

I contrasti

I contrasti vengono utilizzati nell'ambito dell'analisi della varianza (e dei modelli lineari) per confrontare gli effetti fra gruppi diversi (modalità diverse della variabile indipendente, o di raggruppamento).

Un contrasto è una combinazione lineare di medie (μ_i) di fattore con coefficienti (c_i) a somma nulla.

$$\text{contrasto} = \sum c_i \mu_i$$

Il numero di contrasti possibile per una variabile è pari al numero delle sue modalità $- 1$.

Nel modello `aov(data = ChickWeight, weight ~ Diet)` (vedi [Anova a una via](r:analisi_bivariata:anova_one_way)), data la matrice di contrasti di default:

```
contrasts(ChickWeight$Diet)
```

```
##      2  3  4
## 1  0  0  0
## 2  1  0  0
## 3  0  1  0
## 4  0  0  1
```

I contrasti saranno la differenza fra la media di ciascun gruppo e la media del primo, ovvero i coefficienti della regressione:

```
aov.res <- aov(data = ChickWeight, weight ~ Diet)
coef(aov.res)
```

```
## (Intercept)      Diet2      Diet3      Diet4
##  102.64545    19.97121    40.30455    32.61726
```

L'intercetta è infatti la media della classe considerata come base del confronto, e i coefficienti sono gli "effetti", ovvero le differenze fra l'intercetta e le medie di gruppo:

Tab. 1: Medie dei gruppi e coefficienti dell'Anova

Diet	media	coef(aov.res)
1	102,6455	102,64545
2	122,6167	19,97121
3	142,9500	40,30455
4	135,2627	32,61726

Come si vede nella tabella, infatti, il coefficiente del secondo gruppo è in effetti pari a $\{122,6167\} - \{102,6455\}$.

La funzione contrasts

Per definire i contrasti, usiamo la funzione `contrasts()`, che abbiamo usato sopra per visualizzare i contrasti di default. Impostiamo ad esempio i contrasti di trattamento (vedi [oltre](#)):

```
# treatment
contrasts(ChickWeight$Diet) <- contr.treatment(4, base = 3)
```

I contrasti diventano un'attributo della variabile, che possiamo visualizzare con la funzione `contrasts()`:

```
contrasts(ChickWeight$Diet)
```

```
##      1 2 4
## 1 1 0 0
## 2 0 1 0
## 3 0 0 0
## 4 0 0 1
```

Nel comando `contr.treatment(4, base = 3)`, il primo numero indica il numero dei livelli del fattore (dei gruppi).

Tipi di contrasto in R

In R, è possibile definire diversi tipi di contrasti (ovvero di matrici di coefficienti per calcolare i contrasti $\sum \mu_{ic_i}$):

Tab. 2: Tipi di contrasto e funzioni

Metodo	Definizione	Funzione
Dummy	Codifica dummy in cui ciascuna modalità viene confrontata con la prima	
Treatment	Ciascuna modalità viene confrontata con una modalità scelta come base (di default è la prima)	<code>contr.treatment(n, base = 1)</code>
di Helmert	Ciascuna modalità viene confrontata con l'effetto medio di quelle successive (eccetto l'ultima)	<code>contr.helmert(n)</code>
SAS	Ciascuna modalità viene confrontata con l'ultima	<code>contr.SAS(n)</code>
poly	Test per i trend lineari	<code>contr.poly(n, scores = 1:n)</code>

Abbiamo visto sopra i contrasti di default (*dummy*) e i contrasti *treatment*. Vediamo gli altri

tipi comunemente utilizzati:

```
# contrasti di Helmert
contrasts(ChickWeight$Diet) <- contr.helmert(4)
contrasts(ChickWeight$Diet)
```

```
##      [,1] [,2] [,3]
## 1      -1      -1      -1
## 2       1      -1      -1
## 3       0       2      -1
## 4       0       0       3
```

```
# contrasti SAS
contrasts(ChickWeight$Diet) <- contr.SAS(4)
contrasts(ChickWeight$Diet)
```

```
##      1 2 3
## 1 1 0 0
## 2 0 1 0
## 3 0 0 1
## 4 0 0 0
```

