

## TEMA 8 LA FIABILIDAD Y LA VALIDEZ

- Los instrumentos se evaluarán de manera general:

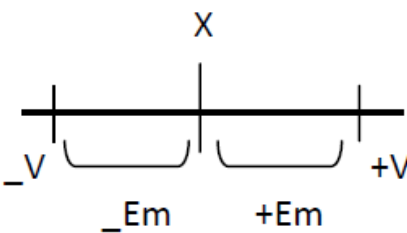
\* fiable ( preciso)

\* valido ( mide lo que dice medir y no otra cosa)

- La fiabilidad: es la característica fundamental de los instrumentos.

Un instrumento es fiable cuando mide algo con precisión , independientemente de los que se está midiendo.

Si un instrumento es bueno te da la nota que deberías de tener , sino recibes la nota que mereces es que comete error, no es preciso.

$$\boxed{X = V \pm E}$$


El diagrama muestra una línea horizontal que representa una escala de medición. En el centro de la línea hay un punto vertical etiquetado como 'X'. A la izquierda del punto 'X' hay un punto etiquetado como '-V', y a la derecha hay un punto etiquetado como '+V'. Dos corchetes debajo de la línea indican los intervalos de error: uno desde '-V' hasta 'X' etiquetado como '-Em', y otro desde 'X' hasta '+V' etiquetado como '+Em'.

Los errores pueden ser:

- Sistemáticos: características internas del instrumento, y por eso afecta a todas personas medidas por el instrumento.
- Aleatorios: obedecen al azar.
- Se considera un instrumento fiable si en su cálculo su valor es mayor de 0.85

Hay muchos tipos de cálculos de la fiabilidad se categorizan en tres grandes grupos:

- Fiabilidad como estabilidad
- Fiabilidad como equivalencia
- Fiabilidad como consistencia interna

La fiabilidad como estabilidad:

- Conocida como método de test retest
- Se aplica la correlación entre las puntuaciones obtenidas por una muestra en dos aplicaciones sucesivas .
- Se aplica la segunda prueba cada veinte o veinticinco días .

La fiabilidad como equivalencia:

- Se denomina de formas paralelas
- Aplico a una muestra mi prueba y otra prueba llamada criterio que se sabe ya buena
- Se compara por medio de una correlación pearson.

**La fiabilidad como consistencia interna**

Procedimiento conocido de las MITADES.

Dividimos la prueba en dos partes (mitades), eligiendo como criterio los ítems pares e impares, ó bien la primera mitad (de 0 al 50%) y la segunda mitad (de 51 al 100%).

Hallando la correlación entre ambas partes estableceremos la consistencia interna del instrumento, si bien en una prueba cuya longitud es la mitad.

Sus procedimientos son :

- Procedimiento de Spearman-Brown
- Procedimiento de Rulon
- Procedimiento de Guttman
- Procedimiento de Kuder-Richardson
- Procedimiento alfa ( $\alpha$ ) de Crombach

PROCEDIMIENTO SPEARMAN BROWN .

- Primero se calcula el coeficiente de correlación pearson entre las partes pares e impares del instrumento.
- Al resultado de la correlación le aplicaré la formula de sperman Brown que aumente la fiabilidad a lo equivalente a una prueba de mayor tamaño.

$$R_{xx} = \frac{2 \cdot r_{xx}}{1 + r_{xx}}$$

PROCEDIMIENTO RULON

Está basado en la **varianza de las diferencias**; es decir, en la relación entre la varianza total del instrumento y la existente entre ambas mitades.

Si las diferencias entre las puntuaciones de dos sujetos en las mitades aumenta, menor será el valor del coeficiente, tal y como se expresa en la fórmula

$$r_{xx} = 1 - \frac{s_d^2}{s_t^2}$$

PROCEDIMIENTO GUTTMAN

Está basado en la **varianza de las mitades**, de forma que a menor valor de la varianza, más elevada será la fiabilidad de la prueba.

$$r_{xx} = 2 \left( 1 - \frac{s_p^2 + s_f^2}{s_t^2} \right)$$

## PROCEDIMIENTO ALFA DE CROMBACH

- Es diferente al resto porque divide al instrumento en tantas partes como ítems tiene, y se trabajara cada ítem por separado y unido al instrumento total.
- Se basan en la varianza de cada ítem y del total y lo relacionan en esta formula

$$\alpha = \frac{n}{n-1} \cdot \left( 1 - \frac{\sum s_{xi}^2}{s_t^2} \right)$$

donde:  $n \rightarrow$  nº de elementos de la prueba o ítems

$s_{xi}^2 \rightarrow$  varianza de cada uno de los ítems  $\rightarrow \sum s_{xi}^2 = s_1^2 + s_2^2 + \dots + s_n^2$

$s_t^2 \rightarrow$  varianza de las puntuaciones totales de la prueba

Esta fiabilidad es buena a partir de 0.7

KUNDER RICHARDSON :

Es el mismo procedimiento que el alfa de crombach pero para ítems de variables dicotómicas.

