F. METELLI

DISPENSE I

CORSO DI PSICOLOGIA DELL'ETA' EVOLUTIVA

FACOLTA' DI LETTERE (TRIESTE)

1961-62

Vicario

A P P U N T I D I P S I C O M E T R I A

dalle lezioni del

Prof. FABIO METELLI

Padova - 1961

Capitolo Primo

RACCOLTA DEI RISULTATI DI UNA PROVA.

1) Distribuzione di frequenza

Immaginiamo di avere un test formato di 200 problemi, e stabiliamo di assegnare un punto per ogni risposta esatta. Se un soggetto ha risposto esattamente a 130 problemi, il suo risultato nel test è di 130 punti. Un tale risultato considerato a se stante non ci dice nulla; esso assume significato solamente se messo in relazione ai risultati del gruppo al quale il nostro soggetto appartiene. E' necessario dunque avere a disposizione i risultati convenientemente ordinati di un gruppo possibilmente numeroso di persone, relativi al test in parola.

Si possono allora disporre i risultati in ordine decrescente, per mettere in evidenza il risultato massimo, ed il minimo; inoltre, si può facilmente stabilire quanti soggetti (o che percentuali di soggetti) hanno un risultato superiore p.es. ai 100 punti, ecc...

Una disposizione di questo genere, però, richiede un lavoro piuttosto lungo se è stato esaminato un numero elevato di soggetti. E' più conveniente allora elencare tutti i risultati possibili dal massimo al minimo, e porre accanto ad ogni risultato il numero di soggetti (frequenza) che hanno ottenuto quel determinato risultato.

Se i risultati possibili sono molti, anzichè incolonnarli ad uno ad uno, conviene raggrupparli in classi o p.es. a 10 a 10, o a 15 a 15. Naturalmente in questa maniera i dati perdono la loro individualità, ma si riuniscono in una tabella che offre un quadro generale della situazione (v. per esempio Tab. 1).

Le classi in cui si raggruppano i risultati devono essere più o meno ampie a seconda dell'uso che si deve fare della tabella. Generalmente è comodo raggruppare i punteggi in più di 12 e meno di 20 classi. Si può allora procedere secondo questa tecnica: si cerca il massimo (M) e il minimo (m) punteggio realizzato, e si calcola la differenza G = M - m, gamma della distribuzione; se per esempio il massimo punteggio M = 187, e il minimo m = 46, G = 187 - 46 = 141. La gamma della distribuzione, divisa per un numero da 12 a 20 (p.es. 15), dà l'ampiezza i da dare agli intervalli di classe, per averne appunto da 12 a 20 (come ci si era proposto). Nell'esempio, i = 141 : 15 = 9; anzichè di 9 punti, si preferisce usare un intervallo di 10, per facilitare la classificazione. Gli intervalli di classe che si adoperano correntemente sono di 1,2,3,5,7,10 e multipli di 5 punti.

A questo punto, si incolonnano gli intervalli di classe dell'ampiezza stabilita; usualmente si parte dall'intervallo superiore, da quello, cioè, che comprende il massimo punteggio realizzato; se i è pari, si prende come limite inferiore del primo intervallo il multiplo di i immediatamente inferiore al massimo punteggio M. Nell'esempio, essendo M = 187 ed i = 10, l'intervallo superiore va dal 180 al 189. Si criverà quindi di seguito:

180 - 189
170 - 179
160 - 169
....

e così via.

Se il numero di punti per ogni intervallo è dispari, si fissano gli intervalli in modo che il valore centrale dell'intervallo sia multiplo di i. Se p.es. i = 5, gli intervalli saranno:

183 - 187 (valore centrale = 185)
178 - 182 (valore centrale = 180)
173 - 177 .... ecc.

Una volta stabiliti gli intervalli, si incomincia la registrazione, spuntando ogni risultato e segnando un'asticina accanto all'intervallo che gli compete. Per facilitare successivamente il computo delle frequenze per ogni intervallo, ogni quinta asticina si segna per traverso sulle altre quattro: p.es. ~~IIII~~, ~~IIII~~, ~~IIII~~, III = 5 + 5 + 5 + 3 = 18.

I limiti indicati dagli intervalli (p.es. 180 - 189) sono i veri limiti solo se si ha a che fare con grandezze discrete (quali risultano dall'operazione del "contare"). Ma trattandosi di grandezze continue, (quali sono in genere le misure), i dati che si

TABELLA I^

Registrazione dei risultati di un test

punteggi	frequenze	
180 - 189	II	2
170 - 179	II	2
160 - 169	IIII	4
150 - 159	~~IIII~~ I	6
140 - 149	~~IIII~~ ~~IIII~~ III	13
130 - 139	~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ III	23
120 - 129	~~IIII~~ ~~IIII~~ ~~IIII~~ IIII	19
110 - 119	~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~	30
100 - 109	~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ ~~IIII~~ II	27
90 - 99	~~IIII~~ ~~IIII~~ ~~IIII~~ II	17
80 - 89	~~IIII~~ ~~IIII~~ I	11
70 - 79	~~IIII~~ II	7
60 - 69	~~IIII~~	5
50 - 59	III	3
40 - 49	I	1

hanno a disposizione sono sempre approssimati: così ad esempio la statura, e il peso ed anche il risultato di un soggetto in una prova sono considerati approssimazioni. Un punteggio di 72, p.es., è un punteggio superiore a 71,5 e inferiore a 72,5, limiti reali entro cui giace il punteggio 72; dell'intervallo 180-189, i limiti reali sono dunque 179,5 e 189,5. In altri casi, i limiti reali sono stabiliti in maniera diversa. P.es. un'età di 17 anni significa un'età da 17 anni esatti a un giorno meno di 18 anni; quindi l'intervallo 17-19 va da 17 (limite inferiore) a 20 anni meno un giorno (limite superiore).

Valore centrale o mediano dell'intervallo è quel valore che sta ad egual distanza dai due limiti reali. Nell'intervallo 180-189 il valore mediano è 184,5; nell'intervallo di età 17-19 anni, è 18 anni e 6 mesi.

2) Istogramma e poligono di frequenza.

Col lavoro sopra descritto, si viene ad avere una classificazione dei risultati in gruppi, dei quali è indicata la relativa frequenza. La somma delle frequenze è ovviamente il numero complessivo di soggetti esaminati (v. Tav. 2, in cui la somma delle frequenze N (numero complessivo dei soggetti) = 170).

Sono noti i vantaggi che una rappresentazione grafica presenta in confronto a una tabella, che deve venir letta completamente per afferrarne il significato.

TAVOLA 2

Distribuzione di frequenza

Punteggio	Valore centrale dell'intervallo	Frequenza	Limite superiore dell'intervallo	Frequenza cumulativa fc	Frequenza cumulativa percentuale fc % = P
180-189	184,5	2	189,5	170	100
170-179	174,5	2	179,5	168	98,8
160-169	164,5	4	169,5	166	97,6
150-159	154,5	6	159,5	162	95,3
140-149	144,5	13	149,5	156	91,8
130-139	134,5	23	139,5	143	84,1
120-129	124,5	19	129,5	120	70,6
110-119	114,5	30	119,5	101	59,4
100-109	104,5	27	109,5	71	41,8
90- 99	94,5	17	99,5	44	25,9
80- 89	84,5	11	89,5	27	15,9
70- 79	74,5	7	79,5	16	9,4
60 -69	64,5	5	69,5	9	5,3
50- 59	54,5	3	59,5	4	2,4
40- 49	44,5	1	49,5	1	0,6
		N = 170			

La più semplice forma di rappresentazione grafica è l'istogramma, che si ottiene affiancando tante colonne quanti sono gli intervalli.

Se per esempio vogliamo rappresentare i dati della Tav. 2 con un istogramma riportiamo in ordinata delle tacche corrispondenti alle frequenze, e in ascissa i limiti reali degli intervalli; innalziamo su di questi delle colonne fino al livello corrispondente alla frequenza che si legge nella tabella, ottenendo l'istogramma di Fig. 1.

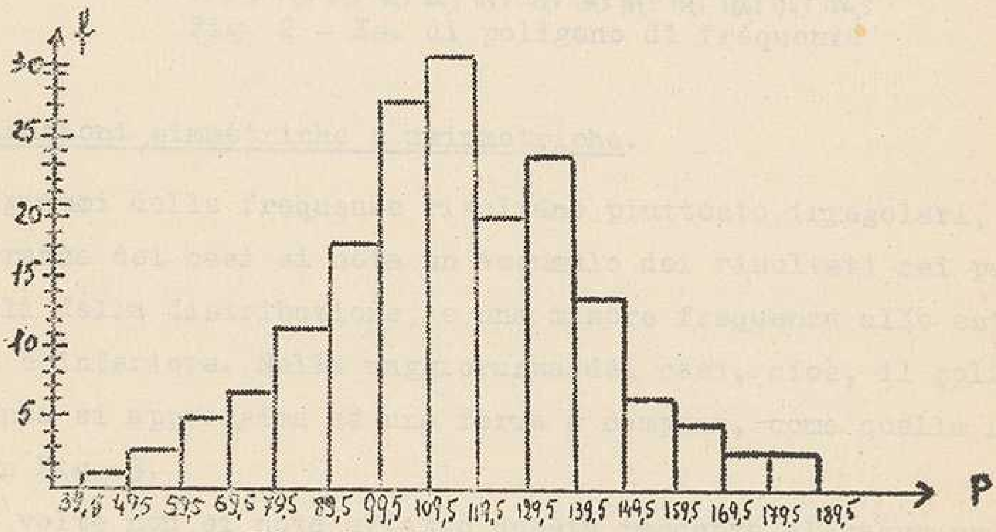


Fig. 1 - Es. di istogramma

Una seconda forma di rappresentazione grafica è il poligono di frequenza: si ottiene congiungendo con una spezzata i valori mediani della sommità delle colonne dell'istogramma. Per la costruzione del poligono di frequenza, dunque si riportano in ascissa i valori mediani degli intervalli (v. fig. 2) e su questi si innalzano le ordinate di altezza proporzionale alla frequenza di ogni singolo intervallo.

