

L'ANALISI DELLA VARIANZA
SU UN MINICALCOLATORE

ANNA PAOLA ERCOLANI <1>

ALESSANDRA ARENI <1>

LUCIA MANNETTI <1>

DARIO SALMASO <2>

GIOVANNI LARICCIA <3>

ANALISI DELLA VARIANZA

SU UN MINICALCOLATORE

<1> ISTITUTO DI PSICOLOGIA DELL'UNIVERSITA' DI ROMA

<2> ISTITUTO DI PSICOLOGIA (C N R)

<3> ISTITUTO PER LE APPLICAZIONI DEL CALCOLO (C N R)

ROMA, GIUGNO 1978

RAPPORTO TECNICO DELL'ISTITUTO DI PSICOLOGIA

DEL C N R N. 246

L'ANALISI DELLA VARIANZA E' UNA TECNICA STATISTICA CHE CONSENTE DI SCOMPORRE LA VARIABILITA' TOTALE DI UN ESPERIMENTO IN VARIABILITA' PARZIALI; ATTRIBUIBILI ALLE FONTI DI VARIAZIONE PRESENTI NELL'ESPERIMENTO STESSO.

IN QUESTO MODO E' POSSIBILE CONTROLLARE IL PESO CHE OGNI FONTE DI VARIAZIONE HA SULLA VARIABILE DIPENDENTE E LA RELATIVA SIGNIFICATIVITA'.

IL PROGRAMMA PROPOSTO E' UN PROGRAMMA GENERALE CHE, ALLO STATO ATTUALE, FORNISCE I RISULTATI DI BASE PER EFFETTUARE TALE CONTROLLO. IL PROGRAMMA, PROPRIO PERCHE' GENERALIZZATO, CONSENTE DI ELABORARE DATI PROVENIENTI DA DISEGNI SPERIMENTALI CON MISURE INDIPENDENTI, DISEGNI CON MISURE RIPETUTE SU TUTTE LE VARIABILI E DISEGNI MISTI.

COME CASO PARTICOLARE, IL PROGRAMMA FORNISCE GLI ELEMENTI PER IL CALCOLO DEL "T" DI STUDENT PER CAMPIONI CORRELATI E PER CAMPIONI INDIPENDENTI (QUEST'ULTIMO CASO SOLO SE IL NUMERO DI SOGGETTI E' IDENTICO NEI DUE CAMPIONI).

ATTUALMENTE IL PROGRAMMA HA UNA CAPACITA' MASSIMA DI SEI FATTORI ($NV = 6$), COMPRENDENDO TRA QUESTI ANCHE IL "FATTORE" SOGGETTI, E DI $M = 2500$ (DOVE M SI OTTIENE DAL PRODOTTO TRA IL NUMERO DEI LIVELLI, AUMENTATO DI UNA UNITA', DI CIASCUN FATTORE, COMPRESO LO PSEUDO-FATTORE SOGGETTI).

PER ESEMPIO: SE L'ESPERIMENTO PREVEDE TRE FATTORI INDIPENDENTI DEI QUALI DUE A TRE LIVELLI (MODALITA') E UNO A DUE LIVELLI; E INOLTRE 10 SOGGETTI PER OGNI COMBINAZIONE DEI LIVELLI DEI FATTORI; ALLORA:

$$M = (3+1) (3+1) (2+1) (10+1) = 528$$

SI FA NOTARE CHE IL NUMERO DI SOGGETTI NELLE VARIE COMBINAZIONI DEI LIVELLI DEI FATTORI DEVE ESSERE COSTANTE (ESPERIMENTO BILANCIATO).

PER UTILIZZARE IL PROGRAMMA E' NECESSARIO:

- COMPILARE DELLE ISTRUZIONI DI CONTROLLO CHE FORNISCONO AL PROGRAMMA STESSO LE INFORMAZIONI SULL'ORGANIZZAZIONE DEI DATI;
- INSERIRE I DATI MEDESIMI.

ORGANIZZAZIONE DEI DATI

.....

CIASCUN RICEPCATORE PUO' ORGANIZZARE I DATI COME CREDE, IN FUNZIONE DELLE VARIABILI CHE SONO IN GIOCO, COSTRUENDO UNA MATRICE A DUE DIMENSIONI (RIGHE PER COLONNE); TENENDO PRESENTE CHE CIASCUNA RIGA PUO' CONTENERE FINO A UN MASSIMO DI 20 UNITA', TRA CIFRE E PUNTI DECIMALI (NON VIRGOLE!!!).

PER ESEMPIO: SUPPONENDO UN ESPERIMENTO A TRE FATTORI (A CON 2 LIVELLI; B CON 3 LIVELLI E C CON 2 LIVELLI) CON 5 SOGGETTI PER OGNI COMBINAZIONE DEI LIVELLI DEI FATTORI, LA MATRICE POTREBBE ESSERE:

TABELLA 1 - ORGANIZZAZIONE DEI DATI IN UN ESPERIMENTO 2 X 3 X 2 X 5

	A1			A2		
	B1	B2	B3	B1	B2	B3
c1	s1	x1111	x2111	x3111	x1211	x2211
	s2	x1121				
	s3					
	s4					
	s5		x3151			x3251
c2	s1					
	s2					
	s3					
	s4					
	s5	x1152	x2152	x3152	x1252	x2252

OPPURE POTREBBE ESSERE:

TABELLA 1A - IDEM

		C1				C2					
B1		B2		B3		B1		B2		B3	
A1	A2	A1	A2	A1	A2	A1	A2	A1	A2	A1	A2

s1	x1111									x2321	
s2											
s3											
s4											
s5	x1115									x2325	

E' EVIDENTE CHE QUESTO SECONDO MODO DI ORGANIZZARE I DATI COM-
 PORTA UNA RIGA PIU' LUNGA, QUINDI POTREBBE NON ESSERE COMPATIBILE CON
 LA CAPACITA' DI 80 UNITA' DELLA RIGA STESSA.

