

Statistik für Digital Humanities

ANOVA – Mehrfaktoriell, Abhängig, Gemischt

Dr. Jochen Tiepmar

Institut für Informatik
Computational Humanities
Universität Leipzig

14. Juni 2021

[Letzte Aktualisierung: 13/06/2021, 10:41]

- 1 Mehrfaktorielle ANOVA
- 2 Zweifaktorielle Unabhängige ANOVA
- 3 ANOVA bei Messwiederholungsdesigns
- 4 Gemischte ANOVA am Beispiel

- H_0 = Alle Mittelwerte sind statistisch ähnlich
- Omnibus Test: Zeigt Effekt an, aber nicht wo er passiert ist
 - $\overline{X_1} = \overline{X_2} \neq \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} = \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} \neq \overline{X_3}$

- $H_0 =$ Alle Mittelwerte sind statistisch ähnlich
- Omnibus Test: Zeigt Effekt an, aber nicht wo er passiert ist
 - $\overline{X_1} = \overline{X_2} \neq \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} = \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} \neq \overline{X_3}$
- Berechnung:
 - $F = \frac{MQE}{MQR} = \frac{\text{Systematische Variation}}{\text{Unsystematische Variation}}$
 - $MQE = \frac{SQE}{k-1}$
 - $MQR = \frac{SQR}{n-k}$
 - $k =$ Anzahl der Gruppen

- Gruppendesign
 - Verschiedene Probanden in Gruppen
 - Gleichzeitige Messung möglich
 - Unabhängiges Design
- Messwiederholungsdesign
 - Gleiche Probanden
 - Wiederholte Messung
 - Abhängiges Design

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments

Jetzt:

- Mehrfaktorielle ANOVA: Experiment durch Veränderung mehrerer Variablen (Gruppenzuordnungen)

- Unabhängig mehrfaktoriell: Mehrere Prädiktoren, jeder gemessen an verschiedenen Probanden
- Abhängig mehrfaktoriell: Mehrere Prädiktoren, jeder gemessen an gleichen Probanden
- Gemischt mehrfaktoriell: Mehrere Prädiktoren, einige mit gleichen Probanden gemessen, einige mit verschiedenen

- Einfaktorielle Unabhängige / One Way Independent ANOVA: 1 Prädiktor, verschiedene Probanden
- Two Way Repeated Measures ANOVA: 2 Prädiktoren, gleiche Probanden
- Zweifaktorielle Gemischte ANOVA: 2 Prädiktoren, 1 gemessen mit denselben Probanden und 1 gemessen mit verschiedenen Probanden
- Three Way Independent ANOVA: 3 Prädiktoren, jeweils verschiedene Probanden
- ...

- 1 Mehrfaktorielle ANOVA
- 2 Zweifaktorielle Unabhängige ANOVA
 - Berechnung
 - Beispiel
 - Visualisierung
 - Effektstärke
- 3 ANOVA bei Messwiederholungsdesigns
 - Sphärizität
 - Messwiederholungs-ANOVA
 - Effektstärke
- 4 Gemischte ANOVA am Beispiel
 - Dateneingabe
 - Datenexploration
 - Kontraste erstellen
 - Modell berechnen

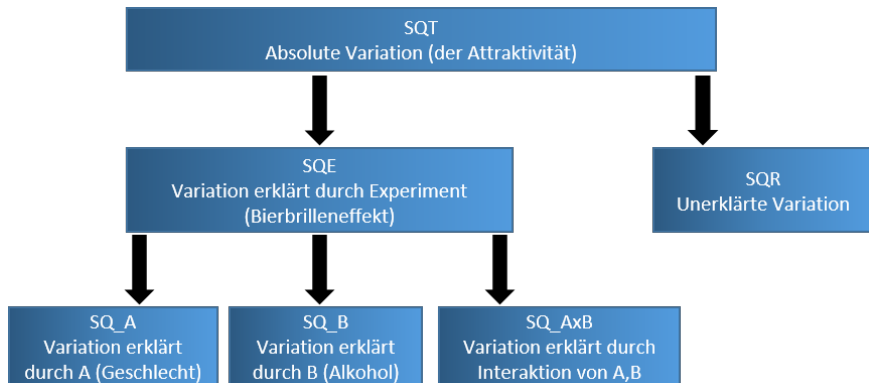
Zweifaktorielle Unabhängige ANOVA

Beispiel Bierbrilleneffekt nach Geschlecht:

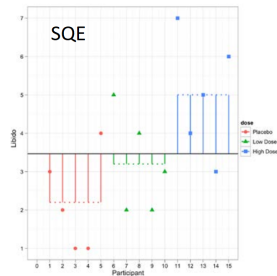
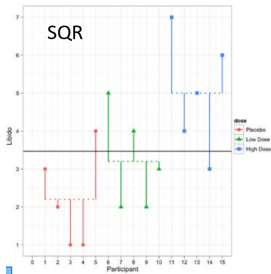
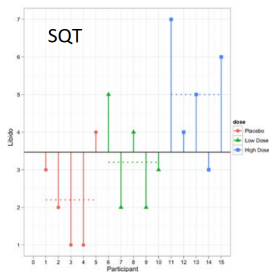
Alkohol	Kein		2 Gläser		4 Gläser	
	Frau	Mann	Frau	Mann	Frau	Mann
Attraktivität des Dates	65	50	70	45	55	30
	70	55	65	60	65	30
	60	80	60	85	70	30
	60	65	70	65	55	55
	60	70	65	70	55	35
	55	75	60	70	60	20
	60	75	60	80	50	45
	55	65	50	60	50	40
Sum	485	535	500	535	460	285
Mean	60.625	66.875	62.50	66.875	57.50	35.625
Var	24.55	106.70	42.86	156.70	50	117.41

Datensatz im Moodle (goggles.csv)

Zweifaktorielle Unabhängige ANOVA



Wiederholung Abweichungsquadrate Mittelwerte



Genau wie vorher bei ANOVA:

- $SQT = \sum (x_i - Grand\ Mean)^2$
- $SQR = \sum (x_i - \overline{group})^2 = \sum s_{group}^2 * (n_{group} - 1)$, $MQR = \frac{SQR}{n-k}$
- $SQE = \sum n_{group} * (\overline{group} - Grand\ Mean)^2$, $MQE = \frac{SQE}{k-1}$

Genau wie vorher bei ANOVA:

- $SQT = \sum (x_i - Grand\ Mean)^2$
- $SQR = \sum (x_i - \overline{group})^2 = \sum s_{group}^2 * (n_{group} - 1)$, $MQR = \frac{SQR}{n-k}$
- $SQE = \sum n_{group} * (\overline{group} - Grand\ Mean)^2$, $MQE = \frac{SQE}{k-1}$

Jetzt neu:

- **Effekt des Prädiktors A**

$$SQ_A = \sum n_{groupA} * (\overline{groupA} - Grand\ Mean)^2$$

$$MQ_A = \frac{SQ_A}{k_A - 1} \text{ mit } k_A = \text{Gruppenanzahl in A}$$

- **Interaktionseffekt**

$$SQ_{A \times B} = SQE - SQ_A - SQ_B$$

$$MQ_{A \times B} = \frac{SQ_{A \times B}}{(k_A - 1) * (k_B - 1)} \text{ mit } k = \text{Gruppenanzahl}$$

- $SQR = \sum (x_i - \overline{group})^2 = \sum s_{group}^2 * (n_{group} - 1), MQR = \frac{SQR}{n-k}$

