

ESERCIZIO 1. Vengono riportati di seguito i risultati di un'analisi discriminante.

Test di uguaglianza delle medie di gruppo

	Lambda di Wilks	F	df1	df2	Sig.
SELF_EFF	,977	1,169	3	147	,324
COLL_EFF	,977	1,134	3	147	,337
COIN_LAV	,946	2,789	3	147	,043
IMPEGNO	,898	5,583	3	147	,001
SODDISF	,941	3,095	3	147	,029
CAP_IST	,993	,323	3	147	,808
COLLEGHI	,989	,529	3	147	,663

1.a. Quali sono le variabili che presentano medie significative nei gruppi ?

Risultati del test

M di Box		150,438
F	Appross.	1,610
	df1	84
	df2	22310,595
	Sig.	,000

Test dell'ipotesi nulla che le matrici di covarianza siano uguali nella popolazione.

1.b. E' possibile accettare l'ipotesi nulla del test M di Box e perché ?

Autovalori

Funzione	Autovalore	% di varianza	% cumulata	Correlazione canonica
1	,143 ^a	54,5	54,5	,354
2	,077 ^a	29,4	83,9	,268
3	,042 ^a	16,1	100,0	,202

a. Per l'analisi sono state usate le prime 3 funzioni discriminanti canoniche.

1.c. Quale è la percentuale di varianza tra i gruppi spiegata da ogni funzione discriminante ?

1.d. Quale è la relazione tra ciascuna funzione discriminante e il gruppo di appartenenza dei soggetti ?

Lambda di Wilks

Test di funzioni	Lambda di Wilks	Chi-quadrato	df	Sig.
Da 1 a 3	,779	36,133	21	,021
Da 2 a 3	,890	16,759	12	,159
3	,959	6,003	5	,306

1.e. Quante sono le funzioni discriminanti statisticamente significative, e perché ?

Matrice di struttura				Coefficienti standardizzati della funzione discriminante canonica			
	Funzione				Funzione		
	1	2	3		1	2	3
IMPEGNO	,881*	,180	-,023	SELF_EFF	-,210	1,042	,108
SODDISF	,617*	-,134	,413	COLL_EFF	,346	-,631	-,550
COIN_LAV	,572*	,273	,313	COIN_LAV	-,097	,405	,716
COLL_EFF	,330*	-,256	-,242	IMPEGNO	1,011	,356	-1,227
COLLEGHI	,029	-,372*	,007	SODDISF	,033	-,879	1,287
SELF_EFF	,287	,353*	,239	CAP_IST	-,612	,541	-,200
CAP_IST	-,146	-,037	-,284*	COLLEGHI	,031	-,535	-,033

Correlazioni comuni entro gruppi tra variabili discriminanti e funzioni discriminanti canoniche standardizzate
Variabili ordinate in base alla dimensione assoluta della correlazione entro la funzione.
* . Correlazione assoluta più grande tra ciascuna variabile e qualsiasi funzione discriminante

1.f. Quali sono le variabili che caratterizzano maggiormente ogni funzione discriminante ?

1.g. Quali sono le variabili che risultano più “ridondanti” per ogni funzione discriminante?

Risultati della classificazione^{b,c}

			Gruppo di appartenenza previsto				Totali
			meno di 5	5-14	15-24	oltre 24	
Originale	Conteggio	meno di 5	14	4	7	3	28
		5-14	11	21	12	13	57
		15-24	7	13	16	8	44
		oltre 24	5	4	2	11	22
		Casi non raggruppati	25	19	20	16	80
	%	meno di 5	50,0	14,3	25,0	10,7	100,0
		5-14	19,3	36,8	21,1	22,8	100,0
		15-24	15,9	29,5	36,4	18,2	100,0
		oltre 24	22,7	18,2	9,1	50,0	100,0
		Casi non raggruppati	31,3	23,8	25,0	20,0	100,0
Cross-validato ^a	Conteggio	meno di 5	11	5	8	4	28
		5-14	12	18	13	14	57
		15-24	8	13	13	10	44
		oltre 24	6	5	3	8	22
		Casi non raggruppati	33	26	26	22	107
	%	meno di 5	39,3	17,9	28,6	14,3	100,0
		5-14	21,1	31,6	22,8	24,6	100,0
		15-24	18,2	29,5	29,5	22,7	100,0
		oltre 24	27,3	22,7	13,6	36,4	100,0
		Casi non raggruppati	33,1	23,8	25,0	20,0	100,0

a. La cross-validazione viene eseguita solo per tali casi nell'analisi. Per la cross-validazione, ogni caso viene classificato dalle funzioni derivate da tutti gli altri casi.

b. 41,1% di casi raggruppati originali classificati correttamente.

c. 33,1% di casi raggruppati cross-validati classificati correttamente.

