TEMA 2
CARACTERÍSTICAS TÉCNICAS DE LOS INSTRUMENTOS DE MEDIDA

Lo que se ve en este capítulo son las formas de analizar los instrumentos de medida que permiten relacionar isomórficamente el objeto con la unidad de medida, poniendo especial atención en los ítem, en lo que se conoce como **teoría clásica de los test,** donde la muestra elegida es importante ya que permite conocer la variabilidad del grupo y por tanto evaluar a partir de ellas las respuestas, y como deben ser evaluados , en su mayoría de manera normativa , salvo algunos instrumentos de referencia criterial que se evalúan con el Kappa Cohen. Entendiendo la evaluación normativa, como aquella donde se establecen unas normas previas para poder evaluar y medir si el individuo posee todos los objetivos marcados, es la típica de la universidad y los institutos . Por otra parte estará la evaluación criterial donde se tomará un criterio de referencia modelo de referencia. Por último también se trabajará con otros modelos diferentes de medida , con **la teoría de respuesta al ítem o del rasgo latente**, cuyos resultados no dependen de la muestra .

LA TEORÍA CLASICA DE LOS TEST

En el trabajo con los test , el punto fundamental en su análisis es el conocer y saber trabajar adecuadamente con los ítems y los dos cálculos por excelencia que nos permiten evaluarlos será:

- El índice de dificultad
- El índice de discriminación
- Análisis de distractores
- Fiabilidad
- Validez
- Dimensionalidad

2.1 Índice de dificultad

La dificultad de un item me indica si la mayoría de personas podrían contestar de forma correcta una pregunta, este cálculo depende también del grupo de sujetos a los que se le aplique la prueba dado que dependiendo de los conocimientos previos, y características intrínsecas del propio grupo serán capaces de contestar correctamente una pregunta o no.

Este tipo de análisis se utiliza sobre todo en test de medida de aptitudes (si son capaces o no) o rendimiento donde se encuentran respuestas correctas o falsas. Claramente la dificultad se ha baremado según unos valores con los que son comparados.

Regla nemotécnica para no confundirse , cuando en un examen sacas un cero es que es un examen muy difícil , pues cuando en un item el valor del índice es cerca de 0 es muy difícil , cuanto más cerca de 1 más fácil es.

Muy fáciles	ID por encima de 0,75
Fáciles	ID comprendido entre 0,55 y 0,75
Normales	ID comprendido entre 0,45 y 0,54
Difíciles	ID comprendido entre 0,25 y 0,44
Muy difíciles	ID por debajo de 0,25

Lo ideal es que la mayoría de los ítems tengan un grado de dificultad media, con una pequeña cantidad de ítems con una dificultad alta y otra pequeña con ítems fáciles. Vamos que estos ítems deben seguir una distribución normal y si los introdujese en una tabla de puntuaciónes típicas tendría la forma de la campana de gauss.

Y donde yo sabiendo el índice de dificultad de un ítem, puedes pronosticar cual será la proporción de gente que suele contestar correctamente a esta pregunta en una muestra normal.

Un índice de dificultad depende también del número de alternativas de respuestas, cuantas más alternativas de respuestas más fácil será el ítem porque menos restaran los errores.

A continuación se presenta un ítem de elección múltiple con 5 alternativas de respuesta. La alternativa A es la correcta:

	Alternatives					
	A	**B**	**C**	**D**	**E**	**Total**
Grup competent	**44**	**36**	**10**	**6**	**4**	**100**
Grup no competent	**28**	**31**	**7**	**18**	**16**	**100**

Cual es el índice de dificultad del ítem. Interpreta los resultados

$$ID = \frac{72 - 128/(5-1)}{200} = 0{,}2$$

Se trata de un índice de dificultad alto porque está muy cercano a 0 y por lo tanto es muy difícil.

El nivel medio de dificultad de un examen debe oscilar entre **0.5 y 0.6**

Es deseable que los valores de ***p*** se distribuyan de la siguiente manera:

5%	de reactivos fáciles
20%	de reactivos medianamente fáciles
50%	de reactivos con una dificultad mediana
20%	de reactivos medianamente difíciles
5%	de reactivos difíciles

2.2. Discriminación

Esta operación nos ayuda a saber si un ítem es bueno o malo a través de distinguir quien sabe sobre el tema en cuestión y por lo tanto saca valores muy altos en la prueba en general o quien no sabe sobre el tema, estos individuos suspenderán la prueba en general y por lo tanto deberían fallar también el ítem. O sea que un ítem discrimina bien cuando hay una relación positiva entre la cantidad de gente que acertó el ítem y que aprobó el examen y viceversa , que suspendió el ítem y el examen.

