

INDICE

SEZIONE PRIMA

Natura dei dati e della statistica

1 Le basi	2		
Statistica: ma a che serve?	2	Intervallo minimo-massimo (o "range")	23
Statistica descrittiva e statistica inferenziale	2	Intervallo interquartile	23
Variabili	3	Intervalli alternativi	24
Tipi di dati	3	Scarto medio	24
Dati discreti e dati continui	3	Varianza e deviazione standard	25
Dati nominali, ordinali, di intervallo e di rapporto	4	Coefficiente di variazione	26
Proporzioni e tassi	5	Asimmetria e curtosi	26
Esercizi	6	E allora, chi è normale?	27
2 Esaminare i dati	7	Box plot	27
<i>Un'occhiata alla rappresentazione grafica dei dati</i>		Cosa dobbiamo usare e quando (e perché)	28
Perché perdere tempo a esaminare i dati?	7	Stimatori robusti	29
Istogrammi, diagrammi a barre e variazioni sul tema	7	Misure alternative per la media	30
Primo movimento e tema principale: grafici a barre	7	Media geometrica	30
Prima variazione sul tema: il dot-plot	8	Media armonica	31
Rappresentare graficamente dati ordinali	8	Esercizi	32
Rappresentare graficamente dati di intervallo e dati di rapporto	9	<i>Come faccio a farlo fare al computer?</i>	32
Rami, foglie e altra vegetazione (gli stem-leaf plot)	10	4 La distribuzione normale	34
Poligoni di frequenza	11	Perché ci interessa la distribuzione normale	34
Poligoni di frequenza cumulativa	13	Valori della distribuzione normale standard, detti anche z-score	35
Come <i>non</i> fare un buon grafico	13	Curva normale	37
Abbiamo veramente bisogno di un grafico?	14	Esercizi	39
Il mistero dello zero scomparso	14	<i>Come faccio a farlo fare al computer?</i>	39
3-D o non 3-D, questo è il dilemma...	14	5 La probabilità	40
La torta in cielo, non nei grafici	15	Avvertenze per l'uso	40
Il peggio dei due mondi	15	Cosa intendiamo per "probabilità"?	40
Grafici sovrapposti (o "ammucchiati"...)	15	Approccio empirico	40
Conclusioni	16	Approccio teorico	41
Come fare tabelle migliori	16	Eventi mutuamente esclusivi ed eventi condizionati	41
Esercizi	18	Eventi mutuamente esclusivi e proprietà additiva della probabilità	41
<i>Come faccio a farlo fare al computer?</i>	18	Eventi condizionati e proprietà moltiplicativa della probabilità	42
3 Descrivere i dati con i numeri	20	Eventi indipendenti	43
<i>Indici di tendenza centrale e di dispersione</i>		Combinazioni e permutazioni	43
Una piccola digressione in merito alle notazioni	20	La regola di "Almeno una volta"	44
Indici di tendenza centrale	21	Distribuzione binomiale	44
Media	21	Che cos'è la distribuzione binomiale?	45
Mediana	21	Impariamo un po' di più sulla distribuzione binomiale	46
Moda	22	Distribuzione binomiale e distribuzione normale	46
Indici di dispersione	22	Riassunto	47
Indice di dispersione	23	Esercizi	48

6 Elementi di statistica inferenziale

Basi	49
Campioni e popolazione	49
Quali differenze?	50
Significato del magico valore di p	50
Un altro po' di nomenclatura	51
Elementi di statistica inferenziale	51
Errore standard e deviazione standard della media	52
Valore critico, zone di rifiuto, errore alfa e di Tipo I	53
Ipotesi alternativa ed errore beta	55
Divagazioni sul tema del "provare" l'ipotesi nulla	57
Da dove viene questo 5%?	58
Riti e leggende del $p < 0.05$	58
Probabilità esatta: l'approccio di Fisher	59
Un altro po' di storia e di filosofia	60
Quando è il caso di dare un'occhiata ai dati? Analisi esplorativa, regole di arresto e galateo statistico	62