1.h. Quali sono i gruppi che sono classificati meglio nella classificazione originale e in quella "cross-validata" ?

1.i. In generale, quale valutazione possiamo dare dei risultati delle due classificazioni riportate sopra?

ESERCIZIO 2. Vengono riportati di seguito i risultati di un'analisi discriminante.

Risultati del test

M di Box		37,396
F	Appross.	1,219
	df1	30
	df2	128147,3
	Sig.	,190

Test dell'ipotesi nulla che le matrici di covarianza siano uguali nella popolazione.

2.a. E' possibile accettare l'ipotesi nulla del test M di Box e perché ?

Autovalori

Funzione	Autovalore	% di varianza	% cumulata	Correlazione canonica
1	,171 ^a	96,7	96,7	,382
2	,006 ^a	3,3	100,0	,076

a. Per l'analisi sono state usate le prime 2 funzioni discriminanti canoniche.

2.b. Quale è la percentuale di varianza tra i gruppi spiegata da ogni funzione discriminante ?

2.c. Quale è la relazione tra ciascuna funzione discriminante e il gruppo di appartenenza dei soggetti?

Lambda di Wilks

Test di funzioni	Lambda di Wilks	Chi-quadrato	df	Sig.
Da 1 a 2	,849	68,952	10	,000
2	,994	2,473	4	,650

2.d. Quante sono le funzioni discriminanti statisticamente significative, e perché ?

Matrice di struttura			Coefficienti standardizzati della funzione discriminante canonica		
	Funzione			Funzione	
	1	2		1	2
FAP_1	,953*	-,126	SEFPA_1	,092	,553
FAS_1	,702*	,397	FAP_1	,805	-,336
PERF_1	,550*	-,484	PROS_1	-,214	,329
PROS_1	,233*	,112	FAS_1	,299	,698
SEFPA_1	,319	,490*	PERF_1	,078	-,769

Correlazioni comuni entro gruppi tra variabili discriminanti e funzioni discriminanti canoniche standardizzate
Variabili ordinate in base alla dimensione assoluta della correlazione entro la funzione.

*. Correlazione assoluta più grande tra ciascuna variabile e qualsiasi funzione discriminante

2.e. Quali sono le variabili che caratterizzano maggiormente ogni funzione discriminante ?

2.f. Quali sono le variabili che risultano più “ridondanti” per ogni funzione discriminante?

Risultati della classificazione^{b,c}

	SES 1	Gruppo di appartenenza previsto			Totali	
		LOW	MEDIUM	HIGH		
Originale	Conteggio	LOW	121	68	0	189
		MEDIUM	73	102	0	175
		HIGH	11	50	2	63
	%	LOW	64,0	36,0	,0	100,0
		MEDIUM	41,7	58,3	,0	100,0
		HIGH	17,5	79,4	3,2	100,0
Cross-validato ^a	Conteggio	LOW	117	72	0	189
		MEDIUM	73	101	1	175
		HIGH	12	49	2	63
	%	LOW	61,9	38,1	,0	100,0
		MEDIUM	41,7	57,7	,6	100,0
		HIGH	19,0	77,8	3,2	100,0

a. La cross-validazione viene eseguita solo per tali casi nell'analisi. Per la cross-validazione, ogni caso viene classificato dalle funzioni derivate da tutti gli altri casi.

b. 52,7% di casi raggruppati originali classificati correttamente.

c. 51,5% di casi raggruppati cross-validati classificati correttamente.

2.g. Quali sono i gruppi che sono classificati meglio nella classificazione originale e in quella “cross-validata” ?

