

PREFAZIONE

Il presente volume, le cui origini risalgono a una semplice raccolta di dispense del corso di Statistica Sociale, poi stampata nel lontano 1977 dall'Editore Cacucci di Bari, trae origine dal desiderio di fornire i principali elementi di metodologia statistica a chi intenda intraprendere ricerche in campo sociale ed abbia una preparazione di Matematica anche solo a livello di scuola media superiore. Esso, perciò, può essere di ausilio non solo agli studenti universitari dei Corsi di laurea in Sociologia, Psicologia, Scienze Politiche e di altri Corsi ove sia contemplato l'insegnamento della Statistica e non quello della Matematica, ma anche agli operatori sociali che non abbiano una specifica preparazione matematica.

Per venire incontro alle esigenze dei ricercatori, il volume, nelle successive edizioni, è stato via via arricchito di argomenti che spesso in letteratura non sono trattati con applicazioni al campo sociale, ma solo a livello strettamente metodologico e talora di difficile comprensione per chi non abbia attitudine alla Matematica.

Ovviamente, pur se nel testo non sono esposti i passaggi matematici per pervenire alla formulazione delle teorie statistiche, si è cercato di evidenziare la logica posta alla base di detta formulazione, i requisiti a cui i dati devono soddisfare per poter essere utilizzati ed, infine, l'interpretazione dei risultati.

Poiché lo studioso dei fatti sociali molto spesso è chiamato ad operare con campionamenti di vario genere (ad es., stratificati, a due stadi, a grappoli, ecc.) si è ritenuto opportuno porre in luce come effettuare sia tali tipi di campionamento, sia le relative stime di alcune grandezze caratteristiche di maggior uso nel campo sociale, valutandone altresì il grado di accuratezza.

Naturalmente, grande importanza è stata data all'inferenza statistica, segnatamente nel caso di piccoli campioni, e ai test non parametrici, a causa del frequente utilizzo di tali tecniche, specie nel campo delle scienze del comportamento.

Il testo dedica molta attenzione ai metodi di analisi delle relazioni e delle interrelazioni fra variabili, sia quantitative che non quantitative. Dato che gran parte dei fenomeni sociali si esprime in risposte non quantitative (ad esempio, possono aversi semplici risposte di tipo binario, positivo/negativo, oppure risposte relative a categorie non ordinabili come "occupato", "disoccupato", "pensionato", "casalinga"), sono qui descritti i modelli log-lineari per l'analisi dell'interdipendenza fra tali variabili, così come i metodi di regressione logit per l'analisi della dipendenza di una variabile non quantitativa da altre variabili di qualsiasi natura.

Sono inoltre descritti e applicati alcuni metodi di quantificazione delle mutabili, dal più semplice al più complesso, utile anche nel caso del confronto di più giudizi medi espressi su scala ordinale. Su suggerimento di alcuni studiosi, sono stati sviluppati altri argomenti quali, ad esempio, l'analisi della mutabilità nel ca-

so di caratteri qualitativi, il test di asimmetria di una distribuzione anche nel caso campionario, il *test robusto su ranghi* di Fligner-Policello per sopperire ai casi in cui non può essere applicato il test di Mann-Whitney, il test di Welch per il confronto di più di due medie quando i campioni sono estratti da popolazioni eteroschedastiche, la tecnica dei confronti multipli, il test di Levene per la verifica robusta dell'omoschedasticità, l'Analisi della Varianza per campioni dipendenti (o misure ripetute), eccetera.

Forse la materia contenuta nel volume è troppo vasta per essere trattata in un corso annuale di lezioni: la struttura del testo, tuttavia, consente al docente di svolgere solo gli argomenti che ritiene più utili per lo studente del proprio corso. Gli altri argomenti possono, però, essere approfonditi o in eventuali corsi successivi, o in discussioni seminariali o durante la stesura di tesi di laurea.

Oltre tutto, a nostro parere, il volume può essere utilizzato soprattutto in ricerche inerenti al lavoro che lo studioso di fatti sociali deve svolgere. È noto, infatti, che non tutto ciò che serve all'esercizio della professione in campo sociale può essere svolto nei corsi universitari e che, una volta immesso in un'attività lavorativa, il neolaureato si trova a dover approfondire metodologie e tecniche per l'analisi dei fenomeni oggetto di studio: è proprio in quell'occasione che può far comodo avere sottomano un manuale ove siano riportate le principali tecniche, in modo che l'ulteriore approfondimento sia un agevole prosieguo di ciò che ha imparato sui banchi universitari. A tale scopo, oltre a numerosi esempi applicativi, sono riportati nel testo schemi e tabelle che aiutano sia lo studente che lo studioso a decidere quale sia la procedura più adatta a soddisfare le proprie necessità.

