

Esercizi sull'ANOVA

Andrea Onofri

22 febbraio 2015

Questa dispensa contiene gli esercizi e le domande che costituiscono l'oggetto della prova di esame. In genere, ad ogni studente viene assegnato un problema relativo all'ANOVA (parti 1 e 5) ed un problema relativo alla modellizzazione (parti seguenti). Salvo diversamente specificato, i dataset sono contenuti nel file Esercizi.xls, disponibile nel sito web del corso.

1 CRD e confronto multiplo

ESERCIZIO n. 1

Un esperimento a randomizzazione completa relativo ad una prova varietale di frumento ha l'obiettivo di porre a confronto la produzione di 5 varietà. Le produzioni (in bushels per acre) osservate siano le seguenti:

Variety	Replicate		
	1	2	3
A	32.4	34.3	37.3
B	20.2	27.5	25.9
C	29.2	27.8	30.2
D	12.8	12.3	14.8
E	21.7	24.5	23.4

Eseguire l'ANOVA, presentare i risultati e commentarli (esempio tratto da Le Clerg *et al.*, 1962)

ESERCIZIO n. 2

Culture di tessuto di pomodoro sono state allevate su capsule Petri trattate con una diversa concentrazione di zuccheri, utilizzando cinque repliche. La crescita colturale è riportata in tabella

Control	Glucose	Fructose	Sucrose
45	25	28	31
39	28	31	37
40	30	24	35
45	29	28	33
42	33	27	34

Calcolare le medie ed eseguire l'ANOVA. Eseguire i test di confronto multiplo con i tre metodi noti (MDS, HSD, Dunnet). Commentare i risultati.

ESERCIZIO n. 3

In un ospedale, è stata misurata la concentrazione di colesterolo nel sangue di otto pazienti. Per ogni paziente, sono stati analizzati due campioni, ottenendo le seguenti concentrazioni:

Paziente	Misura 1	Misura 2
1	167.3	166.7
2	186.7	184.2
3	105.0	107.9
4	214.5	215.3
5	148.5	149.5
6	171.5	167.3
7	161.5	159.4
8	243.6	245.5

Chi sono i pazienti che hanno il colesterolo significativamente più alto? Commentare i risultati.

ESERCIZIO n. 4

E' stato impostato un test di durata su un impianto di riscaldamento, per verificare come la temperatura di esercizio influenza la durata del riscaldatore. Sono state testate 4 temperature, con sei repliche e, per ciascun riscaldatore, è stato rilevato il numero di ore prima della rottura. I risultati sono i seguenti:

Temp.	Hours to failure
1520	1953
1520	2135
1520	2471
1520	4727
1520	6134
1520	6314
1620	1190
1620	1286
1620	1550
1620	2125
1620	2557
1620	2845
1660	651
1660	837
1660	848
1660	1038
1660	1361
1660	1543
1708	511
1708	651
1708	651
1708	652
1708	688
1708	729

ESERCIZIO n. 5

Un entomologo ha contato il numero di uova deposte da un lepidottero sulle foglie di tre varietà di tabacco, valutando 15 femmine per varietà. I risultati sono i seguenti:

Female	Field	Resistant	USDA
1	211	0	448
2	276	9	906
3	415	143	28
4	787	1	277
5	18	26	634
6	118	127	48
7	1	161	369
8	151	294	137
9	0	0	29
10	253	348	522
11	61	0	319
12	0	14	242
13	275	21	261
14	0	0	566
15	153	218	734

Eseguite l'ANOVA. Quali sono le assunzioni necessarie per l'ANOVA? Sono rispettate? Vi sono outliers? Calcolate SEM e SED in modo attendibile.

2 Disegni a blocchi randomizzati

ESERCIZIO n. 1

E' stato impostato un esperimento a blocchi randomizzati per confrontare sei tipi di irrigazione, in un aranceto della Spagna. I risultati sono i seguenti (in pounds per parcella):

Metodo	1	2	3	4	5
Goccia	438	413	375	127	320
Conche	413	398	348	112	297
Aspersione	346	334	281	43	231
Aspersione+goccia	335	321	267	33	219
Sommersione	403	380	336	101	293

Eseguire l'ANOVA. Quali sono le assunzioni necessarie per l'ANOVA? Sono rispettate? Calcolate SEM e SED ed eseguite il confronto multiplo. Qual è il metodo di irrigazione migliore?