2.i. In generale, quale valutazione possiamo dare dei risultati delle due classificazioni riportate sopra?

Statistiche per casi

Numero di caso	Gruppo di appartenenza	Gruppo previsto	Gruppo più alto			
			P(D>d G=g)		P(G=g D=d)	Distanza quadratica di Mahalanobis dal baricentro
			p	df		
Originale 1	3	3	,964	2	,437	,072
2	1	1	,531	2	,585	1,267
3	1	1	,388	2	,654	1,893
4	1	1	,765	2	,555	,537
5	1	1	,370	2	,652	1,987
6	2	1**	,674	2	,519	,788
7	1	3**	,938	2	,370	,128
8	1	1	,173	2	,697	3,508
9	2	1**	,866	2	,485	,288
10	1	2**	,914	2	,383	,180
11	2	3**	,760	2	,491	,550
12	1	1	,896	2	,491	,220
13	1	1	,384	2	,475	1,913
14	1	3**	,957	2	,446	,087
15	1	1	,518	2	,519	1,317
16	1	3**	,437	2	,602	1,655
17	2	1**	,281	2	,592	2,540
18	1	1	,947	2	,442	,108
19	1	2**	,114	2	,437	4,343
20	1	1	,058	2	,758	5,686

** . Caso malclassificato

2.h. Quali è il gruppo di appartenenza e quale il gruppo previsto del caso numero 11?

2.i. Quali è la probabilità che il caso numero 11 venga attribuito al gruppo in cui viene riclassificato?

2.l. Quali sono i 2 casi più vicini e i 2 casi più lontani dai relativi gruppi di appartenenza?

ESERCIZIO 3. Vengono riportati di seguito i risultati di un'analisi multivariata della varianza (MANOVA)

3.a. E' possibile accettare l'ipotesi nulla del test M di Box e perché ?

Test di Box di uguaglianza delle matrici di covarianza ^a

M di Box	40,713
F	1,960
df1	20
df2	20179,380
Sig.	,006

Verifica l'ipotesi nulla per la quale le matrici di covarianza osservate delle variabili dipendenti sono uguali nei gruppi.

a. Disegno: Intercept+TITS

3.b. E' possibile rifiutare l'ipotesi nulla relativa al test multivariato del fattore TITS e perché ?

Test multivariat^d

Effetto		Valore	F	Ipotesi df	Gradi di libertà dell'errore	Sig.	Eta quadrato parziale	Non centralità	Potenza osservata ^a
Intercept	Traccia di Pillai	,987	6335,197 ^b	4,000	333,000	,000	,987	25340,790	1,000
	Lambda di Wilks	,013	6335,197 ^b	4,000	333,000	,000	,987	25340,790	1,000
	Traccia di Hotelling	76,098	6335,197 ^b	4,000	333,000	,000	,987	25340,790	1,000
	Radice di Roy	76,098	6335,197 ^b	4,000	333,000	,000	,987	25340,790	1,000
TITS	Traccia di Pillai	,101	4,437	8,000	668,000	,000	,050	35,495	,997
	Lambda di Wilks	,902	4,424 ^b	8,000	666,000	,000	,050	35,395	,997
	Traccia di Hotelling	,106	4,412	8,000	664,000	,000	,050	35,296	,996
	Radice di Roy	,058	4,828 ^c	4,000	334,000	,001	,055	19,314	,955

a. Calcolato usando alfa = ,05

b. Statistica esatta

c. La statistica è un limite superiore per F che fornisce un limite inferiore per il livello di significatività.

d. Disegno: Intercept+TITS

3.c. E' possibile concludere che l'assunzione di omoschedasticità viene rispettata ?

Test di Levene di uguaglianza delle varianze dell'errore ^a

	F	df1	df2	Sig.
MSQ_F1	2,665	2	336	,071
MSQ_F2	,280	2	336	,756
PIER_F1	2,193	2	336	,113
PIER_F2	1,253	2	336	,287

Verifica l'ipotesi nulla per la quale la varianza dell'errore della variabile dipendente è uguale tra i gruppi.

a. Disegno: Intercept+TITS

3.d. Individuare gli effetti univariati che risultano statisticamente significativi, specificando anche il valore dell'ampiezza dell'effetto e della potenza della verifica degli effetti significativi.