DATI RICHIESTI DAL PROGRAMMA

LE ISTRUZIONI DI CONTROLLO RIGUARDANO:

- IL TITOLO (CONVENZIONALE, DI SEI LETTERE) DEL LAVORO DA ESEGUIRE;
- IL NUMERO DEI FATTORI DEL DISEGNO (COMPRESO LO PSEUDO-FATTORE SOGGETTI);
- LA LETTERA DELL'ALFABETO DISTINTIVA DI OGNI FATTORE CON IL CORRISPONDENTE NUMERO DI LIVELLI (SI CONSIGLIA DI INDICARE SEMPRE CON LA LETTERA 'S' LO PSEUDO-FATTORE SOGGETTI);
- IL FORMATO DI LETTURA DEI DATI.

L'ORGANIZZAZIONE DI QUESTE ISTRUZIONI E' LA SEGUENTE:

SCHEDA 1	COL. 1-6	TITOLO (IN SEI LETTERE, ALFANUMERICO)
	7-8	NUMERO DEI FATTORI
	9-15	BLANK (SPAZI BIANCHI)
	16	LETTERA DISTINTIVA DEL PRIMO FATTORE (♦)
	17-20	NUMERO DI LIVELLI DEL PRIMO FATTORE
	21	LETTERA DISTINTIVA DEL SECONDO FATTORE
	22-25	NUMERO DI LIVELLI DEL SECONDO FATTORE

♦ NELL'INDICARE IL PRIMO, SECONDO, ECC. FATTORE, OCCORRE TENER PRESENTE CHE NELLA LETTURA DEI DATI, IL PROGRAMMA CONSIDERA COME PRIMO FATTORE IL FATTORE DI COLONNA PIU' INTERNO, COME SECONDO FATTORE IL FATTORE DI COLONNA SUCCESSIVO, ECC., ESAURITI I FATTORI DI COLONNA, CONSIDERA COME SUCCESSIVO IL FATTORE DI RIGA PIU' INTERNO E COSI' VIA.

PER ESEMPIO, CONSIDERANDO LA TABELLA 1 (VEDI ORGANIZZAZIONE DEI DATI) I FATTORI VARIANO NELL'ORDINE: B (PRIMO FATTORE), A (SECONDO FATTORE), S (TERZO FATTORE) E C (QUARTO FATTORE). CONSIDERANDO INVECE LA TABELLA 1A, I FATTORI VARIANO NELL'ORDINE: A, B, C, S.

.....

SCHEDA 2 BLANK (SCHEDA BIANCA)

SCHEDA 3 FORMATO DI LETTURA IN F (A COL.1 APERTA PARENTESI; POI
IL FORMATO DI LETTURA; INFINE CHIUDERE LA PARENTESI).
PER ESEMPIO; (12F6.4) INDICA CHE OGNI RIGA CONTIENE 12
NUMERI DEL TIPO 1.4648; OPPURE (10F3.0) INDICA CHE OGNI
RIGA CONTIENE 10 NUMERI DEL TIPO - 24. - .

OUTPUT

IN OUTPUT IL PROGRAMMA PREVEDE LA STAMPA DI:

- TITOLO CONVENZIONALE DEL LAVORO (COME DA INPUT)
- LA LETTERA DISTINTIVA DI CIASCUN FATTORE E IL CORRISPONDENTE NUMERO DI LIVELLI
- LA MEDIA GENERALE
- LA TABELLA DELL'ANALISI DELLA VARIANZA CHE COMPRENDE: LA DEVIANZA, I GDL E LA VARIANZA DI CIASCUNA FONTE DI VARIABILITA'.

PER FONTI DI VARIABILITA' SI INTENDONO L'EFFETTO PRINCIPALE DI CIASCUN FATTORE E TUTTE LE INTERAZIONI POSSIBILI DI PRIMO, SECONDO, ECC. ORDINE, SECONDO IL TIPO DI DISEGNO. PER ESEMPIO, IN UN DISEGNO DI TIPO FATTORIALE COMPOSTO DAL FATTORE A (CON DUE LIVELLI) E DAL FATTORE B (CON TRE LIVELLI) E DA 10 SOGGETTI PER CELLA (PSEUDO-FATTORE 'S'),

SOTTO LA VOCE "FONTI DI VARIABILITA'" IN OUTPUT COMPARIRANNO: A, B, AB, S, AS, BS, ABS E TOTALE.

A QUESTO PUNTO, PER TESTARE LA SIGNIFICATIVITA' DI CIASCUNA FONTE DI VARIAZIONE, E' NECESSARIO L'INTERVENTO DIRETTO DEL RICERCATORE. INFATTI OCCORRE CALCOLARE, PARTENDO DAI RISULTATI OTTENUTI IN OUTPUT, LO SPECIFICO "ERRORE" DI CIASCUNA FONTE DI VARIAZIONE; QUINDI FARE IL RAPPORTO $F = \text{VARIANZA FONTE} / \text{VARIANZA "ERRORE"}$; E INFINE CONTROLLARE LA SIGNIFICATIVITA' DELLA "F" SULLE APPOSITE TAVOLE.