Test multipli ed errori di Tipo I	64
Inferenza statistica e rapporto segnale-rumore	65
Due code o una coda?	65
Intervalli di confidenza	67
Significatività statistica e significatività clinica	68
Entità dell'effetto ("Effect Size")	69
Calcolo delle dimensioni del campione	69
Come presentare i risultati	71
Rappresentare graficamente intervalli di confidenza	71
Forest plot	72
Statistica bayesiana e statistica frequentista	72
Una traccia storica	72
Utilizzi attuali del teorema di Bayes	73
Riassunto	74
Esercizi	74

ALLARMI C.R.A.P.	76
-------------------------	----

SEZIONE SECONDA**Analisi della varianza****7 Confrontare due gruppi**

<i>Il t-test</i>	80
Una visione di insieme	80
Se abbiamo campioni della stessa numerosità	81
Gradi di libertà	82
Formule per la deviazione standard	82
Due gruppi di numerosità diversa: estensione del t -test	83
Stima combinata e non combinata della varianza	83
Occhiometria di precisione	84
Effect Size (entità dell'effetto)	84
Dimensioni del campione e potenza	85
Come presentare i risultati	86
Riassunto	86
Esercizi	87
<i>Come faccio a farlo fare al computer?</i>	87

Minima differenza significativa (Least Significant Difference – LSD) di Fisher	94
Differenza onestamente significativa (Honestly Significant Difference – HSD) di Tukey	94
Test di Neuman-Keuls	95
Differenza completamente significativa di Tukey	95
t di Dunnett	95
Metodo di Scheffé	96
Confronti pianificati ortogonali	96
Trend lineari e quadratici	98
Forza della relazione	98
Dimensioni del campione e potenza	98
Come presentare i risultati	99
Riassunto	99
Esercizi	100
<i>Come faccio a farlo fare al computer?</i>	100

8 Più di due gruppi

<i>ANOVA a una dimensione</i>	88
L'analisi un passo per volta	89
Somme dei Quadrati (Sum of Squares, SS)	89
Gradi di libertà	90
Medie dei Quadrati (Mean Squares, MS)	90
Valori attesi delle medie dei quadrati e distribuzione F	90
Assunzioni dell'ANOVA	91
Confronti multipli	91
Tasso di errore legato all'esperimento e tasso di errore legato alla famiglia di ipotesi	92
Confronti post-hoc	92
Correzione di Bonferroni	92
Modifiche della correzione di Bonferroni	93
Intervallo studentizzato	93

9 ANOVA fattoriale

Somma dei Quadrati e Media dei Quadrati negli studi fattoriali	104
Rappresentare graficamente i dati (ciak 1)	105
Rappresentare graficamente i dati (ciak 2)	107
Fattori casuali e fattori fissi	107
Disegno incrociato e annidato	108
Disegni equilibrati e squilibrati	109
Forza della relazione (versione riveduta e corretta)	110
Calcolo delle dimensioni del campione per le ANOVA fattoriali	111
Come presentare i risultati	111
Esercizi	111
<i>Come faccio a farlo fare al computer?</i>	112

10 Due osservazioni ripetute	113	<i>Esercizi</i>	128
<i>Il t-test per dati appaiati e le sue alternative</i>		<i>Come faccio a farlo fare al computer?</i>	129
Altri utilizzi del t-test per dati appaiati	116		
Effect Size	117		
Occhiometria di precisione	118		
Calcolo delle dimensioni del campione	118		
Come presentare i risultati	118		
Riassunto	118		
<i>Esercizi</i>	119		
<i>Come faccio a farlo fare al computer?</i>	119		
11 ANOVA per misure ripetute	120		
ANOVA per misure ripetute (un solo fattore)	121		
Generalizzazione volta a includere altri fattori nello studio	122		
Fattori “tra-soggetti” (between) e “intra-soggetti” (within)	124		
Altri impieghi dell’ANOVA per misure ripetute	126		
ANOVA per misure ripetute e affidabilità delle misure	126		
Assunti di base e limitazioni delle strutture più complesse di ANOVA	127		
Calcolo delle dimensioni del campione	127		
Come presentare i risultati	128		
Riassunto	128		
		12 ANOVA multivariata (MANOVA)	130
		Che cosa intendiamo per “multivariato”	132
		t per due (o più)	132
		Studio della varianza	133
		Dalla ANOVA alla MANOVA	133
		E finalmente lo si fa	134
		Più di due gruppi	135
		Facciamolo ancora: MANOVA per misure ripetute	136
		Robustezza	138
		Dimensioni del campione e potenza	138
		Quando le cose vanno storte: gestire le osservazioni anomale	138
		Effect Size (entità dell’effetto)	138
		Procedere con cautela	139
		Come presentare i risultati	139
		Tiriamo le somme	139
		<i>Esercizi</i>	139
		<i>Come faccio a farlo fare al computer?</i>	140
		ALLARMI C.R.A.P.	141

