

Psicometria

AC4 - Interpretació de les puntuacions

Entrega 18/12/05

- GUIA MÒDUL IV
- DUBTES DELS MÒDULS
- PRESENTACIÓ PAC 4
- CORRECCIÓ PAC4 (Consultor: Antoni Coscolluela Mas)

GUIA MÒDUL IV - INTERPRETACIÓ DE LES PUNTUACIONS:**1. QUÈ SABEM DE LES PERSONES DESPRÉS D'ADMINISTRAR-LOS UN TEST?**

El càlcul de la puntuació total ens és útil per a resumir el grau de coneixement que la persona avaluada té de la característica que mesura el test.

Quan ordenem a les persones segons la puntuació que han obtingut en un test, les situem en una escala.

1.1. Interpretació absoluta de la puntuació obtinguda en un test

Puntuació total mínima (PT_m): s'obté sumant les respostes donades als ítems del test suposant que en cada un s'hagi respost amb el valor mínim.

Puntuació total màxima (PT_M): és la suma de les respostes donades als ítems del test quan a cada un se li ha donat el valor màxim.

1.2. Interpretació relativa de la puntuació obtinguda en un test

Es tracta de comparar la puntuació obtinguda per una persona amb el conjunt de puntuacions que han obtingut d'altres persones que han contestat el mateix test.

Els barems proporcionen informació sobre el significat i sobre la manera en què s'ha d'interpretar la puntuació obtinguda en un test amb relació al grup de referència.

1.3. Normes per a la interpretació de les puntuacions

El grup normatiu representa la població a la qual va dirigit el test.

Els estudis de baremació exigeixen que, una vegada determinada i definida la població de referència, es disposi d'un grup de persones que sigui:

- a) Representatiu de la població a la qual es dirigeix el test, i
- b) Homogeni respecte a altres variables.

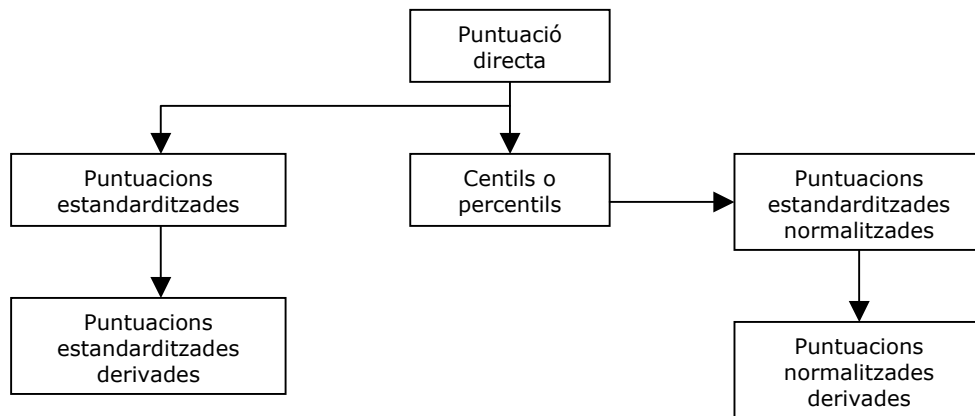
Les dades normatives que s'han de presentar en el manual d'un test són les corresponents als estadístics descriptius, el tipus de transformació aplicada a les puntuacions directes i l'àmbit de representació del grup normatiu.

- Normes nacionals: s'elaboren a partir de mostres representatives d'àmbit nacional. Contenen totes les característiques que influeixen en la interpretació dels resultats del test, és a dir, nivells d'edat, gènere, raça, estrats professionals, nivell socioeconòmic i cultural, etc.
- Normes locals: són les que s'obtenen a partir dels nuclis geogràfics o educatius que representen una subpoblació concreta.
- Normes d'usuari: no s'obtenen a partir d'una població definida per l'usuari del test. A partir de les puntuacions que obtenen les persones en un test que s'ha administrat amb una finalitat determinada i en un espai de temps delimitat, s'elaboren les normes.

2. TRANSFORMACIÓ DE LES PUNTUACIONS

Les tècniques estadístiques conegudes permeten convertir les puntuacions directes (TCT) i les puntuacions vertaderes (TRI) en un conjunt de valors diferents denominats puntuacions derivades. El procés de transformació en puntuacions derivades es denomina escalament de les puntuacions del test.

En l'esquema següent es presenten els tipus més freqüents de transformació de puntuacions utilitzades tant en la TCT com en la TRI:



2.1. Puntuacions estandarditzades

Les puntuacions directes es poden transformar tenint en compte la variabilitat del grup de referència a partir del qual s'ha obtingut. Això permet interpretar les puntuacions de manera que no solament es considera la seva distància respecte a la mitjana sinó en relació amb la variabilitat del grup de referència:

$$z_{xi} = \frac{x_i - \mu_x}{\sigma_x}$$

x_i : puntuació directa

μ_x : mitjana del grup

σ_x : desviació estàndard.

Com que la mitjana de la distribució de les puntuacions estandarditzades és 0 i la seva desviació estàndard és igual a 1, s'obtidran valors positius i negatius, els quals indiquen respectivament si les persones es troben per sobre o per sota de la mitjana del seu grup i fins a quin punt se n'aparten.

2.1.1. Puntuacions estandarditzades derivades

Les puntuacions estandarditzades es poden transformar, al seu torn, en les anomenades puntuacions estandarditzades derivades. L'objectiu d'aquesta transformació és facilitar la interpretació, evitant els valors positius i negatius i els valors decimals de les puntuacions estandarditzades:

$$Y_i(z_{xi}) = \mu_y + \sigma_y(z_{xi})$$

Y_i : puntuació estandarditzada derivada

μ_y : mitjana de la nova distribució o escala

σ_y : la seva desviació típica

z_{xi} : valor de la puntuació estandarditzada.

Les transformacions de z canvien l'origen de l'escala de puntuacions, mantenint les mateixes relacions entre els seus valors.