Per concludere, ringrazio tutti i colleghi che, in questa circostanza come anche in altre precedenti, mi hanno gentilmente fornito consigli e suggerimenti per rendere il volume ciò che è ormai diventato, e che mi auguro potrà essere di aiuto in molteplici situazioni al lettore più attento.

Un ringraziamento di cuore va ai miei affezionati allievi, i professori Ernesto Toma e Francesco d'Ovidio, che non solo hanno insistito perché dessi nuova vita al libro, ma hanno anche collaborato con le loro osservazioni a renderlo sempre attuale. Ringrazio in particolare il prof. d'Ovidio, che con abnegazione ed entusiasmo si è anche sobbarcato alla fatica di curare l'aggiornamento temporale degli esempi e alcune elaborazioni statistico-informatiche, nonché la prima stesura di qualcuno dei nuovi argomenti e l'editing finale del volume.

Ritengo che questo sia un bel modo per festeggiare un 85.mo compleanno.

Bari, 16 febbraio 2015

Francesco Delvecchio
Professore emerito
dell'Università degli Studi di Bari Aldo Moro

INDICE

ALCUNE NOZIONI PRELIMINARI	15
 CAPITOLO I - LA PROGRAMMAZIONE DI UNA RICERCA SOCIALE	
1.- Premessa	19
2.- Definizione del tema e dei sottotemi della ricerca	19
3.- Formulazione delle varie ipotesi da verificare	20
4.- Gli strumenti di valutazione dei fenomeni sociali: scale di misura e indicatori sociali	21
4.1- <i>Le scale nominali</i>	22
4.2- <i>Le scale ordinali</i>	23
4.3- <i>Le scale di intervalli</i>	24
4.4- <i>Le scale assolute o di rapporti</i>	25
4.5- <i>Conclusioni sull'uso delle scale</i>	26
5.- La definizione dell'unità di rilevazione	27
6.- Gli strumenti per conoscere i fenomeni sociali: le rilevazioni statistiche	28
7.- Le modalità di rilevazione delle unità statistiche	28
7.1- <i>Rilevazioni dirette</i>	29
7.1.1- Rilevazioni tramite intervistatore	29
a) L'intervista faccia a faccia	29
b) L'intervista telefonica. La tecnica CATI	29
7.1.2- Rilevazioni autocompilate	31
a) Rilevazioni con questionario postale	31
b) Rilevazioni telematiche, via e-mail oppure online: la tecnica CAWI	32
c) Rilevazioni con libretto (anche telematico)	33
7.2- <i>Rilevazioni indirette</i>	33
7.2.1- Esame dei documenti	34
7.2.2- Metodo del sopralluogo	34
8.- Il questionario	35
8.1- <i>La stesura del questionario</i>	35
8.2- <i>La verifica del questionario</i>	41
8.3- <i>Le norme per l'intervistatore</i>	42
8.4- <i>Le norme per informare l'intervistato</i>	43
9.- Tecniche di rilevazione delle unità statistiche	43
9.1- <i>Campionamenti probabilistici</i>	44
9.1.1- Campionamento bernoulliano	45
9.1.2- Campionamento senza ripetizione	45
9.1.3- Campionamento in blocco	45
9.1.4- Selezione con probabilità diseguali	45
9.1.5- Campionamento stratificato	45
a) Campionamento stratificato proporzionale alle unità di ogni strato dell'universo	46
b) Campionamento stratificato uniforme	46
c) Campionamento stratificato "quasi proporzionale"	46
d) Campionamento stratificato ottimale o di Neyman-Tschuprow	47
9.1.6- Campionamento a grappoli	48
9.1.7- Campionamento areale	48
9.1.8- Campionamento a più stadi	49
9.1.9- Campionamenti ripetuti	50
a) Rilevazioni per "panel"	50
b) Campionamenti "ruotati"	50
9.2- <i>Tecniche di selezione di un campione casuale</i>	55
9.2.1- Selezione con associazione delle unità a numeri aleatori	55
9.2.2- Selezione sistematica delle unità del campione	56

9.3- Campionamenti non probabilistici	56
9.3.1- Campionamenti per quota	56
9.3.2- Campionamenti a scelta ragionata. Le proiezioni elettorali della DOXA	57
9.3.3- Campionamenti bilanciati	58
9.3.4- Campionamenti con adesione volontaria dei rispondenti	58
9.3.5- Campionamenti “a valanga” (<i>Snowball sampling</i>)	58
10.- L'errore	60
10.1- Errori nella costruzione del concetto e dello strumento di valutazione	60
10.2- Errori dovuti alle procedure di campionamento	60
10.3- Errori nella raccolta dei dati	60
10.3.1- Errori attribuibili al comportamento dell'intervistato	60
10.3.2- Errori attribuibili al comportamento dell'intervistatore	61
10.4- Errori dovuti alle mancate risposte (<i>totali o parziali</i>)	61
10.5- Errori nello spoglio dei dati	61
10.6- Errori dovuti ad analisi non corretta dei dati	61
10.7- Errori di stima	61
10.8- Errori di misura	62
11.- Previsione dei costi e dei tempi	62