IL CALCOLO DEGLI SPECIFICI "ERRORI" VARIA SECONDO IL DISEGNO DI ANALISI UTILIZZATO. PER SEMPLIFICARE QUESTA FASE, RIFORTIAMO QUI DI SEGUITO GLI ESEMPI DI ALCUNI SCHEMI DI FREQUENTE USO, DAI QUALI POSSONO ESSERE DERIVATI TUTTI GLI ALTRI SCHEMI SPECIFICI.

E S E M P I O 1

--- PROVE INDIPENDENTI ---

(DA J.L.BRUNING - B.L.KINTZ : COMPUTATIONAL HANDBOOK OF
 STATISTICS GLENVIEW : SCOTT, FORESMAN AND COMPANY, 1968;
 PAG 31)

CASO DI 3 FATTORI A, B, C, CIASCUNO A 2 LIVELLI E 6 SOGGETTI
 PER CIASCUNA COMBINAZIONE DI LIVELLI

T A B E L L A D A T I

c1				c2			
b1		b2		b1		b2	
a1	a2	a1	a2	a1	a2	a1	a2
s 1 15	s 7 23	s13 11	s19 25	s25 14	s31 24	s37 11	s43 39
s 2 8	s 8 16	s14 16	s20 27	s26 6	s32 16	s38 17	s44 32
s 3 9	s 9 17	s15 8	s21 29	s27 3	s33 10	s39 12	s45 38

s 4 7	s10 17	s16 14	s22 34	s28 5	s34 18	s40 14	s46 39
s 5 8	s11 14	s17 20	s23 38	s29 6	s35 13	s41 21	s47 33
s 6 4	s12 11	s18 16	s24 26	s30 7	s36 17	s42 18	s48 31

I N P U T

PROVA1 4 A0002B00002C00002S0006

(8F3.0)

15.23.11.25.14.24.11.39.
 08.16.16.27.06.16.17.32.
 09.17.08.29.03.10.12.38.
 07.17.14.34.05.18.14.39..
 08.14.20.38.06.13.21.33.
 04.11.16.26.07.17.18.31.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA1

LIVELLI DEI FATTORI

A	2
B	2
C	2
S	6

MEDIA GENERALE 17.85417

FONTE DI VARIABILITA'	DEVIANZA	GDL	VARIANZA
A	2093.52083	1	2093.52083
B	1645.02083	1	1645.02083
AB	247.52083	1	247.52083
C	20.02083	1	20.02083
AC	25.52083	1	25.52083
BC	54.18750	1	54.18750
ABC	4.68750	1	4.68750
S	121.10417	5	24.22083
AS	88.10417	5	17.62083
BS	252.60417	5	50.52083
ABS	61.60417	5	12.32083
CS	46.60417	5	9.32083
ACS	28.60417	5	5.72083
BCS	67.93750	5	13.58750
ABCS	24.93750	5	4.98750
TOTALE	4781.97917	47	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE VARIANZE DELLE
SEGUENTI FONTI DI VARIABILITA' : A, B, C, AB, AC, CB, ABC.

PER TROVARE LA DEVIANZA ENTRO I GRUPPI ('ERRORE', 'WITHIN')
OCCORRE SOMMARE LA DEVIANZA DI 'S' CON LE DEVIANZE DELLE INTERAZIONI
DI TUTTE LE FONTI DI VARIAZIONE CON 'S', CIOE' :

$$\text{DEVIANZA 'ERRORE'} = S + AS + BS + CS + ABS + ACS + BCS + ABCS$$

LO STESSO PROCEDIMENTO SI SEGUE PER I GRADI DI LIBERTA' :

$$\text{GDL 'ERRORE'} = S + AS + BS + CS + ABS + ACS + BCS + ABCS$$

NEL CASO ESEMPLIFICATO LA DEVIANZA 'ERRORE' SARA' UGUALE A :

121.10417 +

89.10417 +

252.60417 +

46.60417 +

61.60417 +

28.60417 +

67.93750 +

24.93750 = 691.50060

ANALOGAMENTE PER IL CALCOLO DEI GDL :

$$5+5+5+5+5+5+5 = 40$$

LA VARIANZA DELL'ERRORE E' DATA DAL RAPPORTO :

$$691.50060 / 40 = 17.28751$$

PER TESTARE LA SIGNIFICATIVITA' DI OGNI FONTE DI VARIAZIONE
SI CALCOLANO I RISPETTIVI RAPPORTI 'F' SECONDO LA FORMULA :

$$F = \text{VARIANZA FONTE} / \text{VARIANZA 'ERRORE'}$$

E SI CERCANO SULLE APPOSITE TAVOLE LE CORRISPONDENTI PROBABILITA' .

NELL'ESEMPIO SI AVRA' :

FOONTE DI VARIAZIONE	GDL	VARIANZE	F
.....			
TOTALE	47	4781.97917	
A	1	2093.52083	121.100 ***
B	1	1645.02083	95.157 ***
C	1	20.02083	1.158
AB	1	247.52083	14.318 ***
AC	1	25.52083	1.476
BC	1	54.18750	3.134
ABC	1	4.68750	0.271
ERRORE	40	17.28751	

*** QUESTI VALORI DI F SONO SIGNIFICATIVI AL 0.001

E S E M P I O 2

--- DISEGNO MISTO ---

.....

