

# RIASSUNTO

del libro:

Metodologia della **Ricerca**  
per le Scienze  
Motorie e Sportive.

di J.R. Thomas, J.K. Nelson, S.S. Silverman

prima edizione italiana a cura di P. Bellotti, A. Rainoldi e del Centro Ricerche SUISM

stampato da Calzetti Mariucci / Tivoli - 2012

ISBN: 978-88-6028-239-2

A cura di:

**Antonio Gualtieri**

(Agosto 2012)

## PARTE 1 – SGUARDO D'INSIEME AL PROCESSO DI RICERCA

Pag.	Cap. 1 – Introduzione alla ricerca nell'attività motoria.
13	Domande da farsi per condurre un'INDAGINE SISTEMATICA: qual è il problema → introduzione, come è stato studiato → materiali e metodi, cosa è stato trovato → risultati, cosa significa quanto trovato → discussione.
15	La RICERCA risolve problemi che l'uomo preliminarmente si pone.
16	La RICERCA implica un approccio costituito da 5 caratteristiche: esso è sistematico, logico, empirico, riduttivo e replicabile. La ricerca è <i>l'insieme degli studi e delle indagini che si svolgono nell'ambito delle discipline scientifiche o umanistiche per individuare documenti e fonti, ricostruire eventi o situazioni, scoprire fenomeni, processi, regolarità, leggi, ecc.</i> RICERCA APPLICATA: ha diretta validità per la pratica e per l'uomo di campo, ma nella quale il ricercatore ha un controllo limitato sullo scenario della ricerca.
17	<b>CONTINUUM DELLA RICERCA:</b> ricerca di base e ricerca applicata sono i due terminali di un continuum → la <b>RICERCA DI BASE</b> si indirizza a problemi teorici, spesso in contesti di laboratorio (alto controllo), mentre la <b>RICERCA APPLICATA</b> si confronta con problemi che hanno una considerevole immediatezza, spesso in contesti del mondo reale meno controllati, ed è più strettamente legata all'applicazione pratica.
18	I punti forti della ricerca applicata sono i punti deboli della ricerca di base. <b>VALIDITÀ ECOLOGICA:</b> si riferisce al grado maggiore o minore di emulazione della realtà che presenta la ricerca. Deve far fronte a 2 problemi: come è percepito l'ambiente dal partecipante e se l'ambiente utilizzato possiede un numero sufficiente di caratteristiche del mondo reale.
19	7 step per realizzare una ricerca di qualità/tesi. <b>PERCHÉ STUDIARE MISURE E ANALISI?</b> Per comprendere i risultati della Ricerca: infatti, anche se la maggior parte delle persone ne comprende la necessità, la maggior parte non si informa sui risultati, forse perché non riescono davvero a capirli (solo 1% dei chimici legge pubblicazioni di ricerca).
20	<i>Computerese</i> → linguaggio degli elaboratori, rende statistica di più difficile comprensione.
22	<b>SCIENZA:</b> processo di indagine attento e sistematico (cfr. p. 16); investigazione strutturata (cfr. p. 27) e disciplinata, non un insieme di procedure specifiche → problemi differenti richiedono soluzioni differenti (cfr. p. 33).
23 24	La ricerca ha a che fare con la soluzione di problemi, la qual cosa <u>può</u> portare a nuove conoscenze. Metodi <u>NON</u> SCIENTIFICI di soluzione dei problemi secondo HELMSTADTER: 1. <b>CONVINZIONE OSTINATA:</b> le persone si radicano in alcune credenze indipendentemente dal fatto che esse manchino di conferme evidenti (superstizioni, comportamenti stereotipati, portafortuna). 2. <b>INTUIZIONE:</b> conoscenza intuitiva = senso comune/fenomeno evidente di per sé. Molte delle evidenze sono solo in apparenza verità, poiché esse in seguito verranno trovate false (es. Terra piatta). 3. <b>AUTORITÀ</b> (e autorevolezza della fonte): la qualifica del valore dell'autorità nonché quello dei metodi con cui l'autorità stessa ha acquisito le conoscenze determinano anche la validità di questa fonte di informazione (es. Galileo/Giordano Bruno vs Tolomeo et Chiesa). 4. <b>METODO RAZIONALISTICO:</b> conoscenza viene derivata attraverso il ragionamento (es. sillogismo). Il risultato è affidabile solo se proveniente da premesse che siano vere. Il ragionamento è fondamentale per il metodo scientifico, ma non può essere usato da solo per arrivare alla conoscenza. 5. <b>METODO EMPIRICO:</b> <i>empirico</i> = esperienza e raccolta dei dati, obiettiva osservazione. Ci sono dei limiti di affidabilità (la nostra memoria non è oggettiva, in quanto influenzata da <i>n</i> fattori).

25 - 26	<p>Il <b>METODO SCIENTIFICO DI SOLUZIONE DEI PROBLEMI</b> utilizza i seguenti passaggi:</p> <ol style="list-style-type: none"> <li><b>DEFINIZIONE E DELIMITAZIONE DEL PROBLEMA:</b> individuare la <b>VARIABILE INDIPENDENTE</b> / sperimentale o di trattamento → <u>causa</u> (ciò che il ricercatore manipola, es. metodo di insegnamento) e la <b>VARIABILE DIPENDENTE</b> / rendimento → <u>effetto</u> (effetto della variabile indipendente, es. la misura delle abilità conseguenti al metodo di insegnamento).</li> <li><b>FORMULAZIONE DI UN'IPOTESI:</b> l'ipotesi è il risultato atteso e deve essere <u>verificabile</u> caratteristica necessaria). Se è verificabile, lo studio la confermerà o la confuterà.</li> <li><b>RACCOLTA SISTEMATICA DEI DATI: VALIDITÀ INTERNA</b> → misura di quanto i risultati di uno studio possano essere influenzati dai trattamenti usati nello studio stesso. <b>VALIDITÀ ESTERNA</b> → possibilità di generalizzazione dei risultati di uno studio. La validità interna ed esterna si riferiscono al disegno della ricerca e ai controlli che vengono utilizzati.</li> <li><b>ANALISI E INTERPRETAZIONE DEI DATI RACCOLTI:</b> viene utilizzato il ragionamento <b>INDUTTIVO</b>, mentre quello <b>DEDUTTIVO</b> viene utilizzato nell'enunciazione del problema. Il ricercatore cerca di <u>sintetizzare i dati del suo studio con i risultati di altri studi per contribuire allo sviluppo o alla maggiore consistenza di una teoria.</u></li> </ol>
27	<p>OBIETTIVO FONDAMENTALE DELLA SCIENZA → spiegare le cose o essere in grado di generalizzare e di costruire una teoria.</p> <p><b>PARADIGMA:</b> insieme di teorie, leggi e strumenti che definiscono una tradizione di ricerca all'interno della quale le teorie sono accettate da tutti i cultori (es. Scienza normale).</p> <p><b>SCIENZA NORMALE:</b> maniera obiettiva di affrontare uno studio, basata sulle scienze naturali, con le caratteristiche di essere sistematica, logica, empirica, riduttiva e replicabile.</p> <p>KUHN: storico della scienza, definì <i>paradigma</i> e <i>scienza normale</i>. Rifiutò l'idea di conoscenza oggettiva sostenendo il pieno coinvolgimento dell'osservatore nel progetto di ricerca.</p>
28	<p>TEORIA QUANTISTICA: descrive i sistemi come una sovrapposizione di dati diversi e prevede che il risultato di una misurazione non sia completamente arbitrario, ma sia incluso in un insieme di possibili valori → i fenomeni potrebbero essere statisticamente previsti, ma non spiegati deterministicamente.</p>
29	<p>POPPER → <b>FALSIFICAZIONISMO:</b> ogni esperimento nasce con l'intento di falsificare la teoria in voga. Se la teoria resisterà agli attacchi si rafforzerà e maggiore potrà essere la fiducia in essa. La verità è temporanea, sopravvive fino a quando non venga confutata da altra teoria che spiegherà ciò che la teoria precedente non era in grado di capire, ma <u>includendo</u> le conoscenze precedenti già accettate.</p> <p>PEIRCE → <b>ABDUZIONE:</b> si aggiunge a deduzione (top-down → approccio dogmatico) e induzione (down-top → misura empirica). Consiste nel formulare delle ipotesi plausibili, andando dall'antecedente alle conseguenze (es. dei fagioli bianchi).</p>
30	<p>FEYERABEND: le deviazioni dai codici di "onestà intellettuale" sembrano essere le condizioni preliminari per il progresso. Critica la visione troppo strutturata della filosofia della scienza → per progredire, la scienza ha bisogno della libertà più assoluta.</p>
31	<p>FENOMENO DELLA CRISI DEL PARADIGMA (Kuhn): sviluppo di discrepanze in un paradigma (risultati non più compatibili con previsioni) che porta al proposito di un nuovo paradigma in grado di spiegare meglio i dati.</p> <p><b>RICERCA QUALITATIVA</b> (cfr. p. 37): metodo di ricerca che spesso prevede un'osservazione lunga e approfondita in un ambiente naturale, la registrazione precisa e dettagliata di ciò che si verifica nell'ambiente di studio, l'interpretazione e l'analisi dei dati utilizzando descrizioni, narrazioni, citazioni, carte e tavole. Sono in essa incluse la ricerca etnografica, naturalistica, interpretativa, fondata, fenomenologica, soggettiva ed osservazionale relativamente ai partecipanti allo studio.</p> <p>MARTENS accusa la scienza normale di preferire la <i>possibilità di pubblicazione</i> al significato pratico della ricerca.</p>
32	<p><b>RIDUZIONISMO:</b> caratteristica della scienza normale che assume che i comportamenti complessi possano essere ridotti, analizzati e spiegati come singole parti che possono essere poi messe di nuovo insieme al fine di comprendere il tutto.</p> <p>FIG. 1.1: teoria dei differenti gradi di conoscenza con esempi di tipi differenti di metodi varianti nel grado di affidabilità da <i>non conoscenza a dannatamente sicuro</i> → TEORIA DK (Damn Konfident). Approcci variati ad una ricerca disciplinata sono utili per l'accumulo di conoscenze.</p>
33	<p>I ricercatori potrebbero muoversi fra i livelli di ricerca (da basica ad applicata) e anche fra i paradigmi (da quantitativo a qualitativo) per acquisire conoscenza.</p>