CAPITOLO II - LA RILEVAZIONE DEI DATI

1.- La raccolta dei dati	63
2.- La classificazione e la tabulazione dei dati	63
3.- Il raggruppamento dei dati in classi	67
4.- Frequenze relative - Frequenze cumulate - Frequenze relative cumulate - Funzione di ripartizione	70
5.- Modalità cograduate e modalità associate	71
6.- Le tabelle a doppia entrata	72
7.- Distribuzioni condizionate - Distribuzioni marginali	75
8.- La variabile statistica tripla	77
ESERCIZI	78

CAPITOLO III - LE RAPPRESENTAZIONI GRAFICHE

1.- La rappresentazione grafica di m.s. non ordinabili e di m.s. rettilinee	81
1.1- Grafici di composizione: <i>diagrammi a settori circolari ed a rettangoli</i>	81
1.2- Grafici di comparazione: <i>l'ortogramma e il cartogramma</i>	82
2.- La rappresentazione grafica di mutabili cicliche: il diagramma polare	83
3.- La rappresentazione grafica di v.s.	84
3.1- <i>V.s. discrete: il diagramma a punti ed il diagramma a segmenti</i>	85
3.2- <i>V.s. continue con modalità singole: il diagramma a congiungenti</i>	85
3.3- <i>V.s. continue con modalità raggruppate in classi: l'istogramma</i>	86
4.- Confronto di grafici	87
5.- Diagrammi integrali	88
6.- Poligono di frequenze. Curve di frequenze	89
7.- La rappresentazione grafica delle v.s. doppie	89
7.1- <i>La rappresentazione per modalità singole: stereogramma</i>	90
7.2- <i>La rappresentazione per classi di valori: istostereogramma</i>	90
7.3- <i>I grafici a bolle</i>	91
7.4- <i>La rappresentazione per livelli di frequenza: l'ipsogramma</i>	92

CAPITOLO IV - LE PRIME ELABORAZIONI DEI DATI: LA SINTESI DELLE INFORMAZIONI

1.- I valori medi: indicatori di tipicità	93
1.1- <i>Osservazioni misurate a livello di scala almeno di intervalli: le medie analitiche</i>	93

1.1.1- La media aritmetica	93
a) Media di singoli valori	94
b) Media di una v.s. discreta	94
c) Media di una v.s. divisa in intervalli	96
d) Medie di v.s. doppie	98
e) Media della distribuzione di un carattere qualitativo ordinale	99
f) Media di una v.s. trasformata linearmente	99
1.1.2- La media geometrica	99
1.1.3- La media quadratica	100
1.1.4- La media armonica	100
1.1.5- La scelta della media analitica	100
1.2- Osservazioni misurate a livello di scala almeno ordinale: le medie posizionali	102
1.2.1- Valore centrale di una distribuzione statistica	102
1.2.2- Mediana della distribuzione di un carattere misurato su scala ordinale	103
a) Valori singoli	103
b) Distribuzione di frequenze	103
1.2.3- Mediana della distribuzione di un carattere misurato su scala d'intervalli	104
a) Valori singoli	104
b) Variabile statistica discreta	104
c) Variabile statistica divisa in classi	104
1.2.4- Quartili - Quantili	105
1.3- Valore tipico di osservazioni misurate a livello di scala nominale: la moda	108
2.- I rapporti statistici	109
2.1- Tassi assoluti e relativi d'incremento e di decremento	110
2.2- Rapporti di composizione	111
2.3- Rapporti di coesistenza	111
2.4- Rapporti di derivazione	112
2.5- Rapporto di durata	112
2.6- Rapporti indici	113
2.7- Numeri indici	113
3.- Gli indici di variabilità	115
3.1- Gli indici di dispersione	116
3.2- Il grafico a scatola o boxplot	118
3.3- Gli indici di disuguaglianza	120
3.4- Gli indici relativi di variabilità	121
3.3.1- Indici relativi di variabilità riferiti alla media	121
3.3.2- Indici relativi di variabilità riferiti al massimo	122
3.5- La concentrazione	123
4.- Gli indici di mutabilità	125
4.1- Caratteri qualitativi ordinali - L'indice di diversità	125
4.2- Caratteri qualitativi non ordinali: indici di entropia e di eterogeneità	127
4.2.1- L'indice di entropia	127
4.2.2- L'indice di eterogeneità di Gini	128
5.- Relazione fra le medie e le varianze di v.s. trasformate linearmente	129
6.- La sintesi delle caratteristiche del fenomeno tramite funzione matematica: la rappresentazione analitica	130
6.1- Scelta del tipo di funzione	132
6.2- Determinazione dei parametri	132
6.2.1- Metodo dei punti scelti	132
6.2.2- Metodo dei minimi quadrati	133
6.2.3- Metodo delle somme	134
6.2.4- Metodo delle aree o di Cantelli	135
6.2.5- Metodo dei momenti	135
6.3- Grado di adattamento della funzione ai dati osservati	135
7.- L'asimmetria delle distribuzioni statistiche. L'indice delle terne	136
ESERCIZI	137

