

Transformació i interpretació de les puntuacions

Sergi Valero

PID_00198603

Índex

Introducció	5
1. Interpretació d'una puntuació	7
2. Transformació de les puntuacions	9
2.1. Percentils	9
2.2. Puntuacions estandarditzades	13
2.3. Puntuacions estandarditzades derivades	16
2.4. Puntuacions estandarditzades normalitzades	18
2.5. Normes cronològiques	20
3. Baremació	21
4. Equiparació de puntuacions	25
Resum	29
Bibliografia	31

Introducció

Interpretar el valor obtingut per una persona en una mesura psicològica, i fer-ho de manera raonable i raonada, no solament exigeix conèixer amb certa profunditat les característiques tècniques de l'instrument que és emprat, sinó també conèixer el constructe al qual es fa referència i entendre com aquest s'estructura i s'explica en un marc teòric que té el seu sentit, a la vegada, en un marc disciplinari concret, educacional, clínic o de qualsevol altra tipus.

És evident que un mòdul com aquest no pot abordar satisfactòriament un objectiu d'aquesta magnitud. Per aquesta raó, quan en aquest text es parli d'**interpretar una puntuació** s'haurà de considerar sempre des d'una perspectiva eminentment metodològica i matemàtica i, el més important, inferida d'acord amb la resposta d'un conjunt d'individus que, avaluats amb una mateixa mesura, proporcionaran el model de comportament que serà disposat en forma de norma, de marc de referència respecte al qual la puntuació d'una persona concreta serà interpretada. Sota aquesta perspectiva tindran el seu sentit les puntuacions transformades. Els percentils, les puntuacions estandarditzades i les normes cronològiques es recolliran aquí, exposant-ne les característiques més importants, la manera de calcular-les i, on es consideri oportú, les seves limitacions més rellevants.

Un segon terme de gran relleu, i que serà presentat en el tercer apartat, serà el concepte de *barem*. Es tracta d'una eina de gran importància en la qual s'estableix, de manera estructurada, una connexió entre la puntuació d'un individu i el comportament normatiu proporcionat per una mostra rellevant de subjectes, que, en alguns casos, apareixeran estratificats segons conveniència. Es dedicarà una atenció especial a la determinació de quina ha de ser aquesta mostra normativa. En aquest sentit s'exposaran els criteris de rellevància, representativitat i homogeneïtat com a factors fonamentals per a determinar la qualitat d'aquesta mostra de referència.

El mòdul finalitzarà exposant les diverses estratègies disponibles per a equiparar puntuacions d'instruments diferents que avaluen un mateix tret. Un conjunt d'estratègies, també metodològiques i matemàtiques, que són especialment interessants en els àmbits de l'avaluació en què els processos de memòria o aprenentatge de les persones participants poden esdevenir una limitació rellevant de les inferències que se'n deriven.

1. Interpretació d'una puntuació

Considerem una circumstància gens excepcional en què per a aprovar una assignatura, per exemple llengua anglesa, és necessari obtenir una puntuació de 5, assumint que som davant d'un examen clàssic en què la puntuació mínima possible és 0 i la màxima 10. Alguna vegada us heu preguntat quina propietat intrínseca té el número 5 per a considerar que en ell i per sobre d'ell es té assolida aquesta llengua i que, contràriament, quan s'està per sota, les nostres aptituds no satisfan el nivell exigít? Llevat que ens sentíssim confortables amb algun postulat més propi de la numerologia que no pas de la psicometria, caldria considerar que, *a priori*, el número 5 està mancat de qualsevol atribut intrínsec que el faci especial. I el mateix es podria afirmar d'una puntuació de 4 o de qualsevol altra. Si no se'ns informa de res més, haurem d'assumir que aprovar o suspendre depenent de si se supera o no una puntuació de 5 respon a una simple convenció. S'ha d'establir un punt de tall, perquè moltes vegades és necessari prendre una decisió respecte a qui supera un curs, i el valor de 5, xifra que ocupa la posició central, és el valor de referència emprat més freqüentment.

En d'altres circumstàncies, però, el criteri emprat per a discernir entre persones amb un tret determinat o sense no respon a un judici arbitrari o convencional. El criteri pot respondre a consideracions contrastades empíricament, sorgides generalment de la comparació de poblacions de subjectes ben diferenciats en el tret objecte d'interès i que permeten obtenir un punt de tall, una puntuació de referència dotada de capacitat discriminativa (Ramos-Quiroga *et al.*, 2009).

No obstant això, en moltes situacions no es disposa de criteris externs que proporcionin un marc d'interpretació respecte al qual donar rellevància o significació a una puntuació obtinguda en una mesura. De fet, es pot afirmar que la immensa majoria de les mesures psicològiques estan mancades de criteris externs i independents amb els quals inferir l'estatus actitudinal, aptitudinal o clínic d'un individu. Què es pot inferir llavors de la puntuació d'una persona en una mesura quan aquests referents no estan disponibles?

Mirar **dins** de l'instrument, és a dir, remetre's a alguna de les propietats estructurals de la mesura, és una alternativa possible. Una aproximació senzilla, per exemple, consisteix a atendre a la puntuació mínima i màxima que és possible assolir. Aquesta aproximació absoluta permet contextualitzar una puntuació tot indicant-ne l'excepcionalitat. Qualsevol manual d'ús o publicació de referència d'una mesura hauria de proporcionar informació clara respecte a aquests valors mínim i màxim. No obstant això, la senzillesa evident d'aquesta aproximació esdevé alhora la seva limitació més important: probablement tindrem menys problemes per a emetre un judici sobre una persona que obté en una mesura una puntuació de 17 que no pas sobre una persona que obté una

puntuació de 14 sabent que la puntuació màxima assolible és de 18. Ara bé, com de lluny un valor ha d'estar del valor màxim (o mínim) possible per a considerar que és molt o poc? No és possible donar una resposta operativa i universal a aquesta pregunta.

Una manera alternativa de procedir consisteix a comparar la puntuació obtinguda per una persona amb la puntuació que obtenen altres persones que han estat avaluades amb la mateixa mesura. La mitjana (i també la mediana) és una manera freqüent de resumir el comportament d'aquest grup de persones. D'aquesta manera, si una persona obté en una mesura d'aptituds mnèsiques una puntuació de 7 i la mitjana dels seus col·legues va ser de 9, avaluats seguint la mateixa pauta, podem afirmar que aquesta persona presenta una capacitat de memòria inferior a la que presenten els seus companys. I aquí descansa una de les claus d'aquesta aproximació relativa: la inferència que es fa d'una observació concreta és una funció de la tendència de comportament d'altres persones en la mateixa mesura. Reprendrem aquesta observació més endavant.

2. Transformació de les puntuacions

D'acord amb el que s'ha dit fins ara, és possible disposar d'estratègies que permeten la interpretació d'una mesura i que, a més, no impliquen cap manipulació de les dades obtingudes. No obstant això, no són les úniques estratègies a l'abast per a donar sentit a una puntuació determinada. Altres procediments, menys simples que els exposats fins ara, ofereixen a l'usuari una aproximació més precisa i informativa. No obstant això, aquests procediments impliquen la transformació de les puntuacions observades.

Transformar les puntuacions d'una mesura consisteix a aplicar una estratègia de codificació en què les puntuacions obtingudes en la mesura, les seves puntuacions directes, són recodificades (transformades) en un nou sistema de valors que en faciliten a l'usuari la interpretació.