Test degli effetti fra soggetti

Sorgente	Variabile dipendente	Somma dei quadrati Tipo III	df	Media dei quadrati	F	Sig.	Eta quadrato parziale	Non centralità	Potenza osservata ^a
Modello corretto	MSQ_F1	38,241 ^b	2	19,121	,585	,558	,003	1,171	,147
	MSQ_F2	65,178 ^c	2	32,589	,891	,411	,005	1,782	,203
	PIER_F1	40,402 ^d	2	20,201	8,871	,000	,050	17,743	,971
	PIER_F2	26,737 ^e	2	13,368	7,511	,001	,043	15,023	,942
Intercept	MSQ_F1	176694,780	1	176694,8	5408,733	,000	,942	5408,733	1,000
	MSQ_F2	92347,600	1	92347,600	2524,276	,000	,883	2524,276	1,000
	PIER_F1	6598,142	1	6598,142	2897,613	,000	,896	2897,613	1,000
	PIER_F2	6037,376	1	6037,376	3392,254	,000	,910	3392,254	1,000
TITS	MSQ_F1	38,241	2	19,121	,585	,558	,003	1,171	,147
	MSQ_F2	65,178	2	32,589	,891	,411	,005	1,782	,203
	PIER_F1	40,402	2	20,201	8,871	,000	,050	17,743	,971
	PIER_F2	26,737	2	13,368	7,511	,001	,043	15,023	,942
Errore	MSQ_F1	10976,591	336	32,668					
	MSQ_F2	12292,156	336	36,584					
	PIER_F1	765,104	336	2,277					
	PIER_F2	597,997	336	1,780					
Totale	MSQ_F1	348996,000	339						
	MSQ_F2	186528,000	339						
	PIER_F1	12331,202	339						
	PIER_F2	13185,707	339						
Totale corretto	MSQ_F1	11014,832	338						
	MSQ_F2	12357,333	338						
	PIER_F1	805,506	338						
	PIER_F2	624,734	338						

a. Calcolato usando alfa = ,05

b. R quadrato = ,003 (R quadrato corretto = -,002)

c. R quadrato = ,005 (R quadrato corretto = -,001)

d. R quadrato = ,050 (R quadrato corretto = ,045)

e. R quadrato = ,043 (R quadrato corretto = ,037)

3.e. Considerando il confronto pianificato che esamina la differenza tra il primo e il secondo gruppo commentare i risultati dei test multivariati e del test univariato.

Risultati di test multivariati

	Valore	F	Ipotesi df	Gradi di libertà dell'errore	Sig.	Eta quadrato parziale	Non centralità	Potenza osservata ^a
Traccia di Pillai	,052	4,554 ^b	4,000	333,000	,001	,052	18,215	,943
Lambda di Wilks	,948	4,554 ^b	4,000	333,000	,001	,052	18,215	,943
Traccia di Hotelling	,055	4,554 ^b	4,000	333,000	,001	,052	18,215	,943
Radice di Roy	,055	4,554 ^b	4,000	333,000	,001	,052	18,215	,943

a. Calcolato usando alfa = ,05

b. Statistica esatta

Risultati del test univariato

Sorgente	Variabile dipendente	Somma dei quadrati	df	Media dei quadrati	F	Sig.	Eta quadrato parziale	Non centralità	Potenza osservata ^a
Contrasto	MSQ_F1	38,187	1	38,187	1,169	,280	,003	1,169	,190
	MSQ_F2	59,471	1	59,471	1,626	,203	,005	1,626	,246
	PIER_F1	40,308	1	40,308	17,702	,000	,050	17,702	,987
	PIER_F2	10,445	1	10,445	5,869	,016	,017	5,869	,676
Errore	MSQ_F1	10976,591	336	32,668					
	MSQ_F2	12292,156	336	36,584					
	PIER_F1	765,104	336	2,277					
	PIER_F2	597,997	336	1,780					

a. Calcolato usando alfa = ,05

SOLUZIONI

ESERCIZIO 1

1.a. Quali sono le variabili che presentano medie significative nei gruppi ?