**CAPITOLO V - L'UGUAGLIANZA E LA SOMIGLIANZA DI VARIABILI E
DI MUTABILI STATISTICHE
LA SOMIGLIANZA DI UNITÀ STATISTICHE**

1.- L'uguaglianza di variabili statistiche	139
2.- La somiglianza di variabili statistiche	139
2.1- La somiglianza di v.s. riguardanti lo stesso carattere	140
2.2- La somiglianza di v.s. riguardanti caratteri diversi	142
2.3- La somiglianza di v.s. divise in intervalli	143
3.- L'uguaglianza di mutabili statistiche	143
4.- La somiglianza di mutabili statistiche riferite a caratteri uguali	144
5.- La somiglianza di unità statistiche su cui sono state osservate k variabili	145
6.- La somiglianza di unità statistiche su cui sono stati osservati caratteri misti	146
ESERCIZI	147

CAPITOLO VI - CENNI DI CALCOLO DELLE PROBABILITÀ

1.- Premessa	149
2.- Eventi	149
3.- Le varie concezioni di probabilità	152
3.1- Concezione classica di probabilità	152
3.2- Concezione frequentistica di probabilità. Legge empirica del caso	153
3.3- Concezione soggettivistica di probabilità	154
4.- Teorema della probabilità dell'evento totale	154
5.- Probabilità condizionata - Legge del prodotto	155
6.- Eventi indipendenti	156
7.- Alcuni schemi probabilistici	157
7.1- Lo schema di Bernoulli o problema delle prove ripetute	157
7.2- Lo schema ipergeometrico	157
7.3- Lo schema multinomiale	158
7.4- Lo schema ipergeometrico multivariato	158
8.- Variabili casuali discrete unidimensionali	159
9.- Cenni sulle v.c. unidimensionali continue	161
10.- Variabili casuali bidimensionali. Funzione di verosimiglianza	163
11.- Una notevole proprietà delle variabili casuali indipendenti	164
12.- Teorema di Bienaymé-Tchebycheff	164
13.- Disuguaglianza di Schwarz	165
14.- Covarianza	165
ESERCIZI	165

**CAPITOLO VII - ALCUNE DISTRIBUZIONI DI PROBABILITÀ
LE DISTRIBUZIONI CAMPIONARIE**

1.- La distribuzione normale o di Gauss	167
1.1- La definizione di v. c. normale	167
1.2- Le caratteristiche della curva normale	167
1.3- Famiglia di curve normali	168
1.4- La curva normale standardizzata	169
1.5- Un teorema importante	171
1.6- Applicazione	171
2.- La v.c. di Bernoulli	172
3.- La distribuzione binomiale. La distribuzione di Poisson	172
4.- La distribuzione ipergeometrica	174
5.- La distribuzione uniforme	175
5.1- Distribuzione discreta	175
5.2- Distribuzione continua in $[a,b]$	175

6.- La distribuzione campionaria della media quando è noto σ	176
6.1- <i>Campionamento bernoulliano</i>	176
6.2- <i>Campionamento senza reinserimento</i>	178
6.3- <i>Campionamento in blocco</i>	179
6.4- <i>Campionamento stratificato</i>	181
7.- La distribuzione campionaria di una frequenza relativa quando è noto p	181
7.1- <i>Campionamento bernoulliano</i>	182
7.2- <i>Campionamento senza ripetizione e in blocco</i>	182
8.- La distribuzione campionaria della mediana. La distribuzione campionaria di una generica statistica.....	183
ESERCIZI	184