Este cálculo da resultados entre -1 y +1, y se considera un ítem bueno cuando los valores están entre 0.25 y 0.35 o más .

La fórmula es la siguiente :

$$D = p_{+} - p_{-}$$

Para realizar esta fórmula lo primero que se hace es dividir a los individuos en dos grupos , los que tienen los valores más extremos, El 27% de personas con notas más altas y el 27% con notas más bajas. Donde p+ son la proporción de individuos con mejor nota y p- son la proporción de sujetos con menos aciertos.

Aunque a la hora de la verdad a la hora de calcularlo no se suele realizar esta operación ya que hay otras dos formas de calcularlo. La primera por medio de una correlación entre la nota sacada en la prueba por los alumnos , y la puntuación sacada en el ítem . *AUNQUE SI SE ESTABLECE LA CORRELACIÓN HAY QUE RECORDAR QUE SOLO SERÁ VALIDO SI LA CORRELACIÓN ES CON LA NOTA TOTAL DEL TEST CORREGIDO AL AZAR , O SEA QUE A ESTA NOTA TOTAL SE LE RESTARÁ LA PUNTUACIÓN DEL ITEM AL QUE QUEREMOS COMPARAR.* Si la correlación es elevada o sea superior a 0.7 se podría considerar que el ítem discrimina correctamente

A esta prueba también se le puede llamar índice de homogeneidad, ya que mira si el homogéneo el resultado del test con el resultado del ítem. Y por lo tanto se podrá utilizar esta otra fórmula para calcularlo, formula que lo que busca es ver si los resultados del ítem son similares a los del total. A diferencia de lo visto anteriormente si yo aplico lo que propiamente dicho se denomina índice de Homogeneidad a la hora de establecer la correlación que pide la formula, entre los resultados del ítem y el total , esta correlación no debe estar corregida al azar y por lo tanto no se le debería restar el ítem .

$$IH = \frac{r_{ix} S_x - S_i}{\sqrt{S_x^2 + S_i^2 - 2 r_{ix} S_x S_i}}$$

r_{ix} es la correlación ítem-test

S_x es la desviación típica del test

S_i es la desviación típica del ítem

El calculo de la correlación me decía si dos variables o valores estaban relacionadas entre si , y por cuanto . El coeficiente de correlacion nos mide el valor de covariacion o variaciÓn conjunta de dos series de datos (o sea traducido la relación o no relacion entre dos grupos de valores representadas de manera numerica)

La forma de calcular esta correlación variará dependiendo del tipo de variables o valores con los que trabajemos, y más concretamente dependiendo del nivel de medida de los mismo.

Si son dicotómicas el coeficiente que se debería usar es el coeficiente de correlación de pearson para dos variables dicotómicas,

si una está dicotomizada y la muestra es normal , se utilizaría la tetracórica.

Si una variable es continua , intervalo, y la otra dicotomizada se utiliza la biseraial.

Si hay una continua y otra dicotomica se usa la biserial puntual.

Entre dos variables continuas se utilizará la pearson .

A la hora de la verdad, el calculo que se os permitirá utilizar es el cálculo de la correlación Pearson.

Una vez calculada esta operación la forma de valorarla se realizará a partir de los valores proporcionados por la tabla siguiente.

El ítem discrimina muy bien	IH igual o mayor de 0,4
El ítem discrimina bien	IH comprendido entre 0,3 y 0,39
El ítem discrimina poco	IH comprendido entre 0,20 y 0,29
Ítem límite. Se debe mejorar	IH comprendido entre 0,10 y 0,19
El ítem carece de utilidad para discriminar	IH menor de 0,10

El resultado que proporciona esta fórmula siempre ha de ser mayor que el resultado de haber solo realizado la correlación pearson corregida , donde le restábamos al total el ítem.