33 - 34	<p><b>RICERCA ANALITICA:</b> tipo di ricerca in cui si studiano approfonditamente le informazioni disponibili, al fine di spiegare fenomeni complessi. Può essere categorizzata nel seguente modo:</p> <ol style="list-style-type: none"> <li>1. <b>RICERCA STORICA:</b> si focalizza sugli avvenimenti, le organizzazioni, le istituzioni e le persone.</li> <li>2. <b>RICERCA FILOSOFICA:</b> caratterizzata da un'indagine critica molto rigorosa.</li> <li>3. <b>RASSEGNA (REVIEW):</b> lavoro di ricerca che si presenta come valutazione critica della ricerca su un argomento particolare.</li> <li>4. <b>SINTESI DI RICERCA (METANALISI → <i>where we are and where we want to go</i>):</b> tecnica clinico-statistica che permette di analizzare diversi studi condotti sullo stesso argomento, consentendo una sintesi quantitativa dei risultati. Si presenta particolarmente utile quando i risultati dei singoli studi sono discordanti e in presenza di studi di "potenza ridotta", ovvero effettuati su pochi partecipanti e ritenuti per questo scarsamente affidabili. Consente di realizzare un quadro riassuntivo di un determinato argomento, di aumentare la potenza di un campione, grazie alla dimensione più consistente dello stesso e di effettuare analisi di sottogruppi. I suoi svantaggi consistono in <i>bias</i> di pubblicazione e di selezione.</li> </ol>
34 - 35 - 36	<p><b>RICERCA DESCRITTIVA:</b> si propone di descrivere lo status del punto centrale di uno studio. <u>NON è possibile stabilire un rapporto di causa effetto.</u> Tecniche comunemente utilizzate sono:</p> <ol style="list-style-type: none"> <li>1. <b>QUESTIONARIO:</b> forma scritta per reperire info sulle abitudini attuali, sulle condizioni e sui dati demografici, ecc. Raramente chiede opinioni o indaga su conoscenze.</li> <li>2. <b>INTERVISTA:</b> come sopra, ma in forma orale (di persona o telefonicamente).</li> <li>3. <b>SONDAGGIO NORMATIVO:</b> cerca di riunire dati di capacità o di conoscenza su un campione numeroso e di presentare i risultati sotto forma di standard comparativi o di norme.</li> <li>4. <b>STUDIO DEL CASO:</b> usato per fornire info dettagliate su un singolo → mira a determinare caratteristiche uniche del soggetto o della sua particolare condizione (es. <i>counseling</i> psicologico).</li> <li>5. <b>JOB ANALYSIS:</b> particolare forma di "studio del caso", usata per descrivere la natura di un particolare lavoro.</li> <li>6. <b>RICERCA OSSERVAZIONALE:</b> comportamenti osservati nell'ambiente naturale e proprio dei partecipanti (es. classe).</li> <li>7. <b>STUDI DI SVILUPPO:</b> ad esempio, per determinare gli effetti della crescita su un parametro fisiologico. <b>METODO LONGITUDINALE:</b> gli stessi partecipanti studiati per un periodo di anni. <b>METODO TRASVERSALE:</b> seleziona campioni di partecipanti provenienti da diversi gruppi di età per valutare gli effetti della maturazione.</li> <li>8. <b>STUDIO DELLE CORRELAZIONI:</b> es. correlazione tra misura delle pliche cutanee e % grasso corporeo.</li> <li>9. <b>RICERCA EPIDEMIOLOGICA:</b> vi appartengono le frequenze e le distribuzioni delle condizioni di malattia e salute per varie popolazioni. Non è possibile stabilire causa-effetto sulla base di incidenza e prevalenza, ma una forte inferenza di causa può essere sovente tentata mediante l'associazione.</li> </ol>
36	<p><b>RICERCA SPERIMENTALE:</b> modalità di ricerca in cui i trattamenti vengono manipolati nel tentativo di stabilire relazioni di causa-effetto tra i diversi fenomeni. Vantaggio rispetto agli altri tipi → ricercatore può manipolare i trattamenti che causano gli eventi che accadono.</p>
37	<p><b>RICERCA QUALITATIVA</b> (cfr. p. 31): raramente stabilisce ipotesi; ricercatore = principale strumento di raccolta e analisi dati → presenza intensiva in prima persona del ricercatore. Principali strumenti di raccolta dati sono l'osservazione e le interviste.</p> <p><b>METODI MISTI O MODELLI MISTI DI RICERCA:</b> miscela di metodi quantitativi e qualitativi → modo pragmatico di agire, utili quando si studiano i fenomeni del mondo reale.</p>
38	<p>FIG. 1.2: <i>setting</i> complessivo del processo di ricerca (schema riassuntivo).</p>
39	<p>La rassegna critica della letteratura è strumentale alla formazione di ipotesi ad al ragionamento deduttivo che porta alla definizione precisa del problema che si intende indagare.</p>

## PARTE 2 – FONDAMENTI DI STATISTICA E MISURE PER LA RICERCA

Pag.	Cap. 6 – Familiarizzare con concetti statistici
107	La STATISTICA è uno dei pochi modi in cui dei dati possono essere presentati in maniera uniforme, così da permettere analisi comparative; la statistica è metodica, logica e necessariamente non casuale, contraddittoria o spaventosa (cfr. p. 119).
108	<b>STATISTICA:</b> metodo per interpretare una raccolta di osservazioni. Per <a href="#">descrivere (1)</a> le caratteristiche dei dati, per valutare le <a href="#">relazioni (2)</a> (aspetti comuni → regressione o correlazione) tra gruppi diversi di dati e per misurare le <a href="#">differenze (3)</a> ( → test <i>t</i> e analisi di varianza) tra gruppi vengono utilizzate <u>tecniche statistiche diverse</u> (cfr. p. 118).
109 - 110	Pacchetti SOFTWARE: SPSS → Statistical Package for the Social Sciences; SAS → Statistical Analysis System. Descrizioni e INDUZIONI (o inferenza → generalizzazione dei risultati a una popolazione più ampia) non sono tecniche statistiche. Ogni statistica descrive in realtà solo il campione di partecipanti per cui è stata calcolata. Per questo motivo il metodo per la selezione del CAMPIONE, le procedure e il contesto sono ciò che permetterà o meno le corrette deduzioni (cfr. p. 111). <b>CAMPIONE:</b> gruppo di partecipanti, trattamenti o situazioni selezionati da una POPOLAZIONE più vasta su cui viene condotta la ricerca. Può essere selezionato in diversi modi: 1. SELEZIONE RANDOM: selezionando il campione in modo casuale, questo sarà realmente rappresentativo della popolazione; in questo modo i risultati ottenuti sul campione possono essere applicati tramite un processo <u>induttivo</u> all'intera popolazione. Strumenti: TABELLA DEI NUMERI CASUALI. 2. CAMPIONAMENTO CASUALE STRATIFICATO: metodo di stratificazione di una popolazione su alcune caratteristiche prima della selezione casuale del campione (es. fasce di età). 3. CAMPIONAMENTO SISTEMATICO: può essere applicato all'elenco telefonico, da cui selezionare un campione costituito da 500 persone a caso su 50.000, sistematicamente ogni centesimo nome. 4. ASSEGNAZIONE CASUALE: i gruppi (sperimentale, di controllo, ecc.) all'interno del campione devono essere formati completamente a caso; l'assegnazione casuale permette al ricercatore di assumere che questi siano equivalenti all'inizio della ricerca.
111	Un buon principio di selezione del campione può permettere la generalizzazione dei risultati a una popolazione di cui il campione è sufficientemente rappresentativo (cfr. p. 109). Selezione casuale del campione → campione rappresentativo della popolazione → possibilità di induzione.
112	UNITÀ DI ANALISI o <i>dataset</i> : concetto relativo al campionamento e all'analisi statistica che si riferisce a ciò che è considerata l'unità fondamentale da cui i dati possono essere prodotti. Spesso è identificata con il singolo partecipante, ma può essere anche una classe intera nel momento in cui si confrontino le medie di più classi. Per evitare CONTAMINAZIONI, il <i>dataset</i> può addirittura essere una scuola (la scuola A usufruirà dell'intervento, la scuola B no, altrimenti nella stessa scuola i ragazzi della classe G influenzerebbero quelli della H). Il <i>dataset</i> influenza significativamente la dimensione del campione (confronto tutti i soggetti o le medie delle varie classi).
113 - 114 - 115	<b>1) TECNICHE DESCRITTIVE</b> → vi appartengono: <ul style="list-style-type: none"> <li>▪ <b>CALCOLI DELLA TENDENZA CENTRALE:</b> <ol style="list-style-type: none"> <li>1. <b>MEDIA (M):</b> punteggio medio di un gruppo di valutazioni → <math>M = \Sigma X/N</math>.</li> <li>2. <b>MEDIANA:</b> rappresenta le unità che si trovano nel mezzo di una distribuzione quando gli N valori sono stati ordinati in modo crescente/decrescente → <math>(N+1)/2</math>.</li> <li>3. <b>MODA:</b> rappresenta il valore che compare più frequentemente.</li> </ol> </li> <li>▪ <b>CALCOLI DELLA DISPERSIONE DEI DATI RACCOLTI:</b> <ol style="list-style-type: none"> <li>1. <b>DEVIATIONE STANDARD (SD):</b> stima della variabilità dei valori di un gruppo rispetto alla media → → <math>SD = \sqrt{[\Sigma(X - M)^2 / (N - 1)]}</math>. La M e la SD insieme (es. <math>20 \pm 2</math> anni) costituiscono una descrizione soddisfacente di un gruppo di misure, a patto che la SD non diventi troppo grande rispetto alla M. Per una DISTRIBUZIONE NORMALE (cfr. p. 116) → 68% risultati cade nell'intervallo <math>M \pm 1</math> SD, 95% in <math>M \pm 1,96</math> SD, 99% in <math>M \pm 2,576</math> SD.</li> <li>2. <b>VARIANZA:</b> il quadrato della deviazione standard.</li> <li>3. <b>INTERVALLO DELLE MISURE:</b> la più alta e la più bassa delle misure raccolte.</li> <li>4. <b>INTERVALLO DI CONFIDENZA (CI):</b> fornisce una gamma di valori anziché un singolo valore numerico, in cui è probabile che si trovi il valore medio.</li> <li>5. <b>ERRORE STANDARD:</b> variabilità della campionatura in una statistica.</li> <li>6. <b>DISTRIBUZIONE DI FREQUENZA:</b> distribuzione di punteggi inclusa la frequenza con cui si verificano.</li> <li>7. <b>INTERVALLI DI FREQUENZA:</b> piccole serie di punteggi all'interno di una distribuzione di frequenza in cui i punteggi sono raggruppati → non è possibile conoscere il valore esatto di ciascun individuo all'interno dell'intervallo.</li> <li>8. <b>DIAGRAMMA STEM-AND-LEAF:</b> metodo di organizzazione di punteggi grezzi dove gli intervalli di punteggio sono mostrati sul lato sinistro di una linea verticale e i punteggi individuali che rientrano in ciascun intervallo vengono visualizzati sul lato destro → non c'è perdita di informazione.</li> </ol> </li> </ul>