CAPITOLO VIII - STIME PUNTUALI

LE DISTRIBUZIONI CAMPIONARIE DELLE STIME

1.- Concetti sull'inferenza statistica	185
2.- Stimatori e loro proprietà desiderabili.....	186
3.- Le stime puntuali di μ , σ^2 , p nel campionamento bernoulliano	188
4.- Distribuzione campionaria della media e della frequenza nel caso σ^2 non sia nota e il campionamento sia bernoulliano	189
4.1- <i>Distribuzione campionaria della media</i>	189
4.2- <i>Distribuzione campionaria della frequenza</i>	192
5.- Le stime puntuali di μ , σ^2 , p nel caso di campionamento in blocco e di campionamento senza ripetizione	192
6.- La distribuzione campionaria della media e della frequenza nel caso σ^2 non sia nota e il campionamento sia in blocco o senza ripetizione	193
6.1- <i>Distribuzione campionaria della media</i>	193
6.2- <i>Distribuzione campionaria della frequenza</i>	193
7.- Le stime puntuali di μ e p nel caso di campionamento stratificato	194
8.- Le distribuzioni campionarie della media e della frequenza nel caso di campionamento stratificato quando σ^2 non è nota	194
9.- Le stime puntuali di μ e p nel caso di campionamento a due stadi e le loro distribuzioni campionarie quando σ^2 non è nota.....	196
9.1- <i>Le stime puntuali di μ e di p</i>	197
9.2- <i>Le distribuzioni campionarie di \bar{X} e di \hat{p}</i>	197
10.- Le stime puntuali di μ e p nel caso di campionamento a grappoli e le loro distribuzioni campionarie quando σ^2 non è nota.....	200
10.1- <i>Le stime puntuali di μ e di p</i>	200
10.2- <i>Le distribuzioni campionarie di \bar{X} e di \hat{p}</i>	201
11.- La stima della frequenza di risposte casualizzate mediante domanda incorrelata	202
12.- La correzione delle stime in presenza di non-risposte.....	204
12.1- <i>Integrazione del campione</i>	204
12.2- <i>Stime delle opinioni dei non rispondenti in base a quelle dei rispondenti con solleciti</i>	205
12.3- <i>Metodi particolari di stime in presenza di non risposte parziali</i>	206
13.- Riporto dei risultati campionari all'universo	206
13.1- <i>La rilevazione delle Forze di Lavoro</i>	207

CAPITOLO IX - LA STIMA INTERVALLARE

1.- Premessa	209
2.- Intervallo di confidenza per la media quando il campione è estratto da popolazione normale.....	211
3.- Intervallo di confidenza per la media quando la popolazione da cui è estratto il campione non è normale	212
3.1- <i>Grandi campioni</i>	213
3.2- <i>Piccoli campioni: il teorema di Bienaymé-Tchebycheff</i>	213

4.- Intervallo di confidenza per la frequenza	214
4.1- Grandi campioni.....	214
4.2- Piccoli campioni.....	215
4.2.1- Campionamento bernoulliano.....	215
4.2.2- Campionamento senza ripetizione e in blocco.....	217
4.3- Campionamento con risposte casualizzate mediante domanda incorrelata.....	218
5.- Intervallo di confidenza per un qualsiasi parametro (o caratteristica) dell'universo.....	218
6.- Intervallo di confidenza per la varianza. La v.c. χ^2 (chi-quadrato).....	218
ESERCIZI	222

CAPITOLO X - LA DIMENSIONE DEL CAMPIONE

1.- Il calcolo di n nel caso di inferenza su medie.....	223
1.1- La distribuzione campionaria della media è normale.....	223
1.2- La distribuzione campionaria della media non è normale.....	224
2.- Esempi di calcolo di n nel caso di inferenza su medie.....	225
3.- Il calcolo di n nel caso di inferenza su frequenze relative.....	226
4.- Esempi di calcolo di n nel caso di inferenza su frequenze relative.....	227
5.- Il calcolo di n nel caso di indagini con risposte casualizzate mediante domanda incorrelata.....	228
6.- Il calcolo di n nel caso di stime di più statistiche.....	228
7.- L'ampiezza dei subcampioni.....	228
7.1- Dimensione dei sub-campioni nel campionamento stratificato ottimale.....	229
8.- Correzione della numerosità campionaria in base ai costi dell'indagine.....	230
ESERCIZI	232

CAPITOLO XI - VERIFICHE DI IPOTESI CON UN CAMPIONE

1.- Introduzione.....	233
2.- Errori che si possono commettere nella prova d'ipotesi - Individuazione del test che consente di verificare l'ipotesi.....	234
3.- I requisiti dei test statistici.....	238
4.- Le fasi della verifica di ipotesi.....	238
5.- Esempi di verifica di ipotesi.....	240
6.- Test bilaterali.....	244
7.- Verifica di ipotesi funzionali.....	245
7.1- Test per caratteri misurati a qualsiasi livello.....	246
7.2- Test per m.s. e v.s. discrete.....	247
7.2.1- Campionamento bernoulliano.....	247
7.2.2- Campionamento senza ripetizione o in blocco.....	248
7.3- Il test di Kolmogorov per v.s. continue.....	249
7.4- Ipotesi sulla normalità di una popolazione di parametri ignoti.....	251
7.4.1- Il test di normalità di Shapiro-Wilk per piccoli campioni.....	251
7.4.2- Un indice di normalità per grandi campioni tramite un test di accostamento.....	252
8.- Verifica d'ipotesi sulla simmetria di una distribuzione: il test delle terne (o di Randles-Fligner-Policello-Wolfe).....	253
9.- Ipotesi sulla casualità di un campione: il test delle sequenze (o di Wald-Wolfowitz).....	254
ESERCIZI	256