I contrasti nella tabella Anova

Con l'argomento `split`, possiamo vedere la scomposizione della varianza in base ai contrasti impostati:

```
summary.aov(aov.res,
             split=list(Diet=list(1, 2, 3)))
```

```
##              Df  Sum Sq Mean Sq F value    Pr(>F)
## Diet           3  155863    51954   10.81 6.43e-07 ***
## Diet: C1        1     97         97    0.02  0.887
## Diet: C2        1   74055   74055   15.41 9.71e-05 ***
## Diet: C3        1   81711   81711   17.00 4.29e-05 ***
## Residuals     574 2758693     4806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Nell'output, i contrasti sono numerati in base alla matrice dei contrasti: C1 indica il contrasto fra la dieta 2 e la dieta 1, e dunque corrisponde alla modalità Diet 2. Può dunque essere comodo rinominarli:

```
summary(aov.res,
```

```
split = list(Diet=list("2 vs. 1" = 1, "3 vs. 1" = 2,
                      "4 vs. 1" = 3)))
```

```
##              Df  Sum Sq Mean Sq F value    Pr(>F)
## Diet              3  155863    51954   10.81 6.43e-07 ***
## Diet: 2 vs. 1      1     97      97      0.02  0.887
## Diet: 3 vs. 1      1   74055   74055   15.41 9.71e-05 ***
## Diet: 4 vs. 1      1   81711   81711   17.00 4.29e-05 ***
## Residuals        574 2758693     4806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In caso di analisi fattoriale (vedi l'esempio in `[[anova_two_way_factorial]]`), possiamo inserire nella lista più di una lista di contrasti. Ad esempio:

```
split = list(block = list("3 vs. 1" = 2), N =list("pres. vs. ass." =
1))
```

Impostare i contrasti per l'analisi di regressione

I contrasti possono essere impostati per la singola variabile (vedi anche l'esempio alla voce [Anova a due vie](#)), o come argomento delle funzioni `lm()`, `aov()`, `glm()`. Ad esempio:

```
aov2.res <- aov(data = ChickWeight, weight ~ Diet,
               contrasts = list("Diet" = contr.helmert(4))) # una
lista
```

Confronto dei risultati:

```
rbind("Treatment" = coef(aov.res), "Helmert" = coef(aov2.res))
```

```
##              (Intercept)      Diet2      Diet3      Diet4
## Treatment      102.6455  19.971212  40.30455  32.617257
## Helmert        125.8687   9.985606  10.10631   3.131335
```

Esempio con più variabili (vedi [Anova a due vie e fattoriale](#)):

```
# contrasti di default
res <- aov(yield ~ block + P, npk)
coef(res)
```

```
## (Intercept)      block2      block3      block4      block5
## block6
##  54.616667    3.425000    6.750000   -3.900000   -3.500000
```

```
2.325000
```

```
##          P1
```

```
##    -1.183333
```

```
# contrasti per una variabile
```

```
res <- aov(yield ~ block + P, npk,
           contrasts = list("P" = contr.helmert(2)))
coef(res)
```

```
## (Intercept)      block2      block3      block4      block5
block6
##  54.0250000    3.4250000    6.7500000   -3.9000000   -3.5000000
2.3250000
##          P1
##   -0.5916667
```

```
# contrasti diversi per ciascuna variabile
```

```
res <- aov(yield ~ block + P, npk,
           contrasts = list("block" = contr.treatment(6, 3), "P" =
contr.helmert(2)))
coef(res)
```

```
## (Intercept)      block1      block2      block4      block5
block6
##  60.7750000   -6.7500000   -3.3250000  -10.6500000  -10.2500000
-4.4250000
##          P1
##   -0.5916667
```

Analisi multivariata, Modelli lineari

From:

<https://www.agnesevardanega.eu/wiki/> - Ricerca Sociale con R

Permanent link:

<https://www.agnesevardanega.eu/wiki/r/modelli/contrasti>

Last update: **11/08/2025 14:39**