Si tratta di esaminare nella tabella “Test di uguaglianza delle medie di gruppo” i valori che nella colonna Sig. risultano inferiori a .05. Nella tabella in esame ce ne sono 3, quelli relativi alle variabili: COIN_LAV, IMPEGNO, SODDISF.

1.b. E' possibile accettare l'ipotesi nulla del test M di Box e perché ?

Non è possibile accettare questa ipotesi perché la probabilità (Sig.) associata al test è minore di .05: quindi le matrici di varianza/covarianza dei gruppi non sono omogenee.

1.c. Quale è la percentuale di varianza tra i gruppi spiegata da ogni funzione discriminante ?

Bisogna guardare i valori nella colonna “% di varianza” che sono pari a 54.5 per la prima funzione, 29.4 per la seconda e 16.1 per la terza funzione discriminante.

1.d. Quale è la relazione tra ciascuna funzione discriminante e il gruppo di appartenenza dei soggetti ?

Si tratta di un valore ricavabile dalla correlazione canonica, che rappresenta l'associazione tra la funzione discriminante e le variabili dummy che specificano l'appartenenza del soggetto ad uno dei gruppi definiti dalla variabile dipendente. Essa è uguale a .354 per la prima funzione discriminante, a .268 per la seconda e a .202 per la terza.

1.e. Quante sono le funzioni discriminanti statisticamente significative, e perché ?

Bisogna esaminare i valori nella colonna Sig. relativa alla tabella dei coefficienti Lambda di Wilks. Poiché il valore associato alla riga “Da 1 a 3” è significativo, mentre quelli associati alle righe “Da 2 a 3” e “3” non sono significativi, solo la prima funzione discriminante risulta statisticamente significativa (per l'interpretazione di questi indici si veda il testo a pagina 185).

1.f. Quali sono le variabili che caratterizzano maggiormente ogni funzione discriminante ?

Bisogna esaminare le correlazioni tra variabili e funzioni contenute nella “Matrice di struttura”, avendo in mente che solo la prima funzione discriminante risulta statisticamente significativa, dunque solo la prima va interpretata. Considerando come valore “soglia” per l'interpretazione un coefficiente di .30 (vedi Barbaranelli, 2003), le variabili che caratterizzano maggiormente l'unica funzione significativa sono: IMPEGNO, SODDISF, COIN_LAV, COLL_EFF.

1.g. Quali sono le variabili che risultano più “ridondanti” per ogni funzione discriminante?

In questo caso bisogna esaminare i coefficienti standardizzati della funzione discriminante canonica i quali, essendo coefficienti parziali, tengono conto della correlazione tra le variabili indipendenti, e dunque della eventuale ridondanza tra di esse. Non esiste un criterio assoluto per identificare i coefficienti ridondanti, ma i valori vanno confrontati all'interno di ogni colonna che individua una funzione. Nel caso della prima funzione (che è l'unica ad essere significativa, e quindi ad essere interpretata) i valori più bassi sono ottenuti da SODDISF e COIN_LAV, che risultano avere un'influenza praticamente nulla quando si considerano le loro correlazioni con le altre variabili, a fronte di correlazioni non parziali molto elevate (si veda la matrice di struttura). Verosimilmente questo dipende da una probabile elevata correlazione di queste variabili con la variabile IMPEGNO, che infatti risulta estremamente associata alla funzione.

1.h. Quali sono i gruppi che sono classificati meglio nella classificazione originale e in quella “cross-validata” ?

Bisogna esaminare gli elementi sulla diagonale principale delle due matrici (Originale e Cross-validato), ignorando ovviamente la riga relativa ai “Casi non raggruppati”. Di fatto, il primo e l’ultimo gruppo presentano la percentuale più elevata di casi in cui la classificazione osservata (“Conteggio”) coincide con quella ottenuta tramite i risultati dell’analisi discriminante (“Gruppo di appartenenza previsto”).

1.i. In generale, quale valutazione possiamo dare dei risultati delle due classificazioni riportate sopra?