(DA B.J. WINER STATISTICAL PRINCIPLES IN EXPERIMENTAL DESIGN
NEW YORK : MCGRAW-HILL 1962; PAG 337)

CASO DI 3 FATTORI A, B, C, DEI QUALI 'A' E 'B' A 2 LIVELLI
E 'C' A 4 LIVELLI; CON 'A' E 'B' INDIPENDENTI E 'C' RIPETUTO;
CON 3 SOGGETTI PER CIASCUNA COMBINAZIONE DEI LIVELLI DEI FATTORI
INDIPENDENTI

TABELLA DATI

		c1	c2	c3	c4	
A1		s1	18	14	12	6
	B1	s2	19	12	8	4
		s3	14	10	6	2
		s4	16	12	10	4
	B2	s5	12	8	6	2
		s6	18	10	5	1
		s7	16	10	8	4
	B1	s8	18	8	4	1
		s9	16	12	6	2
A2		s10	19	16	10	8
	B2	s11	16	14	10	9
		s12	16	12	8	8

I N P U T

PROVA204 C000430003B0002A0002

(4F3.0)

18.14.12.06.
19.12.08.04.
14.10.06.02.
16.12.10.04.
12.08.06.02.
18.10.05.01.
16.10.08.04.
18.08.04.01.
16.12.06.02.
19.16.10.08.
16.14.10.09.
16.12.08.08.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA2

LIVELLI DEI FATTORI

C	4	
S	3	
B	2	
A	2	
MEDIA GENERALE		10.00000

Fonte di Variabilità	Devianza	ddl	Varianza
C	991.50000	3	330.50000
S	50.37500	2	25.18750
CS	9.12500	6	1.52083
B	8.33333	1	8.33333
CB	12.16667	3	4.05556
SB	1.54167	2	.77083
CSB	27.95833	6	4.65972
A	10.08333	1	10.08333
CA	8.41667	3	2.80556
SA	7.29167	2	3.64583
CSA	5.20833	6	.86806
BA	80.08333	1	80.08333
CBA	12.75000	3	4.25000
SBA	23.29167	2	11.64583
CSBA	9.87500	6	1.64583
TOTALE	1258.00000	47	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE VARIANZE DELLE SEGUENTI FONTI DI VARIABILITÀ: A, B, AB, C, AC, BC, ABC, LE PRIME TRE RELATIVE AI FATTORI INDIPENDENTI E LE RIMANENTI RELATIVE AL FATTORE RIPETUTO.

IN QUESTO DISEGNO È NECESSARIO CALCOLARE DUE DIVERSE DEVIANZE DI "ERRORE": UNA RELATIVA AI FATTORI INDIPENDENTI E UNA RELATIVA AL FATTORE RIPETUTO.

$$S/AB = S + AS + BS + ABS$$

E

$$SC/AB = SC + SAC + SBC + SABC$$

NELL'ESEMPIO SI OTTIENE, PER S/AB:

$$50.37500 + 7.29167 + 1.54167 + 23.29167 = 8250001$$

ANALOGAMENTE PER IL CALCOLO DEI GDL: $2 + 2 + 2 + 2 = 8$

E, PER SC/AB:

$$\begin{aligned} &9.12500 + \\ &5.20833 + \\ &27.95833 + \\ &9.87500 = 52.16666 \end{aligned}$$

I GDL SARANNO: $6 + 6 + 6 + 6 = 24$

LE DUE VARIANZE SARANNO UGUALI RISPETTIVAMENTE A:

$$\text{VAR}(S/AB) = 82.50001/8 = 10.31250$$

$$\text{VAR}(SC/AB) = 52.16666/24 = 2.17360$$

PER TESTARE LA SIGNIFICATIVITA' DELLE FONTI DI VARIAZIONE

A, B, AB (FATTORI INDIPENDENTI) SI CALCOLANO I RAPPORTI 'F' SECONDO LA FORMULA:

$$'F' = \text{VARIANZA FONTE} / \text{VARIANZA} (S/AB)$$

PER TESTARE LA SIGNIFICATIVITA' DELLE FONTI DI VARIAZIONE

C, AC, BC, ABC (FATTORE RIPETUTO) SI CALCOLANO I RAPPORTI 'F' SECONDO LA FORMULA:

$$'F' = \text{VARIANZA FONTE} / \text{VARIANZA} (SC/AB)$$

NELL'ESEMPIO SI AVRA':

FONTE DI VARIANZA	DEVIANZE	GL	VARIANZA	F
TOTALE	1258.00000	47		
BETWEEN S	181.00100	11		
A	10.08333	1	10.08333	0.977
B	8.33333	1	8.33333	0.808
AB	80.08333	1	80.08333	7.766 *
S/AB	82.50010	8	10.31250	
WITHIN S	1076.99998	36		
C	991.50000	3	330.50000	152.052 ***
AC	8.41667	3	2.80556	1.291
BC	12.16667	3	4.05556	1.866
ABC	12.75000	3	4.25000	1.955
SC/AB	52.16666	24	2.17360	

* QUESTO VALORE DI 'F' E' SIGNIFICATIVO AL 0.05

*** QUESTO VALORE DI 'F' E' SIGNIFICATIVO AL 0.001

(DA J.L.MYERS FUNDAMENTALS OF EXPERIMENTAL DESIGN

BOSTON : ALLYN AND BACON , 1972; PAG 211)

CASO DI 3 FATTORI A, B, C, DEI QUALI 'A' A 3 LIVELLI,
'B' A 2 LIVELLI E 'C' A 4 LIVELLI; CON 'A' INDIPENDENTE E 'B' E 'C'
RIPETUTI ; CON 2 SOGGETTI PER CIASCUN LIVELLO DI 'A'.

