Setup zweifaktorielle Varianzanalyse

Im folgenden betrachten wir zwei diskrete Einflußgrößen $\boldsymbol{C} = \begin{pmatrix} C_1 \\ \vdots \\ C_n \end{pmatrix}$ und $\boldsymbol{D} = \begin{pmatrix} D_1 \\ \vdots \\ D_n \end{pmatrix}$ mit K_C bzw. K_D Ausprägungen. Sei $n_{k,l}$ dabei die Anzahl der Beobachtungen mit $C_i = k$ und $D_j = l$.

! Hier ist die Mittelwertsdarstellung bzw. das Mittelwertsmodell nicht möglich, da dieser davon abhängig ist, welche Variable zuerst kodiert wird.

Modell mit Effekt-Kodierung (mehrfaktoriell)

Das Modell mit Effekt-Kodierung ist gegeben durch

$$Y = (e \;\; z_1^e(C) \cdots z_{K_C-1}^e(C) \;\; z_1^e(D) \cdots z_{K_D-1}^e(D)) \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_{K_C-1} \\ \gamma_1 \\ \vdots \\ \gamma_{K_D-1} \end{pmatrix} + \epsilon$$

mit $\tau_{K_C} = -\sum_{k=1}^{K_C-1} \tau_k$ und $\gamma_{K_D} = -\sum_{k=1}^{K_D-1} \gamma_k$.

Bei dem Modell mit Effekt-Kodierung gibt es einen Intercept μ und die τ_k und γ_l sind die Abweichungen der Gruppe mit $C = k$ bzw. $D = l$ vom Gesamtmittelwert bzw. vom Intercept μ.

Modell mit Referenz-Kodierung (mehrfakt.)

Das Modell mit Referenz-Kodierung ist gegeben durch

$$Y = (e \;\; z_1(C) \cdots z_{K_C-1}(C) \;\; z_1(D) \cdots z_{K_D-1}(D)) \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_{K_C-1} \\ \gamma_1 \\ \vdots \\ \gamma_{K_D-1} \end{pmatrix} + \epsilon$$

mit $\tau_{K_C} = 0$ und $\gamma_{K_D} = 0$.

Bei dem Modell mit Referenz-Kodierung gibt es einen Intercept μ der den Mittelwert der Gruppe mit $C = K_C$ und $D = K_D$ angibt und die τ_k und γ_l sind die Abweichungen der Gruppe mit $C = k$ bzw. $D = l$ vom Mittelwert der Gruppe mit $C = K_C$ und $D = K_D$.

Kodierung Vergleich (mehrfaktoriell) Beispiel

Sei $K_C = 2$ und $K_D = 3$ mit $\underline{n_{k,l} = 2}$ für alle $k = 1, 2, 3$ und $l = 1, 2$. 2 Beob. je Gruppenkombination
Dann erhalten wir als Designmatrix für das Modell mit

Effekt-Kodierung:

$$\boldsymbol{X} = (e \; Z_1^e(\boldsymbol{C}) \; Z_1^e(\boldsymbol{D}) \; Z_2^e(\boldsymbol{D})) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$$

Liegt in beiden Ausgl. kategorien drin

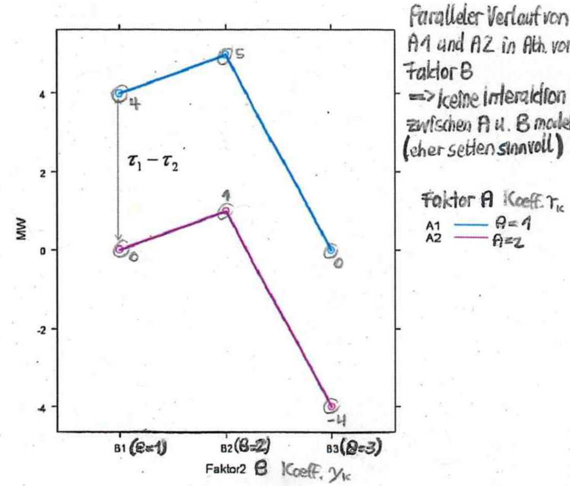
Referenz-Kodierung:

$$\boldsymbol{X} = (e \; Z_1(\boldsymbol{C}) \; Z_1(\boldsymbol{D}) \; Z_2(\boldsymbol{D})) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