Algunes de les transformacions de z més utilitzades en psicologia són les puntuacions T amb mitjana igual a 50 i desviació estàndard igual a 10 i l'escala derivada de quocient intel·lectual (QI) amb mitjana igual a 100 i una desviació estàndard igual a 15.

$$T_i = 50 + 10z_{xi}$$

$$QI_i = 100 + 15z_{xi}$$

2.2. Centils o percentils

La transformació a centils o percentils (P_c) és una de les més habituals. Aquestes puntuacions indiquen el percentatge de persones que es troben per sota i per sobre d'una puntuació directa determinada:

$$P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100$$

f_{aa} : freqüència acumulada anterior a la puntuació directa de la qual volem calcular el percentil

f_i : freqüència absoluta en què es troba la puntuació

N : nombre de persones del grup normatiu.

2.3. Puntuacions estandarditzades normalitzades

Una vegada transformades les puntuacions directes a percentils, podem calcular les puntuacions estandarditzades normalitzades buscant la puntuació z corresponent a cada un dels centils en el supòsit que la distribució fos normal. Els valors obtinguts així es denominen puntuacions estandarditzades normalitzades (z_n) o puntuacions z-normalitzades.

La transformació de les puntuacions directes a puntuacions estandarditzades normalitzades només és aconsellable quan la forma de la seva distribució és estadísticament comprable a la distribució normal.

2.3.1. Puntuacions normalitzades derivades

Les escales normalitzades derivades més habituals en psicologia són les d'estanines o enneatipus i les de decatipus. Les primeres divideixen el conjunt de puntuacions en nou intervals amb una mitjana i gual a 5 i desviació estàndard i gual a 2. Les segones fan una divisió en deu parts amb mitjana igual a 5,5 i amb una desviació estàndard també igual a 2:

$$\text{Enneatipus} = 5 + 2z_n$$

$$\text{Decatipus} = 5,5 + 2z_n$$

z_n : correspon a la puntuació estandarditzada normalitzada.

S'ha de tenir en compte que el valor mínim per a cada escala és 1, de manera que als valor obtinguts inferiors a 1 se'ls assigna un 1. El valor màxim per a l'escala d'enneatipus és 9 i per a la de decatipus és 10; per tant, als valors superior a 9 i 10 se'ls assignarà 9 i 10, respectivament. A més, per als valors intermedis dels enneatipus i decatipus corresponents, s'han d'arrodonir els decimals al valor superior o inferior, segons estigui per sobre o per sota de 0,5 respectivament.

2.4. Transformació de puntuacions en el model de la TRI

La corba resultant o CCT es representa en uns eixos de coordenades. En l'eix d'abscisses els valors de θ i en el d'ordenades els valors obtinguts a partir de la suma de les probabilitats d'encertar cada un dels ítems per a cada valor de θ .

3. EQUIPARACIÓ DE LES PUNTUACIONS

Diem que hi ha correspondència entre les puntuacions de les diferents formes d'un test quan les puntuacions obtingudes són equivalents, de manera que podem usar unes o altres puntuacions indistintament perquè són intercanviables.

Per a equiparar les puntuacions de dos o més tests les formes del test han de complir les següents condicions:

- 1) Han de mesurar la mateixa característica.
- 2) Han de tenir la mateixa fiabilitat.
- 3) Han de complir la condició de simetria.

3.1. Tipus d'equiparacions

Podem establir dos tipus bàsics de correspondències:

- Equiparació horitzontal : es refereix a l'establiment de correspondències entre les puntuacions de formes alternatives d'un mateix test, de vegades també denominades formes paral·leles o equivalents, les quals presenten uns nivells de dificultat i de fiabilitat idèntics. Habitualment aquest tipus d'equiparació és possible quan, anteriorment, els tests es construeixen amb la mateixa dificultat.
- Equiparació vertical : establim correspondències entre puntuacions obtingudes per diferents grups de persones -que, per exemple, difereixen del seu nivell escolar- a les quals s'han administrat tests amb diferents nivells de dificultat per a mesurar una mateixa característica.

3.2. Dissenys i tècniques d'equiparació

Els dissenys més utilitzats per a obtenir les puntuacions directes de les formes alternatives d'un test i dur a terme la seva equiparació són els següents:

- 1) Grup únic: les formes alternatives del test s'administren una després de l'altra al mateix grup de persones, que ha estat seleccionat aleatòriament entre la població a la qual va dirigit el test. Una vegada obtingudes les puntuacions es fa l'equiparació utilitzant alguna de les tècniques que veurem més endavant.

Hem de controlar els possibles efectes d'aprenentatge, pràctica, fatiga i ordre, ja que, si es produeixen, afecten les puntuacions obtingudes en el test aplicat en segon lloc. Precisament per a evitar els efectes esmentats, s'acostuma a utilitzar la tècnica de control de contrabalancejat. Es tracta de dividir el grup aleatòriament i d'administrar a cada subgrup les formes del test en ordres diferents.

- 2) Grups equivalents: tenim més d'un grup de diferents persones escollides aleatòriament a partir de la mateixa població de referència i idèntic respecte a l'aspecte que volem mesurar amb el test. Permeten d'evitar els efectes d'aprenentatge, pràctica, fatiga i ordre. A més, el temps d'administració és més petit que per als dissenys de grup únic ja que cada grup respon una sola de les diferents formes del test. Es considera que l'equiparació feta no presenta biaixos deguts a la diferència entre els grups i, per tant, les puntuacions de les diferents formes del test són intercanviables.

- 3) Test d'ancoratge: El disseny del test d'ancoratge permet de "pal·liar" les possibles diferències aleatòries entre els grups. Es tracta d'administrar, com en els dissenys de grups equivalents, una forma del test a cada grup i, a més, un segon test comú per a tots els grups i en el mateix ordre.

Com que les puntuacions obtingudes en el test d'ancoratge constitueixen punts de referència comuns per a l'equiparació és necessari que els ítems que el constitueixen siguin en contingut i dificultat tan afins com sigui possible als tests dels quals volem equiparar les puntuacions.

Les característiques d'aquest disseny permeten que els grups no siguin equivalents (grups no equivalents són els que, tot i pertànyer a la mateixa població, no són semblants respecte a la característica que avalua el test).

