Contrasti e confronti multipli

Andrea Onofri

25 gennaio 2012

Indice

Sommario

Scopo di questa lezione è illustrare la metodica del confronto multiplo, evidenziandone pregi, difetti e, soprattutto, limiti di impiego.

---

1 Introduzione

Obiettivi

- Se il test F nell'ANOVA è significativo l'effetto del trattamento oggetto di studio (nel suo complesso) è maggiore dell'errore
- Questo non implica una graduatoria di merito tra i diversi livelli. Chi è superiore a chi?
- Rimane quindi da stabilire una graduatoria ed eventualmente confrontare tra di loro i risultati ottenuti da alcuni dei livelli in prova.
- A questa esigenza rispondono i contrasti e i le procedure di confronto multiplo (MCP)

Ammettiamo di aver effettuato una prova con un trattamento sperimentali caratterizzato da quattro livelli qualitativi (tesi di concimazione), con i risultati riportati nella tabella seguente.

Ci chiediamo: quale concimazione è più efficace?

Per rispondere a questa domanda, il primo *step* è, in genere, l'esecuzione dell'ANOVA, con lo scopo di stimare l'errore sperimentale. L'ANOVA può essere eseguita facilmente con un qualunque software statistico.

Esempio: una prova di concimazione

Tesi	Produzione
Non concimato	12
Non concimato	15
Non concimato	13
MEDIA	**13.33**
Concime minerale	20
Concime minerale	21
Concime minerale	23
MEDIA	**21.33**
Concime minerale a lento rilascio	19
Concime minerale a lento rilascio	18
Concime minerale a lento rilascio	16
MEDIA	**17.67**
Concime organico	22
Concime organico	19
Concime organico	20
MEDIA	**20.33**

ANOVA

```
Analysis of Variance Table

Response: Prod
Df  Sum Sq Mean Sq F value    Pr(>F)
Tesi       3 115.000  38.333  16.429 0.0008821 ***
Residuals  8  18.667   2.333
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In questo caso, l'analisi della varianza ed il relativo test di F ci dicono che esiste una differenza significativa tra le medie, ma rimane il problema di classificare le soluzioni concimanti in ordine di efficacia. Questo problema può essere risolto in vari modi, come vedremo in seguito.

2 I contrasti pianificati

Il metodo di base è quello di eseguire dei confronti (contrasti) pianificati tra gruppi di trattamenti.

Contrasti pianificati

Definition 1. Si definisce CONTRASTO una combinazione lineare delle medie dei trattamenti, in modo che i coefficienti k sommati diano 0.

NON E' UN CONTRASTO

$$C = 1 \cdot 13.33 - (1 \cdot 21.33 + 1 \cdot 17.67 + 1 \cdot 20.33)$$

E' UN CONTRASTO

$$C = 1 \cdot 13.33 - \left(\frac{1}{3} \cdot 21.33 + \frac{1}{3} \cdot 17.67 + \frac{1}{3} \cdot 20.33 \right)$$

Si può osservare che in questo caso ho confrontato la media del gruppo A, con la media delle medie dei gruppi B, C e D.

DEVIANZA

Un contrasto ha un grado di libertà e la sua devianza, in caso di egual numero di repliche, è data da:

$$SSC = \frac{C^2}{\sum_{i=1}^{n} k^2}$$

che, nel nostro caso, equivale a

$$SSC = \frac{41.53}{1^2 - ((1/3)^2 + (1/3)^2 + (1/3)^2)} = 93.444$$

Significatività del contrasto

- La varianza del contrasto può essere confrontata con la varianza dell'errore tramite test F, per saggiarne la significatività.

$$F = \frac{93.444}{2.333} = 40.05$$

- Considerando i gradi di libertà (1 e 8), la probabilità di errore è data da (in EXCEL):

$$= distrib.f(40.05; 1; 8)$$

- Il risultato è 0.00023. Quindi possiamo rifiutare l'ipotesi nulla (P $\leq$ 0.05).

Contrasti ortogonali

I contrasti possono essere ortogonali se la somma dei prodotti dei coefficienti è pari a zero. Ad esempio:

$$C = 13.33 - \left(\frac{21.33}{3} + \frac{17.67}{3} + \frac{20.33}{3} \right)$$

$$C = 21.33 - \left(\frac{17.67}{2} + \frac{20.33}{2} \right)$$

Infatti:

$$1 \cdot 0 + 1/3 \cdot 1 - 1/3 \cdot 1/2 - 1/3 \cdot 1/2$$

Più contrasti si dicono mutualmente ortogonali se tutte le coppie di contrasti sono ortogonali. In un disegno sperimentali con p tesi esistono solo *p-1* contrasti ortogonali, con l'importante caratteristica che le loro devianze individuali (una per ogni contrasto) sommate ricostruiscono esattamente la devianza del trattamento.

Contrasti ortogonali - 2

L'uso dei contrasti ortogonali è il modo più potente ed elegante per confrontare tra di loro diversi trattamenti sperimentali, anche se richiede un'apposita pianificazione dell'esperimento.

Nel caso in esempio, si potrebbero ipotizzare i seguenti contrasti:

1. non concimato vs concimato
2. concime organico vs. minerali
3. minerale tradizionale vs lento rilascio.

I coefficienti di ciascun contrasto sono abbastanza semplici da scrivere e sono lasciati per esercizio.

3 Test di confronto multiplo

Nell'esempio precedente abbiamo confrontato i trattamenti a gruppi. Ovviamente, i contrasti possono anche essere utilizzati per confrontare una coppia di trattamenti. In questo caso si utilizza una formula molto più semplice.

Minima Differenza Significativa (MDS)

$$MDS(\alpha) = t_{\alpha/2,\nu} \times SED$$

dove $t_{\alpha/2,\nu}$ è il valore critico della distribuzione di t per un livello di probabilità pari ad $\alpha/2$ e la SED è:

$$SED = \sqrt{s^2\left[\frac{1}{r_1}+\frac{1}{r_2}\right]} \Rightarrow con \quad r_1 = r_2 = r \Rightarrow SED = \sqrt{\frac{2s^2}{r}}$$

dove s^2 è la varianza dell'errore con ν gradi di liberta, ed r è il numero delle repliche rispettivamente per le due tesi a confronto. Quindi:

$$MDS(0.05) = 2.306\sqrt{\frac{2 \times 2.333}{3}} = 2.876$$

Questa formula si presta anche per confrontare tutte le possibili coppie di trattamenti. Si parla quindi di procedure di confronto multiplo (MCP), basate sul calcolo di una differenza critica minima, da utilizzare come base per il confronto tra qualunque coppia di medie. In pratica, due medie sono significativamente diverse quando la loro differenza supera la differenza critica.

Risultati