TABELLA DATI

		B1				B2			
		c1	c2	c3	c4	c1	c2	c3	c4
A1	s1	3	5	7	8	7	6	9	9
	s2	2	5	6	8	8	5	10	12
A2	s3	8	9	9	11	6	7	7	9
	s4	7	7	9	10	7	8	9	11

	s5	10	12	13	15	5	7	8	10
A3	s6	11	11	14	16	8	8	9	11

I N P U T

PROVA304 C0004B00002S00002A0003

(8F3.0)

03.05.07.08.07.06.09.09.
 02.05.06.08.08.05.10.12.
 08.09.09.11.06.07.07.09.
 07.07.09.10.07.08.09.11.
 10.12.13.15.05.07.08.10.
 11.11.14.16.08.08.09.11.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA3

LIVELLI DEI FATTORI

C	4
B	2
S	2
A	3

MEDIA GENERALE 8.58333

FONTE DI VARIABILITA'	DEVIANZA	GDL	VARIANZA
C	115.66667	3	38.55556
B	8.33333	1	8.33333
CB	3.00000	3	1.00000
S	3.00000	1	3.00000
CS	3.00000	3	1.00000
BS	8.33333	1	8.33333
CBS	.33333	3	.11111
A	106.16667	2	53.08333
CA	5.83333	6	.97222
BA	105.16667	2	52.58333
CBA	8.50000	6	1.41667
SA	1.50000	2	.75000
CSA	2.50000	6	.41667
BSA	1.16667	2	.58333
CBSA	3.16667	6	.52778
TOTALE	375.66667	47	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE VARIANZE DELLE
SEGUENTI FONTI DI VARIABILITA' : A, B, AB, C, AC, BC, ABC, LA PRIMA RELA
TIVA AL FATTORE INDIPENDENTE E LE RIMANENTI RELATIVE AI FATTORI RIPETUTI
IN QUESTO DISEGNO E' NECESSARIO CALCOLARE QUATTRO DIVERSE DE
VIANZE DI "ERRORE": UNA RELATIVA AL FATTORE INDIPENDENTE, E LE ALTRE
RELATIVE AI FATTORI RIPETUTI.

$$S/A = S + AS$$

$$SB/A = SB + SBA$$

$$SC/A = SC + SCA$$

$$SBC/A = SBC + SABC$$

NELL'ESEMPIO SI OTTIENE :

$$\text{PER } S/A = 3.00000 + 1.50000 = 4.5$$

$$\text{PER } SB/A = 8.33333 + 1.16667 = 9.49999$$

$$\text{PER } SC/A = 3.00000 + 2.50000 = 5.5$$

$$\text{PER } SBC/A = 0.33333 + 3.16667 = 3.49997$$

ANALOGAMENTE PER I GRADI DI LIBERTA' SI OTTIENE :

PER $S/A = 1 + 2 = 3$
PER $SB/A = 1 + 2 = 3$
PER $SC/A = 3 + 6 = 9$
PER $SBC/A = 3 + 6 = 9$

LE VARIANZE RELATIVE SI OTTENGONO DIVIDENDO LE DEVIANZE
PER I RELATIVI GDL.

PER TESTARE LA SIGNIFICATIVITA' DELLE FONTI DI VARIAZIONE
A, B, AB, C, AC, BC, ABC, SI CALCOLANO I RAPPORTI 'F' DIVIDENDO
LA 'VARIANZA FONTE' PER LA SUA RELATIVA VARIANZA DI 'ERRORE'.

NELL'ESEMPIO SI AVRA' :

Fonte di Variazione	Devianze	GDL	Varianza	F	
TOTALE	375.66667	47			
BETWEEN S	110.66667	5			
A	106.16667	2	53.083	35.389	♦♦
S/A	4.50000	3	1.500		
WITHIN S	265.00000				
B	8.33333	1	8.333	2.632	
AB	105.16700	2	52.583	16.605	♦
SB/A	9.50000	3	3.167		
C	115.66667	3	38.556	62.523	♦♦
AC	5.83333	6	0.972	1.577	
SC/A	5.50000	9	0.611		
BC	3.00000	3	1.000	2.571	
ABC	8.50000	6	1.417	3.643	
SBC/A	3.49997	9	0.389		

♦ QUESTO VALORE DI 'F' E' SIGNIFICATIVO AL 0.025

♦♦ QUESTO VALORE DI 'F' E' SIGNIFICATIVO AL 0.01

E S E M P I O 4.

--- PROVE RIPETUTE ---

CASO DI TRE FATTORI A, B, C, TUTTI RIPETUTI E CIASCUNO A DUE LIVELLI; I SOGGETTI SONO 4.

T A B E L L A D A T I

	A1				A2			
	B1		B2		B1		B2	
	c1	c2	c1	c2	c1	c2	c1	c2
s1	7	12	11	14	6	4	5	3
s2	9	15	20	20	7	8	9	4
s3	9	10	11	15	12	13	11	15
s4	10	11	25	28	10	15	20	18

I N P U T

PROVA404 C0002B00002A00002S0004

<8F3.0>

07.12.11.14.06.04.05.03.

09.15.20.20.07.08.09.04.

09.10.11.15.12.13.11.15.