Una vegada s'han obtingut les puntuacions directes de les formes alternatives d'un test podem procedir a la seva equiparació mitjançant l'ús de diferents tècniques:

1. Equipercantil: es basa en el supòsit que la distribució de les puntuacions obtingudes a partir de les diferents formes del test són iguals.

$$P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100$$

2. Transformació lineal: Quan la correspondència s'estableix entre les puntuacions directes amb el mateix nombre de desviacions estàndard per sobre o per sota de la mitjana.

$$\frac{x_i - \mu_x}{\sigma_x} = \frac{y_i - \mu_y}{\sigma_y}$$

3.3. Equiparació de puntuacions en el model de la TRI

Es diu que un test està calibrat quan s'han estimat els paràmetres dels seus ítems.

El disseny més emprat per l'equiparació és el test d'ancoratge.

4. ESTABLIMENT DE PUNTS DE TALL

4.1. Sensibilitat i especificitat o com s'ha d'avaluar la precisió del criteri seleccionat

Els models més habituals de la TRI analitzen ítems de resposta dicotòmica.

Un cop obtingudes les puntuacions de tots dos grups, podem classificar les persones depenent del criteri o punt de tall que hem determinat. Segons la precisió amb què hàgim decidit el punt de tall, podem cometre dos errors:

- 1) Classificar persones que no presenten el trastorn com a anorèctiques, o cometre el que es denomina falsos positius.
 - 2) Classificar persones anorèctiques com a persones que no pateixen el trastorn, o cometre el que es denomina falsos negatius.
- Quan el criteri establert permet de classificar correctament la majoria de les persones que no pateixen el trastorn diem que el test és molt específic (o poc flexible).

- Quan el criteri establert permet de classificar correctament la majoria de les persones que pateixen el trastorn, diem que el test és molt sensible (o molt flexible).

Per a calcular l'índex d'especificitat n'hi haurà prou d'obtenir la proporció de vertaders negatius (l'especificitat indica la proporció de persones que no posseeixen la característica d'interès que avalua el test i que han estat classificades correctament):

$$\text{Especificitat} = \frac{\text{Vertaders negatius}}{\text{Total}} ; \text{Especificitat} = \frac{\text{Vertaders negatius}}{\text{Vertaders negatius} + \text{Falsos positius}}$$

L'índex de sensibilitat s'obtindrà calculant la proporció de vertaders positius (la sensibilitat indica la proporció de persones que posseeixen la característica d'interès que avalua el test i que han estat classificades correctament):

$$\text{Sensibilitat} = \frac{\text{Vertaders positius}}{\text{Total}} ; \text{Sensibilitat} = \frac{\text{Vertaders positius}}{\text{Vertaders positius} + \text{Falsos negatius}}$$

Quan augmentem el valor del punt de tall, augmentem l'especificitat del test, però això va en detriment de la seva sensibilitat. Inversament, en aplicar un criteri poc estricte o més flexible, disminuïm l'especificitat i augmentem la sensibilitat del test.

4.2. Una mesura global de la precisió del test: la corba ROC

- Un cop obtinguts els valors dels índexs d'especificitat i sensibilitat per a cada punt de tall, els podem representar gràficament mitjançant una corba denominada ROC.
- La corba ROC es construeix a partir de la proporció de vertaders positius i de falsos positius obtinguts en cadascun dels punts de tall possibles del test.
- La proporció de falsos positius és el valor complementari de l'especificitat (1-especificitat).
- Direm que una prova és molt precisa quan el punt de tall triat permet d'obtenir un valor d'especificitat i sensibilitat igual a 1. Una prova no classifica amb precisió les persones quan els índexs d'especificitat i sensibilitat són igual a 0,5.

5. LA COMUNICACIÓ DELS RESULTATS

- Els usuaris dels tests hem de considerar les característiques de les persones a qui va dirigida la informació. S'ha de facilitar la comunicació i comprensió dels resultats a les persones no expertes.

DUBTES DELS MÒDULS:

Data: 21:21:13 01/12/2005

Estic fent l'activitat 5 de la pac 4 i tinc un dubte respecte els percentils, als exercicis sempre surten números rodons, 43... però a mi, fent el càlcul a partir de la fórmula, em surten decimals, estic fent malament els càlculs?

A la pàg. 17 pot haver una errada en el càlcul del percentil 41 ja que posa $f_{aa}=20$ quan pel 41 li tocaria 23, no? I més a baix a on posa $T=50+10$...no tocaria ser $50+14$ (núm. ítems) sinó, d'on surt el 10?

Data: 22:46:50 01/12/2005

Els percentils tenen que ser números enters entre 1 i 100, si a l'hora de calcular-los surten resultats amb decimals, s'aproxima a l'enter més proper.

En el càlcul d'un percentil, la f_{aa} és la freqüència acumulada anterior a la puntuació directa de la qual volem calcular el percentil, i per tant, en el cas de la puntuació directa de 41 de l'exemple de la pàgina 17 del mòdul, la f_{aa} és igual a 20.

Les puntuacions T sempre s'obtenen a partir de la formula: $T_i = 50 + 10 Z_{xi}$,

per tant el valor de la puntuació T per la puntuació de 41 del mateix exemple serà: $T = 50 + 10 (-0,18)$, tal com posa el mòdul.

Data: 01:42:43 02/12/2005

A l'exercici 5, t'hem de ficar un per un, tots els càlculs necessaris o amb la taula ja n'hi ha suficient?

Data: 09:00:53 02/12/2005

A l'exercici 5 potser no cal posar tots els càlculs realitzats, però no estaria de més al menys posar-ne un d'ells (per una de les puntuacions directes), per què així si hi ha un error puc diferenciar si és per què s'aplica una formula equivocada (error més greu), o per què hi ha un error de càlcul que és molt més lleu.