```
             Estimate Std. Error Sign
2 - 1 == 0    8.000      1.247   *
3 - 1 == 0    4.333      1.247   *
4 - 1 == 0    7.000      1.247   *
3 - 2 == 0   -3.667      1.247   *
4 - 2 == 0   -1.000      1.247
4 - 3 == 0    2.667      1.247
```

Gli asterischi indicano le differenze significative (superiori alla MDS).

Display a lettere

- Ordino le medie in senso crescente/decrescente
- Parto dalla prima e aggiungo la lettera A a tutte quelle che non sono significativamente diverse
- Passo alla seconda media e aggiungo la B a tutte quelle che non sono significativamente diverse
- Procedo analogamente

Esempio

Tesi	Produzione	lettera
Concime minerale	21.33	A
Concime organico	20.33	A B
Concime minerale a lento rilascio	17.67	B
Non concimato	13.33	C

Le medie seguite dalla stessa lettera non sono significativamente diverse tra loro.

4 Limitazione delle MCP

MDS e molteplicità

- Con la MDS ogni confronto viene eseguito con una probabilità di errore del 5% (comparisonwise error rate).
- Se i confronti sono molti (family of n comparisons), la probabilità di sbagliarne almeno uno (maximum experimentwise error rate) è data da:

$$\alpha_E = 1 - (1 - \alpha_C)^n$$

- Il numero dei confronti è sempre abbastanza elevato e pari a $k \cdot (k-1)/2$ (k è il numero delle medie).
- Di conseguenza, la probabilità di errore per esperimento può essere molto più alta del valore α prefissato per confronto.

Procedure alternative

1. Test di Duncan;
2. Test di Newman-Keuls;
3. Honest Significant Difference (HSD) di Tukey;
4. Test di confronto multiplo di Tukey;
5. Procedura di Scheffe;

6. Procedura di Bonferroni;

7. Test di Dunnet per il confronto con un testimone;

RICORDA! Il problema della molteplicità si pone anche con i contrasti pianificati!

Per evitare il problema della molteplicità, quando il numero di confronti è elevato (tipicamente maggiore di 10), si preferisce evitare la MDS ed utilizzare delle procedure alternative. Tra queste, la più importante è la HSD di Tukey, simile alla MDS, ma basata sull'impiego di una statistica alternativa, detta 'Studentised range statistics', che sostituisce il t di Student e dipende dal numero delle medie k

HSD

$$HSD(k, \alpha_E) = q_{\alpha,k,\nu} \times \frac{SED}{\sqrt{2}}$$

Il valore di $q_{\alpha,k,\nu}$ si trova in apposite tabelle statistiche. Nel caso specifico, la HSD è:

$$HSD(4, 0.05) = 4.529\sqrt{\frac{2 \times 2.333}{2 \times 3}} = 3.99$$

Le tabelle statistiche per la HSD si trovano (per esempio) in http://cse.niaes.affrc.go.jp/miwa/probcalc/s-range/srng_tbl.html).

La HSD corregge completamente il problema della molteplicità e garantisce un tasso di errore *experimentwise*, indipendentemente dal numero delle medie. Inoltre, come la MDS, la HSD è basata su una differenza critica unica, che risulta molto agevole da impiegare.

Dunnet per confronto con un testimone

Un'altra procedura molto importante è quella di Dunnet, che consente di confrontare tutte le medie con un testimone (o con il migliore/peggiore dei trattamenti). Come la HSD, il test di Dunnet è basato su una particolare forma della 'Studentised Range Statistic', che possiamo trovare in http://biology.ucf.edu/p̃ascencio/classes/Methods/Dun

$$D(k, \alpha_E) = d_{\alpha,k,\nu} \times SED$$

Anche i valori della statistica d sono tabulati, e, per il test a due code portano ai seguenti risultati:

$$D(4, 0.05) = 2.88\sqrt{\frac{2 \cdot 2.333}{3}} = 3.592$$

Per il test ad una coda:

$$D(4, 0.05) = 2.42\sqrt{\frac{2 \cdot 2.333}{3}} = 3.018$$

Anche il test di Dunnet assicura un tasso d'errore per esperimento, ma la differenza critica è più piccola della HSD, perchè viene effettuato un numero di confronti inferiore.

Se volessimo confrontare tutte le medie con la media più alta (o più bassa) potremmo utilizzare il test di Dunnet ad una coda, che utilizza appunto un valore critico tabulato leggermente inferiore (si veda: http://www.watpon. com/ table/dunnetttest.pdf). Test ad una coda significa che andiamo a verificare non solo l'esistenza di una differenza, ma anche la sua direzione. In altre parole, testiamo se una media è significativamente superiore/inferiore ad un'altra (non semplicemente se è diversa).

Altre procedure molto utilizzate (ma non per questo consigliabili) sono quelle di Newman Keuls e Duncan, basate su differenze critiche crescenti al crescere della distanza dei trattamenti in graduatoria. Queste procedure non sono consigliabili perchè non danno nè una protezione *experimentwise*, nè *comparisonwise*.

5 Scegliere la MCP

La cosa fondamentale è muoversi in coerenza con le finalità dell'esperimento. In particolare, bisogna decidere se è necessario adottare un tasso di protezione (errore) per confronto o per esperimento.

Comparisonwise o experimentwise?

- Un approccio *comparisonwise* è accettabile quando il singolo confronto isolato è l'unità concettuale di maggior interesse per il ricercatore
- Un approccio *experimentwise* è preferibile quando l'unità concettuale di maggior interesse per il ricercatore è data da un insieme di confronti
- La gran parte degli esperimenti è pianificata per rispondere ad un insieme di domande e, di conseguenza, i test d'ipotesi non sono mai isolati l'uno dall'altro!

Consigli per l'uso

- Non usare mai contrasti o confronti multipli con serie di dati quantitative. In questo caso la regressione è l'approccio corretto.
- Quando è possibile, pianificare gli esperimenti in modo che sia possibile ottenere le risposte cercate con pochi contrasti di interesse.
- Se il numero dei contrasti aumenta, si pone un problema di molteplicità. Si può ovviare utilizzando la procedura di Bonferroni, che consiste nell'eseguire ogni confronto ad un livello $\alpha_C = \alpha_E/n$ (n è il numero di confronti). In questo modo ci si riporta ad una livello di protezione experimentwise.
- Utilizzare i confronti multipli solo se coerenti con la logica dell'esperimento.

Consigli per l'uso - 2

- Se è giustificabile un tasso di protezione comparisonwise, utilizzare la MDS: è il metodo più potente.

- Se invece è necessario utilizzare un tasso di protezione experimentwise, evitare le procedure di Duncan e Newmann-Keuls: non danno il livello di protezione cercato e, inoltre, non hanno una differenza critica costante e sono quindi difficili da commentare.

- Se la logica dell'esperimento è quella di confrontare le medie con un testimone o con il miglior trattamento, utilizzare la procedura di Dunnet, che garantisce una protezione experimentwise, può essere utilizzata con medie sbilanciate e è caratterizzata da una differenza critica comune.

- Se vogliamo fare tutti i confronti possibili, dobbiamo utilizzare la HSD di Tukey, che ha gli stessi vantaggi del test di Dunnet.