En una prueba de puntuaciónes continuas respuestas dicotómicas en el ítem, se han obtenido los siguientes valores : Una correlación biserial puntual de 0,52 , Una desviación típica del test total es de 3 ,25 y la desviación típica del item dicotómico 0,495 . Este item sería homogeneo o no con el resto de la prueba

$$IH = \frac{0{,}52 \cdot 3{,}25 - 0{,}495}{\sqrt{(3{,}25)^2 + (0{,}495)^2 - 2 \cdot 0{,}52 \cdot 3{,}25 \cdot 0{,}495}} = 0{,}395$$

Como vemos la homogeneidad es muy alto.

VARIABILIDAD Y DISCRIMINACIÓN

La variabilidad de un ítem está relacionada al índice de discriminación, cuando hablamos por tanto de la variabilidad de un ítem podremos conocer si discrimina bien o no.

$$S_x = \sum_{j=1}^{n} S_j r_{jx}$$

donde:

S_x = desviación típica del test

S_j = desviación típica del ítem

r_{jx} = índice de discriminación del ítem j

Así si la desviación típica de un test es nula es 0 , no habrá discriminación alguna , porque todos los individuos han sacado la misma nota y por lo tanto no se discrimina bien entre los que saben y no saben . Esto sería por ejemplo si un item lo suspenden todos eso significa que no es bueno ya que debería haber gente que sabe y gente que no sabe por lo tanto si lo fallan todos entonces es que no ha diferenciado entre los conocimientos de la gente. Y lo mismo pasa si el test lo fallan todos, si el test lo suspenden todos el ítem y el test no discriminarán bien ya que seguro que hay gente que si conocía el tema.

Un ejemplo :

Se supone que un ítem de un test se pasa a 300 estudiantes de los 81 (el 27% de 300) que obtienen mejores puntuaciones en el test , 40 aciertan el ítem y de los 81 que obtienen peores notas en el test , 12 aciertan el ítem , en consecuencia:

$$IH = \frac{40}{81} - \frac{12}{81} = 0,3456$$

Y si se analizan algunos de los 15 items de la prueba:

4d2c	Correlación elemento-total corregida
MAT4	,383
MAT5	,568
MAT6	,386
MAT7	,570
MAT8	,647
MAT9	,399
MAT10	,533
MAT11	,092
MAT12	,420
MAT13	,519
MAT14	,529
MAT15	,382

Según esto el ítem 11 debería de ser eliminado porque tiene un valor muy bajo

PREGUNTAS PRACTICAS QUE HAN CAIDO EN EL EXAMEN DE ESTOS PUNTOS

EXAMEN DE SEPTIEMBRE DE 2012

Un profesor está realizando una prueba de rendimiento en lengua. Pasa dicha prueba a un colectivo de 200 estudiantes de un colegio. Desea saber las características métricas de la misma. Se pide:

a) Calcular el índice de discriminación de un ítem, sabiendo que el 27% que obtiene mejores valoraciones en la prueba 30 acierta el ítem y el 27% con peores valoraciones 10 acierta el ítem.
b) Calcular el índice de dificultad corregido de un ítem de cuatro alternativas que acierta 80 estudiantes suponiendo que lo han respondido todo el colectivo del colegio.

A)

1- En este caso la formula que debería de utilizar es la fórmula concreta del índice de discriminación , dado que el número de sujetos que realizan la prueba es elevado y por tanto ha sido posible, como nos muestran los datos dividir la muestra en el 27% de gente con mas nota y 27% de gente con menos nota en el examen general y ver cuantos aciertan el ítem de cada grupo.

$$D = p_+ - p_-$$

En esta formula la p+ es la proporción o frecuencia relativa de gente que acertó el ítem y está entre los mejores. La frecuencia relativa se calcula como fa/n o sea número de gente que acierta el ítem entre el total del 27% de estos mejores .

La p – por lo tanto es la proporción o frecuencia relativa de los que fallaron el ítem de entre los peores entre el total de estos peores. Así la formula de las frecuencia relativa ya la conocemos.

hi = fa/n

Antes de nada para poder calcular esto por medio de una regla de tres debería poder calcular el total de gente que forma el 27% de los mejores .

27%----------100%

x--------------200

x= (27· 200) / 100 = 54 personas.