Data: 12:37:10 05/12/2005

Algú acaba d'entendre aquesta activitat, ac2 de la web? quan es demana una formula, no tinc ni idea, jo perquè la $PT_m=0$ faria que els valors anessin de 0 a 5, enlloc de 1 a 6. I llavors pregunta que quin valor tindria la PT_m , pues 0, sinó, no ho entenc.

però després a la preg2 torna a preguntar el mateix, no m'entero, torno a estar fosca. algú em pot ajudar?

Data: 21:25:55 06/12/2005

Activitat 2: Porto hores intentant saber que he de fer però ara ja no se res, no recordo res, sembla que no relaciono res. Què estic oblidant?

Data: 01:40:00 07/12/2005

Una pista, una pista.... coeficient alfa per a ítems dicotòmics....

Data: 12:53:11 07/12/2005

El coeficient alfa per ítems dicotòmics molt bé, però d'on obtens la variància de la puntuació total del test?.

Data: 15:59:41 07/12/2005

És més fàcil que això, no busquem la fiabilitat del test, sinó la contribució de cada ítem al test per al que no necessites, en aquest cas, la variància de la puntuació total.

El que hem preguntat es si fa falta buscar el total corregit. Potser seria lia massa la troca, coses que de vegades em porta massa problemes.

Data: 10:34:33 08/12/2005

Jo també estic com vosaltres! Entenc la fórmula que hem d'aplicar però no el perquè.

Data: 11:27:41 08/12/2005

Preguntat, per a que vols aplicar la formula? que és el que et demanen? de vegades els errors venen d'entendre malament els enunciats.

La formula et dona la pista, busca les probabilitats de les proporcions dels ítems dicotòmics i ja ho tens tot.

Data: 18:31:37 09/12/2005

A l'enunciat de la pregunta 3 no acabo de entendre el terme variància dels errors. Al meu entendre la variància de l'error es aquella que no és vertadera, es a dir, que no forma part de la variància comuna ni de la variància específica dels ítems... però aquí em quedo i no sé relacionar això amb la pregunta.

Data: 20:03:59 09/12/2005

Per respondre la pregunta 3, teniu que tenir en compte la profecia de Spearman-Brown, i els supòsits de la TCT de la pàgina 37 del mòdul de fiabilitat.

Data: 13:39:06 10/12/2005

Algú em pot ajudar amb el càlcul de les puntuacions estandarditzades normalitzades, he calculat els percentils i tota la resta però no hi ha manera de trobar el càlcul per obtenir aquestes dades?

Data: 13:44:55 10/12/2005

Per calcular les Zn es pot emprar la funció de l'Excel DISTR.NORMAL.ESTAND.INV i et sortirà un quadre que diu "probabilidad", aquí hem de ficar el percentil dividit per 100 (o sigui la proporció) de la puntuació corresponent.

Data: 13:14:33 11/12/2005

A la pregunta 1.3., quan demana tipus d'inferències, a què es refereix, a les conclusions que podem treure dels resultats del test?

Pel que fa a la validació, segons l'enunciat només es podria fer un tipus de validació? No seria possible fer-ne més d'un, per exemple de contingut i d'estructura interna?

Data: 19:10:51 11/12/2005

Tinc problemes per trobar com es calcula la Distribució de freqüències i de freqüències acumulades. Si algú em pot donar un cop de mà ho agrairé.

Data: 19:29:09 11/12/2005

Jo ho he fet mitjançant l'opció FRECUENCIA de l'Excel.

Data: 09:51:37 12/12/2005

Estic realment encallada a la pregunta 3.

Entenc com afecta el fet d'afegir 5 ítems al test original en la fiabilitat del test (profecia Spearman-Brown) i, per tant, com afectarà al error del test, i també entenc que, d'acord amb els supòsits de la TCT, 2 tests paral·lels tindran la mateixa variància del error.

El que no trobo és com relacionar una cosa amb l'altra. Com afecta un canvi de fiabilitat a la variància del error?

Data: 10:18:40 12/12/2005

Crec, però, que hauríem de tenir en compte que no es tracta de test paral·lels, no tenen el mateix numero d'ítems.

Data: 11:01:31 12/12/2005

Això també ho veig clar... però segueixo sense veure com afecta a la variància del error.

Alguna orientació???

Data: 20:19:47 12/12/2005

El que jo entenc és que si allargant un test augmenta la fiabilitat i suposo que en aquest cas també seria proporcional la variància de l'error, per la qual cosa jo he posat que també la variància de l'error del test original serà menor a 20, però no se si està bé!

Data: 20:48:15 12/12/2005

Jo també he arribat a aquesta conclusió!!

Data: 10:07:52 13/12/2005

Estic encallada amb la 3, i penso es pot solucionar aquesta pregunta amb una regla de tres??? per que jo no acabo de veure com resoldre-ho, per una banda penso que no hi ha dades suficients, i per altra penso que al ser el test més gran hauria de donar un valor superior, però no ho tinc clar, si algú em pot ajudar a entendre-ho ho agrairia molt.

Data: 10:57:48 13/12/2005

Respecte al que comenta sobre la tercera pregunta. Jo penso que segons la seva explicació el test original tindrà més de 20. Ja que menys ítems, menys fiabilitat i per tant més probabilitat d'error. Com ho veieu?

Data: 12:59:37 13/12/2005

Això es el que havia pensat en un començament, però al augmentar els ítems augmentem la variància empírica, que sabem que és igual a la suma de la vertadera i del errors, no podem suposar que només augmenti la variància vertadera, no? el que no se es com demostrar-ho, es a dir en quina relació augmenten.

Data: 15:00:33 13/12/2005

Perquè creus que la variància empírica augmenta quan augmenta el numero de ítems? És que el que no veig és això, com afecten els canvis a la variància?

Data: 16:05:46 13/12/2005

Quan augmenta el nombre d'ítems d'un test, no té per què augmentar la variància empírica.

Data: 18:10:32 13/12/2005

Tal y como yo lo veo nos encontramos con dos test equivalentes, no paralelos, o sea DIFERENTES ¿no?... Para mí aquí esta el tema estructural de esta pregunta...

No creéis que si son diferentes podemos contestar a la pregunta aplicando el 2 y 3 supuesto de la TCT...