116	<p>Due categorie di test statistici:</p> <ol style="list-style-type: none"> <li><b>PARAMETRICI</b>: test basati sull'assunzione che i dati abbiano una <u>distribuzione normale</u>, una <u>varianza uguale</u> a quella della variabile di interesse e che le <u>osservazioni</u> siano <u>indipendenti</u>. Solitamente si applicano per campioni <math>N &gt; 30</math>.</li> <li><b>NON PARAMETRICI</b> o <i>distribution-free</i>: comprendono tutte le tecniche statistiche utilizzate qualora i dati non soddisfino le assunzioni richieste per le tecniche parametriche. <u>Distribuzione libera</u> → distribuzione diversa da quella <u>normale</u>.</li> </ol> <p><b>CURVA NORMALE</b>: distribuzione di dati in cui la media, la mediana e la moda si trovano tutte nello stesso punto (al centro della distribuzione) e in cui l'intervallo <math>\pm 1</math> SD dalla media include il 68% dei dati, <math>\pm 1,96</math> SD include il 95% e <math>\pm 2,576</math> SD includono il 99%.</p> <p><b>POTENZA STATISTICA</b>: la probabilità di scartare un'ipotesi nulla falsa.</p> <p><b>SKEWNESS</b>: descrizione del grado di simmetria della distribuzione, cioè della posizione di picco rispetto alle code. Picco spostato a sinistra → asimmetria <u>positiva</u>; picco a destra → asimmetria <u>negativa</u>.</p> <p><b>KURTOSIS</b>: descrizione della verticalità della distribuzione dei dati, cioè se la curva è più piccata o piatta rispetto a una curva normale.</p>									
117	Figure di curve che meglio spiegano quanto riportato a p. 116.									
118	La correlazione tra due variabili non indica un <b>RAPPORTO DI CAUSALITÀ</b> . Il rapporto tra causa ed effetto non può essere determinato da alcuna statistica o correlazione; cause ed effetti sono stabiliti da teorie, dalla logica e dalla sperimentazione, di cui la statistica è solo una parte. Correlazioni tra variabili possono solo suggerire un legame causale.									
119	La <b>STATISTICA</b> fa fundamentalmente 2 cose: stabilisce la significatività e determina l'importanza. La <u>significatività</u> stabilisce se una relazione o una differenza è affidabile, cioè se ci si può aspettare che si verifichi di nuovo nel caso lo studio fosse ripetuto. L' <u>importanza</u> si riferisce al valore dei risultati.									
Pag.	<b>Cap. 7 – Questioni statistiche nella pianificazione e valutazione di una ricerca</b>									
122 - 123	<p>In questo capitolo sono spiegati i 4 concetti per valutare e pianificare una ricerca: probabilità (<math>\alpha + \beta</math>), <i>effect size</i>, potenza e contesto.</p> <p><b>1) PROBABILITÀ</b>: concetto statistico che cerca di quantificare la possibilità che un certo evento si avveri.</p> <p><b>EVENTI EQUIPROBABILI</b>: concetto di probabilità per il quale vi è la stessa possibilità che eventi diversi si verifichino. <b>FREQUENZA RELATIVA</b>: concetto di probabilità che permette di confrontare che due o più eventi si realizzino.</p> <p><b>ALFA</b>: livello di probabilità fissato dal ricercatore prima dello studio, a volte indicato come livello di significatività. Non può mai essere eliminato. Quantifica l'<b>ERRORE DI TIPO I</b> → rifiuto dell'ipotesi nulla quando essa è vera nella realtà.</p> <p><b>BETA</b>: quantifica l'<b>ERRORE DI TIPO II</b> → mancato rifiuto dell'ipotesi nulla quando in realtà essa è falsa.</p> <p>Fig. 7.1: TABELLA DI VERITÀ → grafico delle decisioni</p> <table border="1"> <thead> <tr> <th></th> <th><math>H_0</math> VERA</th> <th><math>H_0</math> FALSA</th> </tr> </thead> <tbody> <tr> <th>ACCETTAZIONE</th> <td>Decisione corretta</td> <td>Errore di tipo II (<math>\beta</math>)</td> </tr> <tr> <th>RIFIUTO</th> <td>Errore di tipo I (<math>\alpha</math>)</td> <td>Decisione corretta</td> </tr> </tbody> </table> <p>corrette ed errate riguardanti l'errore di tipo I e di tipo II.</p> <p><math>\alpha</math> e <math>\beta</math> sono inversamente proporzionali → meglio commettere errore tipo I, in modo da dare più possibilità, per esempio ad un metodo di allenamento, di dimostrare la propria efficacia rispetto ad un altro. <math>\beta</math> è tipicamente <math>4\alpha</math> (cfr. p. 126).</p> <p><math>\alpha</math> è stabilito a priori come criterio dal ricercatore e i risultati lo soddisfano o meno.</p>		$H_0$ VERA	$H_0$ FALSA	ACCETTAZIONE	Decisione corretta	Errore di tipo II ( $\beta$ )	RIFIUTO	Errore di tipo I ( $\alpha$ )	Decisione corretta
	$H_0$ VERA	$H_0$ FALSA								
ACCETTAZIONE	Decisione corretta	Errore di tipo II ( $\beta$ )								
RIFIUTO	Errore di tipo I ( $\alpha$ )	Decisione corretta								
124	<p><b>2) EFFECT SIZE</b> (importanza e dimensione dell'effetto o <i>Delta</i>): è il valore standardizzato, cioè la differenza tra le medie divisa per la deviazione standard. Serve per valutare l'importanza di una differenza tra due medie, altro elemento di cui preoccuparsi oltre alla significatività (= significato pratico di un effetto o di una relazione) (cfr. p. 152).</p> <p><b>ES</b> = <math>(M_1 - M_2) / SD</math>. 0,2 → piccolo ES; 0,5 → moderato; 0,8 → grande.</p> <p><b>3) POTENZA</b> (DI UN TEST): probabilità di rigettare l'ipotesi nulla quando l'ipotesi nulla è falsa. Varia da 0 a 1 e maggiore è la potenza, maggiore sarà la possibilità di individuare relazioni o differenze reali. Se elevata, significa che ad un secondo tentativo il risultato resterebbe invariato. La potenza è solitamente calcolata come <math>(1 - \beta)</math>, quindi se <math>\beta = 4\alpha</math> e si usa <math>\alpha = 0,05</math> avremo <math>(1 - 0,2 = \mathbf{0,8})</math> → valore di potenza tipico nelle Scienze Motorie.</p>									
125	Attraverso l'esame di studi collegati alla ricerca che sta per iniziare, quando $\alpha$ , ES e potenza sono conosciuti, il <b>NUMERO DI PARTECIPANTI</b> per ogni gruppo necessario per l'esperimento può essere ricavato dalla Fig. 7.3. Se tutte le variabili restano uguali e solamente l' $\alpha$ viene modificato, è necessario un numero maggiore di partecipanti al fine di individuare differenze sostanziali tra i risultati dei diversi gruppi. Se i costi sono uguali, è meglio dare importanza alla potenza, ovvero all'individuazione delle differenze. La dimensione del campione → influenza la SD → influenza la potenza del test in maniera sostanziale (cfr. p. 126). Oltre alle procedure <u>a priori</u> sopra descritte è possibile eseguire dei <u>post hoc</u> .									