Una propietat necessària de les puntuacions transformades és que no alteren l'escalament de les puntuacions directes, és a dir, respecten la diferent disposició dels diferents subjectes d'acord amb la magnitud de les seves puntuacions. I aquest manteniment de l'ordenació original es porta a terme, com s'ha esmentat, atorgant a les noves puntuacions un sentit eminentment pràctic. Els percentils, les puntuacions estandarditzades, les puntuacions estandarditzades derivades i normalitzades i les normes cronològiques seran les diverses estratègies de transformació que es tractaran a continuació.

2.1. Percentils

Per a descriure el concepte de *percentil*, també conegut com a *centil*, s'exposarà en primer lloc el concepte de *mediana*. Un cop ordenats els valors d'un conjunt d'observacions, la mediana serà el valor que ocupa la posició o ordre central. Un exemple senzill ajudarà en aquesta primera aproximació.

Exemple

Es disposa del conjunt de valors següent: 4, 5, 4, 3, 7, 8, 3, 1 i 6. En primer lloc, caldrà ordenar la seqüència: 1, 3, 3, 4, 4, 5, 6, 7, 8. La mediana d'aquest conjunt de valors és el 4 (el segon 4), ja que és el valor que ocupa la posició central de la distribució. La meitat de les persones són disposades a un costat i l'altre d'aquest valor.

En aquest exemple ha estat fàcil identificar el valor central de la distribució atès que el nombre d'observacions és senar. Tota distribució imparella té un valor que ocupa la posició central. I si la distribució de valors fos parella? Afegim a la distribució de valors ja coneguda i ordenada un nou valor, qualsevol: 1, 3, 3, 4, 4, 5, 6, 7, 8 i 9. En una distribució amb un nombre parell d'observacions no són un, sinó dos, els números que s'ubiquen en la posició central. En el cas que ens ocupa són els valors 4 i 5. Aplicant una mitjana aritmètica entre els

dos valors es conclourà que la mediana d'aquesta nova distribució de valors és 4,5. Tot és correcte: una mediana que prové de valors enters pot ser decimal i, a més, no ha coincidir necessàriament amb cap de les observacions originals.

No obstant això, què tenen a veure els percentils amb la mediana? Els percentils són una generalització de la mediana. La transformació basada en els percentils consisteix a assignar a cada puntuació directa una puntuació percentual d'acord amb la posició de les observacions dins del conjunt d'observacions. Si el valor percentual és del 25% s'estarà parlant del percentil 25. I sota aquesta denominació s'estarà identificant la puntuació directa que deixa per sota una quarta part de totes les observacions. O expressat d'una altra manera, el valor que amb ell i per sobre d'ell hi ha el 75% de totes les observacions. Si el percentatge ascendeix a un 50% es parlarà llavors del percentil 50, que es correspondrà amb el valor que ocuparà la posició central de la distribució. El percentil 50 és, efectivament, la mediana.

La primera formulació pràctica dels percentils es remunta al final del segle XIX i prové de l'explorador i científic Francis Galton, cosí de Charles Darwin que, entre altres aficions, es va mostrar força interessat per la mesura de multitud de trets antropomètrics. Avui en dia, i sense abandonar l'àmbit de la mesura de les propietats del cos, els percentils continuen essent una estratègia d'interès, per exemple, per als pares i mares que volen saber en quin grau els seus fills s'ajusten als altres nens d'edat comparable en termes d'alçada i pes.

Els percentils han esdevingut una mesura molt emprada no solament en el context de la mesura del pes o de l'alçada, sinó també en el context de les mesures psicològiques. La seva simplicitat i universalitat, ja que no cal tenir coneixement amplis de matemàtiques ni d'estadística en particular, n'ha estimulat sens dubte la difusió com a estratègia interpretativa. Exposem un cas pràctic que servirà per a exemplificar-ne el sentit i generalitzar-ne el càlcul.

En una mostra de 1.000 persones de nivell acadèmic, edat i gènere diferents (Gomà-i-Freixament *et al.*, 2008) es va administrar el *Zuckerman-Kuhlman Personality Questionnaire* (ZKPQ, Zuckerman, Kuhlman, Joireman, Teta i Kraft, 1993), un instrument que mesura cinc dimensions bàsiques de la personalitat normal.

Entre les cinc dimensions que el model permet mesurar, s'ha centrat l'atenció, per motius de simplificació, en una de les variables de personalitat, el neuroticisme¹. En aquest instrument, aquesta dimensió de personalitat és operativitzada mitjançant un total de 19 ítems de resposta tipus veritable-fals, en què el sumatori de tots (tenint en compte que alguns ítems són directes i altres inversos) esdevé una estimació del grau en què un individu tendeix a presentar aquest tret de personalitat. Cal estar atent a les peculiaritats de cada instrument per a saber com s'ha d'obtenir una puntuació en la variable d'interès. No sempre se segueix una mateixa pauta per a calcular les puntuacions totals.

En el cas que ens ocupa, la puntuació màxima assolible és 19, i la mínima 0. Una puntuació més elevada equival a una intensitat més elevada del tret. La distribució de valors absoluts i relatius d'aquest grup de 1.000 persones en aquesta variable apareix recollida en la taula 1.

⁽¹⁾El neuroticisme expressa la tendència de tota persona a mostrar-se alterada, amb tensió, preocupada, amb indecisió i amb labilitat emocional (estat d'humor freqüentment canviant).

Taula 1. Puntuacions directes, freqüències, percentatges i percentils

Puntuació en l'escala	Freqüència absoluta (fi)	Freqüència acumulada (fa)	Percentatge (%)	Percentatge acumulat (%)	Percentil (Pc)
0	21	21	2,1	2,1	0
1	55	76	5,5	7,6	3
2	49	125	4,9	12,5	8
3	77	202	7,7	20,2	13
4	63	265	6,3	26,5	21
5	70	335	7,0	33,5	27
6	76	411	7,6	41,1	34
7	67	478	6,7	47,8	42
8	68	546	6,8	54,6	48
9	72	618	7,2	61,8	55
10	59	677	5,9	67,7	62
11	60	737	6,0	73,7	68
12	57	794	5,7	79,4	74
13	44	838	4,4	83,8	80
14	43	881	4,3	88,1	84
15	41	922	4,1	92,2	89
16	38	960	3,8	96,0	93
17	23	983	2,3	98,3	97
18	11	994	1,1	99,4	99
19	6	1000	0,6	100,0	100

Si es pren per exemple la puntuació de 8, es pot observar que han estat 68 les persones que han presentat aquesta puntuació (columna de freqüència absoluta). Atès que són un total de 1.000 persones en aquesta mostra, es pot afirmar que un 6,8% dels participants tenen una puntuació directa de 8 (columna de percentatge). La penúltima columna, identificada com a percentatge acumulat, ha exigít que els valors directes fossin disposats de manera ordenada. Els percentatges corresponents a cada puntuació es van sumant i acumulant. En el cas de la puntuació directa 8, el percentatge acumulat és de 54,6. Això vol dir que el 54,6% de les persones de la mostra ha obtingut puntuacions de com màxim 8 unitats (8 o menys). Quin és el percentil associat a aquest valor directe? Atès que ha de ser el percentatge de persones que queden per sota d'aquest valor, caldrà observar que fins a arribar a una puntuació directa de 7 hi ha un percentatge acumulat del 47,8%. Aquest valor, un cop arrodonit a l'enter immediatament superior, passa a ser d'un 48%. La puntuació directa 8 s'associa a un percentil de 48.