Data: 18:51:19 13/12/2005

Estic d'acord en que són dos tests equivalents. I per a respondre m'he centrat en el que va dir el consultor:

1. per un costat en la profecia de Spearman-Brown, en la que m'he quedat amb lo de la lògica d'aquesta, la qual es semblant a la d'una regla de tres clàssica. Per tant els resultats d'aplicar aquesta formula demostren que a més quantitat d'ítems més fiabilitat. Tenint això anem a

2. La teoria clàssica dels tests (TCT) on ens diu que l'error típic de mesura està associat amb el grau de precisió que té un test per a mesurar. La fiabilitat d'un instrument està inversament relacionada amb l'error típic de mesura i, per tant, directament relacionada amb la precisió d'aquest instrument. És a dir a més fiabilitat menys error de mesura i més precisió.

Data: 19:04:54 13/12/2005

L'error típic de mesura, equivaldria a la desviació típica de l'error??? A les hores....

Data: 20:26:37 12/12/2005

La 2.2. pens que podria ésser l'ítem 2 però no estic segura ja que em guiï pel llinar de correlació 0,20-0,30. De la 2.1. no tenc ni idea, per favor algú pot ajudar-me.

Data: 20:52:09 12/12/2005

Respecte a la pregunta 2.1, interpreto que la suma de correlacions haurà de ser 1 i d'aquí n'he tret la conclusió. No sé si m'equivoco...

Respecte la 2.2, si mires la formula del coeficient alfa per a ítems dicotòmics (llegeix el que posa el mòdul amb atenció) i ho relaciones amb les dades que tenim, crec que podràs arribar al que es demana.

Data: 19:33:29 13/12/2005

Ho sento, però jo tampoc no veig res clar a l'activitat 2. Com sabem els ítems que puntuen 1 o 0? si la suma de les correlacions ha de ser 1, com decidim quina correlació és excessiva? com podem calcular la probabilitat que l'ítem puntui 1 o 0 si la formula del coeficient alfa per a ítems dicotòmics calcula la fiabilitat del conjunt dels elements? hem d'aïllar la fórmula d'alguna manera?...

Data: 20:35:35 13/12/2005

Respecte a la pregunta 2 mirat el coeficient alfa per a ítems dicotòmics de la pàgina 25, observa la formula, no fa falta aplicar-la i et sortirà l'exercici. Penso que no t'has de centra en el fet de que el valor erroni pugui estar tan sols en la correlació ítem-PT.

Data: 21:39:29 13/12/2005

Totes dues respostes estan en la formula KR20.

Data: 11:01:44 14/12/2005

Si mires la formula, veuràs que el producte de les proporcions de respostes 1 i de respostes 0 és igual a la variància de cada ítem. A partir d'aquí, has de pensar quines són les possibles proporcions i com afecten a la consistència interna del test.

Data: 13:53:17 14/12/2005

Disculpeu que hi torni amb la 2.2...

Si resulta que no podem calcular la correlació de l'ítem amb el total corregit per així veure la contribució a la consistència interna...

aleshores algú em pot donar una pista de per on he de buscar la resposta?

Data: 16:21:58 14/12/2005

Podem suposar que la correlació del ítem amb el total corregit serà molt similar a la correlació del ítem amb el total.

Data: 00:26:35 15/12/2005

Gràcies Antoni, perquè sense aquesta suposició, per més que hem mirés la fórmula de Cronbach per ítems dicotòmics, jo crec que no es possible de solucionar.

Data: 11:52:29 14/12/2005

No em puc treure del cap el segon i el tercer supòsit de la TCT:

"Amb el segon supòsit s'assumeix que entre les puntuacions vertaderes i els errors de mesura no hi ha cap correlació. El tercer supòsit assumeix que tampoc no hi ha correlació entre els errors que incideixen en l'avaluació de les persones amb dos tests (el test A i el test B) diferents."

Aquests dos tests de què parlem no són paral·lels, en absolut, ja que un tindria 10 ítems i l'altre 15. I vale que el test llarg tindria més fiabilitat, perquè li hem afegit ítems i tal, però pot ser què...

-a) com que entre les puntuacions vertaderes i els errors de mesura no hi ha correlació, ni tampoc n'hi ha entre els resultats de dos tests diferents, el fet que la variància dels errors del test llarg sigui 20 no ens diu res sobre la variància dels errors del test curt.

-b) com que són tests equivalents, el llarg és més fiable. Si el llarg és més fiable, el seu error és més petit. Llavors l'error del curt serà més gran... més gran de 20? una variància més gran indica un valor absolut més gran, o és al revés?

Data: 12:00:59 14/12/2005

Estoy de acuerdo con tus dos posturas... de hecho yo en esta pregunta tengo actualmente dos posibles contestaciones, igual que tú.

Data: 11:41:17 16/12/2005

Estic intentant utilitzar la fórmula FRECUENCIA de l'Excel i no entenc com he de posar les dades. Em podríeu donar un cop de mà? I si no és amb aquesta fórmula, com heu calculat les freqüències?

Data: 12:25:00 16/12/2005

No aconsegueixo calcular ni les freqüències ni el percentil amb el Excel (llavors tampoc puc trobar Zn). Com s'ha de fer? només em surten 0 i 1, per les f, i el mateix valor (de 0 a 6) al percentil.????

Data: 13:10:16 16/12/2005

Us adjunto un document amb les instruccions de com realitzar una taula de freqüències amb l'Excel.

Per altre banda, les puntuacions Zn també es poden calcular amb l'Excel. Amb la funció "DISTR.NORM.ESTAND.INV", i on posa Probabilidad s'ha de posar el percentil dividit per 100.

Data: 19:04:19 16/12/2005

Algú em podria explicar perquè a la taula de la pàg. 17, el percentil del valor 1 és 1?

