

Stephen Bernstein
Ruth Bernstein

Statistica inferenziale

McGraw-Hill

Milano • New York • St. Louis • San Francisco • Auckland • Bogotá
Colombia • London • Madrid • Mexico City • Montreal
New Delhi • San Juan • Singapore • Sydney • Tokyo • Toronto

PREFAZIONE XIII

Capitolo 1

POPOLAZIONI, CAMPIONI E STATISTICA 1

1.1	Popolazione fisica e misurabile	1
1.2	Popolazione finita, infinita e ipotetica	1
1.3	Campioni	2
1.4	Parametri e statistica	2
1.5	La scienza statistica	3
1.6	Problemi di stima e problemi di verifica dell'ipotesi	4
1.7	Ipotesi statistiche e ipotesi di ricerca	4
1.8	Ricerca esplorativa e ricerca di verifica dell'ipotesi	5
1.9	Esperimenti esplorativi	5
1.10	Esperimenti controllati	6
1.11	Studi osservativi	7
1.12	Indagini e censimenti	7
1.13	Tecniche statistiche parametriche e non parametriche	8
1.14	Statistica matematica e statistica generale	9
1.15	Schemi di campionamento	9
1.16	La probabilità nel campionamento; con e senza reinmissione	10
1.17	Campionamento casuale	11
1.18	Campionamento casuale semplice	12
1.19	Campionamento casuale stratificato	12
1.20	Campionamento casuale sistematico	13
1.21	Campionamento a grappoli	13
1.22	Campionamento non casuale	13
1.23	Tavole di numeri casuali	14
	Problemi Risolti	16
	Problemi Supplementari	21

Capitolo 2

DISTRIBUZIONI CAMPIONARIE 23

2.1	Campionamento casuale semplice	23
2.2	Variabili casuali indipendenti	23
2.3	Definizioni matematiche e non matematiche del campionamento casuale semplice	24
2.4	Le ipotesi dei metodi di campionamento	25
2.5	La variabile casuale $\bar{X}$	26
2.6	Distribuzione campionaria teorica ed empirica della media	27
2.7	La media della distribuzione campionaria della media	32
2.8	Precisione di uno stimatore	33

2.9	La varianza della distribuzione della media campionaria: il caso della popolazione infinita o del campionamento con reimmissione	33
2.10	La varianza della distribuzione della media campionaria: il caso della popolazione finita con campionamento senza reimmissione	34
2.11	Lo scarto quadratico medio campionario della media	35
2.12	La variabilità di uno stimatore	37
2.13	Il calcolo delle probabilità attraverso la distribuzione discreta della media campionaria	38
2.14	Calcolo delle probabilità nel caso di una distribuzione della media campionaria normalmente distribuita	39
-2.15	Il teorema limite centrale: campionamento da popolazione finita con reimmissione	40
2.16	Il teorema limite centrale: campionamento da una popolazione infinita	43
2.17	Il teorema limite centrale: campionamento da una popolazione finita senza reimmissione	43
2.18	Quanto grande è il "sufficientemente grande"?	44
2.19	La distribuzione della somma campionaria	45
2.20	Applicazioni del teorema limite centrale alla distribuzione campionaria della somma campionaria	46
2.21	Campionamento da una popolazione binomiale	47
2.22	Distribuzione campionaria del numero di successi	48
2.23	Distribuzione campionaria della proporzione	49
2.24	Applicazioni del teorema limite centrale alla distribuzione campionaria del numero dei successi	50
2.25	Applicazioni del teorema limite centrale alla distribuzione campionaria della proporzione	50
2.26	Il calcolo delle probabilità con l'approssimazione della distribuzione della proporzione campionaria alla normale	52
	Problemi Risolti	53
	Problemi supplementari	66