	STUDI PILOTA → obiettivo è trovare differenze reali → usare $\alpha$ più alto (0,20-0,30).
126 - 127	<b>4) CONTESTO:</b> interrelazioni trovate nel mondo reale. Maggiore è l'ES, minore sarà la sovrapposizione tra le distribuzioni di gruppo sperimentale e gruppo di controllo. $p < 0,05$ e potenza $> 0,8$ → valori arbitrari.
128	RICERCATORI QUANTITATIVI → non forniscono il contesto appropriato per i loro studi, lasciano parlare i dati senza evidenziarne l'importanza o spiegare cosa dicono di nuovo, a chi possono servire o quali ricerche future possano essere previste. RICERCATORI QUALITATIVI → il contesto orienta ogni passo della loro ricerca. Esempio di contesto: viene scoperto un metodo di allenamento che migliora tempo sui 100m di 1 decimo → sarà un'utile conoscenza per un centometrista, non per tutti.
Pag.	<b>Cap. 8 – Relazioni tra variabili</b>
	<b>PRESUPPOSTI PER TUTTE LE TECNICHE (PARAMETRICHE) DISCUSSE NEI CAP. 8-9-10: I DATI DEVONO ESSERE RIPORTATI AD UNA LINEA RETTA E DEVONO ESSERE DISTRIBUITI IN MANIERA NORMALE.</b>
131	<b>2) RELAZIONI:</b> determinate attraverso diverse tecniche: correlazioni <u>univariate</u> (correlazione $r$ di Pearson, regressione lineare, correlazione parziale e semi-parziale, regressione multipla) e correlazioni <u>multivariate</u> (correlazione canonica, analisi fattoriale e modelli di equazioni strutturali come LISREL) (cfr. p. 148). <b>CORRELAZIONE:</b> tecnica statistica per determinare le relazioni tra due o più variabili.
132	COEFFICIENTE DI CORRELAZIONE: valore quantitativo della relazione tra due o più variabili, che va da 0,00 a 1,00 in entrambe le direzioni (positiva e negativa). Correlazione perfetta +1,00 o -1,00 → ogni valore è situato alla stessa distanza relativa dalla media di ciascun gruppo di valori (retta). In natura non esiste mai, quindi si può parlare di evidente relazione, ma non perfetta correlazione (cfr. p. 133).
133	CORRELAZIONE POSITIVA: relazione tra due variabili nella quale un valore piccolo di una variabile è associato con un valore piccolo di un'altra variabile e un valore elevato di una variabile è associato con un elevato valore di un'altra variabile.
134	CORRELAZIONE NEGATIVA: relazione tra due variabili nella quale un valore piccolo per la prima variabile è associato ad un grande valore della seconda variabile e un valore elevato per la prima variabile è associato a un piccolo valore della seconda variabile. Le relazioni possono essere lineari o curvilinee.
135	La correlazione tra due variabili non implica un rapporto causale. La correlazione è condizione necessaria, ma non sufficiente per la causalità → mentre due variabili devono essere correlate perché possa esistere una RELAZIONE CAUSA-EFFETTO, la correlazione da sola non garantisce tale rapporto. <b>1) CORRELAZIONE <math>r</math> DI PEARSON</b> (o correlazione interclasse o correlazione semplice): in questo tipo di correlazione c'è una <u>variabile predittiva</u> (indipendente) e una <u>variabile criterio</u> (dipendente). La relazione tra variabili deve essere lineare. Il calcolo del coefficiente di correlazione coinvolge le distanze relative tra i risultati e i valori medi della distribuzione.
137 - 138	È possibile interpretare $r$ mediante la sua SIGNIFICATIVITÀ (= la fiducia o la probabilità che una statistica si verifichi nuovamente se lo studio venisse ripetuto → $p$ ). La significatività di una correlazione aumenta proporzionalmente al numero dei partecipanti all'esperimento; coefficienti di correlazione molto bassi possono essere significativi in presenza di un alto numero di partecipanti. Il valore del coefficiente di correlazione dovrà essere maggiore con $p$ scelto a 0,01 rispetto a $p = 0,05$ .
138	<b>COEFFICIENTE DI DETERMINAZIONE (<math>r^2</math>):</b> coefficiente di correlazione al quadrato, usato per interpretare la <u>significatività</u> di una correlazione. È la porzione di varianza totale presente in una misura, che può essere spiegata o presa in considerazione dalla varianza di un'altra misura (intersezione di due insiemi). È solitamente espresso come una % di variazione, quindi con $r = 0,70$ → $r^2 = 0,49$ → 49%. Serve una correlazione di 0,71 per costituire almeno la metà della varianza nell'altro test. VARIANZA NON SPIEGATA = $(1 - r^2)$ . Anche detta varianza di errore.
139	Maggiore è la correlazione, più la distribuzione diventa asimmetrica. <b>TRASFORMAZIONE Z DI FISHER:</b> metodo di approssimazione di una distribuzione normale del campione con relazione lineare, dove si trasformano i coefficienti di correlazione in valori Z. I valori Z permettono di combinare campioni (es. fasce di età) di correlazioni per ottenere una stima valida ed affidabile della relazione tra due misure (es. tra $VO_2$ e metri percorsi in 12').

140	<p><b>2) REGRESSIONE</b> o equazione di previsione: formula per prevedere alcuni criteri (per esempio alcune misurazioni di performance) basata sulla relazione tra le variabili da prevedere e i criteri. Nei test di massa, l'analisi delle variabili predittive richiede meno tempo, denaro e fatica rispetto all'analisi delle variabili criterio: viene quindi sviluppata una regressione. Maggiore è la correlazione tra le variabili, maggiore è la precisione con la quale una può essere prevista dall'altra.</p>
142	<p><b>RETTA DI REGRESSIONE OTTIMALE</b> (o metodo dei minimi quadrati): retta di regressione calcolata che risulta nella più piccola somma di quadrati delle distanze verticali di ogni punto dalla curva.</p>
142 - 143	<p><b>3) CORRELAZIONE PARZIALE</b> o spuria: relazione in cui la correlazione tra due o più variabili è dovuta primariamente alla comune influenza di un'altra variabile. In un esperimento è possibile controllare il fattore dell'età in due modi: selezionando solamente bambini della stessa età o <u>parzializzando</u> ( → rimuovere l'influenza di qualche cosa/mantenere qualche cosa costante). <math>r_{12.3}</math> è la correlazione parziale tra le variabili 1 e 2 con 3 mantenuta costante, ovvero dopo che l'influenza della variabile 3 sulle variabili 1 e 2 è stata eliminata.</p> <p>La correlazione parziale permette di sviluppare un'equazione multi-regressiva, con due o più variabili come fattore predittivo.</p>
143 - 144 - 145	<p><b>4) CORRELAZIONE SEMI-PARZIALE:</b> tecnica in cui soltanto una variabile viene parzializzata (= ne vengono rimossi gli effetti) da due variabili in una correlazione. Il simbolo è <math>r_{1(2.3)}</math> → indica che la relazione tra le variabili 1 e 2 è determinata dopo che l'influenza della variabile 3 sulla sola variabile 2 è stata rimossa.</p> <p><b>5) REGRESSIONE MULTIPLA:</b> modello utilizzato per prevedere un criterio da due o più variabili indipendenti, o predittive. L'uso di più di una variabile predittiva solitamente accresce l'accuratezza della previsione.</p> <p><b>R</b> (maiuscola) rappresenta il COEFFICIENTE DI CORRELAZIONE MULTIPLA, che indica la relazione tra un criterio (la variabile dipendente) e una somma pesata delle variabili predittive (variabili indipendenti).</p> <p>REGRESSIONE GERARCHICA: specifica l'ordine e la combinazione delle variabili predittive.</p> <p>Le equazioni di regressione sviluppate con un particolare campione spesso perdono molto della loro accuratezza se applicate ad un campione diverso (concetto di SPECIFICITÀ DI POPOLAZIONE). Questo calo di accuratezza viene chiamato <b>SHRINKAGE</b> (impoverimento, perdita → tendenza della variabilità a diminuire quando la formula è utilizzata con un nuovo campione). Maggiore è l'accuratezza utilizzata nelle procedure di selezione (<i>forward, backward, passo per passo, R<sup>2</sup> massimo</i>) che capitalizzano su caratteristiche specifiche del campione, più è difficile che i risultati possano essere generalizzati.</p> <p>I risultati non avranno nessun significato se quando si legge uno studio che utilizza la regressione multipla il rapporto partecipanti/variabili non è maggiore di 10:1.</p>
146	<p><b>6) CORRELAZIONE CANONICA:</b> estensione della regressione multipla (più fattori predittivi ed un criterio) ad un'analisi avente più fattori predittivi e più criteri. Si formano due strutture miste lineari, una per le variabili predittive ed una per le variabili criterio. La correlazione canonica rappresenta la varianza condivisa dalle due strutture miste lineari di variabili.</p> <p><b>7) ANALISI FATTORIALE:</b> tecnica statistica usata per ridurre un insieme di dati raggruppando variabili simili in componenti base (fattori). L'analisi fattoriale inizia con il calcolo delle inter-correlazioni di tutte le misure utilizzate (correlazione tra tutte le possibili coppie di variabili). Lo scopo è di scoprire i fattori (concetti complessi di base) che meglio spiegano un gruppo di misure e descrivono la relazione di ogni misura rispetto ai concetti complessi base. Se ne distinguono due tipi:</p> <ol style="list-style-type: none"> <li>1. ESPLORATIVA: più variabili sono ridotte ad un insieme base.</li> <li>2. CONFERMATIVA: si supporta o meno una struttura proposta dalla teoria; è spesso utilizzata per <u>validare</u> risultati di strumenti attitudinali, questionari e studi estesi a diversi Paesi.</li> </ol>
147	<p><b>8) MODELLI DI EQUAZIONI STRUTTURALI:</b> tecniche di modellistica strutturale utilizzate per spiegare il modo in cui certe caratteristiche sono legate una all'altra e per cercare di dedurne le cause. Tra queste tecniche ci sono le <b>RELAZIONI STRUTTURALI LINEARI (LISREL)</b>, approccio statistico usato per stabilire le relazioni ed esaminare il modello di equazione strutturale. Permette il raggruppamento di misure connesse in componenti e mostra le relazioni tra le componenti in termini di dimensione e direzione.</p> <p>Stabilendo le direzioni di determinate relazioni, è possibile effettuare deduzioni su quali caratteristiche possano causare altre caratteristiche, ricordando che causa ed effetto non costituiscono risultati statistici, ma logici.</p>