Data: 08:31:10 17/12/2005

Torno a preguntar pels percentils. La meua pregunta sobre la taula de la pag. 17 és perquè si el valor del percentil del primer valor és 1 perquè és el mateix que el de la freqüència absoluta, tinc un problema per fer l'exercici 5 de la PAC. A l'exercici 5, la freqüència absoluta del valor 0 em dona un número més gran de 100, per tant, com calculo el percentil si és el primer valor i no tinc freqüència acumulada?... què faig malament?

Data: 11:47:33 17/12/2005

La fórmula per obtenir un percentil la teniu al final de la pàgina 16. Aplicant aquesta fórmula per la puntuació directa de 35 de l'exemple de la pàgina 17 ens dona:

$$(((0 + (0,5 \times 1)) / 48) \times 100 = 1.$$

O sigui que el percentil del primer valor de la taula de la pàgina 17 no és 1 perquè és la freqüència absoluta d'aquest valor, sinó perquè la fórmula per obtenir el percentil d'aquest valor dona 1.

Data: 02:45:47 17/12/2005

Els càlculs de les freqüències amb l'Excel s'han de fer un per un (calcular cada resultat de cada puntuació) o hi ha alguna manera de fer-los tots alhora?

Data: 08:26:29 17/12/2005

Jo ho he fet un per un. No sé si hi ha una manera de fer-ho tots a l'hora, però per les explicacions crec que no.

Data: 14:17:55 17/12/2005

Estic calculant la taula de freqüències però en el primer pas ja alguna cosa que no em surt bé, perquè he utilitzat l'exemple que ha posat el consultor i ell posa que per comprovar si ho hem fet bé quan apliquem l'autosum hauria de donar un valor de 30 i a mi em surt 48, algú em podria ajudar, que faig de malament?

Data: 18:54:46 17/12/2005

Quan apliquem la formula del percentil, N significa el grup de subjectes del nostre estudi, en aquest cas 497? és així? merci

Data: 19:17:03 17/12/2005

sí, sí, 497.

Data: 23:50:08 17/12/2005

A la pregunta 3 primer he calculat la relació entre ítems i em dona 0.66. Després passo a calcular la profecia de Spearman i trobo que no tinc el valor de la fiabilitat.

He de suposar que al no poder seguir no tinc dades per saber quant val la variància dels errors del test original? O hi ha forma de saber-ho a partir de la variància dels errors, o a partir d'un altre càlcul.

Data: 00:43:05 18/12/2005

Jo no he fet tants càlculs, sinó que m'he deixat guiar per la lògica dels supòsits 2 i 3 de la TCT, i també per un missatge del consultor on deia que la variància empírica no té perquè augmentar encara que un test sigui més llarg...

Data: 01:22:57 18/12/2005

Respecte a 3, també penso que no es tracta de fer càlculs, sinó en fer deduccions a partir de la relació entre grau de fiabilitat i error de mesura.

Data: 23:53:12 17/12/2005

A la pregunta 2 estic utilitzant la formula KR20 i trobo que no se el valor de les proporcions de respostes 1 i 0 (p_j i q_j). Com es calculen

Data: 00:43:05 18/12/2005

L'únic que podem saber de les proporcions entre respostes 1 i 0 és que en cada ítem, si multipliques aquestes proporcions entre elles, t'ha de donar la variància que hi figura. Imagina't que una de les variàncies fos 21: doncs les proporcions podrien ser 3 i 7, o 21 i 1...

Amb aquesta dada em sembla que ja es pot respondre la pregunta, o almenys jo ho he fet així.

Data: 13:05:20 18/12/2005

En la pregunta 4 quan diu digueu 4 formes per millorar la validesa d'un test, es refereix a les que hem après o a inventar-nos nosaltres altres alternatives?

Data: 15:35:51 18/12/2005

En la pregunta 4 em basaria en les directrius que es dona al final del mòdul de validesa, on s'analitza.

Data: 20:05:57 18/12/2005

Estic intentant calcular les puntuacions estandarditzades (z) de l'exercici 5, però ara no sé com es calcula la desviació estàndard. Algú em pot ajudar??

Data: 20:51:08 18/12/2005

Jo l'he calculat amb la funció DESVEST de l'Excel.

PRESENTACIÓ PAC 4 - PROVA GLOBAL:

Objectius

Aquesta prova d'avaluació inclou els continguts corresponents als quatre mòduls didàctics de l'assignatura de "Psicometria". El seu objectiu principal és enfrontar l'estudiant amb una sèrie de situacions pràctiques que tindrà que resoldre amb els coneixements adquirits en aquest temari.

Presentació de la prova

La PAC4 esta composta per una sèrie d'activitats pràctiques. Per la seva correcta realització, serà necessari que l'estudiant faci ús del material de l'assignatura, tant teòric com pràctic. Lògicament per el càlculs dels diferents anàlisis plantejats l'estudiant pot utilitzar el programa Excel. En aquests càlculs, utilitzi només dos decimals.

Si algun aspecte de la PAC4 no queda clar, no dubteu en preguntar al vostre consultor.

Enunciat de la PAC

1. Plantegi i descrigui un estudi de validació (diferent dels plantejats en les PACs anteriors), especificant els següents aspectes:
 - 1.1. Plantegi i descrigui l'objectiu de l'estudi
 - 1.2. Quin tipus de test administrativa?. Esculli un test, indiqui el seu nom, descrigui breument la seva finalitat i contingut, i expliqui el per què d'aquesta decisió.
 - 1.3. Indiqui el tipus d'inferències que l'usuari del test podria realitzar a partir de les puntuacions del test, i senyali el tipus de validació possible a realitzar, raonant adequadament la seva resposta.