ESERCIZIO n. 2

E' stato impostato un esperimento di fertilizzazione secondo uno schema a blocchi randomizzati. I dati ottenuti sono i contenuti percentuali (moltiplicati per 100) in fosforo, in un campione di tessuti vegetali prelevato per parcella:

Trattamento	1	2	3	4	5
Non fertilizzato	5.6	6.1	5.3	5.9	7.4
50 lb N	7.3	NA	7.7	7.7	7.0
100 lb N	6.9	6	5.6	7.4	8.2
50 lb N + 75 lb P ₂ O ₅	10.8	11.2	8.8	12.9	10.4
100 lb N + 75 lb P ₂ O ₅	9.6	9.3	12	10.6	11.6

Eseguire l'ANOVA, considerando il dato mancante. Calcolare SEM e SED. Qual è il trattamento migliore? Aumentare il dosaggio di N senza P₂O₅ è conveniente? E in presenza di P₂O₅?

ESERCIZIO n. 3

Un ingegnere ha studiato l'efficienza di quattro metodi costruttivi (A B C D) per un componente elettronico, in termini di tempo impiegato per la costruzione. Sono stati selezionati cinque tecnici che costruiscono il componente elettronico con i cinque metodi, in ordine random. Tuttavia, è noto che i tecnici si stancano e divengono via via meno efficienti al passare del tempo. Ci sono quindi due fonti di variabilità note, il tecnico (che fa da blocco) e il turno (cioè la posizione d'ordine con cui ogni tecnico usa i quattro metodi di costruzione). Vi sono quindi due effetti di blocco (disegno a quadrato latino). I risultati sono:

Tecnico				
Turno	1	2	3	4
1	90 (C)	96 (D)	84 (A)	88 (B)
2	90 (B)	91 (C)	96 (D)	88 (A)
3	89 (A)	97 (B)	98 (C)	98 (D)
4	104 (D)	100 (A)	104 (B)	106 (C)

Eseguite l'ANOVA. Quali sono le assunzioni necessarie per l'ANOVA? Sono rispettate? Calcolate SEM e SED ed eseguite il confronto multiplo. Qual è il metodo di costruzione più veloce? Qual è il più lento. Se C è il metodo di riferimento, quali sono i metodi migliori di C?

3 ANOVA a due vie

ESERCIZIO n. 1

La biologia di *Sorghum halepense* da rizoma mostra che il peso dei rizomi raggiunge un minimo intorno alla quarta foglia. Di conseguenza, eseguire un trattamento in quest'epoca dovrebbe minimizzare le possibilità di ripresa degli individui trattati, portando anche ad un certo risanamento del terreno. Tuttavia, ci si attende che gli effetti siano maggiori quando le piante provengono da rizomi più piccoli, con un minor contenuto di sostanze di riserva. Per affrontare questi argomenti è stata organizzata una prova in vaso, secondo un disegno a randomizzazione completa con quattro repliche. I risultati sono i seguenti (file: JohnsonGrass.xls):

Rizome Length	Timing (no. of leaves)					Untreated
	2-3	4-5	6-7	8-9	3-4/8-9	
2-nodes	34.03	0.10	30.91	33.21	2.89	41.63
	22.31	6.08	35.34	43.44	19.06	22.96
	21.70	3.73	24.23	44.06	0.10	52.14
	14.90	9.15	28.27	35.34	0.68	59.81
4-nodes	42.19	14.86	52.34	39.06	8.62	68.15
	51.06	36.03	43.17	61.59	0.05	42.75
	43.77	21.85	57.28	48.89	0.10	57.77
	31.74	8.71	29.71	49.14	9.65	44.85
6-nodes	20.84	11.37	55.00	41.77	9.80	43.20
	26.12	2.24	28.46	37.38	0.10	40.68
	35.24	14.17	21.81	39.55	1.42	34.11
	13.32	23.93	60.72	48.37	6.83	32.21

Eseguite l'ANOVA. Verificate il rispetto delle assunzioni parametriche di base e, se necessario, trasformate i dati. Preparate una tabella per le medie marginali e le medie di cella ed aggiungete i rispettivi errori standard (SEMs). Ha senso considerare le medie marginali? Impostate un test di confronto multiplo per gli effetti significativi, coerentemente con la risposta alla domanda precedente.

ESERCIZIO n. 2

Un agronomo ha organizzato un confronto varietale in favino, considerando due epoche di semina: autunnale e primaverile. E' stato utilizzato un disegno a blocchi randomizzati e a parcella suddivisa, con le epoche di semina nelle parcelle principali e le varietà nelle sub-parcelle. I risultati sono (file: fabaBean.xls):

Genotype	autumn sowing				spring sowing			
	1	2	3	4	1	2	3	4
Chiaro	4.36	4.00	4.23	3.83	2.76	2.64	2.25	2.38
Collameno	3.01	3.32	3.27	3.40	2.50	1.79	1.57	1.77
Palombino	3.85	3.85	3.68	3.98	2.24	2.21	2.50	2.05
Scuro	4.97	3.98	4.39	4.14	3.45	2.94	3.12	2.69
Sicania	4.38	4.01	3.94	2.99	3.24	3.60	3.16	3.08
Vesuvio	3.94	4.47	3.93	4.21	2.34	2.44	1.71	2.00

Eseguite l'ANOVA. Verificate il rispetto delle assunzioni parametriche di base e, se necessario, trasformate i dati. Preparate una tabella per le medie marginali e le medie di cella ed aggiungete i rispettivi errori standard (SEMs). Ha senso considerare le medie marginali? Impostate un test di confronto multiplo per gli effetti significativi, coerentemente con la risposta alla domanda precedente.