- **F für Effekt A**

$$SQ_A = \sum n_{groupA} * (\overline{groupA} - Grand\ Mean)^2$$

$$MQ_A = \frac{SQ_A}{k_A - 1} \text{ mit } k = \text{Gruppenanzahl}$$

$$F_A = \frac{MQ_A}{MQR}$$

- **F Interaktionseffekt**

$$SQ_{A \times B} = SQE - SQ_A - SQ_B$$

$$MQ_{A \times B} = \frac{SQ_{A \times B}}{(k_A - 1) * (k_B - 1)} \text{ mit } k = \text{Gruppenanzahl}$$

$$F_{A \times B} = \frac{MQ_{A \times B}}{MQR}$$

Beispiel $F_{\text{Geschlecht}}$

	Frau	Mann
(0 Gl.)	65	50
	70	55

(2 Gl.)	70	45
	65	60

(3 Gl.)	55	30
	65	30

$\overline{Frau} = 60.21$, $\overline{Mann} = 56.46$, $GrandMean = 58.33$

$MQR = 83.036$

Beispiel $F_{Geschlecht}$

	Frau	Mann
(0 Gl.)	65	50
	70	55
(2 Gl.)
	70	45
(3 Gl.)	65	60

(3 Gl.)	55	30
	65	30

$\overline{Frau} = 60.21$, $\overline{Mann} = 56.46$, $GrandMean = 58.33$

$MQR = 83.036$

$$SQ_{Geschlecht} = \sum n_{Frau} * (\overline{Frau} - Grand Mean)^2 + n_{Mann} * (\overline{Mann} - Grand Mean)^2 = 24 * (60.21 - 58.33)^2 + 24 * (56.46 - 58.33)^2 = 168.75$$

Beispiel $F_{Geschlecht}$

	Frau	Mann
(0 Gl.)	65	50
	70	55
(2 Gl.)
	70	45
(3 Gl.)	65	60

(3 Gl.)	55	30
	65	30

$\overline{Frau} = 60.21$, $\overline{Mann} = 56.46$, $GrandMean = 58.33$

$MQR = 83.036$

$$SQ_{Geschlecht} = \sum n_{Frau} * (\overline{Frau} - Grand Mean)^2 + n_{Mann} * (\overline{Mann} - Grand Mean)^2 = 24 * (60.21 - 58.33)^2 + 24 * (56.46 - 58.33)^2 = 168.75$$

$$MQ_{Geschlecht} = \frac{SQ_{Geschlecht}}{k_{Geschlecht} - 1} = \frac{168.75}{1} = 168.75$$

Beispiel $F_{\text{Geschlecht}}$

	Frau	Mann
(0 Gl.)	65	50
	70	55
(2 Gl.)
	70	45
(3 Gl.)	65	60

(3 Gl.)	55	30
	65	30

$$\overline{\text{Frau}} = 60.21, \overline{\text{Mann}} = 56.46, \text{GrandMean} = 58.33$$

$$\text{MQR} = 83.036$$

$$SQ_{\text{Geschlecht}} = \sum n_{\text{Frau}} * (\overline{\text{Frau}} - \text{Grand Mean})^2 + n_{\text{Mann}} * (\overline{\text{Mann}} - \text{Grand Mean})^2 = 24 * (60.21 - 58.33)^2 + 24 * (56.46 - 58.33)^2 = 168.75$$

$$MQ_{\text{Geschlecht}} = \frac{SQ_{\text{Geschlecht}}}{k_{\text{Geschlecht}} - 1} = \frac{168.75}{1} = 168.75$$

$$F_{\text{Geschlecht}} = \frac{MQ_{\text{Geschlecht}}}{\text{MQR}} = \frac{168.75}{83.036} = 2.032$$

Beispiel $F_{Alkohol}$

	Kein	2 Gläser	4 Gläser
(Frau)	65	70	55
	70	65	65

(Mann)	50	45	30
	55	60	30

$\overline{Kein} = 63.75$, $\overline{2\ Gl} = 64.6875$, $\overline{4\ Gl} = 46.5625$, $GrandMean = 58.33$
 $MQR = 83.036$