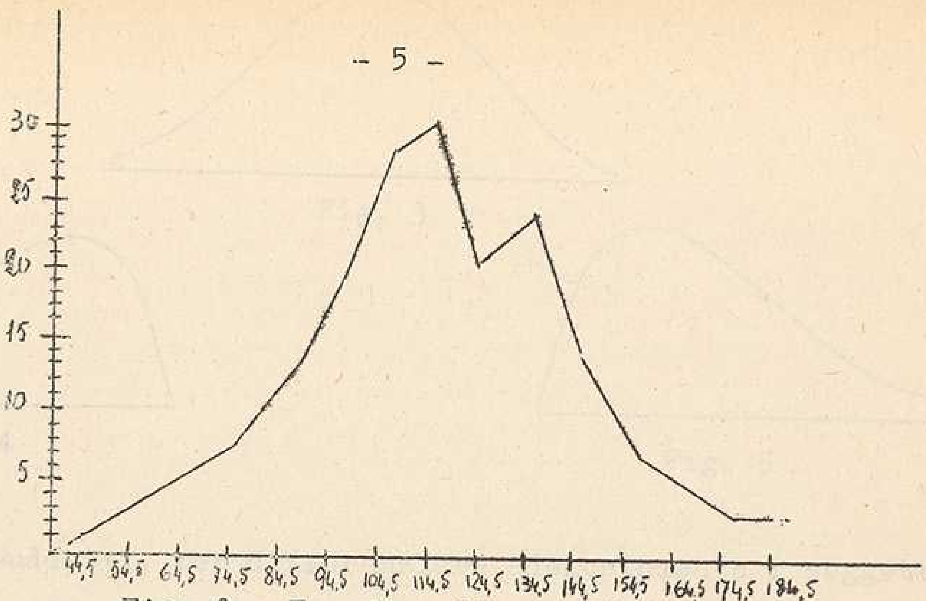


Fig. 2 - Es. di poligono di frequenza

3) Distribuzioni simmetriche e asimmetriche.

I diagrammi delle frequenze risultano piuttosto irregolari, ma nella maggioranza dei casi si nota un accumulo dei risultati nei punteggi centrali della distribuzione, e una minore frequenza alle estremità superiore e inferiore. Nella maggioranza dei casi, cioè, il poligono di frequenza si approssima ad una forma a campana, come quella rappresentata in Fig. 3.

Altre volte non si nota affatto questa tendenza al raggruppamento centrale, perchè la maggior parte dei soggetti realizzano i risultati più elevati (v.Fig.4). E' possibile interpretare simili scostamenti dalla normale distribuzione di frequenza come dovuti al fatto che la prova è troppo facile per il gruppo cui è stata applicata. In questo caso, essa non differenzia adeguatamente tutti i soggetti, poichè qualcuno dei migliori avrebbe probabilmente superato altre difficoltà se la prova ne avesse presentate.

L'opposta distribuzione asimmetrica, in cui da un punteggio minimo realizzato da molti si scende rapidamente a poche frequenze nei risultati superiori (Fig. 5) è molte volte interpretabile come dovuta all'eccessiva difficoltà della prova.

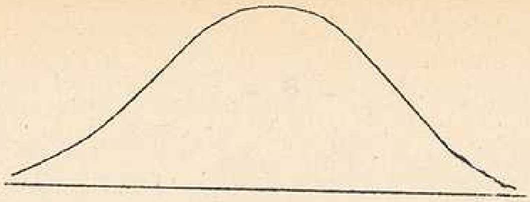


Fig. 3

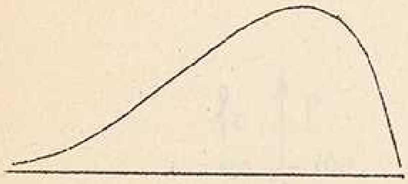


Fig. 4

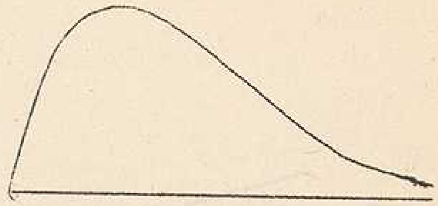


Fig. 5

Nei casi suddetti, per differenziare adeguatamente i soggetti, si dovrà ricorrere ad un'altra prova dello stesso tipo, oppure - se è proprio indispensabile - modificare la prova, procedendo a un accurato studio preliminare.

4) <u>Frequenza cumulativa - Ogiva di Galton</u>

Per scopi particolari (v. cap. III) interessa sapere quanti soggetti hanno ottenuto punteggi inferiori a un punteggi determinato, per esempio (v. Tab. 2) quanti hanno ottenuto un punteggio inferiore al 70: si dovranno allora sommare le frequenze inferiori al 70, che sono 5 + 3 + 1 = 9.

Nella Tav. 2 la penultima colonna riporta le frequenze cumulative, cioè le somme delle frequenze inferiori al limite superiore di ogni intervallo.

La rappresentazione grafica delle frequenze cumulative forma una spezzata generalmente a forma di ogiva e prende appunto questo nome. Per disegnarla, si riporta in ascissa il limite superiore esatto di ogni intervallo ed in ordinata la frequenza cumulativa (vedi fig.6); la frequenza cumulativa dell'intervallo che comprende il risultato più elevato è ovviamente N, cioè il totale dei soggetti esaminati. Oltre alla scala fc (frequenze cumulativa) è utile riportare sull'asse delle ordinate la scala p, delle frequenze cumulative percentuali. La ogiva della frequenza cumulativa si trova talvolta costruita con la scala dei punteggi in ordinata, cioè ruotata di 90 gradi.

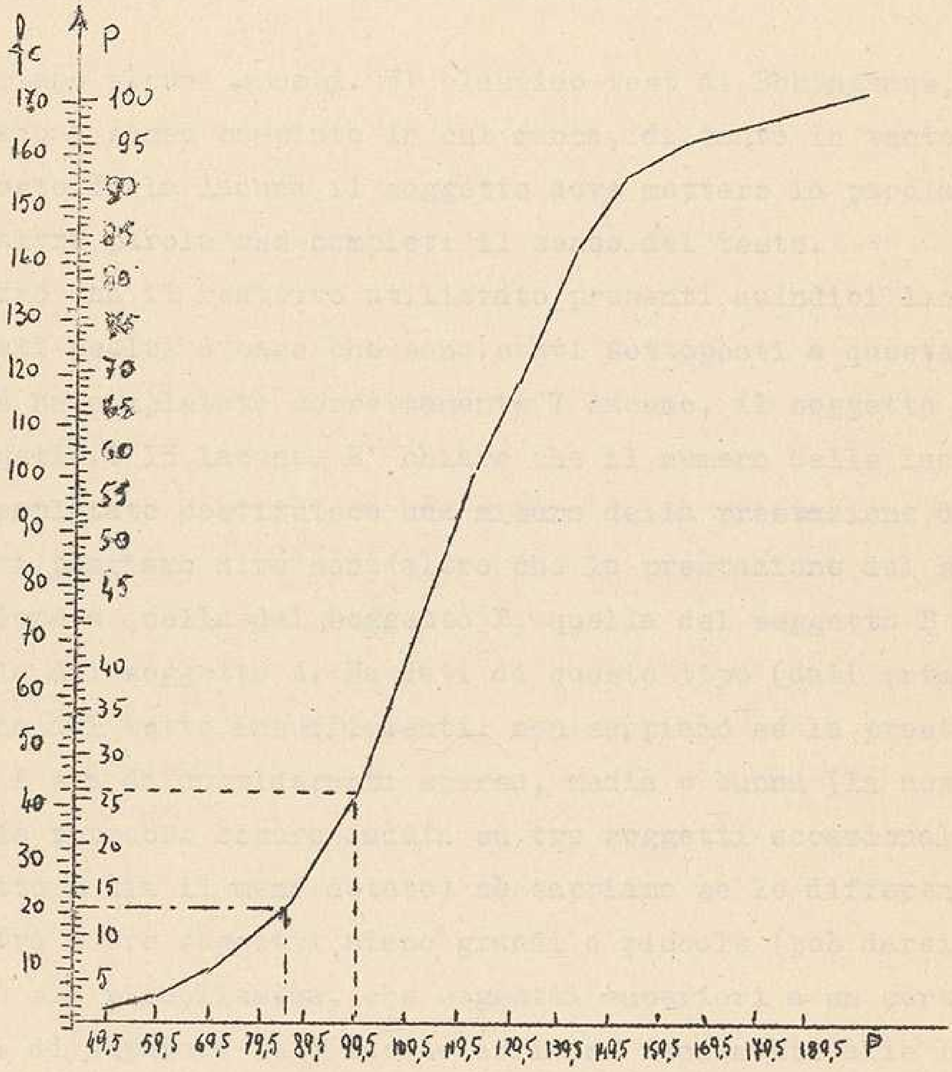


Fig. 6 - Esempio di diagramma della frequenza cumulativa (fc). La scala P si riferisce ai percentili (v. cap. III, pp.23-24): al percentile 25 corrisponde il punteggio 99.

Capitolo II°

PROBLEMA DELLE UNITA' DI MISURA :
ETA' MENTALE E QUOZIENTE INTELLETTUALE

Consideriamo alcuni esempi. Il classico test di Ebbinghaus, consiste in un brano a senso compiuto in cui manca, di tanto in tanto, una parola. Al posto della lacuna il soggetto deve mettere la parola mancante, o un'altra parola che completi il senso del testo.

Ammettiamo che il reattivo utilizzato presenti quindici lacune. Di tre soggetti scelti a caso che sono stati sottoposti a questa prova, il soggetto A ha completato correttamente 7 lacune, il soggetto B 9 lacune, il soggetto C 15 lacune. E' chiaro che il numero delle lacune esattamente completato costituisce una misura della prestazione del soggetto; infatti possiamo dire senz'altro che la prestazione del soggetto C è superiore a quella del soggetto B, quella del soggetto B superiore a quella del soggetto A. Ma dati di questo tipo (dati primitivi o grezzi) sono del tutto insufficienti: non sappiamo se la prestazione del soggetto A sia da considerarsi scarsa, media o buona (la nostra scelta causale potrebbe essere caduta su tre soggetti eccezionali, tra cui il soggetto A sia il meno dotato) nè sappiamo se le differenze di prestazioni tra i tre soggetti siano grandi o piccole (può darsi che la differenza sia piccolissima, che soggetti superiori a un certo livello, capaci di superare le 6 lacune arrivino facilmente alle 15; può darsi invece che la differenza sia grandissima, che cioè tutti i soggetti normali di una data età completino almeno 8 lacune e solo 1 su 1000 arrivi a completarle tutte e 15).

