

Quadratsummentypen

- betrifft Modelle mit mind. 2 Faktoren
- idR sind Faktoren korreliert (ggf. auch schwach und n.s.)
- Ziel der ANOVA ist Zuordnung der Quadratsummen
- bei Überlappung keine eindeutige Zuordnung möglich
- z.B.



- verschieden Typen der ANOVA lösen diese Problematik unterschiedlich
- Beispiel:
 - 2 Faktoren (A und B , $r_{AB} = .1$)
 - ANOVA mit Haupteffekten und Interaktion

Quadratsummentypen

Typ I

- sequentielle QS-Zerlegung
 - $SS(A)$
 - $SS(B|A)$
 - $SS(A*B|A,B)$
- nur bei dieser Form wird Schnittmenge überhaupt berücksichtigt
- Zuordnung aller erklärbarer Varianz zum ersten eingegebenen Faktor A
- dann: Wieviel der Residualvarianz kann noch durch B erklärt werden usw.
- als standard in der `aov()`-Funktion
- Reihenfolge der Faktoren relevant!
- ✓ bei balancierten Designs ohne Interaktionen

Quadratsummentypen

Typ II

- bedingte QS-Zerlegung:
 - $SS(A|B)$
 - $SS(B|A)$
 - $SS(A*B|A,B)$
- Faktoren A und B werden jeweils nur die individuell einmaligen QS zugeordnet
- Also der Anteil der nicht bereits durch die anderen Faktoren erklärt wird
- Die durch mehrere Faktoren erklärbaren Varianzanteile werden *nicht* berücksichtigt
- entsprechend ist die Summe ANOVA-QS nicht gleich der Gesamtquadratsumme
- bei unbalancierten Designs

Quadratsummentypen

Typ III

- bedingte QS-Zerlegung:
 - $SS(A \mid B, A*B)$
 - $SS(B \mid A, A*B)$
 - $SS(A*B \mid A, B)$
- auch hier wird Faktoren A und B jeweils nur die individuelle QS zugeordnet
- aber *zusätzlich* wird auch noch die Interaktion berücksichtigt
- damit noch verzerrender als Typ II
- Die durch mehrere Faktoren erklärbaren Varianzanteil werden *nicht* berücksichtigt
- entsprechend ist die Summe ANOVA-QS nicht gleich der Gesamtquadratsumme
- bevorzugte Variante bei Prüfung der Haupteffekte bei vorhandener Interaktion

Quadratsummentypen

weitere Typen

- Typ IV und Typ V: praktisch nur von historischem Interesse

Quadratsummentypen

in R

- Typ 1:
 - standard in R's `aov()`
 - `aov(yield ~ density + fertilizer, data = df_crop) %>% summary()`
- Typ 2:
 - `car::Anova(lm(yield ~ density + fertilizer, data = df_crop), type=2)`
- Typ 3:
 - `car::Anova(lm(yield ~ density + fertilizer, data = df_crop), type=3)`
- im vorliegenden Fall: perfekt balancierte Daten -> gleiche Ergebnisse
 - siehe `df_crop %>% count(fertilizer, density)`