10.11.25.28.10.15.20.18.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA4

LIVELLI DEI FATTORI

C	2
B	2
A	2
S	4

MEDIA GENERALE 12.09375

FONTE DI VARIABILITA'	DEVIANZA	GDL	VARIANZA
C	16.53125	1	16.53125
B	157.53125	1	157.53125
CB	5.28125	1	5.28125
A	140.28125	1	140.28125
CA	16.53125	1	16.53125
BA	81.28125	1	81.28125
CBA	1.53125	1	1.53125
S	356.34375	3	118.78125
CS	4.59375	3	1.53125
BS	130.09375	3	43.36458
CBS	20.84375	3	6.94792
AS	125.84375	3	41.94792
CAS	14.09375	3	4.69792
BAS	16.84375	3	5.61458
CBAS	9.09375	3	3.03125
TOTALE	1096.71875	31	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE VARIANZE DELLE SEGUENTI FONTI DI VARIAZIONE: A, B, C, AB, AC, BC, ABC; IN QUESTO DISEGNO, TRATTANDOSI DI PROVE TUTTE RIPETUTE, CIASCUNA DELLE FONTI DI VARIAZIONE VIENE TESTATA SU UN SUO SPECIFICO "ERRORE". PER ESEMPIO, LA VARIANZA DI ERRORE RELATIVA AL FATTORE "A" E' "AS", QUELLA RELATIVA A LA FONTE DI VARIAZIONE "BC" SARA' "BCS", E COSI' VIA.

DATO IL PARTICOLARE TIPO DI DISEGNO, NELLA TABELLA DI OUTPUT COMPARIANO DIRETTAMENTE, SENZA NECESSITA' DI ULTERIORI CALCOLI, LE DEVIANZE E LE VARIANZE DI "ERRORE" RELATIVE A CIASCUNA FONTE DI VARIAZIONE DA TESTARE.

NELL'ESEMPIO DI AVRA':

FONTE DI VARIAZIONE	DEVIANZE	GDL	VARIANZE	F
TOTALE	1096.71875	31		
A	140.28125	1	140.28125	3.345
AS	125.84375	3	41.94792	
B	157.53125	1	157.53125	3.633
BS	130.09375	3	43.36458	

C	16.53125	1	16.53125	10.796 *
CS	4.59375	3	1.53125	
AB	81.28125	1	81.28125	14.477 *
ABS	16.84375	3	5.61458	
AC	16.53125	1	16.53125	3.519
ACS	14.09375	3	4.69792	
BC	5.28125	1	5.28125	0.760
BCS	20.84375	3	6.94792	
ABC	1.53125	1	1.53125	0.505
ABCS	9.09375	3	3.03125	

* QUESTO VALORE E' SIGNIFICATIVO AL 0.05

ULTERIORE UTILIZZAZIONE DEL PROGRAMMA A N O V A

DATA LA NOTA RELAZIONE TRA LA 'F' DI FISHER E LA 'T' DI STUDENT NEL CASO DI DUE CAMPIONI (CIOE' $T^2 = \text{SORT } F$), IL PROGRAMMA DI ANALISI DELLA VARIANZA PUO' ESSERE UTILIZZATO PER CALCOLARE LA 'T' DI STUDENT TRA DUE CAMPIONI CORRELATI O TRA DUE CAMPIONI INDIPENDENTI (IN QUEST'ULTIMO CASO SOLO SE IL NUMERO DI SOGGETTI NEI DUE CAMPIONI E' IDENTICO).

E S E M P I O 5

--- DUE CAMPIONI CORRELATI ---

.....

(DA J.L.BRUNING - B.L.KINTZ COMPUTATIONAL HANDBOOK OF
STATISTICS GLENVIEW : SCOTT, FORESMAN AND COMPANY, 1968;
PAG 14)

T A B E L L A D A T I

	A1	A2
s1	89	94
s2	86	94
s3	96	101
s4	100	105
s5	94	100
s6	86	84
s7	81	81
s8	93	96
s9	87	90
s10	89	88
s11	110	115
s12	95	100
s13	107	110
s14	96	102

I N P U T

PROVA502 A000230014

(2F4.0)

089.094.

086.094.

096.101.

100.105.

094.100.

086.084.

081.081.

093.096.

087.090.

089.088.

110.115.

095.100.

107.110.

096.102.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA5

LIVELLI/DEI FATTORI

A 2
S 14
MEDIA GENERALE 95.32143

FONTE DI VARIABILITA'	DEVIANZA	GDL	VARIANZA
A	92.89286	1	92.89286
S	1999.60714	13	153.81593
AS	53.60714	13	4.12363
TOTALE	2146.10714	27	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE VARIANZE DI 'A'
E DI 'AS' NEL MODO SEGUENTE :

$$T' = \text{SQRT} (\text{VARIANZA 'A'} / \text{VARIANZA 'AS'})$$

NELL'ESEMPIO :

$$T' = \text{SQRT} (92.89286 / 4.12363) = 4.74$$

IN QUESTO CASO I GDL SONO PARI AL NUMERO DELLE COPPIE DI
PUNTEGGI MENO UNO : CIDE' 14 - 1 = 13.

NELL'APPOSITA TAVOLA PER LA 'T' DI STUDENT SI CONTROLLERA' LA
SIGNIFICATIVITA' DELL'INDICE TROVATO. NELL'ESEMPIO LA 'T' DI STUDENT
E' SIGNIFICATIVA AL 0.001 .

E S E M P I O 6

--- DUE CAMPIONI INDIPENDENTI ---

CASO DI DUE CAMPIONI INDIPENDENTI, CIASCUNO CON 12 SOGGETTI.