SEZIONE TERZA

Regressione e correlazione

13 Regressione semplice e correlazione	144	<i>F</i> -test parziali	159
Basi dell’analisi di regressione	144	Correlazioni parziali e semiparziali	159
B, beta e test di significatività	147	b e β	160
Retta di regressione: errori e intervalli di confidenza	148	Importanza relativa delle variabili	161
Coefficiente di determinazione e coefficiente di correlazione	148	Cerchiamo un centro...	162
Interpretazione del coefficiente di correlazione	149	Regressione gerarchica e stepwise	163
Intervalli di confidenza, test di significatività ed entità dell’effetto (Effect Size)	151	Regressione stepwise gerarchica	163
Rendere r normale	151	Regressione stepwise ordinaria	164
Intervalli di confidenza	151	Screening preliminare delle variabili	165
Entità dell’effetto (Effect Size)	152	R^2 , R^2 corretto (“adjusted”) e “shrinkage”	165
Calcolo delle dimensioni del campione	152	Interazioni	166
Test di ipotesi	152	Costruire variabili stupide (o “dummy coding”)	167
Come presentare i risultati	153	Quel che resta: esaminiamo i residui	168
Riassunto	153	Quando le cose vanno nel verso sbagliato	169
<i>Esercizi</i>	154	Discrepanza	169
<i>Come faccio a farlo fare al computer?</i>	154	Leverage (“effetto leva”)	169
		Influenza	169
		Multicollinearità	170
		Regressione multipla in pratica	171
14 Regressione multipla	155	Calcolo delle dimensioni del campione	171
Calcoli nella regressione multipla	155	Come presentare i risultati	172
Tipi di variabili	157	Riassunto	172
Relazioni tra le singole variabili	157	<i>Esercizi</i>	172
<i>F</i> -test parziali e correlazioni	159	<i>Come faccio a farlo fare al computer?</i>	173