In beiden Ref. kategorien

Visualisierung Beispiel

Wir können die Effekte visualisieren, indem wir die Mittelwerte der Gruppen betrachten:



Paralleler Verlauf von A1 und A2 in Abh. von Faktor B
=> keine Interaktion zwischen A u. B modelliert
(eher selten sinnvoll)

Note: “Faktor 2” ist hier D und “A1” und “A2” sind hier $C = 1$ und $C = 2$. Auf der y-Achse ist der Mittelwert der Gruppe dargestellt.

In beiden Fällen werden folgende Gleichungen erfüllt:

/ Ref-MW
Gesamt-MW + Koeff$_{\text{gruppe}}$(A) + Koeff$_{\text{gruppe}}$(B) = MW$_{\text{Kombination}}$

$\widehat{\mu} + \widehat{\tau}_1 + \widehat{\gamma}_1 = 4$	$\widehat{\mu} + \widehat{\tau}_2 + \widehat{\gamma}_1 = 0$
$\widehat{\mu} + \widehat{\tau}_1 + \widehat{\gamma}_2 = 5$	$\widehat{\mu} + \widehat{\tau}_2 + \widehat{\gamma}_2 = 1$
$\widehat{\mu} + \widehat{\tau}_1 + \widehat{\gamma}_3 = 0$	$\widehat{\mu} + \widehat{\tau}_2 + \widehat{\gamma}_3 = -4$

Effekt-Kodierung: $\hat{\mu} = (4+5+0+0+1-4)/6 = 1$

erfüllt:
$\hat{\tau}_1 + \hat{\tau}_2 = 0$
$\hat{\gamma}_1 + \hat{\gamma}_2 + \hat{\gamma}_3 = 0$

$\widehat{\mu} = 1$ (geschätzt Gesamt-MW)	$\widehat{\gamma}_1 = 1$ (0+4)/2 = 2 ⤳ 1 [$\hat{\mu} = 1$]
$\widehat{\tau}_1 = 2$ (5+4+0)/3 = 3 ⤳ 2 [$\hat{\mu} = 1$]	$\widehat{\gamma}_2 = 2$ (1+5)/2 = 3 ⤳ 2 ≙ 3-1
$\widehat{\tau}_2 = -2$ $-\tau_1 = -2$ (Ausgl. gruppe)	$\widehat{\gamma}_3 = -3$ $-\gamma_1 - \gamma_2 = -1-2 = -3$ (Ausgl. gruppe)

! Der Verlauf für $C = 1$ und $C = K_C = 2$ ist parallel mit Abstand $\widehat{\tau}_1 - \widehat{\tau}_2$.

Referenz-Kodierung: Ref. kategorie von B ist B3 => $\gamma_3 = 0$
Ref. von A ist A2 => $\tau_2 = 0$

$\widehat{\mu} = -4$ (MW$_{\text{Ref.}}$ ⤳ A2/B3)	$\widehat{\gamma}_1 = 4$
$\widehat{\tau}_1 = 4$	$\widehat{\gamma}_2 = 5$
$\underline{\widehat{\tau}_2 = 0}$ (Ref. gruppe)	$\underline{\widehat{\gamma}_3 = 0}$ (Ref. gruppe)

! Der Verlauf für $C = 1$ und $C = K_C = 2$ ist parallel mit Abstand $\widehat{\tau}_1$. Man beachte auch, dass sich die Schätzer direkt in dem Plot ablesen lassen.