E' certo che sottoponendo numerosi soggetti alla prova si finisce col rendersi conto del valore dei risultati (così ad esempio, in una particolare prova di completamento, si può sapere, per esperienza, che, tranne casi eccezionali, i soggetti di una determinata età completano da 8 a 12 lacune). Ma si tratta di una valutazione approssimativa che non permette di attribuire un significato preciso alla singola misura. Volendo ad esempio, in base ai risultati di una serie di prove, stabilire il livello mentale generale di un soggetto come è possibile procedere ad un calcolo, in base a dati di questo genere: lacune completate esattamente: 8, errori compiuti di una prova di attenzione: 5; sillabe riprodotte esattamente in una prova di memoria: 17; velocità media di reazione ad uno stimolo acustico: 18 centesimi di secondo? Evidentemente non si possono sommare tutti questi valori, data la loro eterogeneità. Come giungere allora ad un valore indicativo del livello mentale di un soggetto?

Si dovette attendere fino al 1908 perchè si giungesse ad una prima soluzione di questa difficoltà che sbarrava la via del progresso alla psicometria.

La soluzione fu dovuta ad Alfredo Binet, direttore dell'Istituto di Psicologia della Sorbona, medico, psicologo e pedagogista ad un tempo, e nei suoi ozi, letterato e drammaturgo.

Nel 1909 il Binet faceva parte di una commissione incaricata della selezione dei frenastenici e dello studio del regime scolastico adatto ad essi. Servirsi di un criterio scientifico per individuare i frenastenici (e in particolare le forme più lievi di frenastenia) era tutt'altro che facile: i limiti tra i vari gradi di deficienza mentale apparivano non soltanto fluidi, ma diversissimi a seconda delle varie classificazioni allora in uso.

Bisognava dunque creare un metodo nuovo. Tre diversi criteri di classificazione si presentavano al giudizio del Binet; un criterio antropologico, un criterio pedagogico e un criterio psicologico. Il criterio antropologico, basato sulla coincidenza fra anomalie somatiche e frenastenia appariva quanto mai infido; il criterio pedagogico, basato sulla constatazione del ritardo scolastico e sul controllo della somma delle conoscenze acquisite poteva servire come punto di partenza, ma si prestava pure a grossolani errori, data la non completa uniformità dell'insegnamento e le varie condizioni che possono influire sul ritardo scolastico; restava il criterio psicologico, l'unico criterio diretto, quantunque di difficile applicazione. E il Binet si accinse a creare ed organizzare una serie di prove psicologiche atte a svelare l'anormalità psichica ed a determinarne il grado.

Preparata la "batteria di prove", si trattava di stabilire come valutarle. Il Binet giunse al criterio di valutazione obbiettivo, che gli diede modo di superare la difficoltà cruciale della psicometria, applicando con consequenzialità il concetto di arretratezza mentale.

Lo sviluppo mentale di un soggetto viene considerato arretrato quando esso corrisponde (grosso modo) ad una fase che i soggetti normali raggiungono ad un'età inferiore. Ecco dunque la soluzione della difficoltà in cui si dibatteva la psicometria: esprimere una differenza di intelligenza mediante una differenza di età.

Per attuare quest'idea bisognava: a) stabilire quali fossero le prestazioni mentali medie di ogni età; b) creare quindi per ogni singola età una serie di prove adatte; c) applicarle infine a scopo di controllo statistico, su un gran numero di soggetti.

Fatto ciò, si era finalmente creato lo strumento necessario per la selezione dei frenastenici; bastava infatti sottoporre un soggetto ad un certo numero di prove, e vedere a quale età corrispondevano le sue prestazioni (determinazione dell'età mentale del soggetto).

L'età mentale è dunque una indicazione precisa e chiara dello stadio di sviluppo mentale raggiunto da un soggetto in un dato momento. Ma nè l'età mentale, nè la differenza fra età mentale ed età cronologica di un soggetto misurano il grado della sua arretratezza mentale.

Consideriamo ad es. il soggetto A, che ha un'età cronologica di 4 anni, ed un'età mentale di 2 anni, e il soggetto B, che ha una età cronologica di 10 anni ed un'età mentale di 8 anni; è evidente che, quantunque ambedue i soggetti abbiano uno sviluppo mentale arretrato di 2 anni, il soggetto A presenta un grado di frenastenia molto più grave del soggetto B; essere in ritardo di due anni a quattro anni è molto più grave che essere in ritardo di due anni a dieci anni.

Per calcolare il grado di frenastenia di un soggetto, bisogna dunque mettere in rapporto l'età mentale con l'età cronologica, ossia calcolare il quoziente intellettuale (Stern).

Il quoziente intellettuale si usa esprimere in percentuale:

$$QI = 100 \frac{E.mentale}{E.cronol.}$$

Esso ha il valore di 100 per i soggetti normali che hanno l'età mentale uguale all'età cronologica: minore di 100 per i tardivi de i frenastenici veri e propri. (Nel nostro caso il soggetto A avrebbe un quoziente intellettuale di 50, il soggetto B di 80).

La formula presenta il vantaggio di misurare il grado di frenastenia, indipendentemente dall'età cronologica del soggetto.

In America i vari gradi di frenastenia sono stati definiti in termini di quoziente intellettuale, secondo la seguente tabella:

Q.I.	grado di frenastenia
90 - 110	normalità
80 - 90	tardività
70 - 80	zona limite
50 - 70	debolezza mentale
25 - 50	imbecillità
⟶ 25	idiozia

Dalla tabella risulta, fra l'altro, che per la normalità, concapita naturalmente anch'essa non come un punto ma come una zona, sono stati fissati, convenzionalmente, come limiti i Q.I. 90 e 110. Quando l'età mentale supera l'età cronologica e cioè il soggetto presenta uno sviluppo mentale superiore a quello caratteristico della propria età, il Q.I. è superiore a 100.

Il metodo di misura escogitato dal Binet non serve dunque soltanto per determinare il grado di frenastenia, ma in genere per stabilire il livello mentale di un soggetto; la tabella continua nel modo seguente:

Q.I.	
110 - 125	intelligenza superiore alla media
125 - 140	intelligenza superiore
140 ⟶	intelligenza eccezionale

Il metodo ebbe grande fortuna e fu tradotto e adattato in tutte le lingue. La morte prematura del Binet gli impedì di perfezionare l'opera. Si trattava anzitutto di compiere un controllo statistico su molte migliaia di soggetti. Tale controllo fu compiuto in America dal

Terman nel 1916 e una seconda volta nel 1938 e portò ad un riordinamento generale: molte prove risultarono o troppo facili o troppo difficili per l'età per la quale le aveva destinate il Binet, e perciò furono destinate ad altre età. Altre prove furono aggiunte dallo stesse Terman per equilibrare il numero delle prove e facilitare il computo dell'età mentale.

Capitolo III°

PERCENTILI, CLASSIFICAZIONE PENTENARIA, PROFILI

Il metodo introdotto dal Binet, consistente nell'esprimere una differenza di sviluppo mentale mediante una differenza di età, trova legittima applicazione finchè si tratta di soggetti nell'età dello sviluppo. Applicato agli adulti, tale sistema di misura richiede l'introduzione di particolari artifici e non sempre dà risultati soddisfacenti. Inoltre l'esprimere il grado di potenzialità di una funzione mentale di un adulto in termini di età, riesce del tutto innaturale; della intelligenza e più ancora della memoria o dell'attenzione di un adulto non ci si chiede a che età corrisponde normalmente, ma se essa sia media (cioè normale), o eccezionale (in senso positivo o negativo); interessa inoltre una misura del grado di questa eventuale eccezionalità.

La misura di una prestazione di un adulto, deve dunque fondarsi sul confronto con quella che risulta essere la prestazione normale degli adulti in condizioni analoghe.

Il primo passo per raggiungere tale scopo consisterà nel determinare la prestazione normale, la quale, com'è noto, si considera corrispondente alla prestazione media di un gruppo omogeneo abbastanza numeroso di soggetti non selezionati, tale cioè da costituire un campione rappresentativo di una determinata popolazione.

Ma conoscere la prestazione media non significa possedere un criterio di misura delle prestazioni; non basta sapere se un soggetto per una determinata attitudine è al di sopra o al di sotto della media, ma di quanto esattamente al di sopra o al di sotto.

Bisogna dunque fissare un'unità di misura: ma non un'unità di misura scelta arbitrariamente, bensì tale che, essendo sempre la stessa per ogni test, permetta di confrontare fra loro i risultati ottenuti dal soggetto in differenti prove attitudinali.

Un tale metodo, atto a risolvere in modo diverso ma egualmente efficace dal metodo di Binet, le difficoltà in cui si dibatteva la psicometria agli albori del secolo, era stato proposto da Francis Galton, e fu introdotto nella prassi psicologica dal Claparède; esso va sotto il nome di metodo dei percentili (o centili).

Ammettiamo di aver sottoposto ad una prova di rapidità di reazione ad uno stimolo acustico un gruppo omogeneo abbastanza numeroso di soggetti non selezionati: in tutto 177. I valori medi ottenuti per i diversi soggetti vanno da un minimo di 11,7 centesimi di secondo ad un massimo di 40,4 centesimi di secondo.

Possiamo considerare il gruppo di soggetti esaminati come un "campione" rappresentativo della categoria alla quale i soggetti appartengono. Possiamo cioè supporre che se i soggetti sono ad esempio operai lombardi, prendendo un altro gruppo di operai lombardi e sottoponendolo alla stessa prova, si otterranno risultati analoghi. In tal caso possiamo servirci dei risultati del gruppo come di una scala, di un sistema di riferimento, per classificare i risultati ottenuti nella stessa prova da altri soggetti della stessa categoria.

Per far ciò dobbiamo ordinare i soggetti secondo il risultato ottenuto nella prova, dal più lento al più rapido, e considerarli come altrettanti gradi della scala.

Ammettiamo ora che si presenti la necessità di sottoporre alla stessa prova un soggetto appartenente alla stessa categoria, e che dai risultati della prova la sua velocità media di reazione semplice ad uno stimolo acustico risulti di 25 centesimi di secondo; non resterà altro che vedere a quale grado della scala corrisponda questa prestazione. Ammettiamo che questa prestazione corrisponda a quella dell'84° soggetto (a partire dal più rapido) del gruppo campione: la prova del neo-esaminato va classificata al rango 84, il che significa che la sua prestazione è la ottantesima quarta su 177 soggetti presi a caso.