Si la correspondència entre puntuacions directes i percentils s'enfoqués de manera inversa, en canvi, es podria formular una pregunta del tipus "a quin valor directe correspondria un percentil de 55?" Segons percentatges acumulats, fins a un valor directe de 8 hi correspon un percentatge acumulat de 54,6%. Un percentatge de 55, tot i que per poc, encara no s'ha produït. El valor directe següent és el 9, i fins a aquest s'ha acumulat el 61,8% de les observacions. En aquest cas sí, el 55% objecte

d'interès ja s'ha produït, ja que està inclòs en el 61,8%. Per tant, la puntuació directa 9 s'associa a un percentil de 55.

D'acord amb aquest raonament caldrà entendre per què una puntuació de 9 és també percentil 56. Però també 57, 58, 59, 60 i 61. Només quan s'arriba a la puntuació directa 10 el percentil passa a ser de 62. Aquesta circumstància, en què un mateix valor directe es vincula a més d'un percentil, és del tot plausible quan la variable objecte d'interès no disposa de prou valors per a cobrir les cent unitats per a les quals el percentil sí està preparat.

Cal tenir en compte que aquesta estratègia per a calcular percentils, basada en els percentatges acumulats, no és l'única possible. Hi ha altres aproximacions que també són acceptades. I diem aproximacions perquè l'estudiant haurà de tolerar que no totes arribin necessàriament a un resultat idèntic. Una estratègia alternativa a la ja exposada per a calcular un percentil, qualsevol, és la que es proposa a continuació:

$$P = \frac{fa + 0,5fi}{N} \times 100$$

Així doncs, i d'acord amb la taula 1, el percentil que s'associa a una puntuació directa de 9 és 58:

$$P_9 = \frac{546 + (0,5 \times 72)}{1000} \times 100 \approx 58$$

Atès que els percentils no tenen decimals, s'ha ignorat el decimal del resultat exacte obtingut (58,2). Cal recordar que, seguint la primera estratègia exposada per a calcular percentils, la puntuació directa 9 era, alhora, els percentils 55, 56, 57, 58, 59, 60 i 61. El valor obtingut en la fórmula anterior és només un dels valors, el central.

El primer dels avantatges de l'ús dels percentils, com ja hem indicat, és la simplicitat conceptual. Un segon avantatge és que per a calcular-lo i interpretar-lo correctament, atès que centra l'atenció només sobre l'ordre en què les observacions són disposades, esdevé irrellevant la manera com es distribueixen els valors objecte d'anàlisi. Això implica que, tot i que hi ha estratègies de transformació que requereixen una distribució de valors essencialment ajustada a una corba normal, més endavant es veurà, en el cas dels percentils és possible abordar qualsevol tipus de distribució. Aquesta característica és especialment interessant en psicologia, en què moltes de les variables d'interès rarament s'ajusten a una distribució normal de valors.

Com s'ha vist, els percentils consisteixen a dividir una distribució de valors en 100 posicions ordinals. No obstant això, hi ha altres abordatges que simplifiquen la quantitat d'ordres possibles i que poden ser suficients per a resumir el comportament d'un conjunt de valors. Els més comuns són:

Lectura de la fórmula

f_g : freqüència acumulada prèvia a la puntuació directa de la qual es vol calcular el percentil.
 f_i : freqüència absoluta en la qual es troba la puntuació directa.
 N : nombre de persones que constitueixen la mostra.

Enllaç recomanat

En la Viquipèdia (en anglès) i en d'altres recursos a Internet, trobareu estratègies alternatives, també vàlides, per a calcular els percentils.

- Els **quartils**. Divideixen una distribució en quatre parts. Se solen identificar com a *quartil 1*, *quartil 2* i *quartil 3*. Coincideixen amb els percentils 25, 50 i 75, respectivament.
- Els **quintils**, que dividint la distribució en cinc parts, s'equiparen als percentils 20, 40, 60 i 80.
- Els **decils**. Representen una divisió en deu parts i es corresponen amb els percentils 10, 20, 30, 40, etc.

2.2. Puntuacions estandarditzades

Si bé la mitjana és un estadístic de tendència central que permet expressar el comportament mitjà d'un conjunt d'observacions, la desviació típica (o estàndard) permet quantificar com de diferents són aquestes observacions. Aquest índex de dispersió expressa la discrepància mitjana entre una observació qualsevol i la mitjana del conjunt d'observacions. La mitjana i la desviació típica esdevenen la matèria primera imprescindible per a calcular una puntuació estandarditzada. Si bé els percentils dirigeixen l'atenció cap a la posició que ocupa un individu respecte a un grup de referència, en el cas de les puntuacions estandarditzades l'interès es focalitza en la discrepància que presenta aquest individu respecte a la forma de comportament mitjà del grup, la seva mitjana.

La puntuació estandarditzada o puntuació típica, és la que permet expressar quantes desviacions típiques una observació se situa per sobre o per sota d'una mitjana. Aquest procés d'estandardització té el seu sentit quan la distribució de valors objecte d'interès s'ajusta essencialment a una corba normal. El valor resultant d'aquesta transformació es constitueix en una nova puntuació que rep el nom de *puntuació estandarditzada z*. El seu càlcul respon a la fórmula següent:

$$Z_x = \frac{X - \bar{X}}{S_x}$$

Cal tenir present que aquesta puntuació *z* tindrà sempre una mitjana de 0 i una desviació estàndard d'1. En la taula 2, i reprenent l'estudi sobre neuroticisme en una mostra de 1.000 persones:

Taula 2. Puntuacions directes, freqüències i puntuacions estandarditzades *Z*

Puntuacions directes	Freqüència absoluta	Puntuacions estandarditzades <i>z</i>
0	21	-1,72
1	55	-1,51
2	49	-1,30
3	77	-1,09

Lectura de la fórmula

X: puntuació directa.
X̄: mitjana de la mostra.
S_x: desviació típica de la mostra.

Puntuacions directes	Freqüència absoluta	Puntuacions estandarditzades z
4	63	-0,88
5	70	-0,66
6	76	-0,45
7	67	-0,24
8	68	-0,03
9	72	0,18
10	59	0,39
11	60	0,61
12	57	0,81
13	44	1,03
14	43	1,24
15	41	1,45
16	38	1,67
17	23	1,88
18	11	2,09
19	6	2,30

Per a poder calcular els valors z, ubicats en la darrera columna de la taula, és necessari saber la mitjana i la desviació típica de la variable en aquesta mostra: la mitjana era 8,14 i la desviació típica 4,72. Així doncs, en el cas concret d'una persona que ha obtingut una puntuació directa de 8, la seva puntuació z serà:

$$-0,03 = \frac{8 - 8,14}{4,72}$$

Què significa aquest valor de -0,03? Atès que la mitjana de z sempre és 0 i que la seva desviació típica és també sempre 1, observar una puntuació típica de -0,03 significa que els subjectes que tenen una puntuació directa de 8, i que són un total de 68 persones, s'ubiquen a 0,03 desviacions típiques de la mitjana de la mostra. Sabem, a més, que aquest valor directe s'ubica per sota de la mitjana perquè la seva puntuació típica és negativa. Necessàriament, tots els subjectes que presenten puntuacions directes que es disposen per sota de la mitjana obtindran puntuacions z negatives, mentre que els que s'ubiquen per sobre de la mitjana presenten puntuacions típiques positives. Si ens fixem en la part final de la taula 2, veurem que les 6 persones que obtenen una puntuació de 19 tenen una puntuació típica de 2,30, la qual cosa implica que aquestes s'ubiquen a més de 2 desviacions típiques per sobre de la mitjana. En canvi, les

21 persones que obtingueren una puntuació directa de 0 presenten una puntuació típica de $-1,72$. Són persones que queden, per tant, a $1,72$ desviacions per sota de la mitjana.