CAPITOLO XII - VERIFICHE DI IPOTESI CON DUE CAMPIONI

1.- Ipotesi di uguaglianza di due varianze.....	257
1.1- Piccoli campioni.....	257
1.1.1- Popolazioni normali: il test F di Snedecor.....	257
1.1.2- Popolazioni non normali: il test di Moses.....	258
1.2- Grandi campioni.....	259
2.- Ipotesi di uguaglianza di due medie.....	259
2.1- Popolazioni normali ed omoschedastiche: il test T di Student.....	260
2.2- Popolazioni normali ed eteroschedastiche: il test ν di Welch-Aspin.....	261
2.3- Popolazioni non normali e grandi campioni ($n_1, n_2 \geq 100$).....	263
2.4- Popolazioni non normali e piccoli campioni ($n_1, n_2 < 100$).....	263

3.- Ipotesi di uguaglianza di due frequenze.....	263
3.1- Grandi campioni	264
3.2- Piccoli campioni: il test esatto di Fisher	265
4.- Ipotesi di uguaglianza delle mediane di due popolazioni non normali	266
4.1- Leggi di distribuzione uguali ed al più traslate. Il test di Mann-Whitney	267
4.2- Leggi di distribuzione simmetriche rispetto alla mediana: il "test robusto su ranghi" di Fligner-Policello	268
4.3- Leggi di distribuzione di qualsiasi tipo: il test della mediana	270
5.- Ipotesi di uguaglianza di due leggi di distribuzione.....	270
5.1- Piccoli campioni.....	270
5.1.1- Livello di misurazione su scala almeno di intervalli: il test di Kolmogorov-Smirnov	271
5.1.2- Livello di misurazione su scala almeno ordinale: il test di Wald-Wolfowitz	273
5.1.3- Livello di misurazione su scala nominale	275
5.2- Grandi campioni	275
6.- Verifiche d'ipotesi con due campioni dipendenti	278
6.1- Carattere misurabile a livello di scala almeno di intervalli con distribuzione delle differenze di forma nota	278
6.2- Carattere quantitativo con distribuzione delle differenze di forma non nota: il test di Wilcoxon per campioni dipendenti	280
6.3- Carattere qualitativo: il test dei segni	282
7.- Conclusioni sull'uso dei test statistici.....	283
ESERCIZI	285

CAPITOLO XIII - VERIFICHE DI IPOTESI CON PIÙ DI DUE CAMPIONI

1.- Uguaglianza di varianze di c popolazioni normali: il test di Bartlett.....	287
2.- Uguaglianza di medie di c popolazioni normali ed omoschedastiche.....	288
2.1- Analisi della varianza ad un criterio di classificazione	288
2.1.1- Aspetti logici dell'Analisi della Varianza.....	289
2.1.2- Dimostrazione.....	291
2.2- Analisi della varianza per classificazione bivalente	292
2.2.1- Una sola osservazione per casella.....	293
2.2.2- Più osservazioni per casella	296
3.- Uguaglianza di medie di c popolazioni quando le medie campionarie sono normalmente distribuite: il test di Welch per c campioni.....	299
4.- Confronti e confronti multipli	300
5.- Ipotesi di uguaglianza di c leggi di distribuzione.....	302
5.1- Misurazione a livello di scala almeno ordinale: il test di Kruskal-Wallis.....	303
5.2- Misurazione a livello di scala nominale	305
5.2.1- Grandi campioni	305
5.2.2- Piccoli campioni: il test di Fisher generalizzato	306
6.-Uguaglianza di varianze di c popolazioni non normali: il test di Levene	308
7.-Verifica di ipotesi con c campioni dipendenti.....	310
7.1- Misurazione a livello di scala almeno intervallare: ANOVA entro i soggetti (ANOVA within)	310
7.2- Misurazione a livello di scala almeno ordinale: il test di Friedman	314
ESERCIZI	317

CAPITOLO XIV - LA MISURA DELLE RELAZIONI DI DIPENDENZA NEL CASO DI DUE VARIABILI

1.- Premessa	319
2.- Definizione di indipendenza statistica fra due caratteri.....	321
3.- La funzione di regressione	322
4.- Il calcolo dei parametri	324
5.- L'attendibilità della stima della retta di regressione	328

6.- La significatività dei parametri della retta di regressione	332
7.- Previsione	335
8.- Regressione non lineare	335
9.- L'attendibilità della stima della funzione di regressione	337
10.- La stima del modello in presenza di eteroschedasticità	340
11.- Ancora sulla trasformazione dei dati	340
ESERCIZI	341