Capitolo 3

STIMA DELLA MEDIA DI UNA POPOLAZIONE ATTRAVERSO UN CAMPIONE 71

-3.1	La stima	71
3.2	Criteri per la scelta dello stimatore ottimo	72
3.3	L'errore standard stimato della media S_x	73
3.4	Stime puntuali	73
3.5	Ripartire e valutare la stima puntuale	74
3.6	Relazione tra stime puntuali e stime a intervallo	75
3.7	Dimostrazione che $P(\bar{x} - z_{\alpha/2} \leq \bar{X} \leq \bar{x} + z_{\alpha/2}) = P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha$	75
3.8	Dimostrazione che $P(\bar{X} - z_{\alpha/2} \sigma_x \leq \mu \leq \bar{X} + z_{\alpha/2} \sigma_x) = 1 - \alpha$	77
3.9	Intervallo di confidenza per la media della popolazione μ per una popolazione distribuita normalmente e con s.q.m. σ noto	78
3.10	Presentazione dei limiti di confidenza	79
3.11	Precisione dell'intervallo di confidenza	80
3.12	Determinazione della numerosità campionaria quando l's.q.m. è noto	82
3.13	Intervallo di confidenza per la media della popolazione μ : s.q.m. σ noto, campione grande ($n > 30$) estratto da popolazione con distribuzione qualsiasi	83
3.14	Determinazione degli intervalli di confidenza per la media μ della popolazione quando l's.q.m. σ della popolazione non è noto	84
3.15	La distribuzione t	84

3.16	Relazione tra la distribuzione t e la distribuzione normale standard	86
3.17	I gradi di libertà	87
3.18	Il termine distribuzione "t di student"	87
3.19	I valori critici della distribuzione t	88
3.20	Tavola A.6: valori critici della distribuzione t	90
3.21	Intervallo di confidenza per la media della popolazione μ : s.q.m. σ non noto, campione piccolo ($n < 30$) estratto da una popolazione distribuita normalmente	92
3.22	Calcolo della numerosità campionaria: s.q.m. non noto, campione piccolo estratto da una popolazione distribuita normalmente	93
3.23	Intervallo di confidenza per la media della popolazione μ : s.q.m. σ non noto, campione grande ($n \geq 30$) da una popolazione normalmente distribuita	95
3.24	Intervallo di confidenza per la media della popolazione μ : s.q.m. non noto, campione grande ($n \geq 30$) da una popolazione che è distribuita non normalmente	97
3.25	Intervallo di confidenza per la media della popolazione μ : campione piccolo ($n < 30$) estratto da una popolazione che non è distribuita normalmente	97
	Problemi Risolti	99
	Problemi Supplementari	108

Capitolo 4

STIMA CAMPIONARIA DELLA VARIANZA, DELLO SCARTO QUADRATICO MEDIO E DELLA PROPORZIONE PER UNA POPOLAZIONE 113

4.1	Stimatori ottimali della varianza, dello scarto quadratico medio e della popolazione	113
4.2	La statistica chi quadro e la distribuzione chi quadro	113
4.3	I valori critici della distribuzione chi quadro	115
4.4	La tavola A.7: i valori critici della distribuzione chi quadro	117
4.5	Costruzione degli intervalli di confidenza per la varianza σ^2 di una popolazione distribuita normalmente	118
4.6	Presentare i limiti di confidenza	119
4.7	L'ampiezza degli intervalli di confidenza per la varianza	120
4.8	Determinare la numerosità campionaria necessaria per ottenere una desiderata qualità della stima della varianza	121
4.9	Usare l'approssimazione normale per determinare gli intervalli di confidenza per la varianza	122
4.10	Usare la distribuzione campionaria della varianza campionaria per approssimare gli intervalli di confidenza per la varianza della popolazione	123
4.11	Intervallo di confidenza per l's.q.m. σ di una popolazione normalmente distribuita	124
4.12	Usare la distribuzione campionaria della deviazione standard campionaria per il calcolo dell'intervallo di confidenza approssimato per l's.q.m. della popolazione	125
4.13	Lo stimatore ottimale per la proporzione p di una popolazione binomiale	126
4.14	Derivare l'intervallo di confidenza approssimato per la proporzione p di una popolazione binomiale	126
4.15	Stima del parametro p	128
4.16	Decidere quando n è "sufficientemente grande" per p non noto	129
4.17	Intervallo di confidenza approssimati per il parametro p della binomiale nel caso di campionamento da una popolazione finita senza reimmissione	129
4.18	L'intervallo di confidenza esatto per il parametro p della binomiale	130
4.19	La precisione delle stime intervallari approssimate per il parametro p della binomiale	130