Bisogna applicare le statistiche z di Huberty e IOCC (vedi Barbaranelli, 2003). Per la classificazione Originale l’indice zH risulta pari a 4.56 che è statisticamente significativo ($p < .01$), ma l’indice IOCC è molto basso, essendo pari a .21. Quindi, pur essendo significativo il miglioramento rispetto ad una classificazione fatta semplicemente a caso, tale miglioramento è mediocre. Per la classificazione cross-validata la situazione è anche peggiore, poiché l’indice zH è pari a 2.30 ($p < .05$), ma l’IOCC è appena .11. Questi calcoli sono stati effettuati tramite il foglio excel “CALCOLI_zh_iocc.xls” presente nel materiale on line.

ESERCIZIO 2

2.a. E’ possibile accettare l’ipotesi nulla del test M di Box e perché ?

Poiché la probabilità (Sig.) associata al test è maggiore di .05, è possibile accettare questa ipotesi: quindi le matrici di varianza/covarianza dei gruppi sono omogenee.

2.b. Quale è la percentuale di varianza tra i gruppi spiegata da ogni funzione discriminante ?

Bisogna guardare i valori nella colonna “% di varianza” che sono pari a 96.7 per la prima funzione, e a 3.3 per la seconda.

2.c. Quale è la relazione tra ciascuna funzione discriminante e il gruppo di appartenenza dei soggetti ?

Si tratta di un valore ricavabile dalla correlazione canonica, che rappresenta l’associazione tra la funzione discriminante e le variabili dummy che specificano l’appartenenza del soggetto ad uno dei gruppi definiti dalla variabile dipendente. Essa è uguale a .382 per la prima funzione discriminante, e a .076 per la seconda.

2.d. Quante sono le funzioni discriminanti statisticamente significative, e perché ?

Bisogna esaminare i valori nella colonna Sig. relativa alla tabella dei coefficienti Lambda di Wilks. Poiché il valore associato alla riga “Da 1 a 2” è significativo, mentre quello associato alla riga “2” e non lo è, solo la prima funzione discriminante risulta statisticamente significativa.

2.e. Quali sono le variabili che caratterizzano maggiormente ogni funzione discriminante ?

Bisogna esaminare le correlazioni tra variabili e funzioni contenute nella “Matrice di struttura”, avendo in mente che solo la prima funzione discriminante risulta statisticamente significativa, dunque solo la prima va interpretata. Considerando come valore “soglia” per l’interpretazione un coefficiente di .30, le variabili che caratterizzano maggiormente l’unica funzione significativa sono: FAP_1, FAS_1, PERF_1, SEFPA_1.

2.f. Quali sono le variabili che risultano più “ridondanti” per ogni funzione discriminante?

In questo caso bisogna esaminare i coefficienti standardizzati della funzione discriminante canonica i quali, essendo coefficienti parziali, tengono conto della correlazione tra le variabili indipendenti, e dunque della eventuale ridondanza tra di esse. Come abbiamo osservato prima, i valori vanno confrontati all’interno di ogni colonna che individua una funzione poiché non esiste un criterio

assoluto per identificare i coefficienti ridondanti. Nel caso della prima funzione (che è l'unica ad essere significativa, e quindi ad essere interpretata) i valori più bassi sono ottenuti da SEFPA_1 e PERF_1, che risultano avere un'influenza praticamente nulla quando si considerano le loro correlazioni con le altre variabili, a fronte di correlazioni non parziali adeguate (si veda la matrice di struttura). Verosimilmente questo dipende da una probabile elevata correlazione di queste variabili con le variabili FAS_1 e FAP_1.

2.g. Quali sono i gruppi che sono classificati meglio nella classificazione originale e in quella "cross-validata" ?

Bisogna esaminare gli elementi sulla diagonale principale delle due matrici (Originale e Cross-validato). Di fatto, il primo e il secondo gruppo presentano la percentuale più elevata di casi per i quali la classificazione osservata ("Conteggio") coincide con quella ottenuta tramite i risultati dell'analisi discriminante ("Gruppo di appartenenza previsto"). L'ultimo gruppo invece presenta una bassissima percentuale di casi classificati correttamente.

2.h. In generale, quale valutazione possiamo dare dei risultati delle due classificazioni riportate sopra?