Una pregunta important que ens podem formular és si distanciar-se en $1,72$ o $2,30$ desviacions típiques d'una mitjana és molt o poc. És clar que allunyar-se d'una mitjana en $2,30$ desviacions típiques és allunyar-se més que $1,72$ vegades. En la taula 2 podeu veure que en aquesta mostra les persones amb la màxima puntuació es disposen més allunyades de la mitjana que les persones que obtenen la puntuació més baixa possible. Ara bé, és molt fer-ho $2,3$ vegades?

En l'estudi de Valero *et al.* (2012) dues mostres de subjectes adults són comparades. Una constituïda per 217 persones adultes afectades per un trastorn per dèficit d'atenció amb hiperactivitat (TDAH), i l'altra per 434 subjectes control aparellats per edat i gènere. Es van contrastar diverses variables de personalitat, però per simplificar l'exposició se centrà de nou l'atenció sobre la dimensió de neuroticisme. La taula 3 exposa les puntuacions típiques en aquesta variable per a cadascun dels dos grups d'estudi.

Taula 3. Puntuacions directes i estandarditzades per a una mostra de persones amb TDAH i controls

Puntuacions directes	Puntuació estandarditzada z	
	TDAH	Controls
0	-2,54	-1,67
1	-2,38	-1,46
2	-2,16	-1,25
3	-1,94	-1,04
4	-1,73	-0,83
5	-1,51	-0,62
6	-1,30	-0,41
7	-1,08	-0,19
8	-0,87	0,02
9	-0,65	0,23
10	-0,43	0,44
11	-0,22	0,65
12	0	0,86
13	0,21	1,08
14	0,43	1,29
15	0,65	1,50
16	0,86	1,71

Puntuacions directes	Puntuació estandarditzada z	
17	1,08	1,92
18	1,29	2,14
19	1,51	2,35
	Mitjana = 12,01	Mitjana = 7,91
	Desviació típica = 4,63	Desviació típica = 4,72

Com es pot veure a peu de taula, els subjectes afectats pel TDAH presenten una mitjana en la mesura de neuroticisme superior a la presentada pels subjectes control (12,01 enfront de 7,91). Es tracta d'un resultat esperable tenint en compte que una és mostra psiquiàtrica i l'altra no. Les desviacions típiques respectives, en canvi, i en aquest estudi, són força comparables. Si per a cada puntuació directa es comparen les puntuacions típiques respectives de les dues mostres es veurà que no coincideixen en cap cas. De fet, es mostren clarament diferenciades.

En el cas de les persones afectades pel TDAH les puntuacions tipificades tendeixen a estar sistemàticament desplaçades cap a la part baixa de la distribució respecte a les presentades pels subjectes control. La discrepància entre les dues mitjanes condiona aquest comportament. Si bé una puntuació directa de 0 en el cas dels subjectes control implica ubicar-se a 1,67 desviacions típiques per sota de la mitjana, en el cas dels pacients afectats de TDAH, la mateixa puntuació directa els allunya més de 2 desviacions típiques i mitja de la mitjana respectiva. Tot depèn del grup en què s'estigui ubicat.

I aquí descansa una de les consideracions més rellevants que cal tenir presents quan s'interpreten puntuacions estandarditzades: les inferències a les quals podem arribar mitjançant l'ús d'aquestes puntuacions depenen de la mostra que és emprada, ja que és d'aquesta que s'obtenen la mitjana i la desviació típiques que permeten el càlcul de les puntuacions transformades. Triar acuradament la mostra de referència esdevé imprescindible. Més endavant, quan es parli del concepte de *barems*, es reprendrà aquesta idea.

2.3. Puntuacions estandarditzades derivades

Encara que intrínsecament no és un problema, les puntuacions z s'expressen amb valors negatius i decimals. El sentit bàsic de les puntuacions estandarditzades derivades és transformar les puntuacions típiques d'acord amb un sistema de codificació que elimina aquests valors negatius i decimals. Una de les puntuacions típiques derivades més comunament emprades és la puntuació *T* de McCall. La seva formulació és senzilla:

$$T = 50 + 10z$$

Consisteix a partir d'una constant, en aquest cas 50, i sumar-hi deu vegades el valor de la puntuació típica z . Per a interpretar degudament aquesta puntuació només cal tenir present que la mitjana d'aquesta puntuació és 50 i la seva desviació típica 10. D'aquesta manera, per exemple, una puntuació z de 2 tindrà un valor de 70 un cop transformada en puntuació T . En tots dos casos (z i T), i necessàriament, seria indicador d'una puntuació directa que s'ubica per sobre de la mitjana corresponent en 2 desviacions típiques.

La taula 4 recull els valors de la taula 2 afegint les puntuacions T corresponents:

Taula 4. Puntuacions directes, freqüències i puntuacions estandarditzades z i T

Puntuacions en l'escala	Freqüència absoluta	Puntuació estandarditzada z	Puntuació estandarditzada derivada T
0	21	-1,72	33
1	55	-1,51	35
2	49	-1,30	37
3	77	-1,09	39
4	63	-0,88	41
5	70	-0,66	43
6	76	-0,45	46
7	67	-0,24	48
8	68	-0,03	50
9	72	0,18	52
10	59	0,39	54
11	60	0,61	56
12	57	0,81	58
13	44	1,03	60
14	43	1,24	62
15	41	1,45	65
16	38	1,67	67
17	23	1,88	69
18	11	2,09	71
19	6	2,30	73

És possible trobar puntuacions típiques derivades de naturalesa diversa, i és que, estrictament, tothom se'n pot crear una de pròpia, ja que una puntuació derivada és formulada d'acord amb criteris arbitraris. Exemples diferents els trobarem en el test d'intel·ligència Stanford-Binet, que empra una mitjana de 100 i una desviació típica de 16; la *Wechsler Adult Intelligence Scale* (WAIS), amb

una mitjana també de 100, però amb una desviació típica de 15; o el *Minnesota Multiphasic Personality Inventory* (MMPI), que, coincidint amb l'escala de McCall, s'interpreta entorn a un mitjana de 50 i una desviació típica de 10.

2.4. Puntuacions estandarditzades normalitzades

Si s'assumeix que la distribució de valors que és objecte d'anàlisi s'ajusta a una corba normal, és possible transformar els percentils en les puntuacions estandarditzades que es correspondrien si la distribució fos, efectivament, normal. Aquestes puntuacions rebrien el nom de *puntuacions estandarditzades normalitzades*.

