Züchtungslehre - Lösung 11

Peter von Rohr

2017-12-08

Gegeben ist der folgende Datensatz.

Tochter	Herde	Vater	Leistung
1	1	C	110
2	1	A	127
3	2	B	110
4	2	A	101
5	2	C	200
6	3	C	170
7	3	C	110
8	3	A	100
9	3	B	150

Aufgabe 1

Um den Einfluss der Herde auf die Leistung der Töchter abschätzen zu können, verwenden wir das folgende fixe Modell.

$$y = Xb + e$$

mit	y	Vektor der Leistungen
	b	Vektor der fixen Herdeneffekte
	X	Inzidenzmatrix von b
	e	Vektor der Reste

Aus der folgenden Regressionsanalyse kennen wir die geschätzten Herdeneffekte

```
## Herde1 Herde2 Herde3
##  118.5  137.0  132.5
```

Ihre Aufgaben

Wir nehmen an, dass die Resteffekte e unkorreliert sind und somit gilt, dass $var(e) = I * \sigma_e^2$, wobei I die Einheitsmatrix darstellt. Schätzen Sie die Restvarianz σ_e^2 aufgrund der Residuen des oben gezeigten Regressionsmodells.

Lösung

Die Residuen r_i erhalten wir durch Subtraktion der Beobachtungen y_i minus die gefitteten Werte $\hat{y}_i = x_i^T\hat{b}$. Für unser Zahlenbeispiel heisst das konkret, dass wir von jedem Beobachtungswert y_i einfach den entsprechenden Herdeneffekt abziehen müssen. Die Herdeneffekte sind oben gegeben. Also ziehen wir von den Beobachtungen y_1 und y_2 den Effekt der Herde 1 ab, von den Beobachtungen y_3 bis y_5 den Effekt der Herde 2 usw. Als Resultat erhalten wir den Vektor der Residuen, der für unser Beispiel wie folgt aussieht.

```
## [1]  -8.5   8.5 -27.0 -36.0  63.0  37.5 -22.5 -32.5  17.5
```

Daraus lässt sich der Vektor der quadrierten Residuen berechnen.

```
## [1]   72.25   72.25  729.00 1296.00 3969.00 1406.25  506.25 1056.25  306.25
```

Summieren wir über die quadrierten Residuen und teilen die Summe durch die Anzahl Freiheitsgrade, dann erhalten wir die gefragte Schätzung der Restvarianz. Die Anzahl Freiheitsgrade hier entspricht den Anzahl Beobachtungen minus die Anzahl Herden.

```
## Restvarianz:  1568.917
```

Aufgabe 2

Den Einfluss der Väter auf die Leistungen der Töchter analysieren wir mit einem Modell, in welchem die Vatereffekte als zufällig betrachtet werden.

Das Modell mit den zufälligen Vatereffekten sieht, wie folgt aus

$$y = 1\mu + Zu + e$$

mit	y	Vektor der Leistungen
	μ	allgemeiner Mittelwert
	1	Vektor, deren Komponenten alle gleich 1
	u	Vektor der zufälligen Vatereffekte
	Z	Inzidenzmatrix für u
	e	Vektor der Reste

Die Varianzen für die zufälligen Effekte u und e sind bestimmt durch

$$var(u) = I * \sigma_u^2 \quad \text{und} \quad var(e) = I * \sigma_e^2$$

Die ANOVA-Tabelle des obigen Modells ist nachfolgend gegeben.

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Vater        2   2499    1250   1.021  0.415
## Residuals    6   7344    1224
```

Ihre Aufgaben

Schätzen Sie aufgrund der gegebenen ANOVA-Tabelle die Varianz (σ_u^2) der zufälligen Vatereffekte und die Varianz (σ_e^2) der Resteffekte.

Lösung

Die Schätzung der Restvarianz $\hat{\sigma_e^2}$ beträgt:

$$\hat{\sigma_e^2} = 1224$$

Die Schätzung der Varianz $\hat{\sigma_u^2}$ der Vatereffekte beträgt:

$$\hat{\sigma_u^2} = \frac{1250 - 1224}{9} = 2.89$$