### ESTUDIO DE LA VALIDEZ

Un instrumento es válido cuando mide lo que dice medir y no otra cosa distinta. Tenemos distintos tipos de validez.

#### Validez de contenido

Debemos comprobar la suficiencia y representatividad de los ítems que componen la prueba.

**Suficiencia**  $\rightarrow$  nº mínimo de elementos que debe incluir el instrumento y que comprenden todos los aspectos esenciales del tema a estudio.

**Representatividad**  $\rightarrow$  definir cada ítem con claridad y precisión, procediendo a la selección de los ítems.

**Proceso de validación**  $\rightarrow$  elaboración de una tabla de especificaciones, realizada por varios expertos, que será sometida a estudio y crítica.

Los procedimientos para calcularlas son las tablas de especificaciones y las opiniones de expertos y jueces con un buen conocimiento del tema.

#### Validez predictiva

Pretendemos establecer predicciones sobre la variable que se mide. Es decir, **qué capacidad tiene la prueba para hacer expectativas de aparición de sucesos futuros.**

Para su cálculo se establece la correlación entre las puntuaciones alcanzadas en la prueba a validar ( $X_i$ ) y las obtenidas en otra prueba denominada criterio ( $Y$ ). Según sean las características de las variables y su nivel de medida, recurriremos al **coeficiente de correlación de Pearson**:

### Validez concurrente

Es una modalidad de la predictiva. La diferencia aquí estriba en que las mediciones de la prueba a validar (Xi) y las del criterio (Y), se realizan al mismo tiempo y que los resultados permiten pronósticos a corto plazo, de utilización inmediata. El cálculo es el mismo.

### Validez de constructo

También denominada de **elaboración ó de construcción**. Es la más compleja y la que requiere en su determinación de procedimientos más complicados. Fox (1987) nos dice que la validez de constructo es la capacidad que tiene el instrumento para distinguir entre aquellos grupos cuyo comportamiento difiere en relación a la variable estudiada.

### Validez aparente

También recibe la denominación de validez didáctica, y se refiere a la serie de aspectos externos de la prueba. Por ejemplo:

- \* un instrumento muy largo puede provocar rechazo o fatiga, y no reflejar opiniones y valoraciones reales.
- \* si el formato es más o menos atractivo, puede motivar a la persona que responde.
- \* el lenguaje ha de ser adecuado a la edad y cultura de las personas que responden.
- \* las instrucciones para las respuestas deben ser claras y fáciles de aplicar.
- \* etc.

### Algunos coeficientes relacionados con la validez

#### (A).- Coeficiente de determinación (d)

$$d = r_{xy}^2$$

Es el coeficiente de validez elevado al cuadrado

Nos indica la proporción de la varianza del criterio que se explica a partir de la prueba. Si se multiplica por 100, se transforma en porcentaje.

$$K = \sqrt{1 - r_{xy}^2}$$

#### (B).- Coeficiente de alienación (K)

Nos indica la proporción de la varianza del criterio que es ajena o extraña a la prueba. Multiplicado por 100, expresa el porcentaje del azar que queda cuando se predice Y a partir de X.

$$E = 1 - K$$

#### (C).- Coeficiente de valor predictivo (E)

Nos indica el grado de seguridad que tenemos en la predicción y si se multiplica por 100, nos indica el porcentaje de seguridad al predecir las puntuaciones de Y a partir de las puntuaciones de X.

#### (A).- Índice de dificultad (ID)

Se expresa numéricamente por el número de sujetos que resuelven satisfactoriamente un determinado elemento. Si solo lo contestan bien unos pocos, el elemento es difícil; si prácticamente todos lo resuelven acertadamente, el elemento es muy fácil.

*Una prueba que se considere bien construida empleará;*

Muy fáciles	10%	0,75 < ID
Fáciles	20%	0,55 < ID < 0,74
Normales (medios)	40%	0,45 < ID < 0,54
Difíciles	20%	0,25 < ID < 0,44
Muy difíciles	10%	ID < 0,24

$$ID = \frac{A - \frac{E}{na - 1}}{n}$$

#### INDICE DE HOMOGENEIDAD

Indica la coherencia de cada uno de los elementos con el total de la prueba. Cada elemento debe medir lo mismo que mide la prueba en su conjunto.

Para calcularlo hallamos la correlación de cada elemento con todos los demás. A mayor correlación, mayor será la homogeneidad. Se dan diferentes casos:

1. ítems dicotómicos → correlación biserial-puntual ( $r_{bp}$ )
2. una variable continua y la otra dicotomizada → correlación biserial ( $r_b$ )
3. dos variables dicotomizadas → correlación tetracórica ( $r_t$ )

pero sea el caso que sea, siempre debemos corregir el valor de índice mediante la fórmula

$$IH = \frac{r_{AB} \cdot s_A - s_B}{\sqrt{s_A^2 + s_B^2 - 2 \cdot r_{AB} \cdot s_A \cdot s_B}}$$

$r_{AB}$  → correlación entre el ítem y el total

$s_A$  → S de las puntuaciones en la prueba

$s_B$  →  $\sqrt{pq}$