Anche in questo caso, tramite il foglio excel "CALCOLI_zh_iocc.xls", abbiamo calcolato gli indici z di Huberty e IOCC (vedi Barbaranelli, 2003). Per la classificazione Originale l'indice zH risulta pari a 13.22 che è statisticamente significativo ($p < .001$), ma l'indice IOCC è piuttosto basso, essendo pari a .37. Quindi, pur essendo significativo il miglioramento rispetto ad una classificazione fatta semplicemente a caso, tale miglioramento è modesto. Per la classificazione cross-validata la situazione è lievemente peggiore, poiché l'indice zH è pari a 12.66 ($p < .001$), e l'IOCC è di .35. Verosimilmente questo risultato dipende dal fatto che i soggetti del terzo gruppo sono classificati in modo quasi completamente errato dalla funzione discriminante.

2.i. Quali è il gruppo di appartenenza e quale il gruppo previsto del caso numero 11?

Bisogna considerare le due colonne omonime nella tabella "Statistiche per casi", quindi il gruppo di appartenenza è il 2 mentre il gruppo previsto è il 3.

2.l. Quali è la probabilità che il caso numero 11 venga attribuito al gruppo in cui viene riclassificato?

Bisogna esaminare la probabilità a posteriori che è riportata nella colonna " $P(G=glD=d)$ " e che in questo caso è uguale a .491.

2.m. Quali sono i 2 casi più vicini e i 2 casi più lontani dai relativi gruppi di appartenenza?

Bisogna considerare l'ultima colonna, nella quale sono riportati i valori della distanza di Mahalanobis del caso dal centroide del gruppo di appartenenza. I due casi più vicini al centroide sono quelli con distanza più piccola (il 1 e il 14), mentre i due più lontani dal centroide sono quelli con distanza più elevata (il 19 e il 20).

ESERCIZIO 3

3.a. E' possibile accettare l'ipotesi nulla del test M di Box e perché ?

Non è possibile accettare questa ipotesi perché la probabilità (Sig.) associata al test è minore di .05: quindi le matrici di varianza/covarianza dei gruppi non sono omogenee.

3.b. E' possibile rifiutare l'ipotesi nulla relativa al test multivariato del fattore TITS e perché?

E' possibile rifiutare l'ipotesi nulla perché la probabilità (Sig.) associata al test è minore di .05: quindi i gruppi risultano significativamente differenti rispetto alla combinazione lineare delle variabili dipendenti utilizzata per il test multivariato.

3.c. E' possibile concludere che l'assunzione di omoschedasticità viene rispettata ?

L'assunzione di omogeneità delle varianze di ciascuna singola variabile dipendente è verificata poiché il test di Levene presenta una probabilità maggiore di .05 per ognuna di esse.

3.d. Individuare gli effetti univariati che risultano statisticamente significativi, specificando anche il valore dell'ampiezza dell'effetto e della potenza della verifica degli effetti significativi.

Per quanto riguarda la significatività dell'effetto, si tratta di esaminare i risultati nella colonna Sig. relativi alla Sorgente TITS (si veda quanto scritto nel documento relativo alla ANOVA per l'interpretazione dei test fra i soggetti). In particolare mentre le prime due variabili (MSQ_F1 e MSQ_F2) non risultano significative ($p > .05$) le ultime 2 (PIER_F1 e PIER_F2) invece lo sono ($p < .05$). L'ampiezza dell'effetto è pari a .05 e .043 rispettivamente per PIER_F1 e PIER_F2, mentre la potenza osservata è uguale rispettivamente a .971 e a .942.

3.e. Considerando il confronto pianificato che esamina la differenza tra il primo e il secondo gruppo commentare i risultati dei test multivariati e del test univariato.

Per quanto riguarda i test multivariati, la differenza tra i due gruppi risulta significativa ($p < .001$), con un'ampiezza dell'effetto di .052 e una potenza pari a .943. Considerando invece il test univariato, non tutte le differenze tra i due gruppi nelle singole variabili dipendente risultano statisticamente significative. In particolare solamente per le due variabili PIER_F1 e PIER_F2 si ha un effetto significativo che si associa ad un'ampiezza rispettivamente pari a .050 e .017 e a una potenza della verifica rispettivamente di .987 e di .976.