CAPITOLO XV - LA MISURA DEL GRADO DI ASSOCIAZIONE FRA DUE CARATTERI

1.- Introduzione	343
2.- La misura del grado di associazione fra due caratteri misurabili a livello di scala almeno di intervalli: il coefficiente di correlazione lineare di Bravais-Pearson	343
3.- Inferenza su ρ in base ai dati campionari	345
4.- La misura del grado d'associazione fra due caratteri misurabili a livello di scala almeno ordinale: l'indice di Spearman	348
4.1- Osservazioni distinte	349
4.2- Osservazioni ripetute	350
5.- La misura del grado di associazione fra due caratteri di cui almeno uno qualitativo	351
5.1- Piccoli campioni	351
5.2- Grandi campioni	352
5.2.1- Il test Q^2 - L'indice di contingenza di Cramer	352
5.2.2- Il test G^2 o del rapporto di massima verosimiglianza	353
ESERCIZI	355

CAPITOLO XVI - LA MISURA DELLE RELAZIONI FRA PIÙ VARIABILI

1.- La relazione di dipendenza fra variabili misurate a livello di scala almeno di intervalli	357
1.1- Il modello lineare classico	357
1.1.1- La stima dei parametri e la loro interpretazione	358
1.1.2- La significatività del modello stimato	359
1.1.3- L'individuazione di dati anomali	363
1.2- Il modello lineare in caso di perturbazioni autocorrelate	364
1.3- Errore di specificazione	367
1.4- Un esempio di funzione linearizzabile: il modello moltiplicativo	369
1.5- La correlazione parziale	369
1.6- La scelta delle variabili	370
1.6.1- Procedura all'indietro (backward elimination)	370
1.6.2- Procedura in avanti (forward selection)	371
1.6.3- Procedura passo a passo (stepwise regression)	371
1.7- Cenni sulle variabili dummy	373
2.- Misura della relazione tra più caratteri ordinali: l'indice medio di cograduazione di Spearman e il coefficiente di concordanza di Kendall	374
2.1- L'indice medio di cograduazione $s\bar{\rho}$ di Spearman	374
2.2- Il coefficiente di concordanza W di Kendall	375
ESERCIZI	377

CAPITOLO XVII - CENNI SUI MODELLI LOG-LINEARI

1.- Premessa	379
2.- Modello log-lineare d'indipendenza per tabelle di contingenza bidimensionali	379
3.- Modello log-lineare saturato per tabelle di contingenza bidimensionali	380
3.1- Stima dei parametri del modello log-lineare saturato per tabelle di contingenza bidimensionali	381
4.- Verifica dell'esistenza di associazione tra due caratteri	382

4.1- Il test G^2	382
4.2- La significatività dei parametri stimati	383
5.- Interpretazione dei parametri d'interazione dei modelli log-lineari per tabelle di contingenza bidimensionali	384
6.- Altri modelli per tabelle di contingenza bidimensionali	387
6.1- Assenza dell'effetto riga	387
6.2- Assenza dell'effetto colonna	387
6.3- Tutte le coppie (x_i, y_j) sono equiprobabili	388
7.- I modelli d'indipendenza per una tabella di contingenza tridimensionale	389
7.1- Indipendenza mutua	390
7.2- Indipendenza congiunta	390
7.3- Indipendenza condizionata di due variabili da ogni livello dell'altra	390
7.4- Indipendenza di ogni coppia di variabili dalla terza	391
8.- Il modello saturato per una tabella di contingenza a tre dimensioni	392
8.1- Stime dei parametri dei modelli log-lineari a tre dimensioni	392
8.2- La scelta del modello	392
9.- Interpretazione dei parametri dei modelli log-lineari per tabelle tridimensionali	394
ESERCIZI	395

CAPITOLO XVIII - CENNI SUI MODELLI LOGIT

1.- Premessa	397
2.- Modello logit a risposta dicotomica con variabili esplicative categoriche	397
3.- Modello logit a risposta dicotomica con variabili esplicative quantitative	399
4.- Modelli logit a risposta politomica su scala nominale con variabili esplicative categoriche	402
5.- Modelli logit a risposta politomica su scala nominale con variabili esplicative quantitative	405
6.- Modelli logit a risposta politomica su scala ordinale con variabili esplicative quantitative	408
6.1- Modelli logit per categorie adiacenti	408
6.2- Modelli logit continuation-ratio	409
6.3- Modelli logit cumulativi	415