Pag.	Cap. 9 – Differenze tra gruppi
	<p>Il test statistico è effettuato sempre sull'ipotesi nulla (<math>H_0</math>). La soluzione del test può accettare o rifiutare l'<math>H_0</math>, ma non può confermare l'ipotesi di ricerca, confermabile solo dal ragionamento logico, un dettagliato progetto sperimentale e una profonda conoscenza della teoria alla base del fenomeno indagato. Con la statistica si possono dunque rilevare solo le <b>DIFFERENZE (3)</b> tra gruppi di campioni, ma non i motivi di tali differenze.</p>
150	<p>L'uso delle funzioni statistiche <math>t</math> e di <i>distribuzione F</i> (v. dopo) richiede l'assunzione di 5 corollari:</p> <ol style="list-style-type: none"> <li>1. osservazioni prelevate da una popolazione con distribuzione normale;</li> <li>2. le osservazioni rappresentano un campione casuale della popolazione;</li> <li>3. il numeratore e il denominatore di <math>t</math> e <math>F</math> sono stimati dalla stessa varianza di popolazione;</li> <li>4. il numeratore e il denominatore di <math>t</math> e <math>F</math> sono indipendenti;</li> <li>5. le varianze (quadrato della SD) dei due gruppi sono equivalenti (cfr. p. 153).</li> </ol>
151 - 152 - 153 - 154	<p><b>1) TEST <math>t</math> PER UN CAMPIONE E UNA POPOLAZIONE DI RIFERIMENTO:</b> per stabilire, ad es., se un campione di studenti differisca da una popolazione più estesa. Il test <math>t</math> è un test sull'ipotesi nulla che suppone che non ci siano differenze tra la media campionaria (<math>M</math>) e quella della popolazione (<math>\mu</math>), cioè <math>M - \mu = 0</math>.</p> $t = \frac{M - \mu}{S_M / \sqrt{n}} \quad df = n - 1$ <p><b>2) TEST <math>t</math> PER CAMPIONI INDIPENDENTI:</b> determina le differenze reciproche tra due campioni significativamente differenti e dalle medie conosciute.</p> $t = \frac{M_1 - M_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \quad df = n_1 + n_2 - 2$ <p><math>t(21) = 4,66 \rightarrow</math> il numero in parentesi (21) è il <math>df</math>, cioè <math>N - 2</math>, con <math>N =</math> numero totale di partecipanti.</p> <p><b>3) TEST <math>t</math> PER CAMPIONI DIPENDENTI:</b> test per valutare la significatività delle differenze tra le medie di due campioni che sono in relazione tra loro, così come quando alcuni soggetti sono valutati in due occasioni differenti.</p> $t = \frac{M_{post} - M_{pre}}{\sqrt{\frac{(S_{post}^2 + S_{pre}^2) - (2r_{pp} \cdot s_{post} \cdot s_{pre})}{(N - 1)}}} \quad df = N - 1$ <p><math>N \rightarrow</math> numero di coppie delle osservazioni.  <math>r_{pp} \rightarrow</math> 2 volte la correlazione fra la condizione pre- e quella post-.</p>
155	<p>È possibile STIMARE L'EFFETTO DELL'ALLENAMENTO dividendo la "differenza delle medie" (<math>M_D</math>) con la "media del pre-test" (<math>M_{pre}</math>): <b>% di incremento = <math>M_D / M_{pre}</math></b>.</p>
156	<p><b>TEST <math>t</math> A UNA CODA:</b> test in cui si assume a priori che le differenze tra le due medie si dirigano verso una sola direzione. È molto difficile riscontrare situazioni in cui si è sicuri dei propri risultati al punto da impiegare un test <math>t</math> a una coda.</p> <p><b>TEST <math>t</math> A DUE CODE:</b> test in cui si assume a priori che le differenze tra le due medie potrebbero essere a favore di entrambi i gruppi.</p>
156 - 157	<p><b>POTENZA DELL'EFFETTO (o Effect Size, ES):</b> modo per quantificare l'EFFICACIA DEL TRATTAMENTO (cfr. p. 124). Qualora ci sia un gruppo di controllo, allora dovrebbe essere utilizzata la SD di tale gruppo, altrimenti si dovrebbe operare con la <math>SD_p</math> (= RAGGRUPPATA o mediata) (cfr. p. 152).  Come AUMENTARE LA POTENZA DI UN TEST <math>t</math>: (tenere presente la formula del punto 2, p. 152)</p> <ol style="list-style-type: none"> <li>1. incrementare la differenza tra le medie <math>\rightarrow</math> incrementando l'efficacia dei trattamenti, per esempio prolungando a 12 settimane un trattamento progettato per 6.</li> <li>2. Ridurre la SD e la varianza <math>\rightarrow</math> personalizzando il trattamento i soggetti reagiranno in maniera omogenea.</li> <li>3. Incrementare il numero di partecipanti.</li> <li>4. Variare <math>\alpha \rightarrow</math> da 0,05 a 0,10 la potenza aumenta, ma si rischia di imbattersi in errori del tipo I.</li> <li>5. Utilizzare il test <math>t</math> per campioni dipendenti invece che quello per campioni indipendenti.</li> </ol> <p>L'efficacia del trattamento si deve stimare solo dopo che si è rifiutata l'ipotesi nulla: <math>t</math> deve essere superiore a 1 affinché la VARIANZA REALE superi la VARIANZA RESIDUA per una quantità significativa. <math>R^2</math> rappresenta la frazione di varianza totale che è dipendente dai trattamenti (varianza reale).  <b>R QUADRO (<math>R^2</math>)</b> O COEFFICIENTE DI DETERMINAZIONE LINEARE: % della varianza della variabile dipendente spiegata dalla variabile indipendente <math>\rightarrow</math> serve per interpretare la significatività del grado di relazione tra variabile dipendente e indipendente (cfr. p. 150). <b><math>R^2 =</math> varianza reale / varianza totale.</b>  Anche L'ES è uno stimatore della forza della significatività delle differenze tra gruppi o fra trattamenti.</p>
158 - 160	<p>La suddivisione delle tecniche statistiche in 2 + 1 categorie è solo un artefatto didattico, perché sia correlazioni che differenze si basano su un MODELLO LINEARE GENERALE, ovvero una proporzione tra le differenze e la totalità dei casi.</p>

160	<p>ANALISI DELLA VARIANZA (ANOVA) → metodo per determinare le differenze tra due o più gruppi.</p> <p><b>4) ANOVA SEMPLICE</b> (o a una via): test che valuta l'ipotesi nulla tra le medie di due o più gruppi. È un'estensione del test <math>t</math> per gruppi indipendenti, tant'è che <math>t</math> è proprio un caso particolare di ANOVA, dove ci sono solo due gruppi.</p>
161	<p>L'algoritmo alla base del calcolo dell'ANOVA è chiamato METODO ABC.</p> <p>SOMMA DI QUADRATI SS (o variabilità): misura della variabilità dei dati; la somma dei quadrati degli scarti rispetto alla media dei dati.</p> <p>COEFFICIENTE DI FISHER <math>F</math>: ricavato dal rapporto tra varianza reale e varianza residua (<math>F = MS_{rea} / MS_{res}</math>).</p>
162 - 163	<p><b>POST HOC</b>: per conoscere le differenze reciproche tra 3 gruppi (ovvero, gruppo 1 e 2 differiscono dal gruppo 3, ma non tra loro?) si deve eseguire un test a posteriori di comparazioni multiple, un post hoc. Se ne distinguono diversi tipi, tutti concettualmente riferibili al test <math>t</math>:</p> <ol style="list-style-type: none"> <li>1. Metodo di Scheffé → il più conservativo (rileva poche differenze statistiche); il più raccomandato. Produce una costante definita VALORE CRITICO (CV) per il confronto di tutte le medie; il CV si confronta con le differenze tra le medie dei vari gruppi (es. <math>M_1 - M_2 = 10 - 4 = 6</math> → se è superiore al CV, la differenza tra i gruppi 1 e 2 è significativa).</li> <li>2. Metodo di Tuckey → un po' meno del precedente.</li> <li>3. Metodo di Newman-Kalls → una via di mezzo.</li> <li>4. Metodo di Duncan → il più permissivo (rileva molte significatività).</li> </ol>
164	<p>CONFRONTI PIANIFICATI → confronti fra i gruppi eseguiti prima dell'esperimento, anziché a posteriori tramite l'ANOVA.</p> <p><math>\omega^2</math> = la formula dell'omega quadro di Tolson è una stima più accurata del grado di variabilità dei risultati ottenuti dai trattamenti. Un <math>\omega^2</math> pari a 0,54 indica che il 54% della varianza residua è spiegata dai trattamenti.</p> <p><math>t^2 = F</math> (cfr. p.161)</p>
164 - 165 - 166 - 167	<p><b>5) ANOVA FATTORIALE</b>: analisi della varianza in cui sono presenti più di una variabile indipendente, le cui interazioni influenzano statisticamente i caratteri di una singola variabile dipendente.</p> <p>Esempio: due variabili indipendenti (es. intensità e condizione fisica), di due caratteri (es. alta e bassa intensità, buona e scarsa condizione fisica) ciascuno → <math>2 \times 2</math> ANOVA, perché presenta 2 fattori, una sola interazione (il prodotto "x"), e 2 effetti principali. EFFETTO PRINCIPALE SEMPLICE: test eseguito su una variabile indipendente quando tutte le altre indipendenti sono mantenute ad un valore costante</p> <p>Nella maggior parte dei casi accade che solo i caratteri di una variabile indipendente riversino il loro contributo con significatività statistica nei caratteri delle altre variabili indipendenti.</p> <p><b>PRINCIPIO KISS: Keep It Simple, Stupid</b> → rendere sempre lo studio il più semplice possibile.</p>
168 - 169 - 170	<p><b>6) ANOVA PER MISURE RIPETUTE</b>: analisi dei dati registrati dagli stessi soggetti, ma in successive occasioni, come una batteria di prove ripetute; anche chiamata split-plot ANOVA o ANOVA gruppi x prove. Il test <math>t</math> per campioni dipendenti è il più semplice caso di ANOVA per misure ripetute.</p> <p>Le procedure per misure ripetute permettono di incrementare la potenza del test, individuando le differenze tra le misure e distinguendole dai termini di errore, in caso di monitoraggio temporale, anche se composti da pochi rilevamenti.</p> <p>VANTAGGI: controlla le differenze individuali che si riscontrano fra i partecipanti; più economica in quanto richiede una ridotta numerosità del campione; le procedure statistiche per misure ripetute sono in grado di sostenere indagini di fenomeni attraverso il tempo.</p> <p>PROBLEMI: effetti per accumulo (trattamenti somministrati prima influenzano i successivi), di apprendimento, di affaticamento, di adattamento (aumento resistenza organica agli effetti dei trattamenti).</p> <p><b>SFERICITÀ</b>: condizione nella quale le misure ripetute non sono correlate ed hanno la stessa varianza. Più una variabile risente dell'assunto di sfericità, meglio è stimata dal descrittore statistico <math>\epsilon</math> (valori da 1 a 0, dove 1 = sfericità perfetta). Nelle procedure statistiche per misure ripetute è auspicabile un <math>\epsilon &gt; 0,75</math>. La violazione dell'assunto di sfericità rientra negli errori di tipo I. 2 importanti considerazioni:</p> <ol style="list-style-type: none"> <li>1. Quando l'assunzione di sfericità è completamente rispettata, l'ANOVA per misure ripetute è più potente delle tecniche multivariate.</li> <li>2. Se il numero di partecipanti è piccolo, allora si può usare solo l'ANOVA per misure ripetute.</li> </ol> <p>COEFFICIENTE CORRETTIVO DI GEISSER/GREENHOUSE: procedimento statistico conservativo che corregge la stima dell'<math>\epsilon</math> nelle ANOVA per misure ripetute e calcola i gradi di libertà vicariati per trovare il rapporto <math>F</math> per determinare la significatività → <math>\theta = 1 / (k-1)</math>.</p>
171	<p><b>7) ANCOVA</b>, Analisi della Covarianza: combinazione di regressione e ANOVA per controllare statisticamente la <u>variabile di disturbo (covariata)</u> che affligge la variabile dipendente. Procedura che si effettua a due livelli: 1) correzione tramite correlazione fra variabile dipendente e covariata; 2) calcolo dell'ANOVA semplice. Frequentemente utilizzata quando si effettua un pre-test, si somministra un trattamento e si esegue un post-test <math>t</math>, solo quando i gruppi non variano durante le prove.</p>