15 Regressione logistica e di Poisson	174	Terzo passo: costruiamo il modello a livello 1	212
Procedure di maximum likelihood estimation	176	Quarto passo: costruiamo il modello a livello 2	212
Maximum likelihood estimation non condizionata e condizionata	177	E infine, un passo di lato: mettiamo le equazioni insieme	213
Calcoli sul nostro campione	177	Alcuni vantaggi di HLM	214
Interpretare rischi relativi e odds ratio	178	Come gestire soggetti "raggruppati" ("clusters")	214
"Goodness of fit" e test complessivi di significatività	179	Calcolo delle dimensioni del campione	215
Falsi R^2	180	Come presentare i risultati	215
Regressione logistica stepwise e test parziali	180	Riassunto	216
Situazioni più articolate	181	<i>Come faccio a farlo fare al computer?</i>	216
Regressione di Poisson	181	19 Componenti principali e analisi fattoriale	217
Un esempio per capire	183	<i>Giocherellando con i fattori</i>	
Statistiche "devianti"	183	Cosa sono i "fattori"?	217
Se i dati sono strambi	184	Come si fa	219
Calcolo delle dimensioni del campione	185	Matrice di correlazione	219
Come presentare i risultati	185	Estrazione dei fattori	220
Riassunto	185	Analisi dei componenti principali e fattorializzazione dell'asse principale	221
Esercizi	185	Mantenere o scartare i fattori	223
<i>Come faccio a farlo fare al computer?</i>	186	Matrice dei pesi fattoriali	225
16 Problemi avanzati di regressione e ANOVA	188	Rotazione dei fattori – Ma perché poi ruotare?	225
Analisi della covarianza	190	Interpretare i fattori	229
Utilizzare ANCOVA per controllare le differenze al tempo 0	192	Diamo nomi ai nomi	230
Come gestire disegni più complessi	193	Utilizzo dei fattori	230
Regressione non lineare	194	Tipi di dati da usare	231
Modello generalizzato lineare	194	Analizzare le domande – Alcune altre precauzioni	232
Assunti di base dell'ANCOVA	195	Calcolo delle dimensioni del campione	232
Calcolo delle dimensioni del campione	196	Come presentare i risultati	233
Riassunto	197	Esercizi	233
Esercizi	197	<i>Come faccio a farlo fare al computer?</i>	234
<i>Come faccio a farlo fare al computer?</i>	197	20 Path analysis e modelli a equazioni strutturali	235
17 Misurare i cambiamenti nel tempo	199	Path analysis (analisi dei "sentieri")	235
Problemi nella misura del cambiamento	200	Interpretare i numeri	236
Affidabilità delle differenze	200	Trovare la propria strada attraverso i sentieri	237
Regressione verso la media	201	Path analysis e causalità	237
Regressione alla media e ANCOVA	202	Variabili endogene ed esogene	238
Quando abbiamo più di una osservazione di follow-up: ANCOVA con contrasti	204	Tipologie di modelli di path analysis	238
Osservazioni multiple, tempo-dipendenti	205	Scusate il disturbo	239
Tiriamo le somme	207	Modelli ricorsivi e non-ricorsivi	240
Esercizi	207	K.I.S.S.	240
<i>Come faccio a farlo fare al computer?</i>	207	Meglio di così non si può: indicatori della bontà dell'adattamento (goodness-of-fit)	241
18 Analisi di dati longitudinali: modelli lineari gerarchici	208	Che cosa assumiamo	242
Modelli lineari gerarchici	209	Qualche parola sulla dimensione del campione	242
Cosa vuol dire?	209	Modelli a equazioni strutturali (SEM)	243
Cosa ci serve?	209	SEM e analisi fattoriale	243
L'analisi passo dopo passo	210	Specificazione del modello	245
Primo passo: esaminiamo i dati	210	Identificazione	246
Secondo passo: costruiamo rette di regressione individuali	211	Stima	247
		Valutare l'adattamento del modello	247
		Ri-specificazione	249

Analisi fattoriale confermativa	249	Riassunto	254
Confrontare due analisi fattoriali	251	<i>Esercizi</i>	254
Un modello SEM completo	252		
Come presentare i risultati	253	ALLARMI C.R.A.P.	255

SEZIONE QUARTA

Statistica non parametrica

21 Test di significatività per frequenze di dati categorici

Test del chi-quadrato	261
Scomponiamo il chi-quadrato	263
Effect Size	264
Numeri piccoli: correzione di Yates e test esatto di Fisher	264
Correzione per la continuità di Yates	264
Test esatto di Fisher	264
Osservazioni appaiate e corrispondenti – Il chi-quadrato di McNemar	267
Due fattori: il chi-quadrato di Mantel-Haenszel	268
Molti fattori: analisi log-lineare	270
Differenze tra proporzioni indipendenti	272
Test del segno	273
Calcolo delle dimensioni del campione	274
Jackknife, bootstrapping e ricampionamento	275
Come presentare i risultati	276
Riassunto	276
<i>Esercizi</i>	276
<i>Come faccio a farlo fare al computer?</i>	277

22 Misure di associazione per dati categorici

Misure di associazione per le tabelle 2 × 2 e altre di ordine superiore	279
Coefficiente phi, coefficiente di contingenza, Q di Yule e V di Cramer	280
Kappa di Cohen	280
Concordanza parziale e kappa pesato	282
Relazione tra kappa e correlazione intraclasse	284
Calcolo delle dimensioni del campione	284
Riassunto	285
<i>Esercizi</i>	285
<i>Come faccio a farlo fare al computer?</i>	285

23 Test di significatività per dati trasformati in ranghi

Due gruppi indipendenti	287
Più di due gruppi	288
Confronti multipli	289
Effect Size	289
Mediane ordinate: il test di Jonckheere	289
Due fattori di raggruppamento	290