Il passaggio dal dato primitivo (espresso in centesimi di secondo) all'ordine di merito (o rango) offre il vantaggio di dare un'idea chiara del valore della prestazione; mentre del risultato in centesimi di secondo preso in sè non potevano dire neppure se era medio o eccezionale in uno o nell'altro senso, il risultato espresso in ordine di rango, ci dice che si tratta di una prestazione non lontana da quella del rango mediano (89°) ma superiore al risultato mediano; e sappiamo anche esattamente di quanto la prestazione sia superiore al rango mediano;la prestazione si distacca dal rango mediano per 5 degli 88 ranghi che coprono la distanza dal rango mediano al rango massimo.

In questo modo si otterrebbe già una classificazione dei risultati. Ma il confronto fra le classificazioni dei risultati di due diversi reattivi sarebbe quanto mai disagevole; siccome il numero dei soggetti di cui si dispone per costituire il gruppo campione è, per ovvie ragioni, diverso di volta in volta, ci si troverebbe nella necessità di fare ogni volta un calcolo per stabilire se un soggetto sottoposto a due diverse prove funzionali (p.es. ad una prova dei tempi di reazione e ad una prova di memoria), abbia dato un risultato superiore, e di quanto superiore, nell'una o nell'altra prova.

Si impone quindi una semplificazione; essa consiste semplicemente nel trasformare i ranghi in ranghi percentuali.

La trasformazione è semplicissima poichè non si tratta che di risolvere una proporzione per stabilire a quale rango percentuale corrisponda un determinato rango.

Così, restando all'esempio finora considerato, per calcolare a quale rango percentuale corrisponda il rango 84 su 177 soggetti, basta risolvere la proporzione

$$88 : 177 = X : 100$$

$$X = 100 \frac{84}{177} = 47,46$$

cioè

84	:	177	=	47,46	:	100
rango		n° dei soggetti		rango percentuale		

Al rango 84° su 177 soggetti corrisponde dunque un rango percentuale 47; il che equivale a dire che quel soggetto risulta in quella particolare prova (rapidità di reazione) 47° su 100 soggetti presi a caso.

Il vantaggio del rango percentuale è evidente: esso permette di confrontare direttamente i risultati di un soggetto in diversi tests.

La misura comunemente usata, il percentile, corrisponde sostanzialmente al rango percentuale, ma inverte l'ordine dei ranghi. Infatti il soggetto che dà il migliore risultato, e che perciò è primo nell'ordine di rango, raggiunge il centesimo percentile (P_{100}). Ma anche operando tale inversione vi è un leggero sfasamento fra percentili e ranghi percentuali, in quanto il percentile è definito quel valore al do sotto del quale sta un corrispondente percentuale di risultati (p.es. percentile 52 è quel valore al di sotto del quale sta il 52% dei risultati).

Perciò la formula per calcolare il percentile che corrisponde ad un determinato rango

$$P = 100 \frac{n-r+0,5}{n} \quad (1)$$

(in cui P = percentile, r = rango, n = numero dei soggetti) è un po' meno semplice della proporzione usata per calcolare il rango percentuale.

Diamo alcuni esempi di calcolo:

a) A quale percentile corrisponde la prestazione precedentemente considerata, pari al rango 84 del gruppo campione di 177 soggetti?

(1) Gli autori francesi usano generalmente la formula di Claparède, che non dà risultati perfettamente uguali a quella riportata nel testo:

$$P = \frac{n-r}{n-1}$$

Applicando la formula si avrà :

$$P = 100 \frac{177 - 84 + 0,5}{177} = 52,8$$

Il percentile corrispondente al rango 84 è il 53°.

b) A quale rango corrisponde nel nostro esempio il percentile 50?

$$50 = 100 \frac{177 - r + 0,5}{177} \qquad r = 89$$

Il percentile 50 corrisponde, nel nostro esempio, al rango 89, che è dunque il rango mediano (il che era fin dal principio evidente, in quanto 50 è il rango mediano, in una serie di ranghi che va dal rango 0 al rango 100).

c) Lo stesso soggetto la cui prestazione psicomotrice era risultata corrispondente al rango 84 del gruppo campione di 177 soggetti, risulta in una prova di intelligenza logica 30° rispetto ad un gruppo campione di 71 soggetti. In quale delle due prove risulta superiore, e di quanto?

La soluzione si ottiene calcolando i percentili relativi alle due prestazioni.

Per la prestazione psicomotrice conosciamo già il percentile, che è il 53°. Per l'altra prestazione basta applicare la formula:

$$P = 100 \frac{71 - 30 + 0,5}{71} = 58,4$$

Il soggetto risulta dunque superiore nella prestazione intellettuale: fra questa e la prestazione psicomotrice vi è una differenza di 5 percentili. Le due prestazioni sono ambedie superiori alla prestazione di rango mediano, rappresentata dal 50° percentile; ma nè l'una nè l'altra sono prestazioni eccezionali.

Si considerano infatti generalmente come normali le prestazioni che vanno dal 25° al 75° percentile, e che comprendono il 50% dei soggetti.

Mentre il 50° percentile rappresenta la mediana, il 25° percentile va sotto il nome di quartile inferiore, e il 75° percentile è denominato quartile superiore.

Comunque, la formula sopra riportata è di uso limitato: la si applica soltanto quando il numero di casi è così piccolo che non conviene procedere alla costruzione di una tabella delle frequenze.

Quando invece si dispone di una tabella, i percentili si determinano in uno dei due modi seguenti:

a) utilizzando il diagramma delle frequenze cumulative (fig.6), nel quale sono riportate anche le frequenze cumulative percentuali.

b) utilizzando la tabella delle frequenze cumulative.

Consideriamo successivamente i due tipi di problemi che si presentano:

1. Determinazione del rango percentile R_p (cioè determinazione del valore che nella scala percentilare corrisponde ad un determinato punteggio grezzo X).

2. Determinazione del percentile P (cioè determinazione del punteggio grezzo corrispondente ad un determinato rango percentile.

Utilizzando il diagramma delle frequenze cumulative si procede nel modo seguente:

Problema 1 : Dal punto dell'asse delle ascisse corrispondente al punteggio X si innalza una perpendicolare fino ad incontrare il diagramma; dal punto d'incontro si prosegue lungo la parallela all'asse delle ascisse fino ad incontrare l'asse delle ordinate nel quale si legge, nella scala centesimale (frequenze cumulative percentuali) il rango percentile (vedi es. in fig. 6).

Problema 2 : Per il punto dell'asse delle ordinate corrispondente al rango percentile R_p (nella scala delle frequenze cumulative percentuali) si fa passare una parallela fino ad incontrare il diagramma. Dal punto d'incontro si segue la parallela all'asse delle ordinate fino ad incontrare l'asse delle ascisse. A questo punto si legge il punteggio grezzo.

Utilizzando invece la tabella delle frequenze cumulative ci si serve di uno speciale metodo di calcolo che ci offre un risultato più preciso di quello ottenuto utilizzando il diagramma delle frequenze cumulative. Detto calcolo si basa sulla convenzione che la frequenza relativa ai punteggi contenuti in una classe di intervallo sia uniformemente distribuita nell'intervallo medesimo.

Sarebbe scomodo dover applicare la formula ogni qualvolta si sottopone un soggetto ad una prova, per determinare il rango percentile corrispondente al risultato ottenuto. Perciò si calcolano fin dal principio i ranghi percentili corrispondenti ai dati ottenuti col gruppo campione, raccogliendoli in una tabella; da questa si può leggere approssimativamente il rango percentile corrispondente al risultato ottenuto in quella stessa prova da un soggetto appartenente alla stessa categoria alla quale appartiene il gruppo-campione.

Siccome una scala centesimale rappresenta una graduazione troppo sottile per le misure che si possono ottenere mediante i tests, ci si accontenta in genere di una scala decimale: si calcolano cioè soltanto i percentili 10,20,30,40,50,60,70,80,90, aggiungendo talvolta i due sistemi estremi, che si usano talora indicare come percentili 0 e 100.

Per la prova dei tempi di reazione applicata al succitato gruppo campione si è ottenuta così la seguente tabella:

Intervalli	Percentili	Media dei tempi di reazione a stimoli acustici (centesimi di secondo)
I	(0)	40.4
II	10	19.7
III	20	18.5
IV	30	17.7
V	40	17.1
VI	50	16.5
VII	60	15.9
VIII	70	15.5
IX	80	14.8
X	90	14.1
	(100)	11.7

Tav. 3 - Percentili relativi alla prova dei tempi di reazione.

Spesso si usano tabelle che comprendono solo i percentili 0,25 - 50 - 75 - 100, come la seguente tabella, relativa ad un test di riconoscimento di forme geometriche :

Percentili	14 anni	15 anni	16 anni
100	14	11	11
75	6	8	8
50	4	6	6
25	3	4	5
0	1	1	0

Tav. 4

L'uso delle tabelle di graduazione è estremamente semplice: esse servono per passare direttamente, senza alcun calcolo, dal risultato ottenuto nell'applicare una prova ad un soggetto, al rango percentile corrispondente. Siccome per lo più i risultati non corrispondono esattamente ai dati contenuti nella tabella, ci si limita a stabilire l'intervallo entro il quale cadono (p.es. fra i percentili 10 e 20, fra il 60° e il 70°, ecc.). Per una illecita estensione vengono talora chiamati decili gli intervalli anzichè i valori che limitano gli intervalli; ed allora la tabella di graduazione è redatta nel modo seguente:

Tav. 5 - Decili-intervalli, relativi alla prova dei tempi di reazione.

Decili-intervalli	Media dei tempi di reazione
I	19.7 ⟶
II	18.5-19.6
III	17.7-18.4
IV	17.1-17.6
V	16.5-17
VI	15.9-16.4
VII	15.5-15.8
VIII	14.8-15.4
IX	14.1-14.7
X	⟶ 14

Basta un'occhiata alla tavola di graduazione per rendersi conto di una importante sua caratteristica: gli intervalli corrispondenti ai vari decili sono diversi.