Beispiel $F_{Alkohol}$

	Kein	2 Gläser	4 Gläser
(Frau)	65	70	55
	70	65	65

(Mann)	50	45	30
	55	60	30

$\overline{Kein} = 63.75$, $\overline{2\ Gl} = 64.6875$, $\overline{4\ Gl} = 46.5625$, $GrandMean = 58.33$

$MQR = 83.036$

$$SQ_{Alkohol} = \sum n_{Kein} * (\overline{Kein} - Grand\ Mean)^2 + n_{2\ Gl} * (\overline{2\ Gl} - Grand\ Mean)^2 + n_{4\ Gl} * (\overline{4\ Gl} - Grand\ Mean)^2 = 16 * (63.75 - 58.33)^2 + 16 * (64.6875 - 58.33)^2 + 16 * (46.5625 - 58.33)^2 = 3332.292$$

Beispiel $F_{Alkohol}$

	Kein	2 Gläser	4 Gläser
(Frau)	65	70	55
	70	65	65
(Mann)
	50	45	30
	55	60	30

$\overline{Kein} = 63.75$, $\overline{2\ Gl} = 64.6875$, $\overline{4\ Gl} = 46.5625$, $GrandMean = 58.33$

$MQR = 83.036$

$$SQ_{Alkohol} = \sum n_{Kein} * (\overline{Kein} - Grand\ Mean)^2 + n_{2\ Gl} * (\overline{2\ Gl} - Grand\ Mean)^2 + n_{4\ Gl} * (\overline{4\ Gl} - Grand\ Mean)^2 = 16 * (63.75 - 58.33)^2 + 16 * (64.6875 - 58.33)^2 + 16 * (46.5625 - 58.33)^2 = 3332.292$$

$$MQ_{Alkohol} = \frac{SQ_{Alkohol}}{k_{Alkohol} - 1} = \frac{3332.292}{2} = 1666.146$$

Beispiel $F_{Alkohol}$

	Kein	2 Gläser	4 Gläser
(Frau)	65	70	55
	70	65	65
(Mann)
	50	45	30
	55	60	30

$\overline{Kein} = 63.75$, $\overline{2\ Gl} = 64.6875$, $\overline{4\ Gl} = 46.5625$, $GrandMean = 58.33$

$MQR = 83.036$

$$SQ_{Alkohol} = \sum n_{Kein} * (\overline{Kein} - Grand\ Mean)^2 + n_{2\ Gl} * (\overline{2\ Gl} - Grand\ Mean)^2 + n_{4\ Gl} * (\overline{4\ Gl} - Grand\ Mean)^2 = 16 * (63.75 - 58.33)^2 + 16 * (64.6875 - 58.33)^2 + 16 * (46.5625 - 58.33)^2 = 3332.292$$

$$MQ_{Alkohol} = \frac{SQ_{Alkohol}}{k_{Alkohol} - 1} = \frac{3332.292}{2} = 1666.146$$

$$F_{Alkohol} = \frac{MQ_{Alkohol}}{MQR} = \frac{1666.146}{83.036} = 20.065$$

Beispiel $F_{Alkohol \times Geschlecht}$

$$SQ_{Alkohol} = 3332.292$$

$$SQ_{Geschlecht} = 168.75$$

$$SQE = 5479.167$$

$$MQR = 83.036$$

Beispiel $F_{Alkohol \times Geschlecht}$

$$SQ_{Alkohol} = 3332.292$$

$$SQ_{Geschlecht} = 168.75$$

$$SQE = 5479.167$$

$$MQR = 83.036$$

$$SQ_{Alkohol \times Geschlecht} = SQE - SQ_{Alkohol} - SQ_{Geschlecht} =$$
$$5479.167 - 18.75 - 332.292 = 1978.125$$