CAPITOLO XIX - LA QUANTIFICAZIONE DELLE MUTABILI ORDINALI

1.- Premessa	419
2.- Quantificazione determinata direttamente	420
3.- Quantificazione determinata indirettamente mediante ipotesi sulla forma della distribuzione	420
3.1- Confronto dei giudizi di più gruppi di individui su uno stesso fenomeno	421
3.2- Confronto dei giudizi di uno o più gruppi di individui sul complesso di più fenomeni ritenuti omogenei	422
4.- Quantificazione basata sull'ottimizzazione di una funzione obiettivo	423
4.1- Notazioni di base	424
4.2- Confronto tra i giudizi medi espressi dagli individui di due gruppi nei riguardi di uno stesso carattere	426
4.3- Confronto tra i giudizi medi espressi dagli individui di più gruppi nei riguardi di uno stesso carattere	428
4.4- Arrotondamento delle soluzioni	429
5.- Metodo basato sulla propensione dei soggetti a quantificare	433
6.- Riduzione di più caratteri ordinabili ma non omogenei in uno complessivo	435

APPENDICE

Tav.1 - Aree α sottese alla curva normale standardizzata oltre z , ove $\alpha = P(Z \geq z)$	439
Tav.2 - Intervalli di confidenza ottimali per $p=10.000$ in caso di campionamento bernoulliano, al livello $\alpha = 0,05$ e $\alpha = 0,01$	440
Tav.2 bis - Valori soglia x_c per coda sinistra della distribuzione binomiale	442
Tav.3 - Valori soglia $t_{g,\alpha}$ della distribuzione "T _g di Student" per vari valori di α	443

Tav.4 - Valori soglia $\chi_{g,\alpha}^2$ della distribuzione del χ_g^2	444
Tav.4 bis - Quantili ${}_1\chi_g^2$ e ${}_2\chi_g^2$ della v.c. χ_g^2 tali che $P({}_1\chi_g^2 < \chi_g^2 < {}_2\chi_g^2) = 1-\alpha$ e l'intervallo $]\text{ns}^2/{}_2\chi_g^2, \text{ns}^2/{}_1\chi_g^2[$ sia minimo, per $1-\alpha=0,99$ e $1-\alpha=0,95$	446
Tav.5.a - Valori soglia $f_{g_1, g_2, \alpha}$ della distribuzione "F $_{g_1, g_2}$ di Snedecor" per $\alpha = P(F_{g_1, g_2} \geq f_{g_1, g_2, \alpha}) = 0,05$	448
Tav.5.b - Valori soglia $f_{g_1, g_2, \alpha}$ della distribuzione "F $_{g_1, g_2}$ di Snedecor" per $\alpha = P(F_{g_1, g_2} \geq f_{g_1, g_2, \alpha}) = 0,025$	450
Tav.5.c - Valori soglia $f_{g_1, g_2, \alpha}$ della distribuzione "F $_{g_1, g_2}$ di Snedecor" per $\alpha = P(F_{g_1, g_2} \geq f_{g_1, g_2, \alpha}) = 0,01$	452
Tav.5.d - Valori soglia $f_{g_1, g_2, \alpha}$ della distribuzione "F $_{g_1, g_2}$ di Snedecor" per $\alpha = P(F_{g_1, g_2} \geq f_{g_1, g_2, \alpha}) = 0,005$	454
Tav.6 - Valori soglia v_α del test di Welch-Aspin per $\alpha=0,05$ e $\alpha=0,01$	456
Tav.7 - Valori soglia $d_{n,\alpha}$ del test D_n di Kolmogorov (distribuzione a una coda),.....	458
Tav.8 - Coefficienti a_i del test di Shapiro-Wilk, per $3 \leq n \leq 50$ e $1 \leq i \leq k = \text{Int}[(n+1)/2]$	459
Tav.8 bis - Valori soglia $w_{n,\alpha}$ del test W di Shapiro-Wilk (a una coda), per $3 \leq n \leq 50$	460
Tav.9 - Valori soglia u_α del test U di Mann-Whitney.....	461
Tav.10 - Valori soglia $\hat{u}_{n_1, n_2, \alpha}$ del test robusto su ranghi \hat{U}_{n_1, n_2} di Fligner-Policello.....	463
Tav.11 - Valori soglia $d_{n_1, n_2, \alpha}$ del test D_{n_1, n_2} di Kolmogorov-Smirnov.....	467
Tav.12 - Valori soglia v_α del test di Wald-Wolfovitz.....	469
Tav.13 - Valori soglia w_α del test di Wilcoxon per campioni dipendenti.....	470
Tav.14 - Valori soglia k_α del test di Kruskal-Wallis, per c campioni ($c=3, 4, 5$).....	471
Tav.15 - Valori soglia r_α del test di Friedman.....	472
Tav.16 - Limiti inferiori d_L e superiori d_U dei valori soglia del test di Durbin-Watson.....	473
Tav.17 - Valori soglia d_α^* approssimati da Theil-Nagar del test di Durbin-Watson.....	475
Tav.18 - Valori soglia $r_{s,\alpha}$ del test di Spearman.....	476
Cenni sui logaritmi.....	477
Sul concetto di funzione.....	477
Sul concetto di limite.....	478
Sul concetto di derivata.....	478