4.20	Determinare la numerosità campionaria per l'intervallo di confidenza per il parametro p della binomiale	130
4.21	Intervallo di confidenza approssimato per la percentuale di una popolazione binomiale	132
4.22	Intervallo di confidenza approssimato per il numero totale di "successi" di una popolazione binomiale	133
4.23	Il metodo di cattura e ricattura per la stima della numerosità della popolazione n	133
	Problemi Risolti	134
	Problemi supplementari	144

Capitolo 5

VERIFICA DELLE IPOTESI SULLA BASE DI UN SOLO CAMPIONE 147

5.1	La verifica delle ipotesi statistiche	147
5.2	L'ipotesi nulla e l'ipotesi alternativa	147
5.3	Verificare l'ipotesi nulla	148
5.4	Verifica di una ipotesi unidirezionale contro una ipotesi bidirezionale	149
5.5	Test delle ipotesi riguardo la media della popolazione μ : s.q.m. σ noto, popolazione distribuita normalmente	149
5.6	Il P -value	150
5.7	Errore di I tipo contro errore di II tipo	152
5.8	Valori critici e regioni critiche	152
5.9	Il livello di significatività	154
5.10	Regole di decisione per i test di ipotesi statistiche	155
5.11	Scegliere le ipotesi statistiche	156
5.12	La probabilità dell'errore di II tipo	157
5.13	Il rischio del consumatore e il rischio del produttore	158
5.14	Perché non è possibile provare l'ipotesi nulla	158
5.15	L'inferenza classica e l'inferenza bayesiana	158
5.16	Una procedura per la verifica dell'ipotesi nulla	159
5.17	Verifica delle ipotesi usando $\bar{X}$ come statistica test	161
5.18	La potenza del test, la curva operativa caratteristica e la curva di potenza	162
5.19	Verifica delle ipotesi riguardo la media della popolazione μ : s.q.m. σ non noto, campione piccolo ($n < 30$) estratto da una popolazione distribuita normalmente	163
5.20	IL P -Value per la statistica t	164
5.21	Regole di decisione per la verifica delle ipotesi con la statistica t	165
5.22	β , $1 - \beta$, curva di potenza e curva OC	166
5.23	Verifica delle ipotesi sulla media della popolazione μ : campione grande ($n \geq 30$) estratto da una popolazione con distribuzione qualunque	166
5.24	Le assunzioni della verifica di ipotesi parametriche basate su un solo campione	167
5.25	Cosa accade quando le assunzioni sono violate	168
5.26	La verifica delle ipotesi sulla varianza σ^2 di una popolazione distribuita normalmente	169
5.27	La verifica delle ipotesi sull's.q.m. σ di una popolazione distribuita normalmente	171
5.28	La verifica delle ipotesi sulla proporzione p di una distribuzione binomiale: il caso di un campione grande	171
5.29	La verifica delle ipotesi sulla proporzione p di una distribuzione binomiale: il caso di un campione piccolo	173
	Problemi Risolti	173
	Problemi supplementari	188