Beispiel $F_{Alkohol \times Geschlecht}$

$$SQ_{Alkohol} = 3332.292$$

$$SQ_{Geschlecht} = 168.75$$

$$SQE = 5479.167$$

$$MQR = 83.036$$

$$SQ_{Alkohol \times Geschlecht} = SQE - SQ_{Alkohol} - SQ_{Geschlecht} =$$
$$5479.167 - 3332.292 - 168.75 = 1978.125$$

$$MQ_{Alkohol \times Geschlecht} = \frac{SQ_{Alkohol \times Geschlecht}}{(k_{Alkohol} - 1) * (k_{Geschlecht} - 1)} = \frac{1978.125}{2} = 989.062$$

Beispiel $F_{Alkohol \times Geschlecht}$

$$SQ_{Alkohol} = 3332.292$$

$$SQ_{Geschlecht} = 168.75$$

$$SQE = 5479.167$$

$$MQR = 83.036$$

$$SQ_{Alkohol \times Geschlecht} = SQE - SQ_{Alkohol} - SQ_{Geschlecht} =$$
$$5479.167 - 3332.292 - 168.75 = 1978.125$$

$$MQ_{Alkohol \times Geschlecht} = \frac{SQ_{Alkohol \times Geschlecht}}{(k_{Alkohol} - 1) * (k_{Geschlecht} - 1)} = \frac{1978.125}{2} = 989.062$$

$$F_{Alkohol \times Geschlecht} = \frac{MQ_{Alkohol \times Geschlecht}}{MQR} = \frac{989.062}{83.036} = 11.911$$