T A B E L L A D A T I

	A1		A2
s1	107	s13	109
s2	96	s14	94
s3	88	s15	127
s4	131	s16	76
s5	109	s17	115
s6	84	s18	121
s7	79	s19	87
s8	105	s20	92
s9	108	s21	91
s10	92	s22	98
s11	96	s23	104
s12	101	s24	96

I N P U T

PROVA602

A0002S0012

(2F4.0)

107.109.

096.094.

088.127.

131.076.

109.115.

084.121.

079.087.

105.092.

108.091.

092.098.

096.104.

101.096.

O U T P U T

ANALISI DELLA VARIANZA.....PROVA6

LIVELLI DEI FATTORI

A 2

S 12

MEDIA GENERALE

100.25000

FONTE DI VARIABILITA'	DEVIANZA	GDL	VARIANZA
A	8.16667	1	8.16667
S	1251.50000	11	113.77273
RS	3294.83333	11	299.53030
TOTALE	4554.50000	23	

DALLA TABELLA DELL'OUTPUT SI UTILIZZANO LE DEVIANZE DI 'S'
E 'RS' CHE SOMMATE DANNO LA DEVIANZA DI 'ERRORE' RELATIVA AD 'A'

NEL NOSTRO CASO :

DEVIANZA DI 'ERRORE' = 1251.50000 + 3294.83333 = 4546.33333

ANALOGAMENTE PER I GDL :

GDL 'ERRORE' = 11 + 11 = 22

LA VARIANZA DI 'ERRORE' CORRISPONDERA' ALLORA A :

$4546.33333 / 22 = 206.652$

QUINDI $T = \text{SQRT} (8.16667 / 206.652) = 0.199$

IN QUESTO CASO I GDL SONO PARI AL NUMERO TOTALE DEI SOGGETTI
MENO 2. NEL NOSTRO CASO SONO UGUALI A 22 .

IN QUESTO ESEMPIO IL VALORE DELLA 'T' DI STUDENT NON E'
SIGNIFICATIVO.

A P P E N D I C E

LISTA DEI PROGRAMMI FORTRAN INSTALLATI SUL MINICALCOLATORE
TEXAS 990/10 (SISTEMA OPERATIVO 3.0)

C.....

C "ANOVA.FTN"..... P.ERCOLANI, S.ARENI, APRILE 1978

C L.MANNETTI D.SALMASO..G.LARICCIA.....

C PROGRAMMA GENERALE DI ANALISI DELLA VARIANZA ISPIRATO ALLO SCIENTIFIC
C AND STATISTIC PACKAGE (SSP) DELL'IBM

C UTILIZZA LE SUBROUTINES

C \$AVDAT

C \$AVCAL

C \$MEANO

C

C UNIT 1 ----- FILE DI INPUT

C UNIT 5 ----- STAMPANTE (D FILE)

C.....28/6/78

DOUBLE PRECISION X(2500), GMEAN, SUMSQ(63), SMEAN(63), SUM

DIMENSION FMT1(72), FMT2(72)

DIMENSION HEAD(6), LEVEL(6), ISTEP(6), KOUNT(6), LASTS(6)

~~DIMENSION~~ NDF (63)

DIMENSION FNT (15)

```

1  FORMAT (A4,A2,I2,A4,3X,11(A1,I4))
2  FORMAT (1H1,' ANALISI DELLA VARIANZA .....', A4,A2//)
3  FORMAT (' LIVELLI DEI FATTORI'/(3X,A1,7X,I4))
4  FORMAT (1H1,' MEDIA GENERALE',F20.5//)
5  FORMAT (' FONTE DI ',18X,'DEVIANZA',10X,'GDL',11X,'VARIANZA',6X,'F
♦  '/// VARIABILITA'///')
6  FORMAT (1H ,15A1,F20.5,6X,I6,F20.5,7X,' .')
7  FORMAT (' TOTALE',10X,F20.5,6X,I6)
8  FORMAT (72A1)
9  FORMAT (A4,A2,I2,I4,(11(A1,I4)))
10  FORMAT (' I DATI, NON SONO LEGGIBILI SECONDO QUEL FORMATO')
100  READ (1,1,END=999) PR,PR1,K,BLANK,(HEAD(I),LEVEL(I),I=1,K)
      WRITE (5,2) PR,PR1
      WRITE (5,3) (HEAD(I),LEVEL(I),I=1,K)
      READ (1,8) (FMT1(L),L=1,72)
      READ (1,8) (FMT2(L),L=1,72)
      N=LEVEL(1)
      DO102I=2,K
102  N=N+LEVEL(I)
      READ (1,FMT1,ERR=998) (X(I),I=1,N)

```

CALL AVIDAT (K, LEVEL, N, X, L, ISTEP, KOUNT)

CALL AVCAL (K, LEVEL, X, L, ISTEP, LASTS)

CALL MEANO (K, LEVEL, X, GMEAN, SUMSQ, NDF, SMEAN, ISTEP, KOUNT, LASTS)

WRITE (5, 4) GMEAN

WRITE (5, 5)

LL=2**K-1

ISTEP(1)=1

DO105I=2,K

105 ISTEP(I)=0

DO110I=1,15

110 FNT(I)=BLANK

NN=0

SUM=0.0

120 NN=NN+1

L=0

DO140I=1,K

FNT(I)=BLANK

IF(ISTEP(I))130,140,130

130 L=L+1

FNT(L)=HEAD(I)