Capitolo 6

**METODI DI STIMA E DI VERIFICA DELLE IPOTESI
NEL CASO DI DUE CAMPIONI 193**

- 6.1 Campioni indipendenti e campioni appaiati 193
- 6.2 Lo stimatore ottimo della differenza tra le medie di due popolazioni ($\mu_1 - \mu_2$) 194
- 6.3 La distribuzione campionaria teorica della differenza tra due medie 194
- 6.4 Intervalli di confidenza per la differenza tra le medie ($\mu_1 - \mu_2$):
s.q.m. noti (σ_1 e σ_2), campioni indipendenti estratti da popolazioni
distribuite normalmente 195
- 6.5 Verifica delle ipotesi circa la differenza tra le medie ($\mu_1 - \mu_2$):
s.q.m. noti (σ_1 e σ_2), campioni indipendenti estratti da popolazioni
distribuite normalmente 196
- 6.6 La stima dell'errore standard della differenza fra due medie 198
- 6.7 Intervalli di confidenza per la differenza fra due medie ($\mu_1 - \mu_2$): s.q.m. non noti
ma uguali ($\sigma_1 = \sigma_2$), numerosità campionarie piccole ($n_1 < 30$ e $n_2 < 30$),
campioni indipendenti estratti da popolazioni distribuite normalmente 199
- 6.8 Verifica delle ipotesi sulla differenza fra due medie ($\mu_1 - \mu_2$): s.q.m. non noti
ma uguali ($\sigma_1 = \sigma_2$), numerosità campionarie piccole ($n_1 < 30$ e $n_2 < 30$),
campioni indipendenti estratti da popolazioni normalmente distribuite 200
- 6.9 Intervalli di confidenza per la differenza fra due medie ($\mu_1 - \mu_2$): s.q.m. non noti
(σ_1 e σ_2), numerosità campionarie grandi ($n_1 \geq 30$ e $n_2 \geq 30$),
campioni indipendenti estratti da popolazioni qualunque 202
- 6.10 La verifica delle ipotesi sulla differenza fra due medie ($\mu_1 - \mu_2$): s.q.m. non noti
(σ_1 e σ_2), numerosità campionarie grandi ($n_1 \geq 30$ e $n_2 \geq 30$),
campioni indipendenti estratti da popolazioni qualunque 203
- 6.11 Intervalli di confidenza per la differenza fra le medie ($\mu_1 - \mu_2$): campioni appaiati 203
- 6.12 La verifica delle ipotesi sulla differenza fra le medie ($\mu_1 - \mu_2$): campioni appaiati 207
- 6.13 Le assunzioni per la stima e la verifica parametrica riguardo alle medie
per due campioni 208
- 6.14 Quando le assunzioni vengono violate 209
- 6.15 Il confronto fra i metodi per campioni indipendenti e per campioni appaiati
in termini di precisione e potenza 210
- 6.16 La statistica F 210
- 6.17 La distribuzione F 211
- 6.18 I valori critici della distribuzione F 213
- 6.19 Tavola A.8: valori critici della distribuzione F 215
- 6.20 Intervalli di confidenza per il rapporto delle varianze (σ_1^2/σ_2^2):
parametri ($\sigma_1^2, \sigma_1, \mu_1$ e $\sigma_2^2, \sigma_2, \mu_2$) non noti, campioni indipendenti estratti
da popolazioni distribuite normalmente 215
- 6.21 Verifica delle ipotesi sul rapporto delle varianze (σ_1^2/σ_2^2):
parametri ($\sigma_1^2, \sigma_1, \mu_1$ e $\sigma_2^2, \sigma_2, \mu_2$) non noti, campioni indipendenti estratti
da popolazioni distribuite normalmente 218
- 6.22 Quando è necessario verificare l'omogeneità delle varianze 220
- 6.23 Lo stimatore ottimo della differenza tra le proporzioni ($p_1 - p_2$):
campioni indipendenti grandi 220
- 6.24 La distribuzione campionaria teorica della differenza fra due proporzioni 221
- 6.25 Intervalli di confidenza approssimati per la differenza fra proporzioni
di due popolazioni binomiali ($p_1 - p_2$): campioni grandi e indipendenti 222
- 6.26 Verifica delle ipotesi sulla differenza fra proporzioni di due popolazioni
binomiali ($p_1 - p_2$): campioni grandi e indipendenti 224
- Problemi Risolti 226
- Problemi supplementari 241