Misure ripetute: il test dei ranghi con segno di Wilcoxon e l'ANOVA a due dimensioni di Friedman	290
Calcolo delle dimensioni del campione	291
Riassunto	292
<i>Esercizi</i>	292
<i>Come faccio a farlo fare al computer?</i>	293

24 Misure di associazione per i ranghi

Rho di Spearman	294
Significatività del rho di Spearman	295
Correlazione punto-biseriale	295
Tau di Kendall	296
Calcolo di tau	296
Test di significatività per tau	297
Interpretare rho piuttosto che tau	297
W di Kendall	297
Calcolo di W	297
Test di significatività per W	298
Gamma	298
Quello che non vi diremo	299
Riassunto	300
<i>Esercizi</i>	300
<i>Come faccio a farlo fare al computer?</i>	300

25 Analisi della sopravvivenza

Quando utilizzare l'analisi della sopravvivenza	302
Come sintetizzare i dati	303
Sopravvivenza media	303
Tasso di sopravvivenza (survival rate)	304
Uso degli anni-persona	304
Tecniche di analisi di sopravvivenza	304
Approccio attuariale	305
Metodo di Kaplan-Meier per l'analisi della sopravvivenza	307
Funzione dell'hazard	308
Misure di sintesi	308
Errore Standard	309
Assunti di base dell'analisi della sopravvivenza	309
Confrontare due (o più) gruppi	310
Test z	311
Test di Mantel-Cox (log-rank test) e altri test	311
Controllo per le covariate	313
Verifichiamo l'assunzione di proporzionalità	315
Intervalli di confidenza	316

Dimensioni del campione e potenza	316	Rischio relativo	321
Come presentare i risultati	317	Riduzione assoluta e relativa del rischio	321
<i>Esercizi</i>	317	Odds ratio	322
<i>Come faccio a farlo fare al computer?</i>	318	OR versus RR	323
26 Misure di impatto	319	Problemi con i rapporti	323
Contare i vivi: incidenza e prevalenza	319	Numero necessario da trattare (NNT)	323
Incidenza	319	Intervalli di confidenza per RR, OR e NNT	324
Prevalenza	320	Contare i morti: indici di esiti negativi	325
Quando usarle	320	Letalità e tassi di mortalità	325
Indici di rischio: rischio relativo, odds ratio e altri	320	Riassunto	326
		ALLARMI C.R.A.P.	327

SEZIONE QUINTA

Epilogo

27 Studi di equivalenza e non inferiorità	332	29 Mettiamo tutto insieme	351
Introduzione alla Filosofia I	332	Statistica descrittiva	351
Razionale statistico	333	Statistica univariata	352
Valutare l'assenza di differenza	333	Statistica parametrica	352
Intervallo di equivalenza	333	Statistica non parametrica	354
Studi di equivalenza	334	Statistica multivariata	355
Non-inferiorità	335	Dimensioni campionarie su due piedi	355
Errori di Tipo I e di Tipo II	336	Differenze tra proporzioni	355
Dimensione campionaria ed effetti di <i>I</i>	336	Differenze tra molte medie	355
Altre opportunità per i bricconi	336	Come scrivere i risultati di un'analisi statistica	356
		Nel frattempo, a casa	357
28 Problemi, stranezze e altri capricci della scienza	338	30 Primi passi con SPSS	358
<i>Osservazioni anomale, dati mancanti e trasformazioni</i>		Mettetevi alla prova	368
Scovare i dati anomali	338	<i>Una raccolta di domande e risposte</i>	
Tipi di dati mancanti ("missing data")	340	Risposte agli esercizi dei Capitoli	373
Riempiamo i vuoti	341	Bibliografia e Letture consigliate	383
Facciamo finta che non ci siano proprio	341	Un Dizionario senza Vergogna dei Termini Statistici	392
Nel dubbio, tirate a indovinare	342	Appendice	395
Trasformazioni dei dati	344	Indice analitico	419
Trasformare o non trasformare	344		
Se avete proprio deciso di trasformare	347		
Dopo la trasformazione	349		
<i>Come faccio a farlo fare al computer?</i>	349		