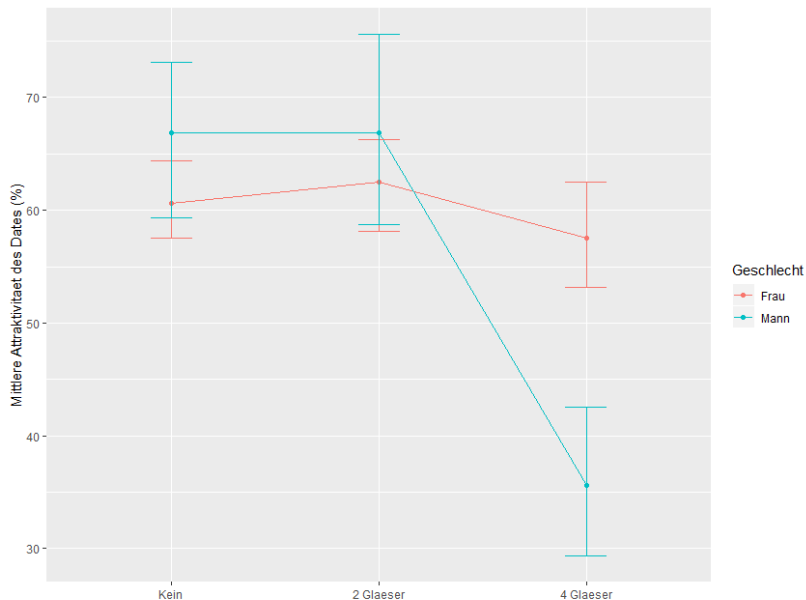
$$F_{\text{Geschlecht}} = 2.032$$

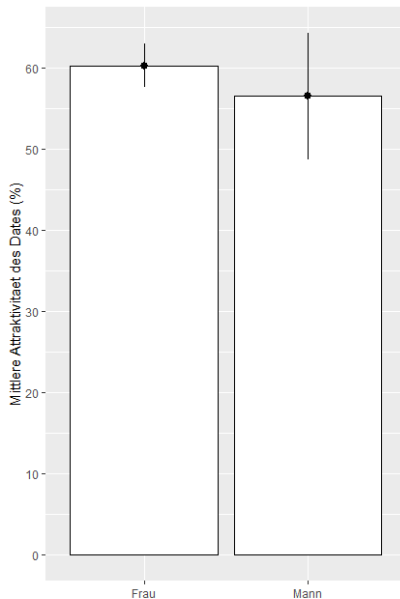
$$F_{\text{Alkohol}} = 20.065$$

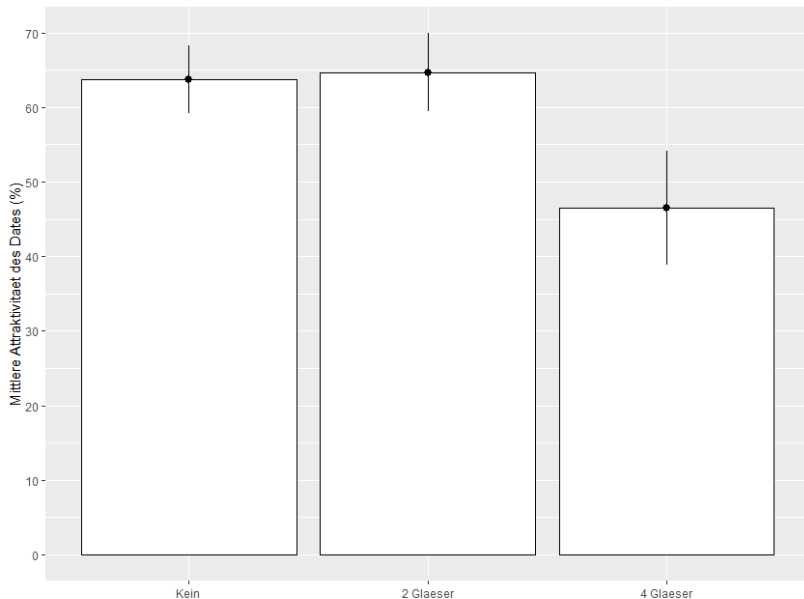
$$F_{\text{Alkohol} \times \text{Geschlecht}} = 11.911$$

- Interpretation analog zu ANOVA
 - eigentlich machen wir nichts anderes als je 1 ANOVA für jeden Effekt und die Interaktion
- Kontrastierung und Post-Hoc Tests anwendbar
- Einfache Effektanalyse möglich indem man Unterschiede bei Variation eines Faktors analysiert (Siehe Unterschiede im Interaktionsgraph)

Interaktionsgraph







Bestimmbar mittels ω^2 , aber sehr umständlich. Deshalb R Skript:

```
omega_factorial<-function(n,a,b, SQa, SQb, SQab, SQr)
{
  MQa<-SQa/(a-1)
  MQb<-SQb/(b-1)
  MQab<-SQab/((a-1)*(b-1))
  MQr<-SQr/(a*b*(n-1))
  varA<-((a-1)*(MQa-MQr))/(n*a*b)
  varB<-((b-1)*(MQb-MQr))/(n*a*b)
  varAB<-((a-1)*(b-1)*(MQab-MQr))/(n*a*b)
  varTotal<-varA + varB + varAB + MQr

  print(paste("Omega-Squared A: ", varA/varTotal))
  print(paste("Omega-Squared B: ", varB/varTotal))
  print(paste("Omega-Squared AB: ", varAB/varTotal))
}

omega_factorial(8,2,3,169,3332,1978,3488)
```

- 1 Mehrfaktorielle ANOVA
- 2 Zweifaktorielle Unabhängige ANOVA
 - Berechnung
 - Beispiel
 - Visualisierung
 - Effektstärke
- 3 ANOVA bei Messwiederholungsdesigns
 - Sphärizität
 - Messwiederholungs-ANOVA
 - Effektstärke
- 4 Gemischte ANOVA am Beispiel
 - Dateneingabe
 - Datenexploration
 - Kontraste erstellen
 - Modell berechnen

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments
- Mehrfaktorielle ANOVA: Experiment durch Veränderung mehrerer unabhängigen Variablen (Gruppenzuordnungen)