In tutte le tavole di graduazione compilate ottemperando alle relative norme (soggetti omogenei, non selezionati, in numero sufficiente, reattivo adatto, cioè nè troppo facile nè troppo difficile per la categoria di soggetti esaminati) gli intervalli estremi (I° e X°) sono i più ampi, e quelli intermedi (V° e VI°) i più ristretti; gli intervalli diminuiscono cioè dai decili estremi ai decili intermedi.

In altre parole, una differenza di 10 ranghi percentili è una differenza piccola se si tratta di percentili centrali, grande se si tratta di percentili estremi. E quanto si è detto vale per qualsiasi altra classificazione fondata sui ranghi.

Questo è indubbiamente un difetto di questo sistema di graduazione, difetto che sta ad indicare che si tratta di un metodo di classificazione, ma non di un sistema di misura: in una scala metrica la unità di misura è costante, mentre nelle graduazioni qui considerate essa varia nei diversi intervalli della scala.

Vi è modo di ovviare praticamente al citato inconveniente, senza abbandonare il comodo sistema dei percentili.

Comunemente per classificare i risultati ci si può accontentare di cinque intervalli. Allora, anzichè passare dalla classificazione a decili ad una classificazione a quintili (prendendo cioè i ranghi percentili per ogni nuovo intervallo) si procede ad una classificazione pentenaria i cui intervalli comprendono un numero diverso di ranghi percentili, cioè quaranta l'intervallo centrale, venti ciascuno dei due intervalli contigui all'intervallo centrale, e dieci ciascuno dei due intervalli estremi. In tal modo la diversa estensione delle distanze interpercentili viene ad essere approssimativamente compensata (v. tav. 6).

TABELLA DI GRADUAZIONE DEI TEMPI DI REAZIONE

MEDI A STIMOLAZIONE ACUSTICA

Intervallo	Percentuale dei soggetti del gruppo campione compresi in ogni intervallo	Ranghi percentili	Limiti dell'intervallo
I	10%	0 - 10	19.7 ⟶
II	20%	10 - 30	17.7-19.6
III	40%	30 - 70	15.5-17.6
IV	20%	70 - 90	14.1-15.4
V	10%	90 -100	⟶ 14

Tav. 6 - Graduazione pentenaria a intervalli compensati

Non sarà inutile richiamare alla memoria il significato e la utilità delle tabelle di graduazione del tipo di quella qui riportata.

Il I° e il V° intervallo comprendono rispettivamente i soggetti che hanno dato un risultato eccezionalmente scadente (I°), e quelli che hanno dato un risultato eccezionalmente buono (V°) (1); il II° e il IV° intervallo, rispettivamente i soggetti i quali, senza che la loro prestazione sia eccezionale, risultano, per quella prestazione, nettamente inferiori (II° intervallo) o nettamente superiori (IV° intervallo) alla media; il III° intervallo comprende i soggetti le cui prestazioni si differenziano tanto poco dalla media da poter essere praticamente considerate prestazioni medie.

Le prestazioni che rientrano nei cinque intervalli possono quindi venire denominate "ottime", "buone", "medie", "scarse", "scarsissime". Ma non va dimenticato il preciso significato statistico della graduazione. Quando la prestazione di un soggetto rientra p.es. nel V° intervallo (2) si può cioè dire, grosso modo, che, prendendo a caso cento sog-

---

(1) Va tenuto presente che talvolta le tabelle di graduazione comprendono invece nel I° intervallo i soggetti eccezionalmente buoni, e nel V° intervallo i soggetti eccezionalmente scadenti.

(2) Se nel V° intervallo sono comprese le prestazioni migliori, poichè come si è detto, il fatto di far rientrare nel V° o nel I° intervallo le prestazioni migliori è del tutto convenzionale.

getti della stessa categoria di quelli del gruppo campione (categoria alla quale appartiene il soggetto stesso) la prestazione del soggetto verrà a cadere fra quelle dei dieci migliori.

Una tabella di graduazione pentenaria del tipo di quella precedentemente esaminata si può compilare con sufficiente precisione, senza bisogno di applicazione di formule, procedendo nella maniera seguente.

Applicando il reattivo al gruppo campione, si ordinano secondo grandezza i risultati (cioè dal minimo al massimo o dal massimo al minimo). Si suddivide quindi la serie così ordinata in dieci parti uguali, ognuna delle quali costituisce in tal modo un "decile-intervallo". Si raggruppano quindi i "decili-intervalli" corrispondentemente a quanto è stato fatto nella tabella 6, e si fissano i limiti per i cinque intervalli della classificazione pentenaria.

L'esempio che si riporta qui sotto comprende, per brevità, un gruppo campione di soli 30 soggetti. Va notato però che un gruppo campione dovrebbe essere molto più numeroso, e comunque non dovrebbe mai andare al di sotto dei 50 soggetti.

I 30 soggetti di un gruppo campione sottoposti ad una prova dell'intelligenza consistente in 45 problemi (prova del Bobertag ridotta) presentano dei risultati che variano da un minimo di 4 ad un massimo di 39 punti (ogni punto corrisponde ad un problema risolto esattamente;la durata della prova è di 30 minuti).

Ordinati dal minimo al massimo tali risultati vengono divisi in 10 gruppi comprendenti ciascuno tre ranghi (v.tav.7). Ogni gruppo costituisce un "decile-intervallo". Si raggruppano quindi i decili corrispondentemente agli intervalli della graduazione pentenaria (1,2,4,2,1) ed infine si indicano i limiti di questi nuovi intervalli. Tutto questo procedimento risulta chiaramente dalla tav. 7.

Naturalmente si può giungere allo stesso risultato anche partendo da una tabella di distribuzione di frequenza o da una curva di frequenza: si tratta allora di calcolare (o di leggere nel diagramma) i punteggi corrispondenti ai percentili 10,30,70,90, che costituiscono i limiti tra i 5 intervalli.

I metodi di graduazione basati sui percentili, trovano tuttora largo impiego in psicometria. Altri metodi, che saranno presentati nei prossimi capitoli, perchè richiedono lo sviluppo di alcuni concetti statistici, sono privi dei difetti inerenti al metodo dei percentili, e tendono ormai a sostituirsi alle scale percentilari.

Ma se la scelta fra i due metodi è facoltativa, va rilevato che la graduazione è condizione indispensabile della utilizzabilità di un reattivo.

Una prova psicologica graduata arbitrariamente, dà risultati illusori ed è quindi peggio che inutile.

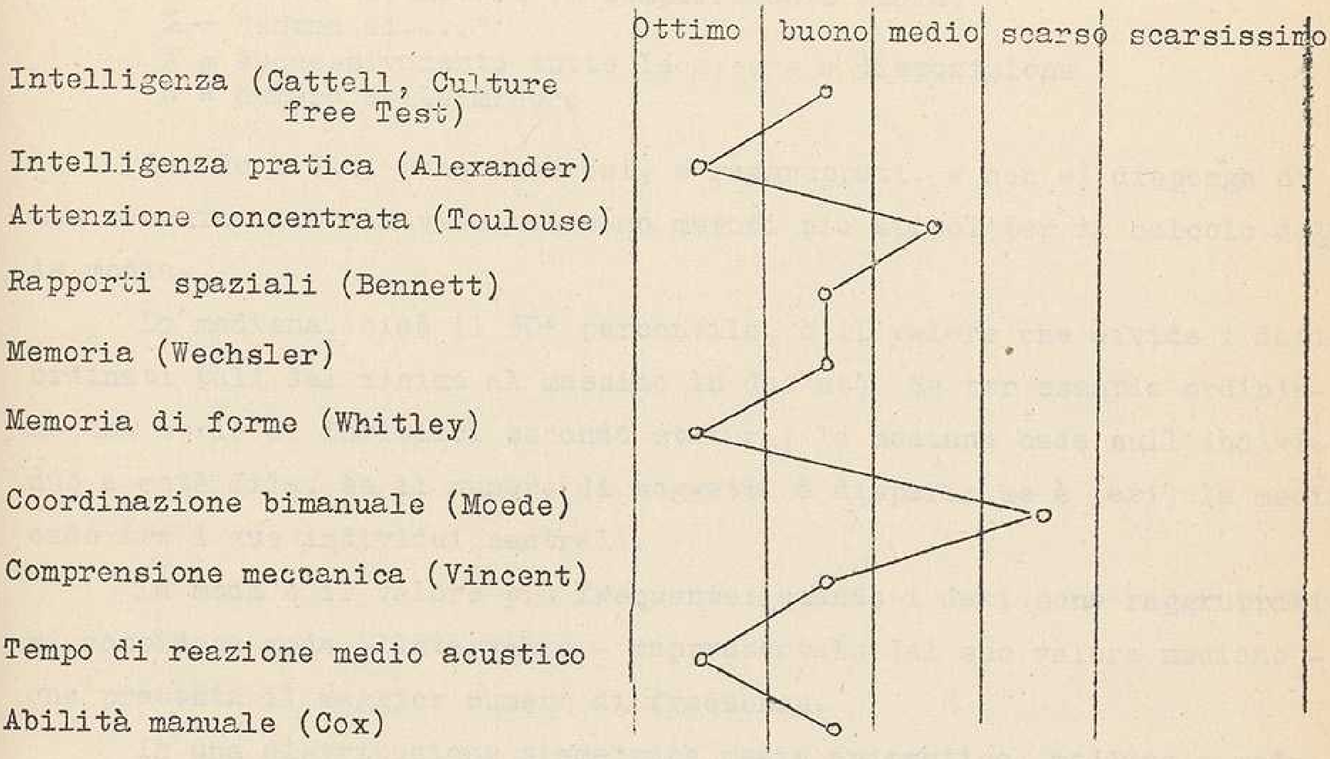
L'unitarietà del sistema di graduazione consente di confrontare direttamente i risultati conseguiti dallo stesso soggetto in differenti prove.

Si usano a tale scopo dei diagrammi (profili psicologici) che permettono di constatare a colpo d'occhio le debolezze e le qualità più spiccate rivelate dal soggetto all'esame psicologico (Tav. 8).

Punti	4-6-8	9-10-11	12-13-14	14-14-15	15-16-16	17-18-18	19-19-20	21-22-23	24-24-26	29-33-39
Rango dei soggetti	1-2-3	4-5-6	7-8-9	10-11-12	13-14-15	16-17-18	19-20-21	22-23-24	25-26-27	28-29-30
Decile-intervallo	I°	2°	3°	4°	5°	6°	7°	8°	9°	10°
Graduazione pentenaria	I°	II°		III°				IV°		V°
Limiti degli inter. pen.	fino a 8 punti	da 9 a 13 punti		da 14 a 20 punti				da 21 a 27 punti		da 28 punti in poi

Tav. 7

Tav. 8
Profilo psicologico di un soggetto
(graduazione pentenaria ad intervalli compensati)



Capitolo IV°

VALORI RAPPRESENTATIVI DI UNA COLLETTIVITA' DI RISULTATI

1) - Misure della tendenza centrale: media aritmetica, mediana e moda.