Taula 5. Puntuacions directes, freqüències, puntuacions estandarditzades z i T i normalitzades

Puntuació en l'escala	Freqüència absoluta (f_i)	Freqüència acumulada (F_o)	Percentatge acumulat	Percentil (P_c)	Puntuació estandarditzada z	Puntuació estandarditzada normalitzada
0	21	21	2,1	2	-1,72	-2,05
1	55	76	7,6	3	-1,51	-1,88
2	49	125	12,5	8	-1,30	-1,41
3	77	202	20,2	13	-1,09	-1,13
4	63	265	26,5	21	-0,88	-0,81
5	70	335	33,5	27	-0,66	-0,61
6	76	411	41,1	34	-0,45	-0,41
7	67	478	47,8	42	-0,24	-0,20
8	68	546	54,6	48	-0,03	-0,05
9	72	618	61,8	55	0,18	0,13
10	59	677	67,7	62	0,39	0,31
11	60	737	73,7	68	0,61	0,47
12	57	794	79,4	74	0,81	0,64
13	44	838	83,8	80	1,03	0,84
14	43	881	88,1	84	1,24	0,99
15	41	922	92,2	89	1,45	1,23
16	38	960	96,0	93	1,67	1,48
17	23	983	98,3	97	1,88	1,88
18	11	994	99,4	99	2,09	2,33
19	6	1.000	100,0	100	2,30	—

D'acord amb la taula 5, si es pren per exemple la puntuació directa 9 s'observarà que, com ja se sabia, la seva puntuació estandarditzada z és de 0,18, i el seu percentil de 55. El pas següent és cercar una taula estandarditzada de puntuacions normals, també coneguda com a *taula de puntuacions z*. A Internet en trobareu moltes². La taula 6 és l'extracte d'una d'aquestes taules. Si s'hi cerca una proporció de 0,55 es localitzarà el valor z corresponent.

⁽²⁾En un cercador teclegeu, per exemple, *z tables*.

En el cas que ens ocupa, la proporció més propera és 0,5517, i en conseqüència el valor z que s'hi vincula és 0,13. Qualsevol percentil, en forma de proporció, té associat un valor z i viceversa.

Taula 6. Extracte taula puntuacions z

z	0,0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224

Reiterem que el valor obtingut en una taula de valors z assumeix que la distribució de valors que és objecte d'interès s'ajusta bàsicament a una corba normal. La discrepància observada entre les puntuacions 0,18 i 0,13 (vegeu la taula 5), o genèricament entre qualsevol puntuació estandarditzada calculada i la corresponent puntuació estandarditzada normalitzada, respon a la discrepància entre la distribució de valors empírics i una distribució teòrica normal. Cal tenir present que la discrepància entre les dues distribucions de valors tipificats no serà l'excepció, sinó la norma. Com ja s'ha esmentat amb anterioritat, són moltes les variables que no s'ajustaran a una (perfecta) distribució normal. Per a les que ho facin essencialment, heus aquí una manera possible de transformar una puntuació.

De la mateixa manera que en el cas de les puntuacions estandarditzades derivades és possible disposar també de puntuacions estandarditzades normalitzades derivades. En psicologia, les més habituals són les estannines o enneatipus i els decatipus. La formulacions respectives són:

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

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

En què la z és la puntuació estandarditzada normalitzada.

Assumint una distribució normal de valors, si cliqueu en un cercador termes com *z distribution percentile* podreu veure de manera gràfica la correspondència que s'estableix entre alguns dels procediments de transformació presentats en aquest mòdul.

2.5. Normes cronològiques

Una altra possible transformació de les puntuacions directes consisteix a fer-ho en forma de normes cronològiques o d'edats. L'estratègia a seguir consisteix a administrar a diversos grups d'edat, representatius d'una determinada població, una mateixa mesura, recollint per a cada grup d'edat la mitjana respectiva. Posteriorment, quan un individu és avaluat, la seva puntuació és transformada en una edat, l'edat que en la població de referència s'aproximaria a la puntuació obtinguda. La idea bàsica que justifica aquesta fórmula és simple: assumir que hi ha formes de comportament que tendeixen a ser típiques entorn d'una edat. Així doncs, si després de seleccionar una mostra representativa d'una determinada població, s'administra una determinada mesura als nens de sis anys i la mitjana obtinguda és de 12, i la mateixa mesura administrada als nens de vuit anys és de 17, si un nen concret obté una puntuació de 17 s'afirmarà que en aquesta mesura el nen té una edat de vuit anys (encara que cronològicament en tingui set).

Un dels contextos en què aquesta estratègia ha esdevingut més popular ha estat en parlar del quocient d'intel·ligència. La seva formulació és:

$$QI = \frac{EM}{EC} \times 100$$

La transformació en edats és una estratègia antiga, probablement tant com ho ha estat la mesura dels constructes psicològics. No obstant això, no es pot deixar d'esmentar que aquesta transformació s'acompanya de limitacions que no es poden menystenir i que al llarg del temps n'han anat reduint la presència en la bibliografia. Una de les seves limitacions més rellevants és que el model assumeix que els canvis que es van produint al llarg del temps, segons s'analitzen diversos grups d'edat, són canvis essencialment lineals, és a dir, que el canvi produït entre les persones de sis a set anys és comparable als canvis que es produeixen entre les persones de vuit respecte les de nou anys. Malauradament, assumir aquest principi és excessiu per a molts processos psicològics. Canvis produïts entre determinades edats cronològiques no tenen necessàriament la mateixa rellevància o significat al llarg de totes les edats mentals ni, per descomptat, a l'inrevés.

Lectura de la fórmula

EM: edat mental.
EC: edat cronològica.

Lectura recomanada

A Fancher (1987) trobareu un text interessant en el qual s'aborda no solament el concepte d'edat mental, sinó que també es fa un repàs històric del concepte d'intel·ligència i també d'algunes controvèrsies entorn de la seva mesura.

3. Baremació

La baremació, també coneguda com a *escalament d'una mesura*, és un element de gran rellevància en el procés d'elaboració i ús d'un instrument de mesura. Com s'ha vist en aquest mòdul, la transformació d'una puntuació ajuda l'usuari a injectar significat a un determinat valor d'acord amb el comportament d'un conjunt d'observacions. Si l'instrument disposa d'un manual, aquest hauria de reportar els barems de la mesura, amb les transformacions necessàries, si s'escau, a més de les indicacions bàsiques per a interpretar degudament qualsevol puntuació. En d'altres casos no hi haurà manual de referència, però sí probablement una o més publicacions periòdiques a les quals ens podrem remetre i trobar la informació necessària.

Independentment del tipus de publicació amb què el barem d'una mesura es reporti, perquè tota interpretació que se'n derivi es faci segons un estàndard de qualitat elevat, cal que la mostra que proporciona les dades que permeten satisfer una transformació (per exemple, mitjanes i desviacions típiques) i que rebrà el nom de *mostra normativa*, sigui una mostra rellevant, representativa i homogènia.

Rellevant perquè cal garantir que la mostra que conforma els barems d'un instrument fa referència a una població que és significativa d'acord amb la satisfacció d'uns objectius específics. Tal com s'ha vist durant l'exposició de les puntuacions estandarditzades, el sentit que pot tenir una determinada puntuació tipificada ha de ser entès sempre dins del marc de la població que proporciona els paràmetres de referència. No hi ha una població *a priori* ideal sobre la qual es pugui desenvolupar el barem d'un instrument de mesura. La població ideal depèn del tipus de persona que es vol descriure i, sobretot, dels termes en què es proposi fer-ho.

Probablement s'estaria d'acord que si es vol inferir alguna cosa rellevant sobre el grau de coneixement dels joves d'un institut sobre la Revolució Industrial seria convenient poder disposar d'una mostra àmplia d'altres joves provinents d'altres instituts amb els quals contrastar-ne el rendiment. No obstant això, no sempre és possible identificar una única població respecte a la qual es pugui inferir satisfactòriament l'estat d'un individu.