Con esa cantidad de gente que es el 27% de los mejores y los peores , debería calcular sus frecuencias relativas.

hi = 30/ 54 =0.555…

hi = 10/54 =0.185…

Para concluir solo habrá que restarle a esta frecuencia relativa de gente que mejor nota sacan le restaremos la frecuencia relativa de la gente que acertó el ítem- Así al 0.555- 0.185 =0.37 podemos observar que el ítem es bueno porque es superior de 0.25.

B) A la hora de calcular el índice de dificultad como estamos trabajando con un ítem con alternativas de respuesta debería utilizar la fórmula de índice de dificultad corregido al azar.

$$I.D. = \frac{A - \frac{E}{n_a - 1}}{n}$$

En este caso los aciertos serán 80 y si en total había 200 personas podremos saber el número de errores , ya que lo han respondido todos, entonces total – aciertos = errores.

200 -80 = 120 y ya podemos calcular

(80 – (120/3)) / 200 = 0.2

Esto es un ítem difícil porque esta cerca de 0.

EXAMEN DE 2013 PRIMERA SEMANA

1) Un profesor recoge información de la realización de una prueba objetiva cuyos 5 ítems están en escala dicotómica de 0 o 1 siendo 0 (no acierta el ítem) y 1 (lo acierta). Los resultados de 4 estudiantes fueron los siguientes:

	Item1	Item2	Item3	Item4	Item5
Juan	0	1	0	0	1
Pedro	1	1	1	0	1
Luis	1	0	1	0	1
Enrique	0	0	1	1	0

Se pide:

a) Calcular el índice de dificultad de cada ítem

El índice de dificultad de cada ítem será la formula del índice de dificultad normal dado que no hablan de alternativas de respuesta la formula por tanto debe ser:

$$I.D. = \frac{A}{n}$$

Y donde el numero de aciertos se sabrá sumando el número de 1 de cada ítem. Y el total será mirando el número de valores que hay.

Item 1:

Id= 2/4= 0,5 como el valor va de 0 a 1 estará entre 0,45 y 0.54 se considera un ítem de dificultad normal

Item 2

Id= 2/4= 0.5 , como el anterior es de dificultad normal.

Item 3

Id= ¾ = 0.75 como esta entre 0.55 y 0.75 se considera un ítem fácil.

Item 4

Id= ¼ =0.25 como está entre 0.25 y 0.54 se considera difícil

Item 5

Id= ¾ = 0.75 va entre 0.55 y 0.75 y se considera un ítem fácil.

En este test no es del todo adecuado ,ya que un buen test debe tener sobretodo ítems de dificultad normal y el resto distribuido de manera equivalente entre valores fáciles y difíciles.

EXAMEN DE RESERVA DE SEPTIEMBRE DE 2013

2) La siguiente tabla presenta los resultados de 5 sujetos en 3 ítems.

		1	2	3	T
Sujetos	1	1	3	4	8
	2	5	1	0	6
	3	3	4	5	12
	4	1	0	1	2
	5	4	2	3	9

a) Se pide, sin realizar cálculos, sólo indicar el proceso de cálculo de índice de discriminación del ítem 1 como correlación ítem con el total **(T)**. Como alternativa al proceso manual, se puede indicar cómo se realizaría con R.

Este es un examen muy habitual, te mandan explicar que no calcular el proceso para desarrollar el índice de discriminación.

En este caso debería de explicar que de los tres procesos que se pueden realizar para conocer el índice de discriminación del ítem 1 se debe realizar según el enunciado el de correlación entre el ítem y el total. Claramente no hubiese podido calcular el método de

$$D = p_{+} - p_{-}$$

Entonces podría escoger entre la formula del índice de homogeneidad propiamente dicha o como nos pide que la realicemos.

Debemos realizar la correlación Pearson para una variable dicotómica y otra de intervalo, o la formula de la correlación biserial puntual que es apropiada para estos tipos de variables.