2. Els indicadors de 3 ítems dicotòmics obtinguts en una mostra de 10 subjectes, han estat els següents:

	Ítem 1	Ítem 2	Ítem 3
Variància	0,15	0,25	0,09
Correlació Ítem-PT	0,90	0,30	0,40

PT: Puntuació Total

- 2.1. Hi ha un valor erroni (no possible) en la taula. Indiqui'l, justificant la seva resposta i ofereixi el valor correcte més proper a aquest valor erroni.
- 2.2. Quin és l'ítem que menys contribueix a la consistència interna del test? Per què?
3. Un test té 10 ítems i se li afegeixen 5 ítems paral·lels. Si la variància dels errors val 20 en el test allargat, quant valdrà en el test original?
 - a) 20
 - a) més de 20
 - a) menys de 20
 - a) No hi ha dades per a saber-ho

Indiqui l'alternativa correcta i justifiqui la teva resposta.
4. Indiqui 4 formes de millorar la validesa d'un test, justificant la seva resposta.
5. A partir de les dades obtingudes en la passació del EPQRA (veure fitxer "DadesEPQRA(05-06_1)"), obtingui en una única taula i **només per les puntuacions totals directes (de 0 a 6) de l'escala N de Neuroticisme** les següents distribucions:
 - Distribució de freqüències, de freqüències acumulades, de puntuacions estandarditzades (z), de puntuacions T, de percentils, de puntuacions estandarditzades normalitzades (z_n), de Enneatipus i de Decatipus.
6. Com a resum de totes les activitats d'avaluació continuada realitzades amb el test EPQRA, faci un comentari general **breu** de les dades obtingudes en la nostre passació d'aquest test, on constin les característiques principals de la mostra de subjectes utilitzada, i de les dades obtingudes respecte a las característiques psicomètriques del test (de forma similar però més resumit a l'estudi de Ibañez i col·laboradors).
 Considera, en general, que el test EPQ-RA compleix el requisit psicomètriques adequats que tot test té que complir?. Raoni la seva resposta

Criteris d'avaluació

La valoració de la prova es realitzarà tenint en compte:

- La coherència i el contingut del document, i
- La correcta integració dels diferents aspectes analitzats

En quant a l'avaluació de la prova, es realitzarà tenint en compte:

% Result	Avaluació
95 % - 100 %	A: Molt bé
76 % - 94 %	B: Bé
56 % - 75 %	C+: Correcte
41 % - 55 %	C-: No arriba als mínims
0 % - 40 %	D: Incorrecte

Format de lliurament

Totes les respostes s'inclouran en un document de Word. El nom de l'arxiu de Word tindrà el següent format: "cognom1_cognom2_PAC4.doc". Els cognoms s'escriuran sense accents. Per exemple, un estudiant anomenat "Pep Alba Pastor" posaria el següent nom al seu arxiu "alba_pastor_PAC4.doc"

Data de lliurament

La data límit pel lliurament d'aquesta activitat és el **18 de desembre de 2005**

CORRECCIÓ PAC4 (Consultor: Antoni Coscolluela Mas):

1. Plantegi i descriu un estudi de validació (diferent dels plantejats en les PACs anteriors), especificant els següents aspectes:
 - 1.1. Plantegi i descriu l'objectiu de l'estudi
 - 1.2. Quin tipus de test administrativa?. Esculli un test, indiqui el seu nom, descriu breument la seva finalitat i contingut, i expliqui el per què d'aquesta decisió.
 - 1.3. Indiqui el tipus d'inferències que l'usuari del test podria realitzar a partir de les puntuacions del test, i senyali el tipus de validació possible a realitzar, raonant adequadament la seva resposta.

En aquest primer exercici hi ha moltes possibles respostes i totes elles poden ser igualment adequades.

2. Els indicadors de 3 ítems dicotòmics obtinguts en una mostra de 10 subjectes, han estat els següents:

	Ítem 1	Ítem 2	Ítem 3
Variància	0,15	0,25	0,09
Correlació Ítem-PT	0,90	0,30	0,40

PT: Puntuació Total

- 2.1. Hi ha un valor erroni (no possible) en la taula. Indiqui'l, justificant la seva resposta i ofereixi el valor correcte més proper a aquest valor erroni.
 El valor erroni és la variància de l'ítem 1. La variància d'un ítem dicotòmic (amb respostes 0 i 1), és igual a p (proporció de zeros) per q (proporció de uns). Si tenim una mostra de 10 subjectes, no hi ha possibilitat de que la variància d'un ítem sigui 0,15 per què no hi ha cap proporció de zeros i uns que doni aquest resultat. El valor correcte més proper a 0,15 serà el 0,16, per què aquest valor correspon a una $p=0,8$ i una $q=0,2$ (o a l'inrevés) ($0,8 \times 0,2 = 0,16$).
- 2.2. Quin és l'ítem que menys contribueix a la consistència interna del test? Per què?
 L'ítem que menys contribueix a la consistència interna del test és l'ítem 2, per què és el que té una correlació més baixa amb la puntuació total del test.
3. Un test té 10 ítems i se li afegixen 5 ítems paral·lels. Si la variància dels errors val 20 en el test allargat, quant valdrà en el test original?
 - a) 20
 - b) més de 20
 - c) menys de 20
 - d) No hi ha dades per a saber-ho

Indiqui l'alternativa correcta i justifiqui la teva resposta.

Contestar aquesta qüestió únicament amb la informació de que disposem als mòduls pot dur a respostes que de fet no son correctes encara que les justificacions puguin semblar adequades. Per aquest motiu serem molt laxes en la correcció d'aquesta pregunta, i valorarem més la coherència del raonament de justificació de la resposta que la resposta en sí.

Tenim en compte només la informació dels mòduls, la resposta més coherent seria la següent:

La resposta correcta és l'alternativa b), ja que el test allargat tindrà, segons la profecia de Spearman-Brown una fiabilitat més alta que el test original. Per tant, el test original tindrà una fiabilitat menor que el test allargat, i en conseqüència la variància dels errors serà major que en el test allargat, per

$$\rho_{xx'} = 1 - \frac{\sigma_e^2}{\sigma_x^2}$$

què, segons la Teoria clàssica dels tests, el coeficient de fiabilitat és igual a: Així, al augmentar la fiabilitat d'un test disminueix la variància dels errors, i al disminuir la fiabilitat augmenta la variància dels errors. Si el test allargat té una variància dels errors de 20, el test original té que tenir una variància del errors més elevada.