D'acord amb la mateixa línia argumental exposada en parlar del neuroticisme i TDAH, si es vol interpretar la intensitat de la simptomatologia depressiva d'una persona mitjançant un determinat instrument clínic, una puntuació concreta no té la mateixa consideració si els barems emprats provenen de població general, que si provenen de diversos centres d'atenció primària o

d'unitats d'hospitalització psiquiàtrica. *A priori* les tres poblacions són igualment vàlides, en tant que en qualsevol d'elles pot ser pertinent per a obtenir dades normatives i interpretar la gravetat de la simptomatologia depressiva d'aquesta persona. Ara bé, no és el mateix ubicar-se en un percentil de 70 prenent com a mostra normativa població general que població clínica. L'excepcionalitat (o no) de la simptomatologia d'aquest individu dependrà necessàriament de quina sigui la població de referència que s'expressa mitjançant els barems, i això condicionarà la interpretació que se'n pugui derivar i les decisions que caldria prendre respecte a aquesta (McIntire i Miller, 2007). Cal delimitar degudament, en forma d'objectius de treball clars, què ha de ser avaluat i, sobretot, amb quina finalitat.

S'ha afirmat també que la mostra constitutiva dels barems ha de ser **representativa**. Un cop definida quina ha de ser la població diana, l'estratègia emprada per a seleccionar les persones de la població que formaran part de la mostra normativa esdevé ara el focus d'atenció principal. D'acord amb els estàndards que proporciona l'American Psychological Association (APA, 1999), cal exposar amb cura i claredat quina ha estat la població objectiu i, alhora, la mostra que se n'ha seleccionat. Només així és possible delimitar el tipus d'inferència a la qual es pot arribar quan s'empren uns determinats barems.

Un mostreig probabilístic és la millor estratègia metodològica disponible per a assegurar l'obtenció d'una mostra representativa, ja que només és en aquest context aleatori en el qual qualsevol individu de la població disposa d'una probabilitat idèntica de formar part de la mostra. I tal com exposa Aiken (1996), el mostreig aleatori estratificat hauria de ser l'estratègia d'elecció. Aquest tipus de mostreig hauria de proporcionar la deguda representativitat per a tots els paràmetres que han estat considerats rellevants. En aquest sentit, l'edat i el gènere són alguns dels factors que més freqüentment articulen un barem, però qualsevol altre variable pot ser objecte d'interès i per tant d'estratificació. L'àmbit d'estudi i els objectius que han de ser resolts condicionaran el tipus de variables que se sotmeten a escalament i el grau d'estratificació amb el qual aquestes variables s'haurien d'expressar.

Malauradament no sempre s'explicita en els textos especialitzats, però no es pot deixar d'esmentar que no sol ser fàcil satisfer un mostreig estrictament probabilístic. Són moltes les condicions que, en context aplicat, minven l'èxit d'un mostreig aleatori. Considerem per exemple la presència de persones institucionalitzades a les quals difícilment es tindrà accés (presos, persones hospitalitzades, etc.), les que no tenen les aptituds necessàries per a donar resposta a una determinada mesura (analfabetes, amb problemes de comprensió, etc.), les avaluades en contextos en què hi pot haver una alt desig social (selecció de personal, presència de preguntes compromeses, etc.), persones que simplement no volen col·laborar, etc. Els barems de la majoria d'instruments d'avaluació, si no tots, estan subjectes a circumstàncies que limiten la generalització dels seus valors. L'esforç de tot investigador hauria de dirigir-se a mi-

nimitzar o atenuar en la mesura del possible aquests biaixos i, en tot cas, a informar sempre detingudament els usuaris de l'instrument sobre què s'ha fet i com s'ha resolt.

Quan es parla de l'**homogeneïtat** de la mostra es fa referència a la necessitat de controlar la presència d'efectes de confusió o estranys que puguin esbiaixar o fins i tot invalidar les interpretacions que se'n derivin. Les persones que seran avaluades mitjançant una determinada mesura haurien de ser comparables a les persones que van formar part de la mostra de participants que en van constituir el barem. Discrepàncies consistents entre les dues mostres poden derivar en inferències que sobreestimarien o infraestimarien l'estat real de les persones avaluades.

Aquestes mateixes consideracions justifiquen la necessitat que hi hagi, també, una actualització contínua de les dades que estructurin els barems. Caldria poder disposar de barems que se sotmetin a revisions amb una determinada regularitat. I és que caldria assumir per defecte que les dades contingudes en un barem caduquen. Les variables que són objecte d'interès són dinàmiques i, per tant, canviant, modulades pel decurs dels temps. Són sensibles a les transformacions científiques, socials i culturals que es van esdevenint i, per aquesta raó, tot barem hauria de poder reflectir de manera fidel l'estat actual del concepte que és objecte de consideració. Tot i que les consideracions teòriques i empíriques són diferents, emprar actualment barems de qualitat de vida en mostres i sota criteris de fa trenta anys, per exemple, seria comparable a emprar barems d'alçada incloent-hi persones avaluades durant els anys seixanta. En qualsevol dels dos casos s'estaria cometent un error o, en el millor dels casos, un biaix difícilment assumible.

L'adaptació de l'instrument a les característiques lingüístiques i culturals de les persones que han de ser avaluades ha de ser també motiu d'una necessària atenció. En la majoria de casos, els instruments d'avaluació que són emprats en psicologia en el nostre entorn cultural tenen l'origen en el món anglosaxó. L'avantatge que presenta aquesta circumstància és que permet que, un cop traduït un instrument, esdevingui més fàcil l'homologació de l'instrument des d'un punt de vista internacional, alhora que facilita la satisfacció d'estudis de tipus transcultural en els quals un mateix tret es contrasta entre diversos grups culturalment diferents. No obstant això, cal tenir present que traduir un instrument és condició necessària però no sempre suficient. Seguint les indicacions dels estàndards de l'American Psychological Association (APA, 1999), cal que hi hagi una adaptació cultural adequada que sigui sensible a les peculiaritats dels individus que seran avaluats amb l'objectiu. Una adaptació correcta de la mesura constitueix un objectiu dirigit, de nou, a proporcionar estimacions sense biaix i que, per tant, permeti inferir amb garanties suficients de validesa i fiabilitat l'estat dels individus.

En la configuració del barem d'un instrument és possible emprar qualsevol de les estratègies exposades en aquest mòdul. Les dades descriptives, disposades en taules, es poden expressar de manera transformada, per exemple mitjançant percentils, i agrupades segons estratificacions per edat (Rodríguez *et al.*, 2012). En d'altres casos, els percentils es disposaran sense estratificació (Alegret *et al.*, 2012; Romero *et al.*, 2000). En algunes publicacions s'oferiran mitjanes i desviacions típiques que permetran, si es vol, una transformació en puntuacions estandarditzades. I ho poden fer d'acord amb dos factors, per exemple, l'edat i el gènere (Gomà-i-Freixenet i Valero, 2008), o bé sense cap estratificació (Aluja *et al.* 2008). Altres proporcionaran puntuacions estandarditzades derivades, en forma de puntuacions *T*, i de manera estratificada (Yueh-Hsien *et al.*, 2012). Les combinacions són evidentment múltiples, no són més que alguns exemples dels molts que poden ser reportats. Queda en mans de l'investigador proporcionar la informació necessària per a justificar una opció o una altra.