ESERCIZIO n. 3

Gli erbicidi mostrano sempre un certo grado di persistenza nel terreno. Di conseguenza, se la coltura fallisce subito dopo il diserbo, la scelta delle colture di sostituzione può essere condizionata dal diserbo già eseguito. Per questo motivo, è stato impostato un esperimento di pieno campo volto a valutare se tre erbicidi del mais (rimsulfuron, imazethapyr and primisulfuron) erano in grado di danneggiare quattro colture (soia, girasole, rapa e sorgo) seminate 20 giorni dopo il trattamento. Gli erbicidi sono stati distribuiti su terreno nudo, seguendo un disegno a blocchi randomizzati, su parcelle di elevate dimensioni. Per ogni blocco, la semina è stata eseguita su strisce trasversali, perpendicolari ai trattamenti eseguiti (schema a strip-plot). I risultati sono (file: recrop.xls):

Herbicides	Block	Crop (grams per square meter)			
		sorghum	rape	soyabean	sunflower
Untreated	1	180	157	199	201
	2	236	111	257	358
	3	287	217	346	435
	4	350	170	211	327
Imazethapyr	1	47	10	193	51
	2	43	1	113	4
	3	0	20	187	13
	4	3	21	122	15
primisulfuron	1	271	8	335	379
	2	182	0	201	201
	3	283	22	206	307
	4	147	24	240	337
rimsulfuron	1	403	238	226	290
	2	227	169	195	494
	3	400	364	257	397
	4	171	134	137	180

Eseguite l'ANOVA. Verificate il rispetto delle assunzioni parametriche di base e, se necessario, trasformate i dati. Preparate una tabella per le medie marginali e le medie di cella ed aggiungete i rispettivi errori standard (SEMs). Ha senso considerare le medie marginali?

Impostate un test di confronto multiplo per gli effetti significativi, coerentemente con la risposta alla domanda precedente.

4 Sottocampionamento, esperimenti ripetuti e misure ripetute

ESERCIZIO n. 1

E' stato eseguito un confronto tra sei varietà di frumento duro, secondo uno schema sperimentale a blocchi randomizzati. Al termine dell'esperimento, sono stati prelevati due sottocampioni di granella per parcella e sono stati sottoposti alla determinazione del peso ettolitrico. I risultati ottenuti sono riportati in tabella. Elaborare correttamente i risultati dell'esperimento e stabilire se vi è una differenza significativa tra le varietà. Se appropriato, preparare una tabella riassuntiva delle medie, insieme alle lettere di significatività.

Varietà	Blocco	Peso ett. 1	Peso ett. 2
Achille	1	78.2	77.7
Achille	2	75.0	75.5
Achille	3	74.1	74.1
Alemanno	1	74.8	74.8
Alemanno	2	77.7	77.5
Alemanno	3	79.0	78.6
Anco Marzio	1	73.4	73.9
Anco Marzio	2	79.3	79.0
Anco Marzio	3	78.8	79.0
Arnacoris	1	64.2	64.0
Arnacoris	2	62.0	61.8
Arnacoris	3	62.1	62.5
Casanova	1	72.8	73.0
Casanova	2	74.8	74.3
Casanova	3	67.6	67.6
Chiara	1	66.7	66.9
Chiara	2	62.3	62.3
Chiara	3	69.4	69.4

ESERCIZIO n. 2

Un esperimento varietale di frumento duro è stato ripetuto in sette anni con lo stesso disegno sperimentale a blocchi randomizzati. Il dataset è contenuto nel file WinterWheat.xls. Eseguire l'ANOVA cumulata, stabilire se l'interazione genotipo-ambiente è significativa. Commentare i risultati. E' lecito considerare l'effetto ambiente come 'random'? In questo caso, ripetere l'ANOVA ed eseguire un test di confronto multiplo tra le varietà. Quale è la varietà più affidabile? Quali varietà non sono significativamente diverse dalla migliore? Calcolare un indice di stabilità produttiva e commentarlo.