Quando è stato applicato un test a un gruppo di persone, è molto utile disporre di un unico valore per rappresentare sinteticamente l'insieme dei risultati. In psicometria si usa a tale scopo generalmente la media aritmetica, più raramente la mediana o la moda. Esse caratterizzano una popolazione, o un gruppo più o meno ristretto di persone, oppure un individuo se la misura è stata ripetuta più volte, al fine di ridurre gli errori.

La media aritmetica è definita dalla formula:

$$M = \frac{\Sigma X}{N} \qquad (2)$$

in cui M = media aritmetica (o semplicemente media)
Σ = "somma di...."
X = successivamente tutte le misure a disposizione
N = numero delle misure

Quando i dati sono numerosi, e raggruppati, e non si disponga di una macchina calcolatrice, vi sono metodi più spicci per il calcolo della media.

La mediana, cioè il 50° percentile, è il valore che divide i dati ordinati tali dal minimo al massimo in due età. Se per esempio ordiniamo una serie di individui secondo statura, la mediana cade sull'individuo a metà fila, se il numero di soggetti è dispari; se è pari, la mediana cade tra i due individui centrali.

La moda è il valore più frequente: quando i dati sono raggruppati, si considera moda l'intervallo - rappresentato dal suo valore mediano - che presenta il maggior numero di frequenze.

In una distribuzione simmetrica media aritmetica, mediana e moda coincidono esattamente.

2) Vantaggi e svantaggi delle tre misure centrali.

La moda si legge direttamente nella tabella o sul diagramma di frequenza, per cui si adopera quando si desidera una stima immediata della tendenza centrale, o quando si vuol conoscere il risultato più frequente.

Le altre due misure presentano aspetti particolari che le fanno preferire a seconda dei casi. Consideriamo ancora la fila di stature e immaginiamo p.es. che il più alto cresca; la serie nel suo insieme rimane la stessa, e invariata rimane la mediana, ma la media si sposta. La mediana non risente dei valori estremi, che possono essere eccezionali. Essa, è preferibile se si ha a che fare con un piccolo numero di risultati, in cui può non essere compensato un caso eccezionale da un altro nell'altro senso. Ed ancora, se un certo numero di soggetti raggiunge il massimo risultato possibile, non è corretto calcolare la media aritmetica, ma la mediana.

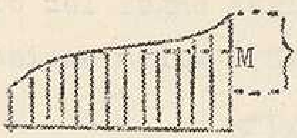
Nel caso in cui non si disponga di un'unità di misura, ma con altri metodi si possa stabilire un ordine di rango (p.es. su un aspetto del carattere, come sull'autodominio o nella valutazione di disegni o di calligrafia), si può determinare la mediana o la moda, mentre evidentemente non vi è la possibilità di calcolare la media aritmetica.

La mediana non è però un risultato che si presti ad ulteriori calcoli; normalmente è senz'altro da preferire la media aritmetica.

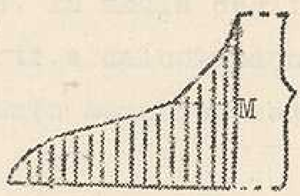
3) Misure della dispersione, o variabilità, di una distribuzione.

I valori medi, media mediana o moda, danno un solo valore rappresentativo dei risultati conseguiti dal gruppo esaminato. E' però frequente il caso di due distribuzioni con lo stesso valore medio, ma con differente dispersione dei risultati.

Consideriamo per esempio i risultati conseguiti dagli scolari di due classi in una prova di matematica, valutata col sistema usuale di 10 punti. Diamo il caso che le due medie siano eguali, 6 in tutte e due le classi, ma che la prima classe (classe A) sia piuttosto uniforme per rendimento, i risultati si distribuiscano vicini alla media, ovvero siano stati assegnati solo i voti 5,6,7; mentre nella classe B troviamo maggiori differenze di rendimento, cioè anche punteggi più lontani dalla media, come 2,3,8.9,10.... Rappresentiamo graficamente le due serie di risultati; mettiamo in ascissa i soggetti disposti in ordine dirango (dal meno bravo al più bravo) e in ordinata il punteggio conseguito da ognuno.



classe A



classe B

Fig. 7

La differenza fra i punteggi del primo e dell'ultimo (segnata dalla grafia) è maggiore nella classe B che nella classe A.

Questo fenomeno della diversa dispersione dei dati attorno al valore medio si chiama appunto dispersione o variabilità.

Una volta osservato il fenomeno, sorge il problema di trovare una misura della variabilità.

Una tale misura si può anche ottenere calcolando la differenza fra il massimo e il minimo risultato (ampiezza della distribuzione); si rischia così però di tener troppo conto dei risultati eccezionali, di avere una misura poco rappresentativa della distribuzione nel suo insieme.

Una misura della variabilità di uso comune è la differenza dei quartili, vale a dire fra il 75° e il 25° percentile.

Anzichè la differenza, si usa adoperare la semidifferenza, per uniformarsi all'ordine di grandezza delle altre misure della variabilità. Si indica con la lettera $Q = \frac{Q_3 - Q_1}{2}$ in cui Q_3 e Q_1 indicano il 3° e il I° quartile, e si chiama semiinterquartile o anche semplicamente quartile. Questa misura ha le stesse caratteristiche della mediana, con i difetti dei valori enumerativi.

In psicometria come misura della variabilità si usa spesso la variazione media o scostamento semplice medio.

Trovata la media aritmetica della distribuzione, si definisce come scarto o scostamento dalla media la differenza di un risultato individuale (X) dalla media (X-M). Si può pensare di calcolare la media di questi scarti per ottenere una misura della variabilità, senonchè, per la definizione di media aritmetica, la media degli scarti è nulla. Non tenendo conto del segno degli scarti e calcolandone la media, otteniamo la Variazione media o Scostamento semplice medio,

$$V_m = \frac{\sum |X - M|}{N} \qquad (3)$$

che si legge media degli scarti $|X - M|$ presi nel loro valore assoluto.

Tale indice di uso comune in psicologia applicata, non è utilizzabile negli scarti teorici, dove si adopera invece la media dei quadrati degli scarti, o Varianza

$$V = \frac{\sum(X - M)^2}{N} \qquad (4)$$

oppure la radice quadrata della media degli scarti quadrati, chiamato impropriamente scarto quadratico medio (Standard Deviation) o scarto tipo, o anche col nome della lettera greca <u>sigma</u> che lo indica:

$$\sigma = \sqrt{\frac{\sum(X - M)^2}{N}} \qquad (5)$$

Il σ è dello stesso ordine di grandezza della V_m e confrontabile con la media. Esso ha una grande importanza in statistica per il suo significato nella curva di distribuzione normale, ed è tra le misure della variabilità, quella più comunemente usata.

Capitolo V°

PROBLEMA DELLE UNITA' DI MISURA - GLI SCARTI RIDOTTI

In origine ogni prova ha una graduazione propria e diversa dalle altre, per cui al fatto che un punteggio ottenuto da un unico soggetto in un dato test ci dice ben poco, si aggiunge che non abbiamo possibilità di confronto di questo risultato col risultato dello stesso soggetto in un altro test. Abbiamo dato una prima soluzione al problema che ne sorge con l'introduzione deiranghi percentili, che permettono di evitare i difetti propri delle unità di misura grezze. Tuttavia, passando dalle misure ai ranghi, si perdono le caratteristiche individuali delle misure, cioè le differenze date dalle diverse distanze tra i soggetti. Diamo come esempio una serie di punteggi conseguiti da alcuni soggetti in una prova:

150	149	148	140	138	136	135

Ordinando per rango, abbiamo in corrispondenza

1	2	3	4	5	6	7

In punteggi, la distanza tra il secondo e il terzo soggetto è 149 - 148 = 1, quella tra il terzo ed il quarto 148 - 140 = 8, cioè la differenza tra questi due è di 8 punti, tra il quarto ed il quinto di 2, ecc. Passando alla classificazione per rango, queste differenze vengono perdute, sappiamo soltanto che il quarto è migliore del terzo, il terzo del secondo, ecc., ma non di quanto.

Analogamente il sistema dei ranghi percentili considera eguali grandezze che son diverse. Esso ci fornisce un ordine di classificazione, e non una serie di misure. Se praticamente i ranghi percentili talvolta possono essere sufficienti, non sono però adatti ad ulteriori calcoli.

Si dovrà quindi introdurre una misura vera e propria che sia sempre la stessa per ogni tipo di test.

A tale scopo si assume come origine del sistema di misure la media aritmetica della distribuzione. In altre parole, si misura in scarti dalla media; la media assume il valore zero, e gli scarti saranno positivi se superiori alla media, negativi se inferiori (gli scarti si indicheranno con x minuscola, mentre i valori grezzi si indicheranno con X maiuscola).

Il vantaggio che porta l'introduzione di una origine comune per qualsiasi test e per qualsiasi distribuzione è che dal segno dello scarto si vede immediatamente se un risultato conseguito in un test è superiore o inferiore al risultato medio del gruppo esaminato. Tuttavia gli scarti dalla media non costituiscono un sistema comune alle diverse prove. A tale scopo si deve tenere conto non solo della media aritmetica della distribuzione ma anche della variabilità. E' chiaro infatti che un punteggio 65 ha diverso significato non solo se la media del gruppo sia 30 o 60, ma anche se - essendo la media 60 - i punteggi siano raggruppati tra 50 e 70 o se invece si distribuiscono tra 20 e 100.

Perciò il risultato di un soggetto in un test, espresso in scarti dalla media $x = X - M$, viene diviso per l'indice di variabilità σ o in altre parole l'indice di variabilità della distribuzione viene assunto come unità di misura. Si ottiene in tal modo la così detta scala zeta, in scarti ridotti.

$$z = \frac{x}{\sigma} = \frac{X - M}{\sigma} \qquad (6)$$

Se un risultato è eguale alla media, $z = 0$, se superiore z è positivo, negativo se inferiore alla media; σ rappresenta l'unità di misura.