4. Equiparació de puntuacions

Equiparar les puntuacions de dues o més mesures consisteix a establir una correspondència entre les seves puntuacions. Aquesta estratègia ha de permetre que, recollida la puntuació en una de les mesures, sigui possible estimar la puntuació de les altres mesures.

Un dels contextos en què pot ser especialment interessant resoldre equiparacions entre tests és aquell en què és necessari mesurar les aptituds de persones de manera repetida i en què, per tant, es poden produir processos d'aprenentatge o memòria, o simplement cansament, que podrien conduir cap a una més que probable estimació esbiaixada de les capacitats de les persones. En aquestes condicions seria interessant poder disposar de dues o més formes alternatives de la mesura que en permetessin l'ús indistintament. Tanmateix, les condicions que s'hauran de satisfer per a una equiparació adequada són les següents:

- a) Que es mesuri la mateixa característica.
- b) Que les dues mesures siguin igualment fiables.
- c) Que la transformació sigui invertible, és a dir, la transformació ha de ser possible de A a B , però també de B a A .

Dos tipus de transformacions són possibles, l'equiparació horitzontal i la vertical. El primer cas s'esdevindria quan el tret d'interès hagués de mantenir la seva dificultat entre les diferents mesures alternatives. Un sol grup de subjectes és el que, generalment, proporciona les dues o més mesures. En el cas vertical, en canvi, s'assumeix que la dificultat entre formes alternatives és diferent. En aquesta darrera circumstància, són avaluats dos o més grups de persones que presenten capacitats diferenciades.

Els diversos dissenys que permeten crear equiparacions entre mesures són els següents:

- els de grup únic,
- els de grups equivalents,
- els tests d'ancoratge.

En el cas dels dissenys de **grup únic**, un mateix grup de persones extret aleatòriament de la població destinatària de les mesures equivalents és avaluat mitjançant les diverses mesures d'interès. Com a estratègia de control, per a evitar processos d'aprenentatge i memòria, i per tant un possible biaix com

a conseqüència d'haver administrat primer una forma alternativa i posteriorment l'altra, es recomana que les mostres responguin a les mesures de manera contrabalancjada. Així, es pot dividir la mostra aleatòriament en dos grups i administrar a cada subgrup les diverses formes alternatives en un ordre diferent, per exemple, en un grup la seqüència *A-B*, i en l'altre la *B-A*.

En els dissenys que impliquen **grups equivalents** són dos els grups que se seleccionen aleatòriament, però en aquest cas s'administra una mesura a la primera mostra, i a la segona mostra la segona mesura. Poder assumir amb garanties que les dues mostres de subjectes són comparables, per exemple, mitjançant una assignació aleatòria dels subjectes, és condició necessària per a assegurar la validesa del procediment.

L'avantatge d'emprar dos grups de persones és que, en comparació amb l'estratègia anterior, redueix l'efecte de cansament i no exposa els participants a processos d'aprenentatge o memòria que podrien ser arrossegats entre les diverses mesures, ja que els participants només s'exposen a una d'aquestes. La limitació més important d'aquesta aproximació és que, en general, exigeix incloure una quantitat de participants superior a la necessària en el cas del primer disseny.

Els **tests d'ancoratge** parteixen, també, de disposar de dos grups de participants en què cada un rep una de les mesures dirigides a ser equiparades. En aquest cas, però, a cada mostra se l'administra alguns ítems que formen part de la mesura que s'administra a l'altra mostra. Aquests ítems comuns esdevenen així ítems d'ancoratge. L'estratègia és especialment interessant quan no és possible garantir que els dos grups de persones siguin comparables, situació freqüent en un context aplicat en què no ha estat possible generar els dos grups de persones de manera aleatòria. Sota aquesta circumstància, els ítems d'ancoratge proporcionaran una estimació del grau en què les dues mostres esdevenen comparables.

Independentment de quina de les tres estratègies s'hagi emprat per a satisfer l'equivalència entre mesures, dos són els mètodes més comuns que es poden executar per a resoldre l'equiparació entre mesures: l'equipercentil i la transformació lineal.

L'**equiparació equipercentil**, que sol ser l'estratègia més habitual, consisteix a considerar equivalents les puntuacions directes que es vinculen a un mateix percentil. Així doncs, si les puntuacions directes 10 i 14 obtingudes amb la mesura *A* i *B*, respectivament, s'associen a un percentil 30, es considerarà que les dues puntuacions són equivalents. El mateix raonament seria aplicable a la resta de parells de puntuacions directes que convergissin en un mateix percentil.

Lectura recomanada

En el text de Muñiz (2003) trobareu més informació respecte a aquesta estratègia.

Les consideracions que s'han de tenir en compte a l'hora d'aplicar aquesta estratègia són les mateixes que ja s'han exposat en parlar dels percentils. Es tracta, doncs, d'una estratègia de fàcil comunicació i execució.

La **transformació lineal** consisteix a equiparar puntuacions estandarditzades. Concretament consisteix a equiparar les puntuacions directes que convergeixen en una mateixa puntuació típica. Expressat d'una altra manera, i d'acord amb la fórmula que permet calcular una puntuació típica z , es pot dir que:

$$\frac{X - \bar{X}}{S_x} = \frac{Y - \bar{Y}}{S_y}$$

D'on si s'aïlla Y :

$$Y = \frac{S_y}{S_x}(X - \bar{X}) + \bar{Y}$$

Exemple

Considerem l'exemple de dues mostres equivalents de persones extretes aleatòriament d'una determinada població (disseny de grups equivalents). La primera mostra obté en una mesura d'aptituds aritmètiques (X) una mitjana de 10 i una desviació típica de 4. La segona mostra, en la segona mesura també d'aptituds aritmètiques (Y), obté una mitjana de 8 i una desviació típica de 2. Segons aquesta informació, una puntuació de 9 en la mesura A , a quina puntuació directa correspondria en la mesura B ?

$$Y = \frac{4}{2}(9 - 10) + 8 = 6$$

Obtenir un 9 en la mesura A és equivalent a obtenir una puntuació de 6 en la mesura B .

En el cas del disseny d'un sol grup en el qual s'ha emprat una estratègia de contrabalanceig, cal tenir en consideració l'ordre en què es van administrar les dues mesures (ordre $A-B$ i ordre $B-A$). Sota aquesta circumstància es calcula la desviació típica i mitjana en cadascun dels grups de balanceig. Així doncs, començant per la desviació típica:

$$S_x = \sqrt{S_{x1}^2 + S_{x2}^2}$$

$$S_y = \sqrt{S_{y1}^2 + S_{y2}^2}$$

En què els estadístics inclosos en els radicals són les variàncies corresponents al primer i segon ordre en l'administració de les dues mesures.

En el cas de les mitjanes:

$$\bar{X} = \frac{\bar{X}_{x1} + \bar{X}_{x2}}{2}$$

$$\bar{Y} = \frac{\bar{X}_{y1} + \bar{X}_{y2}}{2}$$

Essent els estadístics dels numeradors les mitjanes corresponents al primer i segon ordre de les dues mesures. Calculades les mitjanes i desviacions típiques, només resta aplicar la fórmula ja coneguda que permet establir la correspondència entre les puntuacions directes equivalents de les dues mesures.