ESERCIZIO n. 3

Un esperimento varietale di erba medica è stato seguito per tre anni consecutivi. Il dataset è contenuto nel file AlfaAlfa.xls. Eseguire l'ANOVA cumulata, stabilire se l'interazione

genotipo-ambiente è significativa. Commentare i risultati. E' possibile redigere una graduatoria media delle varietà nei diversi anni? Come? Quale è la varietà più affidabile? Quali varietà non sono significativamente diverse dalla migliore?

5 Altri esercizi di riepilogo

Per tutti gli esercizi, elaborare i risultati nel modo più opportuno e commentarli.

ESERCIZIO n. 1

Un orticoltore ha organizzato un esperimento per valutare l'effetto della concimazione azotata sulla lattuga, utilizzando un disegno a blocco randomizzato.

N level	B1	B2	B3	B4
0	104	114	109	124
50	134	130	154	164
100	146	142	152	156
150	147	160	160	163
200	133	146	156	161

(da Kuhel, modificato)

ESERCIZIO n. 2

E' stato condotto un esperimento parcellare per valutare l'interazione tra il momento dell'applicazione dell'azoto al terreno (early, optimum, late) e due livelli di un inibitore della nitrificazione (none, 5 lb/acre). L'inibitore ritarda la nitrificazione e riduce le perdite per lisciviazione profonda. L'azoto è stato somministrato in forma marcata (^{15}N) e i dati raccolti riguardano la percentuale di azoto assorbito dalla pianta.

Blocco	None			5 lb/acre		
	Early	Opt	Late	Early	Opt	Late
1	21.4	50.8	53.2	54.8	56.9	57.7
2	11.3	42.7	44.8	47.9	46.8	54.0
3	34.9	61.8	57.8	40.1	57.9	62.0

ESERCIZIO n. 3

Un ricercatore ha condotto un esperimento di nutrizione del bestiame con quattro trattamenti consistenti in diverse tipologie di acque di abbeveramento. L'esperimento è stato condotto in due stagioni diverse (primavera e inverno) ed è stato ripetuto in due anni. I dati seguenti rappresentano gli incrementi medi di peso corporeo in 112 giorni. Considerate trattamento e stagione come effetti fissi e l'anno come effetto random e considerate anche due blocchi, consistenti in due stalle diverse.

Trattamento	2010		2011	
	SPRING	WINT	SPRING	WINT
A	1.81	2.14	2.06	2.17
	1.88	2.32	1.91	2.55
B	1.77	2.27	1.57	2.06
	1.60	2.02	1.32	2.20
C	1.85	2.13	1.51	2.25
	1.59	1.93	1.49	1.94
D	1.51	1.85	1.31	1.83
	1.56	1.95	1.20	2.15

ESERCIZIO n. 4

E' stato organizzato un esperimento per valutare l'effetto della temperatura di lavaggio sulla riduzione di lunghezza di alcuni tessuti. I risultati sono espressi in percentuale di riduzione e sono stati ottenuti in un disegno sperimentale a randomizzazione completa, con quattro tessuti e altrettante temperature.

Tessuto	210°F	215°F	220°F	225°F
A	1.8	2.0	4.6	7.5
	2.1	2.1	5.0	7.9
B	2.2	4.2	5.4	9.8
	2.4	4.0	5.6	9.2
C	2.8	4.4	8.7	13.2
	3.2	4.8	8.4	13.0
D	3.2	3.3	5.7	10.9
	3.6	3.5	5.8	11.1

ESERCIZIO n. 5

Un processo di sintesi chimica prevede due reazioni, la prima richiede un alcohol e la seconda richiede una base. Viene impostato un esperimento fattoriale 3 x 2, con tre alcohols e due basi, con uno schema sperimentale completamente randomizzato a quattro repliche. Quali sono le vostre raccomandazioni per la prima e la seconda reazione, sulla base dei risultati dell'esperimento. La variabile rilevata mostra la produzione percentuale del processo

Base	Alcohol					
	C		D		E	
A	91.3	89.9	89.3	88.1	89.5	87.6
	90.7	91.4	90.4	91.4	88.3	90.3
B	87.3	89.4	92.3	91.5	93.1	90.7
	91.5	88.3	90.6	94.7	91.5	89.8

ESERCIZIO n. 6

E' stato disegnato un esperimento a quadrato latino per valutare l'effetto della concimazione sulla produzione della lattuga. Sono stati utilizzati quattro tipi di concimi con quattro repliche.

Concime	Riga	Colonna	Produzione
A	1	1	104
B	1	2	114
C	1	3	90
D	1	4	140
A	2	1	134
B	2	2	130
C	2	3	144
D	2	4	174
A	3	1	146
B	3	2	142
C	3	3	152
D	3	4	156
A	4	1	147
B	4	2	160
C	4	3	160
D	4	4	163