L'unico svantaggio delle misure z è che per differenziare fra loro i risultati bisogna usare i decimali, perchè è eccezionale trovare risultati oltre ± 3 z.

Perciò si usa spesso una scala analoga, in cui però la media è fissata a 50 e il σ è 10 anzichè 1. Il passaggio alle nuove unità, indicate col simbolo Z maiuscolo si realizza tenendo presente che

$$Z = 10_z + 50$$

Capitolo 6°

LA CORRELAZIONE

1) Concetto di correlazione

Consideriamo due distribuzioni di dati, relative ad uno stesso gruppo di soggetti, per esempio i voti relativi ad uno stesso gruppo di scolari in matematica e in latino, oppure i rilievi di peso e di statura. Spesso si nota che le due serie variano in senso più o meno concorde, cioè che gli scolari migliori in matematica sono pressappoco gli stessi che sono migliori in latino, i peggiori in matematica sono scadenti anche in latino, e i medi in una materia, rimangono medi anche nell'altra; che i più alti di statura sono quasi sempre quelli che pesano di più, e viceversa. In altre coppie di dati statistici, non vi è assolutamente modo di scoprire un simile andamento, per es. tra le serie delle stature e i risultati in latino è del tutto improbabile che ci sia una qualche concordanza. Infine, fra due altre serie può scoprirsi una più o meno spinta concordanza inversa, che cioè nella maggioranza dei casi gli individui superiori nell'una serie siano inferiori nell'altra, e viceversa.

E' questo il fenomeno della correlazione.

Allorchè si scopre una certa concordanza tra due serie di dati, si dice che v'è una certa correlazione; correlazione diretta o positiva, se le due serie variano più o meno nello stesso senso (come i voti di matematica e quelli di latino), correlazione negativa o inversa se variano in senso opposto (p.es. c'è in genere una correlazione inversa tra l'intelligenza e l'età per gli scolari di una stessa classe, poichè, tranne casi particolari, i maggiori di età sono i ripetenti per incapacità). Se la concordanza è perfetta, se p.es. il migliore in un test è il migliore anche nell'altro, il secondo nell'uno è secondo anche nell'altro, e così via, si parla di perfetta correlazione positiva. Se il migliore in un test è il peggiore nell'altro, il secondo penultimo, il terzo terzultimo ecc., si parla di correlazione negativa perfetta. Se non appare nessuna concordanza, ma c'è completa indipendenza tra le due serie, si dice che non cè correlazione, o che la correlazione è nulla.

Per chiarire il concetto di correlazione, di importanza fondamentale in psicometria, prendiamo in considerazione due serie di dati relativi agli stessi soggetti, nel test X (con punteggi da 10 a 75) e nel test Y (con punteggi da 20 a 50). Si tratterà di procedere a una classificazione rispetto a due modalità contemporaneamente; si dovrà costruire perciò una tabella meno semplice di quella per la distribuzione di frequenza di un test, precedentemente considerata.

La nuova tabella si chiama tabella a doppia entrata; essa si costruisce segnando orizzontalmente i punteggi possibili di un test (generalmente raggruppati in intervalli di frequenza) in ordine crescente da sinistra a destra, e verticalmente i punteggi dell'altro test in ordine crescente dal basso all'alto (v.tav.9). Se il soggetto A ha ottenuto 44 nel test X e 40 nel test Y, si segna un trattino nella casella all'incrocio dell'intervallo 40-49 di X con 38-42 di Y. Se il soggetto B ha 21 in X e 36 in Y, un trattino nella casella che gli compete, e così via per ogni soggetto. Esauriti i soggetti, si conta il numero dei trattini segnati in ogni singola casella, e si segna il numero corrispondente.

La densità nelle varie zone della tabella a doppia entrata può non rivelare niente di particolare: si può riscontrare un addensamento nelle caselle centrali, e minore frequenza nelle caselle della periferia, ma ciò è conforme all'osservazione generale che i punteggi medi sono i più comuni. In altri casi si nota invece una netta preferenza per le caselle che stanno sulla diagonale sinistra basso-destra alto, la qual cosa significa che i soggetti con punteggi superiori in una prova sono superiori anche nell'altra, gli inferiori nell'una hanno punteggi inferiori nell'altra, e i medi rimangono medi: ci sono allora le condizioni per affermare che c'è una correlazione positiva tra le due prove. Viceversa l'eventuale addensamento lungo l'altra diagonale sinistra alto-destra basso è segno di una correlazione inversa, o negativa.

Naturalmente, l'addensamento lungo una diagonale può essere più o meno netto, e può andare da una distribuzione esattamente rettilinea (correlazione perfetta) ad una distribuzione di forma grossomodo circolare, a densità degradante verso l'esterno (correlazione nulla). Nei casi intermedi, i dati si dispongono all'incirca in una forma ovale;

l'asse maggiore dell'ovale va da sinistra in basso a destra in alto se la correlazione è positiva, da sinistra in alto a destra in basso se negativa; più schiacciato è l'ovale, tanto più elevata è la correlazione.

Sogg.	Test X	Test Y		X	Y		X	Y
C	44	38	I	39	30	O	48	34
D	41	37	J	52	27	P	70	31
E	40	33	K	33	20	Q	36	32
F	64	35	L	50	41	R	62	43
G	59	44	M	37	37	S	49	49
H	20	26	N	53	29	....		

Test Y \ Test X	10-19	20-29	30-39	40-49	50-59	60-69	70-79
48-52				1			
43-47					1	1	
38-42				1	1		
33-37		1	1	3		1	
28-32			2	1	1		1
23-27		1			1		
18-22			1				

Tav. 9 - Tabella a doppia entrata

(per semplicità raggruppiamo i punteggi in soli 7 intervalli)

Il concetto di correlazione è qualche cosa di diverso dal concetto di funzione, e non deve essere confuso con esso.

Si dice che una grandezza variabile è funzione di un'altra grandezza variabile, e viceversa, quando a un valore dell'una corrisponde, secondo una data legge, un ben determinato valore dell'altra.

Per rappresentare graficamente una funzione, si scelgono due rette ortogonali (gli assi cartesiani) e - fissato un segmento come unità di misura, e l'origine al punto d'incontro delle due rette - si riportano su uno degli assi (p.es. sull'orizzonte, che viene a chiamarsi ascissa) successivamente i valori della variabile che si considera indipendente; da ognuno di questi valori ci si sposta lungo una parallela all'altro asse (asse della funzione, detto delle ordinate) fino alla distanza equivalente al valore corrispondente dell'altra variabile, considerata dipendente, e si segna un punto. L'insieme di questi punti costituisce generalmente una linea, che rappresenta la funzione. Una funzione può essere esprimibile mediante una equazione algebrica, ma può anche non esserlo.

Consideriamo alcuni tipi di funzioni:

a) Tra due serie di numeri x ed y: ad ogni numero x compreso in un dato intervallo corrisponde un preciso valore di y: per esempio la funzione esprimibile mediante la equazione $y = ax + b$ (in cui a e b sono numeri, costanti al variare della x e della y) che è l'equazione di una retta.

b) Tra due entità geometriche, per esempio la relazione tra il raggio e la circonferenza di un cerchio: dato il raggio r, è univocamente determinata la circonferenza c (e viceversa, secondo l'equazione $c= 2\pi r$.

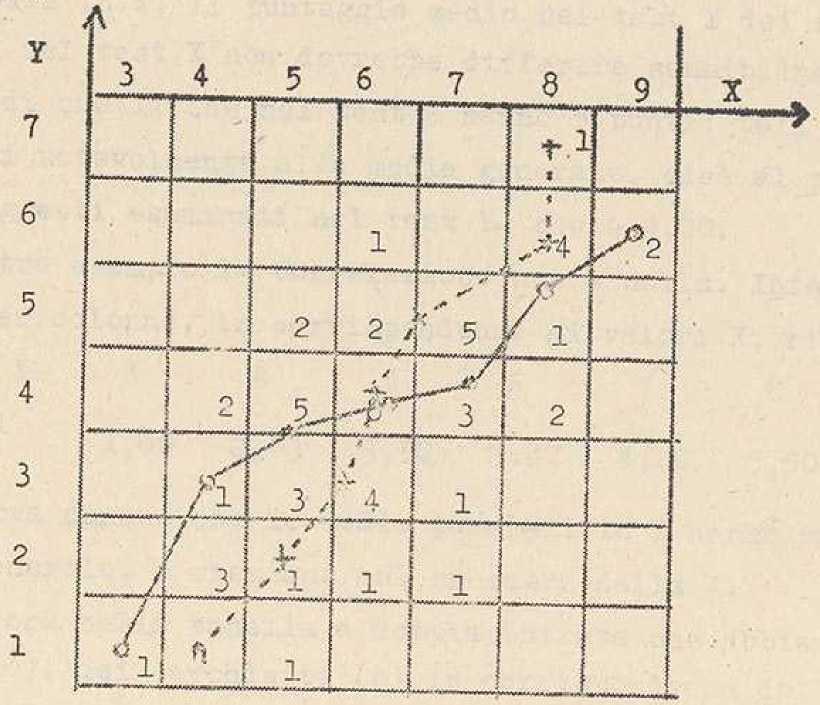
c) Funzione empirica, pure univocamente determinata, ma non esprimibile con una equazione; in cui cioè, dato un valore di una variabile, non si ha modo di indicare una operazione algebrica per determinare il valore dell'altra. Un esempio di funzione empirica è dato dai rilievi di temperatura in un intervallo di tempo t: ad ogni valore di t corrisponde un valore della temperatura ambiente.

d) Relazione indotta, come la legge di caduta dei gravi di Galileo o la legge di Boyle-Mariotte dei gas perfetti. Si tratta di leggi deter minate a posteriori dopo numerose sperimentazioni, ritenendo di aver eliminato gli errori di misura.

Un concetto più vasto, in quanto comprende il concetto di funzione, è quello di correlazione. La correlazione - come abbiamo detto - può presentarsi in gradi diversi; una correlazione perfetta tra due variabili è semplicemente una relazione di funzione tra di esse. Nel caso generale in cui tra due variabili esista un certo nesso di correlazione, ad un valore di una variabile corrispondono più valori dell'altra, i quali però tendono a concentrarsi attorno ad un valore medio. Quanto meno elevata è una correlazione, tanto più dispersi sono questi valori.