171 -	<p><b>TASSO D'ERRORE SPERIMENTALE</b> o inflazione cumulativa delle probabilità → contributo della correzione del livello di significatività.</p> <p><b>TECNICA DI BONFERRONI:</b> divide il livello <math>\alpha</math> per il numero di confronti da effettuare, in modo da correggere il livello di significatività <math>\alpha</math>, necessario quando si calcoli l'ANOVA per ogni singola variabile dipendente di un gruppo <math>&gt; 2</math>.</p> <p><b>TECNICHE MULTIVARIATE:</b> considerano più di una variabile dipendente. Sono tecniche multivariate l'analisi dei discriminanti, la MANOVA e la MANCOVA.</p>																								
172 -	<p><b>8) ANALISI DEI DISCRIMINANTI:</b> quando si hanno una variabile indipendente (a due o più livelli) e due o più variabili dipendenti. Tali analisi sono un'unione tra regressione multipla e ANOVA semplice, per combinare le variabili dipendenti e decidere di accettare o escludere alcuni livelli delle variabili indipendenti come, per esempio, i componenti di un gruppo. L'analisi dei discriminanti usa gli stessi metodi delle regressioni multiple per identificare i predittori più importanti → tecniche di selezione:</p> <ol style="list-style-type: none"> <li><b>FORWARD:</b> dispone le variabili dipendenti in ordine di importanza, in modo che la prima sia quella che contribuisce maggiormente alla separazione dei gruppi (la discriminante più incisiva o il miglior predittore di appartenenza a gruppi).</li> <li><b>BACKWARD:</b> dispone le variabili dipendenti in ordine discendente rispetto ai contributi che ognuna di loro è in grado di riversare sulla separazione dei gruppi.</li> <li><b>STEPWISE:</b> simile alla precedente, ma ad ogni passaggio tutte le variabili dipendenti sono rivalutate in termini di contributi per la separazione.</li> <li><b>TECNICA F STEPDOWN:</b> iterata per ogni variabile dipendente, considerando le successive variabili come covariate.</li> </ol> <p>L'analisi dei discriminanti non è in grado di supportare gruppi randomizzati, che devono quindi avere una composizione costante.</p>																								
174	<p><b>9) MANOVA</b>, analisi della varianza multivariata: si realizza una variabile dipendente per ogni combinazione ottimale delle variabili dipendenti (di tipo lineare), di modo da massimizzare la varianza spiegata da associare alle variabili indipendenti.</p> <p><i>Quali variabili dipendenti contribuiscono significativamente alla separazione tra i gruppi?</i> Questa è la domanda alla quale si è cercato di rispondere con la descrizione della tecnica statistica dell'analisi dei discriminanti e della tecnica <math>F</math> stepdown per i test a posteriori, ma entrambe le tecniche non sono in grado di valutare le differenze tra interazioni, mentre la MANOVA è in grado di assegnare dei gradi di importanza ai livelli delle variabili indipendenti.</p>																								
175	<p><b>10) MANCOVA</b>, analisi multivariata della covarianza: ci sono una o più variabili indipendenti, due o più variabili dipendenti e una o più covariate. È una tecnica statistica speciale, che molti ricercatori usano in modo impreciso poiché non è conosciuta in modo approfondito.</p> <p>MISURE RIPETUTE CON PIÙ VARIABILI DIPENDENTI:</p> <p><b>11) MMM ANALISYS (Multivariate Mixed Model):</b> modello di analisi multivariata incrociata → tratta la variabile indipendente come un tipico caso multivariato, formando una serie di combinazioni lineari dalle <math>n</math> variabili dipendenti.</p> <p><b>12) DM ANALISYS (Double Multivariate):</b> non richiede la conferma del criterio di sfericità. L'analisi si rivolge alla variabile indipendente. Sono realizzate due combinazioni lineari: una per le variabili dipendenti e una per le <math>n</math> prove.</p>																								
176	<p><b>Ricordarsi</b> che c'è una corrispondenza tra tecniche per relazioni tra variabili e per differenze fra gruppi, come ad esempio tra <math>t</math> ed <math>r</math> e tra l'ANOVA e la regressione multipla.</p> <p>Ogni tecnica valuta due questioni di base:</p> <ol style="list-style-type: none"> <li>L'effetto della relazione è significativo?</li> <li>Qual è l'intensità dell'effetto della relazione?</li> </ol> <p>Di seguito una tabella riassuntiva che confronta le tecniche descritte nei cap. 8 e 9 (<math>VI =</math> Variabile Indipendente; <math>VD =</math> Variabile Dipendente).</p> <table border="1"> <thead> <tr> <th>Descrizione</th> <th>Differenze fra gruppi</th> <th>Differenze fra variabili</th> </tr> </thead> <tbody> <tr> <td>1 VI (2 livelli) → 1 VD</td> <td>test <math>t</math> per campioni indipend.</td> <td></td> </tr> <tr> <td>1 causa → 1 effetto</td> <td></td> <td><math>r</math> di Pearson</td> </tr> <tr> <td>1 o 2 VI → 1 VD</td> <td>ANOVA fattoriale</td> <td></td> </tr> <tr> <td>2 o più cause → 1 effetto</td> <td></td> <td>regressioni multiple</td> </tr> <tr> <td>1 VI (2 o più livelli) → 2 o più VD</td> <td>analisi discriminataria</td> <td></td> </tr> <tr> <td>2 o più VI → 2 o più VD</td> <td>MANOVA</td> <td></td> </tr> <tr> <td>2 o più cause → 2 o più effetti</td> <td></td> <td>correlazione canonica classica</td> </tr> </tbody> </table>	Descrizione	Differenze fra gruppi	Differenze fra variabili	1 VI (2 livelli) → 1 VD	test $t$ per campioni indipend.		1 causa → 1 effetto		$r$ di Pearson	1 o 2 VI → 1 VD	ANOVA fattoriale		2 o più cause → 1 effetto		regressioni multiple	1 VI (2 o più livelli) → 2 o più VD	analisi discriminataria		2 o più VI → 2 o più VD	MANOVA		2 o più cause → 2 o più effetti		correlazione canonica classica
Descrizione	Differenze fra gruppi	Differenze fra variabili																							
1 VI (2 livelli) → 1 VD	test $t$ per campioni indipend.																								
1 causa → 1 effetto		$r$ di Pearson																							
1 o 2 VI → 1 VD	ANOVA fattoriale																								
2 o più cause → 1 effetto		regressioni multiple																							
1 VI (2 o più livelli) → 2 o più VD	analisi discriminataria																								
2 o più VI → 2 o più VD	MANOVA																								
2 o più cause → 2 o più effetti		correlazione canonica classica																							

