladas para su neutralización. EN Estadística Descriptiva se suele trabajar con variables dependientes a las que denominaremos medidas o simplemente variable.

En otras ocasiones, las variables se clasifican en función de **su naturaleza**. Desde este punto de vista podemos decir que hay variables cualitativas, cuasi-cuantitativas y cuantitativas. Esta clasificación está directamente relacionada con las escalas que hemos mencionado anteriormente, y podemos decir también que, **según la escala**, las variables pueden ser de escala nominal (naturaleza cualitativa), de escala ordinal (naturaleza cuasi-cuantitativa), de escala de intervalo (naturaleza cuantitativa), o de escala de razón (naturaleza también cuantitativa).

Merece la pena indicar que **las variables cuantitativas pueden ser discretas o continuas**. Las variables discretas son aquellas cuyas medidas son numerables mediante números naturales, y las continuas son no numerables o incontables porque entre dos medidas cualesquiera siempre hay infinitas medidas. Por ejemplo, el número de errores en una prueba es una variable en escala de razón, discreta, porque entre dos medidas dadas hay un número finito de medidas (entre 3 errores y 10 errores hay 7 errores por ejemplo); mientras que en una prueba, el tiempo de reacción es una variable en escala de razón, continua, porque entre dos tiempos cualesquiera hay un número infinito de posibles tiempos (entre 100 y 101ms están 100'1-100'2...101'0 ms; entre 100'1 y 100'2 están 100'11-100'12...100'20 ms, etc.). Esto en realidad significa que las escalas cuantitativas continuas son fraccionables mientras que las discretas no lo son.

Naturaleza	Nivel de medida	Rol en una investigación
------------	-----------------	--------------------------

Cualitativa	Nominal	Independiente
Cuasi-cuantitativa	Ordinal	Dependiente
Cuantitativa	Intervalo	Extraña
	Razón	

Tabla 1. Cuadro resumen tipos de variables

En general, las variables se denotan o escriben como letras mayúsculas con un subíndice que indica la posición de un valor determinado de la variable (son series de números). Los valores se suelen poner entre llaves indicando que son un conjunto de números o bien entre corchetes

- Por ejemplo: $X_i = \{X_1, X_2, X_3, X_4, \dots, X_n\}$ para $i=1,2,3,\dots,n$

El valor de n indica el número total de puntuaciones, que indicará a su vez el número de participantes en el estudio.

6. Notación estadística

El sumatorio es la operación más común en estadística y consiste en realizar una suma múltiple de elementos. Se denota con la letra griega sigma mayúscula en forma de expresión matemática (suma de los términos de una sucesión). Para indicar que queremos sumar todos los valores de la variable X, ponemos:

$$\sum_{i=1}^n X_i = x_1 + x_2 + \dots + x_n$$

Donde i es el límite inferior de la sucesión y n es el límite superior.

Hay tres propiedades simples muy útiles para comenzar a utilizar el sumatorio que conviene que repasemos. Fíjate especialmente en los términos que no son iguales:

El **sumatorio de una suma** es igual a la suma de dos sumatorio (igual ocurre con la resta):

$$\sum_{i=1}^n (X_i + Y_i) = \sum_{i=1}^n X_i + \sum_{i=1}^n Y_i$$

El sumatorio de un producto es distinto al producto de sus sumatorios:

$$\sum_{i=1}^n (X_i \cdot Y_i) \neq \sum_{i=1}^n X_i \cdot \sum_{i=1}^n Y_i$$

El sumatorio del cuadrado es distinto al cuadrado de su sumatorio:

$$\sum_{i=1}^n X_i^2 \neq \left(\sum_{i=1}^n X_i \right)^2$$

Otras propiedades útiles del sumatorio son:

Suponiendo que K y C son constantes numéricas (números que no cambian)

Desarróllalas en clase con ayuda del profesor

- $\sum_{i=1}^n K = \sum_{i=1}^3 3 = 3+3+3=3*3=9$
- $\sum_{i=1}^n (X_i \cdot K) = K * \sum_{i=1}^n (X_i)$
- $\sum_{i=1}^n (CX_i + K) =$
- $\sum_{i=1}^n (X_i + K)^2 =$
- $\sum_{i=1}^n (X_i - K)^2 =$