Per maggior chiarezza, considieriamo la tabella a doppia entrata (Tav. 10) nella quale sono distribuite le coppie di risultati di 55 soggetti in due prove, X e Y, i cui punteggi vanno rispettivamente dal 3 al 9, e dall'1 al 7.

96 =



Tav. 10 - Esempi di tabella a doppia entrata

Valori rappresentativi delle due distribuzioni:
$M_x = 6,09$; $\sigma_x = 1,38$; $M_y = 4,00$; $\sigma_y = 1,33$

O : media per colonne
+ : media per righe

Osserviamo per esempio la colonna corrispondente al punteggio 8 del test X: ci sono 8 soggetti che hanno ottenuto 8 punti nel test X, e di questi uno ha ottenuto 7 punti nel test Y, mentre 4 hanno ottenuto 6 punti, uno 5 punti, due 4 punti. Osserviamo ancora la colonna relativa al punteggio 5 nel test X: pure questi 12 soggetti che sono pari quanto alla prova X danno risultati diversi nella prova Y.

Proviamo ora a calcolare la media aritmetica dei risultati in Y di ogni gruppetto di soggetti che sono pari quanto al risultato nel test X. Per esempio, per gli 8 soggetti con 8 punti in X il punteggio medio in Y è 1.7 + 4.6 + 1.5 + 2.4 = 44; 44 : 8 = 5,50; i 12 soggetti con 5 punti nel test X hanno per media 3,50 nel test Y. Se la correlazione fosse nulla, cioè se vi fosse completa indipendenza tra i risultati in X e quelli in Y, il punteggio medio nel test Y dei soggetti che hanno 8 punti nel test X non dovrebbe differire sensibilmente, dal punteggio medio di quelli che nel test X hanno 5 punti: tali medie dovrebbero avvicinarsi notevolmente alla media generale, cioè al punteggio medio di tutti i soggetti esaminati nel test Y- che è 4,00.

Nel nostro esempio la correlazione non è nulla. Infatti la serie delle medie per colonne, in corrispondenza ai valori X, risulta:

Valori di X	3	4	5	6	7	8	9
corrispondenti medie in Y	1,00	2,83	3,50	3,87	4,20	5,50	6,00

Si osserva dunque che le medie parziali in Y hanno valori diversi dalla media generale, e crescono col crescere della X.

Segnamo ora nella tabella a doppia entrata che abbiamo portato ad esempio (tav.10), dei cerchietti (o) in corrispondenza delle medie parziali nel test Y relative ai valori di X da 3 a 9, e congiungiamo i cerchietti con una spezzata (a segno continuo, ———).

La spezzata si approssima in genere ad una retta, che passa per il punto d'incontro delle due medie generali, retta che si può determinare con un metodo grafico, o analitico. Essa prende il nome di <u>retta di regressione</u> di Y su X.

Ripetiamo ora per le righe quello che abbiamo fatto per le colonne della tabella a doppia entrata. Calcoliamo cioè la media dei risultati in X dei soggetti che sono pari nel test Y. Risulta (vedi tav. 10):

Y	1	2	3	4	5	6	7
X	4,00	5,00	5,55	5,90	5,90	6,50	8,00

Anche qui si osserva che i punteggi medi in X crescono col crescere dei punteggi in Y.

Segnando nella tabella i punti corrispondenti alle medie per righe (con una crocetta, +) e congiungendo le crocette (con segno tratteggiato, - - -), otteniamo una spezzata, che pure si avvicina a una retta che si può determinare con uno dei metodi sopra citati. Essa è la retta di regressione di X su Y (1).

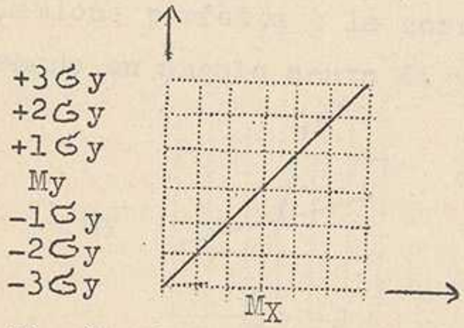


Fig.8 - Correlazione perfetta positiva

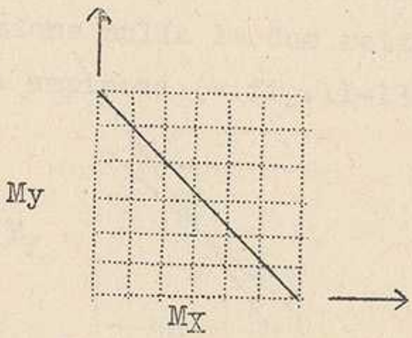


Fig. 9 - Correlazione perfetta negativa

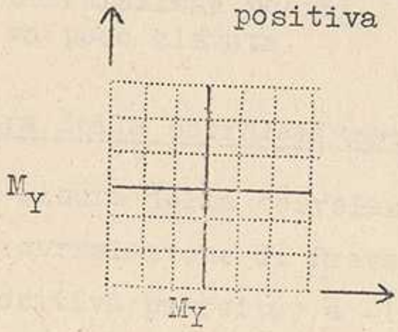


Fig.10 - Correlazione nulla

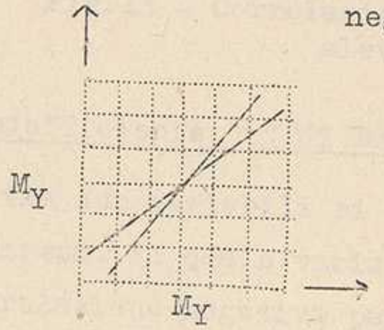


Fig. 11 - Correlazione positiva elevata

---

(1) Il termine regressione deriva dall'osservazione, compiuta da Francis Galton verso la fine del secolo scorso, che i figli di genitori molto alti di statura erano in media più alti del normale, ma meno di quanto lo fossero i genitori, e che i figli di genitori molto bassi di statura erano in media un po' più alti dei genitori; cioè che, in media, le stature dei figli regrediscono verso la media generale in confronto alle stature dei genitori. A Galton era sembrato che ciò fosse in contrasto con la teoria dell'evoluzione delle specie di suo cugino Charles Darwin.

La posizione delle due rette di regressione, indica il grado e la natura della correlazione. Nel caso di correlazione perfetta, esse coincidono; infatti ad un valore dell'una variabile corrisponde sempre lo stesso valore nell'altra. In particolare, se la correlazione perfetta è positiva le due rette coincidono in direzione sinistra basso-destra alto (v.fig.8); se la correlazione perfetta è negativa, si avrà un'unica retta in direzione alto-destra basso (v.fig.9). Se la correlazione è nulla, le due rette sono perpendicolari e giacciono sugli assi corrispondenti alle due medie generali (v.fig.10). Nei casi intermedi tra la correlazione perfetta e la correlazione nulla le due rette di regressione formano un angolo acuto di varia ampiezza (v.fig.11-13).

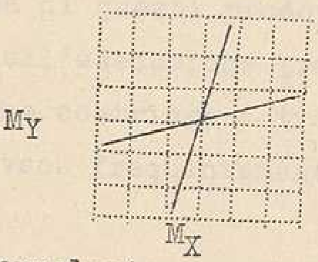


Fig.12 - Correlazione positiva poco elevata

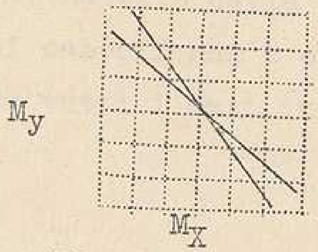


Fig.13 - Correlazione negativa elevata

2) - <u>Misura della correlazione: il coefficiente "r" di Bravais-Pearson.</u>

Come misura della correlazione fra due variabili si usa il coefficiente di correlazione di Bravais-Pearson, il quale varia da +1 (correlazione positiva perfetta) a -1 (correlazione negativa perfetta); lo zero indica assenza di correlazione. Il simbolo usato per il coefficiente di correlazione di Bravais-Pearson è r_{xy} (correlazione fra le due variabili x e y).

Il coefficiente di correlazione tra le due coppie di dati x ed y (nel nostro caso, generalmente tra due tests) è definito dalla formula

$$r_{xy} = \frac{\sum z_x z_y}{N} \qquad (7)$$

nella quale i dati sono espressi in unità confrontabili, e precisamente in scarti ridotti (z_x e z_y), ed N è il numero di coppie di dati che si prendono in considerazione (se si tratta di correlazione tra due tests, N è il numero dei soggetti esaminati tanto nel test X che nel test Y).

Per calcolare r_{xy} secondo la formula (7) è quindi necessario aver prima calcolato media e sigma di ciascuna delle due distribuzioni, aver trasformato i valori grezzi in scarti dalla media, e averli rapportati al sigma della distribuzione. A questo punto si eseguiranno i prodotti tra z_x e z_y relativi ad ogni soggetto, e si calcola infine la mèdia aritmetica di questi prodotti, come indicato dalla (7).

Ma nell'esecuzione pratica dei calcoli non è conveniente usare la (7), che costringe a trasformare tutti i dati in scarti ridotti. Si usa invece frequentemente la

$$r_{xy} = \frac{\Sigma xy}{N \sigma_x \sigma_y} \qquad (8)$$

che si ottiene dalla (7) sostituendo a z_x, e a z_y, rispettivamente $\frac{x}{\sigma x}$ e $\frac{y}{\sigma y}$, o la

$$r_{xy} = \frac{\Sigma xy}{\sqrt{\Sigma x^2 \; \Sigma y^2}} \qquad (9)$$

che si ottiene dalla (8) sostituendo a σ_x e a σ_y, rispettivamente

$$\sqrt{\frac{\Sigma x^2}{N}} \qquad \sqrt{\frac{\Sigma y^2}{N}}$$

Le formule (8) e (9) impongono tuttavia anch'esse, la trasformazione dei dati grezzi (X, Y) in scarti dalla media ($x = X - M_x$; $y = Y - M_y$). Perciò, per l'esecuzione pratica dei calcoli si suole preferire la

$$r_{xy} = \frac{\frac{\sum XY}{N} - M_x M_y}{\sigma_x \sigma_y} \qquad (10)$$

che permette di utilizzare direttamente i dati grezzi, e si ricava dalla (8) sostituendo a x e y rispettivamente $X - M_x$ e $Y - M_y$, o altre formule equivalenti.