Pag.	Cap. 10 – Tecniche non parametriche
179	Le statistiche non parametriche sono ideate per il <u>trattamento di classificazioni, ranghi e categorie</u> (es. valutazione del comportamento affettivo e questionari con risposte aperte); hanno minor potenza rispetto alle statistiche parametriche.
180	<b>RANGO:</b> posizione dei dati in un insieme ordinato. Es. vengono costruiti 180 ranghi (il totale di ciclisti appartenenti a 3 gruppi sperimentali), ordinati dal più piccolo (rango pari a 1) al più grande (rango pari a 180) (cfr. p. 188).
181 - 182 - 183 - 184	<p>I dati come età, genere, livello di istruzione e tipologia di trattamento sono MISURE NOMINALI (o categoriali), ovvero possono essere differenziati in categorie.</p> <p><b>CHI QUADRO</b> → test statistico sulla significatività delle discrepanze fra i dati sperimentali e quelli di riferimento:</p> $\chi^2 = \sum[(O - E)^2/E],$ <p>dove <math>O</math> è la frequenza campionaria osservata, mentre <math>E</math> è la frequenza campionaria attesa (<i>Expected</i>).  <math>df = (c - 1)</math>, dove <math>c</math> è il numero di campi (celle).</p> <p>In alcuni casi le frequenze attese (<math>E</math>) per le misure nominali possono essere ottenute da sorgenti di informazioni già presenti: es. <math>O</math> → studenti di un determinato professore e i loro voti, <math>E</math> → distribuzione di riferimento del dipartimento.</p> <p>Restrizioni d'uso: le osservazioni per ogni campione devono essere tra loro indipendenti e le categorie mutuamente esclusive → le singole osservazioni appartenenti a una categoria non devono essere dipendenti da nessuna delle osservazioni appartenenti alle altre categorie. In ogni categoria, il totale delle frequenze attese deve essere lo stesso delle frequenze osservate. Il numero totale delle osservazioni trattate non deve essere inferiore a 20.</p> <p><b>TABELLA DI CONTINGENZA:</b> CLASSIFICAZIONE A DUE VIE ( → problema che lega due o più categorie di eventi/soggetti a due o più gruppi di possibili rilevazioni) di eventi/soggetti associati a gruppi di osservazioni, idonea a calcolare la significatività delle differenze tra i valori osservati e quelli previsti.</p> <p>CORREZIONE PER LA CONTINUITÀ DI YATES: metodo di correzione utilizzato nelle tabelle di contingenza di tipo 2 x 2, sottraendo 0,5 dalla differenza fra le frequenze osservate e quelle attese, prima dell'elevamento al quadrato.</p> <p>COEFFICIENTE DI CONTINGENZA (C): metodo per calcolare la relazione fra variabili nominali dicotomiche come genere e razza. Se <math>\chi^2</math> è statisticamente significativo, allora lo sarà anche C.</p>
184 - 185	<p><b>TECNICHE DI ORDINAMENTO DEI DATI PER RANGO:</b></p> <ol style="list-style-type: none"> <li>1. Test U di Mann-Whitney: analogo del test <math>t</math> parametrico per campioni indipendenti.</li> <li>2. Test di Wilcoxon per dati appaiati o test del segno per ranghi: analogo del test <math>t</math> parametrico per campioni dipendenti.</li> <li>3. ANOVA di Kruskal-Wallis per ranghi: analogo alla parametrica ANOVA a una via.</li> <li>4. ANOVA di Friedmann a due vie per ranghi: analogo alla parametrica ANOVA per misure ripetute.</li> <li>5. Correlazione dei ranghi di Spearman: analogo del coefficiente di correlazione di Pearson in ambito parametrico.</li> </ol> <p>Procedura standardizzata di ordinamento dei dati per ranghi: sostituire i valori dei dati con il rango; eseguire le procedure statistiche con i ranghi dei dati; valutare il livello di significatività del test calcolando il valore critico <math>L</math> e confrontandolo con i valori di riferimento di <math>\chi^2</math> (tab. A7).</p> <p><b>TEST STATISTICO NON PARAMETRICO L:</b> può essere usato al posto di tutte le tecniche precedentemente illustrate, assumendo però che tutti i dati si dispongano lungo una retta. <math>L = (N-1) r^2</math></p>
190	<b>SOMMARIO:</b> <i>in questo capitolo sono stati presentati i test statistici che trattano dati con caratteristiche differenti dalle condizioni parametriche (cap. 8 e 9). Questi test comprendono il chi quadro per frequenze campionarie e una serie di procedure che trattano i ranghi dei valori dei dati originari, allo stesso modo delle omologhe procedure parametriche. Se la distribuzione dei dati non è normale, allora è possibile considerare i ranghi dei valori dei dati. I ranghi dei dati possono essere trattati attraverso applicativi informatici di statistica. Alla fine, per stabilire la significatività del test, viene restituito un coefficiente L da comparare con un valore critico tabellare del chi quadro. Queste procedure possono essere estese a tutti i modelli lineari parametrici.</i>

Pag.	Cap. 11 – Validare le variabili sperimentali
192	<p><b>VALIDITÀ DELLA MISURA:</b> bontà della corrispondenza tra il dato ricavato dal test e la variabile da misurare.</p> <p><b>1) VALIDITÀ LOGICA</b> (di facciata o di immagine): quanto una misura sia in relazione con la grandezza che si vuole indagare.</p> <p><b>2) VALIDITÀ DI CONTENUTO:</b> quanto un test (di solito in un contesto pedagogico) rilevi effettivamente i parametri che si era presupposto indagare. Logica e contenuto non possono essere valutate tramite un'evidenza statistica, ma serve un gruppo di 20 o più esperti che ha il compito di giudicare l'attinenza → serve un'approvazione pari all'80-85%.</p> <p><b>3) VALIDITÀ DI CRITERIO:</b> quanto i valori ottenuti da un test siano in relazione con un criterio di riferimento. <b>VALIDITÀ CONCORRENTE:</b> tipo di validità di criterio che identifica il grado di correlazione tra lo strumento di misura e un criterio (<i>gold standard</i>) che concorre contestualmente al fenomeno che si vuole indagare.</p>
193 - 194	<p><b>VALIDITÀ PREDITTIVA:</b> grado di accuratezza col quale una serie di variabili predittive possono predire un andamento dei dati. In questo caso il criterio è una condizione futura e si deve quindi conoscere un valore o un andamento di riferimento della data condizione. Poiché diversi predittori possono competere per garantire una maggiore attinenza tra i dati e il criterio di riferimento (&gt; validità), le regressioni multiple sono più usate delle correlazioni.</p> <p><b>CRITERIO DI RESTRIZIONE (SHRINKAGE):</b> la validità predittiva tende a diminuire quando si viene ad applicare a popolazioni di soggetti tanto differenti da quella di origine; inversamente proporzionale al numero del campione di origine e dal numero di predittori → tanti predittori quanti i campioni utilizzati = <u>perfetta predizione e massima perequazione</u>. <i>Perequazione:</i> processo per rendere uguale.</p> <p><b>CROSS-VALIDAZIONE:</b> tecnica per determinare l'accuratezza di una perequazione quando questa è applicata a un campione diverso dal campione di realizzazione.</p> <p><b>TABELLA DELLE FREQUENZE RELATIVE:</b> tabella a doppia entrata che mostra la probabilità che alcuni valori sperimentali si comportino allo stesso modo di un determinato criterio.</p>
195	<p><b>4) VALIDITÀ DI COSTRUTTO:</b> grado di corrispondenza tra una misura sperimentale e il costrutto ipotetico che si vuole valutare, in relazione ad un'associazione tra i risultati del test e i relativi caratteri del costrutto. Può essere rilevata tramite correlazioni o analisi fattoriale. <u>Costrutti → non direttamente misurabili</u> (misurabili attraverso indicatori, NdC).</p> <p>In <b>CONCLUSIONE</b>, tutte le validità descritte servono per rafforzare principalmente la validità di costrutto relativa: infatti, avere una buona stima di tutte le validità presentate permette di convalidare la progettazione di uno strumento di misura (<u>anche i TEST sono strumenti di misura</u>) e di conseguenza l'uso dei suoi risultati.</p>
196 - 197 - 198	<p><b>AFFIDABILITÀ</b> (= ATTENDIBILITÀ): principio alla base della validità, principio che indica l'immutabilità, la ripetibilità di una misura. <b>VALORE OSSERVATO</b> → nella <i>Classical Test Theory</i> (CTT) è il valore misurato inteso come somma tra il valore effettivo e gli errori di misura. <b>VALORE REALE</b> → parte del valore osservato che rappresenta il valore effettivo della misura, esente da ogni forma di errore. <b>ERRORE DI MISURA</b> → parte del valore osservato attribuita alle <b>SORGENTI DI ERRORE:</b> soggetti da valutare, protocollo sperimentale, trattamento dei dati, strumenti di misura.</p> <p><b>STIMA DELL'AFFIDABILITÀ TRAMITE CORRELAZIONE:</b></p> <ol style="list-style-type: none"> <li>1. Correlazione interclasse → metodo di correlazione più utilizzato per confrontare due variabili, chiamato coefficiente di correlazione di Pearson <i>r</i>.</li> <li>2. Correlazione intraclasse (ICC, <i>Intra Class Correlation</i>) → variante dell'ANOVA per stimare l'affidabilità di una misura.</li> </ol> <p>Tre principali motivi non permettono sempre l'uso della correlazione interclasse per il calcolo del coefficiente di affidabilità: Pearson <i>r</i> è un descrittore statistico bivariato e non univariato; il calcolo è limitato a sole due variabili; non è in grado di rilevare le sorgenti di variabilità dei test che prevedono prove multiple. È dunque preferibile usare l'ICC, che provvede alla stima degli errori sistematici e della varianza non spiegata, nei casi in cui si vogliono confrontare le differenze sistematiche tra le prove di uno stesso protocollo → tramite una ANOVA è possibile valutare la qualità del test prova per prova e definirne i limiti di affidabilità (con l'ANOVA si individuano le differenze significative nelle prove di un test e si decide se scartarle o considerare quelle prove affette da errori di misura).</p> <p>Le tecniche di correlazione usate per stimare i coefficienti di affidabilità sono <u>diverse</u> da quelle usate per definire la <u>validità</u>.</p> <p><b>AFFIDABILITÀ INTEROPERATORE (OGGETTIVITÀ):</b> grado di differenza che diversi operatori possono ottenere effettuando la stessa misura. <b>CONCORDANZA INTEROSSERVATORE (IOA, <i>Inter Observer Agreement</i>):</b> metodo di stima dell'affidabilità tra diversi osservatori utilizzando una formula che divide il numero di concordanze per il numero totale dei pareri (concordanze + discordanze) (cfr. p. 201).</p>