Capitolo 7

TEORIA DELLA STIMA E VERIFICA DELLE IPOTESI**NEL CASO DI CAMPIONI MULTIPLI 247**

- 7.1 Inferenza nel caso di campioni multipli 247
- 7.2 L'analisi della varianza 247
- 7.3 *ANOVA*: a un fattore, a due fattori, a più fattori 248
- 7.4 *ANOVA* a un fattore: effetti fissi o effetti casuali 248
- 7.5 *ANOVA* a un fattore con effetti fissi: le assunzioni 249
- 7.6 *ANOVA* a un fattore con effetti fissi e campioni di uguale numerosità: H_0 e H_1 249
- 7.7 *ANOVA* a un fattore con effetti fissi e campioni di uguale numerosità:
organizzare i dati 249
- 7.8 *ANOVA* a un fattore con effetti fissi e campioni di uguale numerosità:
la logica fondamentale 251
- 7.9 $SST = SSA + SSW$ 252
- 7.10 Le formule per il calcolo di SST e SSA 253
- 7.11 Gradi di libertà e medie dei quadrati 253
- 7.12 Il test F 255
- 7.13 La tabella dell'*ANOVA* 257
- 7.14 I test per i confronti multipli 257
- 7.15 Il test dei range multipli di Duncan 258
- 7.16 Calcolo degli intervalli di confidenza secondo i confronti multipli 259
- 7.17 La verifica dell'omogeneità delle varianze 261
- 7.18 *ANOVA* a un fattore con effetti fissi: numerosità campionarie uguali o diverse 262
- 7.19 La procedura generale dell'*ANOVA* a un fattore con effetti fissi:
l'organizzazione dei dati 263
- 7.20 La procedura generale dell'*ANOVA* a un fattore con effetti fissi:
la somma dei quadrati 264
- 7.21 La procedura generale dell'*ANOVA* a un fattore con effetti fissi:
gradi di libertà e medie dei quadrati 264
- 7.22 La procedura generale dell'*ANOVA* a un fattore con effetti fissi: il test F 265
- 7.23 La procedura generale dell'*ANOVA* a un fattore con effetti fissi:
i confronti multipli 266
- 7.24 La procedura generale dell'*ANOVA* a un fattore con effetti fissi: il calcolo
degli intervalli di confidenza e la verifica dell'omogeneità delle varianze 267
- 7.25 Violazione delle assunzioni dell'*ANOVA* 268
- Problemi Risolti 269
- Problemi supplementari 280

Capitolo 8

REGRESSIONE E CORRELAZIONE 285

- 8.1 Analizzare le relazioni fra due variabili 285
- 8.2 Il modello di regressione lineare semplice 286
- 8.3 La retta dei minimi quadrati 287
- 8.4 Lo stimatore della varianza $\sigma^2_{Y.X}$ 290
- 8.5 La media e la varianza dell'intercetta $\hat{a}$ di y e del coefficiente angolare $\hat{b}$ 290
- 8.6 Intervalli di confidenza per l'intercetta a di y e del coefficiente angolare b 292
- 8.7 Intervalli di confidenza per la varianza $\sigma^2_{Y.X}$ 293
- 8.8 Intervalli di previsione per i valori di Y 294
- 8.9 Verifica delle ipotesi sul coefficiente angolare b 295
- 8.10 Il confronto tra equazioni di regressione lineare semplice di due o più campioni 296
- 8.11 Regressione lineare multipla 296

8.12	Correlazione lineare semplice	296
8.13	Derivazione del coefficiente di correlazione r	298
8.14	Intervallo di confidenza per il coefficiente di correlazione della popolazione ρ	301
8.15	La verifica delle ipotesi sul coefficiente di correlazione della popolazione ρ tramite la distribuzione r	302
8.16	La verifica delle ipotesi su ρ con la distribuzione t	304
8.17	Uso della distribuzione Z per verificare l'ipotesi $\rho = c$	304
8.18	Interpretare il coefficiente di correlazione campionario r	305
8.19	La correlazione multipla e la correlazione parziale	306
	Problemi risolti	306
	Problemi supplementari	325

Capitolo 9

METODI NON PARAMETRICI 331

9.1	Metodi non parametrici e metodi parametrici	331
9.2	Il test del chi quadrato	331
9.3	Il test del chi quadrato per la bontà dell'adattamento	332
9.4	Test del chi quadrato per l'indipendenza: le tabelle di contingenza	333
9.5	Il test del chi quadrato per l'uguaglianza fra k proporzioni di binomiali	335
9.6	I test sui ranghi	337
9.7	Test per un solo campione: il test dei ranghi con segno di Wilcoxon	337
9.8	Test per due campioni: il test dei ranghi con segno di Wilcoxon per campioni dipendenti	340
9.9	Test per due campioni: il test U di Mann-Whitney per campioni indipendenti	341
9.10	I test per più campioni: il test H di Kruskal-Wallis per k campioni indipendenti	344
9.11	Il test di Spearman sulla correlazione fra i ranghi	347
	Problemi Risolti	351
	Problemi supplementari	371

APPENDICE 377