Exemple

Les mesures X i Y han estat administrades a dos grups de persones de manera contrabalancejada. La mesura X , quan es va administrar en primer lloc, va presentar una mitjana de 10 i una desviació típica de 2. Quan es va administrar en segon lloc, la mitjana va ser de 8 i la desviació típica d'1. En el cas de la mesura Y , administrada en primer lloc, va presentar una mitjana de 12 i una desviació típica de 3, mentre que quan es va administrar en un segon ordre la seva mitjana va ser de 14 i la desviació típica de 4. Quina puntuació correspondria a la mesura Y si en la mesura X s'obtingués una puntuació de 7?

Calculant les desviacions típiques de les mesures X i Y segons l'ordre en què es van administrar, respectivament:

$$2,24 = \sqrt{4+1}$$

$$5 = \sqrt{9+16}$$

En el cas de les mitjanes, respectivament:

$$9 = \frac{10 + 8}{2}$$

$$13 = \frac{12 + 14}{2}$$

Establint la correspondència entre les dues mesures, s'arribaria a la conclusió que la puntuació en la mesura Y hauria de ser 9:

$$\frac{5}{2,24}(7 - 9) + 13 = 8,54 \approx 9$$

L'aplicació d'aquesta forma d'equiparació entre mesures ha d'ajustar-se a les mateixes consideracions que es van recollir en parlar de les puntuacions estandaritzades. Implica, per tant, disposar l'interès sobre la tendència de resposta dels individus en les dues mesures i assumir que les dues distribucions de valors objecte d'equiparació s'ajustin bàsicament a una distribució normal.

Resum

En la primera part d'aquest mòdul s'ha parlat dels percentils, les puntuacions estandarditzades i les normes cronològiques, puntuacions transformades comunament emprades, dirigides a atorgar interpretabilitat a les puntuacions obtingudes en una mesura. El percentils, calculats d'acord amb la posició d'una persona en un continu i expressats com a percentatges, es converteixen en una aproximació fàcil d'entendre i per tant de difondre, fins i tot a persones poc acostumades a l'ús d'aquest tipus d'estratègies matemàtiques. A més, el fet de ser essencialment insensibles als grups de valors no ajustats a una distribució normal permet que es puguin considerar com una transformació d'ampli espectre, una estratègia útil i aplicable per a la majoria de variables quantitatives que són d'interès en psicologia.

Les puntuacions estandarditzades presenten més limitacions en termes d'aplicabilitat. Requereixen que les observacions que són objecte d'anàlisi s'ajustin bàsicament a una distribució normal. I tal com s'ha recollit, són moltes les variables psicològiques que, avaluades en una mostra concreta de persones, acaben no satisfent aquesta forma de distribució. No obstant això, quan aquesta circumstància pot ser assumida, l'estandardització representa una aproximació diferenciada de la dels percentils que pot resultar especialment interessant. A diferència dels percentils, es focalitzen en la tendència de resposta, la mitjana de comportament que proporciona el grup normatiu. La desviació típica, el com de diferents són els individus també del grup normatiu, esdevé la manera de quantificar la discrepància entre la puntuació d'una persona concreta i aquest comportament mitjà.

Concloent el conjunt d'estratègies de transformació, s'ha exposat la puntuació en forma de normes cronològiques o edat. En aquest cas, i d'acord amb la tendència de resposta observada de nou en un grup normatiu, constituït per edats diferents, la puntuació d'un individu passa a ser transformada i interpretada d'acord amb aquesta variable cronològica. El seu raonament, com en el cas dels percentils, és també simple, fàcilment comunicable. No obstant això, i tal com s'ha fet notar, presenta inconvenients que no es poden menysprear. Tot i aquestes limitacions, ha esdevingut una estratègia molt emprada en l'àmbit de la mesura de la intel·ligència, la qual cosa ha justificat l'espai dedicat en aquest mòdul.

El barem ha estat un dels conceptes fonamentals del mòdul. Gràcies a aquest, l'usuari d'un instrument pot fer inferències respecte a la puntuació d'un individu d'acord amb el perfil de comportament d'una mostra normativa, i emprar taules que estableixen aquesta correspondència, moltes vegades mitjançant l'ús de puntuacions transformades. No obstant això, més enllà de les consideracions que es deriven de l'ús d'una estratègia de transformació o altra,

l'element més rellevant que s'ha de tenir en consideració és la necessària i adequada justificació de quina ha de ser aquesta mostra normativa que estructuri aquest barem. Aquesta mostra determinarà fortament les inferències que se'n puguin derivar i podrà condicionar, quan s'estigui treballant en un context aplicat, el sentit de les decisions que s'hagin de prendre.

El mòdul conclou amb algunes de les estratègies emprades amb més freqüència dirigides a equiparar puntuacions entre mesures diferents que mesuren un mateix objecte psicològic. L'equiparació equipercenil o la transformació lineal esdevenen estratègies que poden ser especialment útils, per exemple, en els contextos de treball en què una avaluació repetida d'una mateixa variable en un mateix individu pot conduir a una estimació esbiaixada. Els avantatges i limitacions que l'usuari ha de tenir presents quan es vulgui emprar alguna d'aquestes dues estratègies són homòlegs dels descrits en el cas dels percentils o de les puntuacions estandarditzades.

Bibliografia

- Aiken, L. R. (1996). *Tests psicológicos y evaluación* (8a. ed.). Mèxic: Prentice Hall.
- Alegret, M., Espinosa, A., Vinyes-Junqué, G., Valero, S., Hernández, I., Tárraga, L., et al (2012). Normative data of a brief neuropsychological battery for Spanish individuals older than 49. *Journal of Clinical and Experimental Neuropsychology*, 34, 209-219.
- Aluja, A., Blanch, A., Solé, D., Dolcet, J. M., i Gallart, S. (2008). Validez convergente y estructural del NEO-PI-R. Baremos orientativos. *Boletín de Psicología*, 92, 7-25.
- American Educational Research Association, American Psychological Association, i National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Gomà-i-Freixanet, M. i Valero, S. (2008). Spanish normative data of the Zuckerman-Kuhlman Personality Questionnaire in a general population sample. *Psicothema*, 20, 324-330.
- Muñiz, J. (2003). *Teoría clásica de los tests*. Madrid: Pirámide.
- Ramos-Quiroga, J. A., Draigue, C., Valero, S., Bosch, R., Gómez-Barros, N., Nogueira, M., Palomar, G., Roncero, C., i Casas, M. (2009). Validación al español de la escala de cribado del trastorno por déficit de Atención/hiperactividad en adultos (ASRS v. 1.1): una nueva estrategia de puntuación. *Revista de Neurología*, 48, 449-452.
- Rodríguez, C., Jiménez, J. E., Díaz, A., García, E., Martín, R., i Fernández, S. (2012). Datos normativos para el Test de los Cinco Dígitos: desarrollo evolutivo de la flexibilidad en Educación Primaria. *European Journal of Education and Psychology*, 5, 27-38.
- Valero, S., Ramos-Quiroga, J. A., Gomà-i-Freixanet, M., Bosch, R., Gómez-Barros, N., Nogueira, M., Palomar, G. et al. (2012). Personality profile of adult ADHD: The alternative five factor model. *Personality Research*, 198, 130-134.
- Yueh-Hsien, L., Chwen-Yng, S., Wei-Yuan, G., i Yee-Pay, W. (2012). Psychometric validation and normative data of the second Chinese version of the Hooper Visual Organization test in Children. *Research in Developmental Disabilities*, 33, 1919-1927.
- Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P., i Kraft, M. (1993). A comparison of the three structural models for personality: the big three, the big five, and the alternative five. *Journal of Personality and Social Psychology*, 65, 757-768.