199 - 200	<p><b>METODI PER STABILIRE L’AFFIDABILITÀ.</b> L’affidabilità è espressa da tre tipi di coefficienti:</p> <ol style="list-style-type: none"> <li>1. STABILITÀ → coefficiente di affidabilità stimato con il METODO DEL TEST-RETEST in più giorni: il test viene somministrato allo stesso modo in due giorni differenti.</li> <li>2. FORME RECIPROCHE (metodo delle forme parallele o dell’equivalenza) → metodo per stabilire l’affidabilità di un test, formato da due prove (test) differenti, ma in grado di valutare la stessa caratteristica o il suo reciproco, somministrate entrambe allo stesso campione.</li> <li>3. CONSISTENZA (immutabilità) INTERNA → ulteriori coefficienti di affidabilità possono essere ottenuti da alcuni metodi che vengono classificati sotto il nome di tecniche di consistenza interna: <ul style="list-style-type: none"> <li>▪ Metodo del TEST-RETEST INFRAGIORNALIERO: ripetizione di due prove effettuate nello stesso giorno. Si per test fisici, no per questionari (sensibilizzazione alle domande).</li> <li>▪ Tecnica dello SPLIT-HALF: stima l’affidabilità mediante la suddivisione in due del test, generalmente considerando le righe pari e le righe dispari di un questionario separatamente e correlandole tra loro successivamente. Quando si utilizza il metodo split-half, è possibile usare: <ul style="list-style-type: none"> <li>▪ la FORMULA PREDITTIVA DI SPEARMAN-BROWN: per correggere il coefficiente di affidabilità: affidabilità parziale → affidabilità intero questionario;</li> <li>▪ il METODO DI FLANAGAN: divide il test in due parti e ne analizza la varianza rispetto alla varianza totale;</li> <li>▪ il METODO DI KUDER-RICHARDSON: stima l’affidabilità tramite unica somministrazione;</li> <li>▪ la TECNICA DEL COEFFICIENTE <math>\alpha</math>: per test a prove multiple; la forma più generale e versatile di coefficiente di affidabilità.</li> </ul> </li> </ul> <p>Si per i questionari, raramente per test fisici con molte prove ripetute.</p> </li></ol>
201	<p><b>ERRORE STANDARD DELLA MISURA:</b> è buona norma pensare che ogni misura sia errata e per tanto deve sempre essere considerata come una stima del valore vero e presentata sotto forma di intervallo all’interno del quale c’è anche, con una certa probabilità, il valore cercato.</p> <p><math>S_{Y.X} = s \sqrt{1,00 - r}</math> → è dipendente dalla variabilità della misura e dalla sua affidabilità.</p>
202	<p>Un modo per confrontare le diverse prestazioni è quello di convertire ogni scala di valori in una <b>SCALA STANDARDIZZATA</b> → scala di valori espressa in termini di deviazioni standard rispetto alla media.</p> <p>SCALA DEI PUNTI Z O Z-SCORE: scala standardizzata che converte i dati grezzi in frazioni delle deviazioni standard, con la media dei dati che corrisponde allo 0 e con la SD pari a 1. Se due prestazioni giacciono sulla stessa posizione lungo la scala dei punti Z (es. 1 SD sopra la media), è possibile asserire che nelle due prove si sono ottenuti gli stessi livelli di prestazione fisica.</p> <p>SCALA DEI PUNTI T: scala standardizzata derivata dalla scala dei punti Z, che pone la media dei valori originali a 50 e la SD a 10, rimuovendo i decimali e rendendo positivi tutti i valori. <math>T = 50 + 10Z</math> → vengono così rimossi i decimali e i valori diventano tutti positivi. Più del 99% dei valori ricade in <math>M \pm 3</math> SD → raro trovare valori di T al di sotto di 20 (<math>Z = -3</math>) e al di sopra di 80 (<math>Z = +3</math>).</p>
203	<p><b>VALUTARE I SONDAGGI.</b></p> <p>COMPORAMENTO AFFETTIVO composto da diversi aspetti: carattere, personalità, stato di ansia, autovalutazione, inserimento sociale, sportività.</p>
204	<p>EFFETTO RETROATTIVO DEL TEST: il primo test polarizza l’attenzione dei soggetti sulla psicomelia indagata, creando una maggiore focalizzazione sull’autovalutazione e alterandone la percezione nei test successivi → l’autovalutazione porta di solito a falsare il grado di veridicità delle risposte.</p>
205	<p>SCALA LIKERT: tipo di questionario usato per la catalogazione dei tratti caratteriali composto da domande a risposte chiuse che comprendono più risposte ( &gt; 5 è meglio) disposte in ordine equidistante per differenze di definizioni.</p> <p>SCALA DEL DIFFERENZIALE SEMANTICO: scala utilizzata per misurare la propensione comportamentale in base alla distanza che intercorre tra due aggettivi di significato opposto e suddivisa in sette intervalli equidistanti → vicino al valore 7 si trova l’aggettivo polarizzato positivamente e verso l’1 l’aggettivo polarizzato negativamente.</p> <p>SCALA DI VALUTAZIONE: misura soggettiva delle caratteristiche del comportamento che si basa su una lista di classificazioni (es. commissione di esperti che valuta la tecnica di un tuffo).</p> <p>SCALA DELLO SFORZO PERCEPITO (RPE): scala di autovalutazione realizzata da Borg nel 1962, che serve per misurare la percezione soggettiva dello sforzo (da 6 a 20) durante l’esercizio fisico.</p>

206 - 207	<p><b>ERRORI DI VALUTAZIONE:</b></p> <ol style="list-style-type: none"> <li>1. Imprecisione → tendenza degli osservatori a essere troppo tolleranti sui valori registrati delle misure.</li> <li>2. Errori di tendenza alla centralità → orientamento dei valori raccolti della stessa misura verso la loro media, allontanandosi dai valori più estremi. Es. giudice che si considera esperto → dà voti bassi per dimostrare (anche inconsciamente) la sua abitudine alla perfezione; per non essere considerato troppo severo → alza i voti; sperando in miglioramenti futuri (anche nel corso di una gara) → eccessiva indulgenza.</li> <li>3. Effetto volano → tendenza a riversare un giudizio determinato per una caratteristica di un soggetto su tutte le altre: accumulo di impressioni negative su un soggetto tendono ad abbassare il giudizio delle sue prestazioni future o viceversa.</li> <li>4. Errori di prossimità → tendenza di un osservatore a considerare simili i comportamenti che sono catalogati vicini su una scala di classificazione, rispetto a quando sono disposti lontani (es. estroverso e intraprendente).</li> <li>5. Errori di giudizio dell'operatore → tendenza degli operatori a farsi influenzare dalle proprie opinioni.</li> <li>6. Errori di aspettative dell'operatore → tendenza di un operatore a interpretare un comportamento in base alla direzione delle aspettative supposte.</li> </ol> <p><b>SPERIMENTAZIONE A DOPPIO-CIECO:</b> indicata per controllare questi tipi di errore. Nella sperimentazione a doppio-cieco, gli sperimentatori non conoscono quali soggetti ricevono il trattamento e quali no, tantomeno i valori che ognuno dei soggetti raggiunge nelle fasi di pre-test e in quelle di post-test.</p>
207	<p>Per determinare se il livello di difficoltà del questionario sia appropriato per lo scopo previsto è possibile avviare l'<b>ANALISI DEGLI ITEM</b> (<i>item</i> = una domanda del questionario) → nei test della valutazione delle capacità cognitive, è il processo di stima della pertinenza delle domande e la loro rispettiva abilità discriminatoria. 2 aspetti importanti:</p> <ol style="list-style-type: none"> <li>1. valutare per ogni domanda il grado di difficoltà → valutazione effettuata attraverso la divisione tra il numero delle risposte corrette attribuite alla domanda in un sondaggio, fratto il totale numero delle risposte assegnate alla domanda. Più la domanda è difficile, più l'indice di difficoltà è basso. Eliminare le domande con indici inferiori a 0,1 o superiori a 0,9 → situazioni in cui nessuno o tutti rispondono alla domanda in modo corretto non aggiungono nessuna nuova informazione circa le differenze tra le persone all'interno di scale di misurazione standardizzate.</li> <li>2. valutare il grado di discriminazione delle domande → ovvero il contributo che una data domanda rivela nella capacità del test di distinguere i soggetti che hanno risposto correttamente da quelli che hanno risposto in modo errato. È anche detto indice discriminatorio.</li> </ol>
208 - 209	<p><b>TEORIA CLASSICA DEI TEST</b> (CTT, <i>Classical Test Theory</i>): teoria che evidenzia come ogni misura sperimentale è la somma tra il valore reale della misura e una componente d'errore. La maggior parte delle informazioni riguardanti la validità, affidabilità e analisi dei questionari fa parte della CTT.</p> <p><b>TEORIA DELLE RISPOSTE</b> (IRT, <i>Item Response Theory</i>): teoria che si focalizza sulle caratteristiche intrinseche della domanda e le sue potenziali risposte nel contesto dell'intero questionario, al fine di valutare esclusivamente le abilità del candidato/campione esaminato. È anche chiamata teoria dell'abilità nascosta. <b>ITEM BANKING:</b> raccolta di domande che viene utilizzata per la realizzazione di questionari specifici per la valutazione delle abilità nascoste. <b>ADAPTIVE TESTING</b> (o personalizzazione): processo di gestione degli item che permette di selezionare le domande migliori per indagare nel modo più specifico possibile il livello di abilità di un soggetto.</p> <p>L'IRT può essere utilizzata per scegliere il numero ottimale di domande che serve per fornire maggiori e precisi dettagli sull'abilità soggettiva da valutare → per ogni individuo, differenti gradi di abilità corrispondono a un numero differente di domande.</p>