140 CONTINUE

WRITE (5, 6) (FNT(I), I=1, 15), SUMSQ (NN), NDF (NN), SMEAN (NN)

SUM=SUM+SUMSQ (NN)

IF (NN-LL) 145,170,170

145 DO160I=1,K

IF (ISTEP(I)) 147,150,147

147 ISTEP(I)=0

GOTO160

150 ISTEP(I)=1

GOTO120

160 CONTINUE

170 N=N-1

WRITE (5,7) SUM,N

GOTO100

999 STOP 1

998 WRITE (5,10)

STOP 2

END

8.4

SUBROUTINE RYDAT(K,LEVEL,N,X,L,ISTEP,KOUNT)

DOUBLE PRECISION X(1)

DIMENSION LEVEL(1),ISTEP(1),KOUNT(1)

M=LEVEL(1)+1

DO 105 I=2,K

```
105  M=N*(LEVEL(I)+1)
      N1=M+1
      N2=N+1
      DO 107 I=1,N
      N1=N1-1
      N2=N2-1
107  X(N1)=X(N2)
      ISTEP(1)=1
      DO 110 I=2,K
110  ISTEP(I)=ISTEP(I-1)*(LEVEL(I-1)+1)
      DO 115 I=1,K
115  KOUNT(I)=1
      N1=N1-1
      DO 135 I=1,N
      L=KOUNT(1)
      DO 120 J=2,K
120  L=L+ISTEP(J)*(KOUNT(J)-1)
      N1=N1+1
      X(L)=X(N1)
      DO 130 J=1,K
      IF (KOUNT(J)-LEVEL(J)) 124, 125, 124
124  KOUNT(J)=KOUNT(J)+1
```

60 / TO 135

125 KOUNT(J)=1

130 CONTINUE

135 CONTINUE

RETURN

END

8.4

SUBROUTINE AWCAL(K,LEVEL,X,L,ISTEP,LASTS)

DOUBLE PRECISION X(1),SUM

DIMENSION LEVEL(1),ISTEP(1),LASTS(1)

LASTS(1)=L+1

DO 145 I=2,K

145 LASTS(I)=LASTS(I-1)+ISTEP(I)

150 DO 175 I=1,K

L=1

LL=1

SUM=0.0

NN=LEVEL(I)

FN=NN

INCRE=ISTEP(I)

LAST=LASTS(I)

155 DO 160 J=1,NN

```

      SUM=SUM+X(L)
160  L=L+INCRE
      X(L)=SUM
      DO 165 J=1,NN
      X(LL)=FN*X(LL)-SUM
165  LL=LL+INCRE
      SUM=0.0
      IF(L-LAST) 167, 175, 175
167  IF(L-LAST+INCRE) 168, 168, 170
168  L=L+INCRE
      LL=LL+INCRE
      GO TO 155
170  L=L+INCRE+1-LAST
      LL=LL+INCRE+1-LAST
      GO TO 155
175  CONTINUE
      RETURN
      END

```

8.4

```

      SUBROUTINE MEANDOK, LEVEL, X, GMEAN, SUMSQ, MDF, SMEAN, MSTEP, KOUNT,
1  LASTS)
      DOUBLE PRECISION X(1), GMEAN, SUMSQ(1), SMEAN(1), FN1

```

```
DIMENSION LEVEL(1),MDF(1),MSTEP(1),KOUNT(1),LASTS(1)
N=LEVEL(1)
DO 150 I=2,K
150 N=N+LEVEL(I)
LASTS(1)=LEVEL(1)
DO 178 I=2,K
178 LASTS(I)=LEVEL(I)+1
MM=1
LL=(2+K)-1
MSTEP(1)=1
DO 180 I=2,K
180 MSTEP(I)=MSTEP(I-1)+2
DO 185 I=1,LL
185 SUMSQ(I)=0.0
DO 190 I=1,K
190 KOUNT(I)=0
200 L=0
DO 260 I=1,K
IF(KOUNT(I)-LASTS(I)) 210, 250, 210
210 IF(L) 220, 220, 240
220 KOUNT(I)=KOUNT(I)+1
IF(KOUNT(I)-LEVEL(I)) 230, 230, 250
```



```

230  L=L+MSTEP(I)
      GO TO 260
240  IF(KOUNT(I)-LEVEL(I)) 230, 260, 230
250  KOUNT(I)=0
260  CONTINUE
      IF(L) 265, 265, 270
270  SUMSQ(L)=SUMSQ(L)+X(NN)*X(NN)
      NN=NN+1
      GO TO 200
265  FN=N
      GMEAN=X(NN)/FN
      DO 310 I=2,K
310  MSTEP(I)=0
      NN=0
      MSTEP(1)=1
320  ND1=1
      ND2=1
      DO 340 I=1,K
      IF(MSTEP(I)) 330, 340, 330
330  ND1=ND1+LEVEL(I)
      ND2=ND2+(LEVEL(I)-1)
340  CONTINUE
    
```

$FN1 = N + ND1$
 $FN2 = ND2$
 $NN = NN + 1$
 $SUMSQ (NN) = SUMSQ (NN) / FN1$
 $MDF (NN) = ND2$
 $SMEAN (NN) = SUMSQ (NN) / FN2$

IF (NN-LL) 345, 370, 370

345 DO 360 I=1,K

IF (MSTEP(I)) 347, 350, 347

347 MSTEP(I)=0

GO TO 360

350 MSTEP(I)=1

GO TO 320

360 CONTINUE

370 RETURN

END