De totes maneres la resposta anterior no és correcta, per què de fet al allargar el test augmenta la variància de les puntuacions empíriques (σ_x^2) per què augmenta la mitjana de les puntuacions en el test i la amplitud total d'aquestes puntuacions (penseu que si en lloc dels 6 ítems de l'escala N del EPQ-RA ni hagués 9, l'amplitud total de les puntuacions aniria de 0 a 9, i per tant la seva variabilitat

augmentaria). Al augmentar la variància de les puntuacions empíriques també augmenta la variància tant de les vertaderes com dels errors, però el que succeeix és que aquesta variància dels errors augmenta en menys magnitud que les altres dues, i en conseqüència la proporció de variància del error de la

empírica $\left(\frac{\sigma_e^2}{\sigma_x^2}\right)$ disminueix, que és el que pronostica la profecia de Spearman-Brown.

4. Indiqui 4 formes de millorar la validesa d'un test, justificant la seva resposta.

També aquí hi ha moltes respostes possibles, sobre tot en quan a la seva expressió concreta, per tant la resposta que donem (una de les recollides de les vostres PACs), es una de les moltes possibles.

Per millorar la validesa d'un test tenim que tenir en compte quins factors afecten a aquesta validesa i de quina manera ho fan.

Les diferents maneres de millorar-la poden ser:

- conèixer la validesa de cada un dels ítems, i això ens determinarà la validesa de tot el test complet. Una manera de fer-ho és corregint i millorant cada un dels ítems o seleccionant els millors, d'aquesta manera s'identifica biaix i es pot corregir.

Les decisions preses a partir de l'anàlisi de continguts i del comportament empíric de cada ítem permet d'augmentar la validesa del test.

- Conèixer la fiabilitat del test, és a dir, conèixer l'absència d'errors en les puntuacions que ofereix el test. S'ha de tenir en compte que la validesa d'una prova sempre estarà limitada per la seva fiabilitat.

Per aquest motiu incrementar la llargada del test contribueix d'una manera indirecta de millorar la validesa.

- Conèixer les característiques de la mostra analitzada, de manera que és important fer un anàlisi de dues mostres independents, tenint en compte que ha de dur-se a terme amb el mateix mètode i amb la mateixa població de subjectes però canviant la mostra.

Quan duem a terme aquest procés el que estem fent és fer un disseny de validació creuada, i ens permet calcular l'estabilitat del coeficient de validesa, quan aquest té un resultat de coeficients estables llavors els resultats de validesa són generalitzables respecte a la mostra.

- Conèixer les condicions específiques d'avaluació, canvis en les condicions específiques d'avaluació poden canviar els resultats dels estudis de validació. Ja sigui la presència d'examinadors diferents, les condicions d'aplicació, instruccions ... és per aquest fet que és imprescindible utilitzar metodologia rigorosa en els estudis de validació.

Conèixer la metaanàlisi, és una eina útil per conèixer els límits de la generalització dels resultats de validació.

5. A partir de les dades obtingudes en la passació del EPQRA (veure fitxer "DadesEPQRA(05-06_1)"), obtingui en una única taula i **només per les puntuacions totals directes (de 0 a 6) de l'escala N de Neuroticisme** les següents distribucions:

- Distribució de freqüències, de freqüències acumulades, de puntuacions estandarditzades (z), de puntuacions T, de percentils, de puntuacions estandarditzades normalitzades (z_n), de Enneatipus i de Decatipus.

X	f _i	F _a	Z	T	P _c	z _n	E	D
0	118	118	-1,20	38	12	-1,17	3	3
1	87	205	-0,69	43	32	-0,47	4	5
2	69	274	-0,18	48	48	-0,05	5	5
3	78	352	0,33	53	63	0,33	6	6
4	51	403	0,84	58	76	0,71	6	7
5	53	456	1,35	63	86	1,08	7	8
6	41	497	1,86	69	96	1,75	9	9

Escala N	
Media	2,36
Desviación estándar	1,96
Cuenta	497

X: Puntuació directa

f_i : Freqüència absoluta

F_a : Freqüència acumulada

z: Puntuació estandarditzada: $Z_{xi} = \frac{X_i - \mu_x}{\sigma_x}$

Puntuació de 0: $Z_{xi} = \frac{X_i - \mu_x}{\sigma_x} = \frac{0 - 2,36}{1,96} = -1,20$

i així amb les altres puntuacions.....

T: Puntuació T: $T_i = 50 + 10 \times Z_{xi}$

Puntuació de 0: $T_i = 50 + 10 \times Z_{xi} = 50 + 10 \times (-1,20) = 38$

i així amb les altres puntuacions.....

P_{ci} : Percentil: $P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100$

Puntuació de 0: $P_{ci} = \frac{f_{aa} + 0,5f_i}{N} \times 100 = \frac{0 + 0,5 \times 118}{497} \times 100 = 11,87 \rightarrow 12$

i així amb les altres puntuacions.....

Z_n : Puntuació estandarditzada normalitzada

Puntuació de 0: Puntuació z que deixa per sota una $P = 0,12 \rightarrow Z_n = -1,17$

i així amb les altres puntuacions.....

E: Enneatipus: $Enneatipus = 5 + 2Z_n$

Puntuació de 0: $Enneatipus = 5 + 2 \times (-1,17) = 2,66 \rightarrow 3$ (arrodonint)

i així amb les altres puntuacions.....

D: Decatipus: $Decatipus = 5,5 + 2Z_n$

Puntuació de 0: $Decatipus = 5,5 + 2 \times (-1,17) = 3,16 \rightarrow 3$ (arrodonint)

i així amb les altres puntuacions.....

6. Com a resum de totes les activitats d'avaluació continuada realitzades amb el test EPQRA, faci un comentari general **breu** de les dades obtingudes en la nostre passació d'aquest test, on constin les característiques principals de la mostra de subjectes utilitzada, i de les dades obtingudes respecte a las característiques psicomètriques del test (de forma similar però més resumit a l'estudi de Ibañez i col·laboradors).

Considera, en general, que el test EPQ-RA compleix el requisit psicomètriques adequats que tot test té que complir?. Raoni la seva resposta

També en aquest cas hi ha moltes respostes possibles i adequades