La correlación pearson debería de ser realizada entre los resultados que tienen los sujetos en el ítem en cuestión y el test en general corregido al azar, lo que significa que a los resultados de cada individuo en todos los test les restaría el valor del ítem en cuestión. Para que este ítem no influya en los resultados totales de los individuales. Por lo tanto a la nota total que nos dan le restaría el valor del ítem 1 en cada sujeto y con esos valores establecería la siguiente fórmula

$$r_{xy} = \frac{n \cdot \sum X \cdot Y - \sum X \cdot \sum Y}{\sqrt{\left[n \sum X^2 - \left(\sum X\right)^2\right]\left[n \sum Y^2 - \left(\sum Y\right)^2\right]}}$$

El resultado se valoraría como si fuese una correlación no con los valores del índice de homogeneidad

Si por el contrario lo quisiese realizar el proceso por el r , lo primero que debemos hacer es abrir el paquete de calculo estadístico del programa que nos interesaría para ello si ya esta instalado el paquete solo debería escribir:

> library(Rcmdr)

Después debería introducir los valores para ello debería crear los vectores (línea de valores) introduciendo primero por ejemplo los valores del ítem y del total menos el ítem

>”ítem 1”< - c (1,5,3,1,4)

>”total-1”<-c (7,1,9,1,5)

>cor (c(“item1”),c(“total-1))

EXAMEN 2014 FEBRERO DE PRIMERA SEMANA

1) Un profesor recoge información de la realización de una prueba objetiva cuyos 5 ítems están en escala dicotómica de 0 o 1 siendo 0 (no acierta el ítem) y 1 (lo acierta). Los resultados de 4 estudiantes fueron los siguientes:

	Item1	Item2	Item3	Item4	Item5
Juan	0	1	0	0	1
Pablo	1	1	1	0	1
Raúl	1	0	1	0	1
Oscar	0	0	1	1	0

Se pide:

a) Calcular el índice de discriminación del ítem2 como correlación ($r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$) entre la puntuación de los sujetos en el ítem y la puntuación total

b) Calcular el índice de dificultad de cada ítem

c) Indicar los comandos de R necesarios para calcular el índice de dificultad del ítem3.

a) Nos pide calcular el indice de discriminación como en el examen anterior como correlación por eso debemos calcular la pearson entre el valordel ítem 2 y el total menos el ítem dos.
Sería lo siguiente

	Item1	Item2	Item3	Item4	Item5
Juan	0	1	0	0	1
Pablo	1	1	1	0	1
Raúl	1	0	1	0	1
Oscar	0	0	1	1	0

Lo primero que debemos hacer es calcular el total menos el item dos que es el que queremos discriminar.
así el total sería:
Juan 1
pablo 3
Raul 3
Oscar 2

Ahora calcularemos la formula superior con ese total y el valor del item dos.
Para ello crearemos una tabla donde desarrollaremos los calculos necesarios:

item 2	total-2	itemx total	item al cuad	total al cuad
1	1	1	1	1
1	3	2	1	4
0	3	0	0	9
0	2	0	0	9
2	9	3	2	23

$$= \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$$

$$\frac{4\cdot 3 - 2\cdot 9}{\sqrt{(4\cdot 2 - 2^2)(4\cdot 23 - 9^2)}} = \frac{12 - 18}{\sqrt{(8-4)(92-81)}} = \frac{-6}{\sqrt{4\cdot 11}}$$

$$\frac{-6}{\sqrt{44}} = \frac{-6}{6{,}63} = -0{,}904$$

b) Calcula el índice de dificultad de cada ítem :
El índice de dificultad debe ser realizado con la formula sencilla porque no hablan de alternativas de respuesta por lo tanto el resultado sería :
Ítem 1: 2/4=0.5 como está entre 0.45 a0.54 es normal
Ítem 2: 2/4 =0.5 como está entre 0.45 a 0.54 se considera normal
Ítem 3: ¾=0.75 como esta entre 0.55 y 0.75 se considera fácil
Ítem 4: ¼=0.25 está entre 0.25 y 0.54 se considera difícil
Ítem 5: ¾=0.75 como esta entre 0.55 y 0.75 se considera fácil

c) Para mostrar los comandos del r para calcular el índice de dificultad del ítem 3 por el r lo primero será abrir el paquete de calculo estadístico del programa que nos interesaría para ello si ya esta instalado el paquete solo debería escribir:

```
> library( Rcmdr)
```

a) Después debería introducir los valores para ello debería crear los vectores (línea de valores) introduciendo primero el valor del ítem

```
>"item3"<-c(0.1.1.1)
>id<-sum("item3")/4
```

O por el contrario poner

> id<(“item3”)