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments
- Mehrfaktorielle ANOVA: Experiment durch Veränderung mehrerer unabhängigen Variablen (Gruppenzuordnungen)

Jetzt:

- Abhängige ANOVA: Experiment durch Veränderung einer abhängigen Variable , die nicht das Outcome ist

- H_0 = Alle Mittelwerte sind statistisch ähnlich
- Dieselben Probanden in jeder Gruppe
- Omnibus Test: Zeigt Effekt an, aber nicht wo er passiert ist
 - $\overline{X_1} = \overline{X_2} \neq \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} = \overline{X_3}$
 - $\overline{X_1} \neq \overline{X_2} \neq \overline{X_3}$

Problem:

- ANOVA ist parametrisch
- Dieselben Probanden in jeder Gruppe
- →

Problem:

- ANOVA ist parametrisch
- Dieselben Probanden in jeder Gruppe
- → Ergebnisse der Gruppen sind abhängig voneinander

Problem:

- ANOVA ist parametrisch
- Dieselben Probanden in jeder Gruppe
- → Ergebnisse der Gruppen sind abhängig voneinander

Lösung Sphärizität:

- Verhältnis zwischen Gruppenpaaren ungefähr gleich
- Stärke der Abhängigkeit ist etwa gleich für alle Probanden
- Varianz der Differenzen zwischen den Paaren verschiedener Gruppen müssen ungefähr gleich sein
 - → Erst ab 3 Gruppen relevant!

Gruppe A	Gruppe B	Gruppe C	A-B	A-C	B-C
10	12	8	-2	-2	4
15	15	12	0	3	3
25	30	20	-5	5	10
35	30	28	5	7	2
30	27	20	3	10	7
		Var:	15.7	10.3	10.7

Gruppe A	Gruppe B	Gruppe C	A-B	A-C	B-C
10	12	8	-2	-2	4
15	15	12	0	3	3
25	30	20	-5	5	10
35	30	28	5	7	2
30	27	20	3	10	7
		Var:	15.7	10.3	10.7

- Lokale Sphärizität gegeben bei B-C und A-C
- Gegebene Abweichung der Sphärizität noch im Rahmen
- Signifikanztest möglich mit Mauchly's Test: Signifikant ($p < 0.05$) →

Gruppe A	Gruppe B	Gruppe C	A-B	A-C	B-C
10	12	8	-2	-2	4
15	15	12	0	3	3
25	30	20	-5	5	10
35	30	28	5	7	2
30	27	20	3	10	7
		Var:	15.7	10.3	10.7

- Lokale Sphärizität gegeben bei B-C und A-C
- Gegebene Abweichung der Sphärizität noch im Rahmen
- Signifikanztest möglich mit Mauchly's Test: Signifikant ($p < 0.05$) → signifikante Unterschiede →

Gruppe A	Gruppe B	Gruppe C	A-B	A-C	B-C
10	12	8	-2	-2	4
15	15	12	0	3	3
25	30	20	-5	5	10
35	30	28	5	7	2
30	27	20	3	10	7
		Var:	15.7	10.3	10.7

- Lokale Sphärizität gegeben bei B-C und A-C
- Gegebene Abweichung der Sphärizität noch im Rahmen
- Signifikanztest möglich mit Mauchly's Test: Signifikant ($p < 0.05$) → signifikante Unterschiede → Sphärizität verletzt
- Mehr dazu und zu den Konsequenzen im Begleitmaterial

- Lokale Sphärizität gegeben bei B–C und A–C
- Gegebene Abweichung der Sphärizität noch im Rahmen
- Signifikanztest möglich mit Mauchly's Test: Signifikant ($p < 0.05$) → signifikante Unterschiede → Sphärizität verletzt
- Mehr dazu und zu den Konsequenzen im Begleitmaterial

Umgang mit verletzter Sphärizität

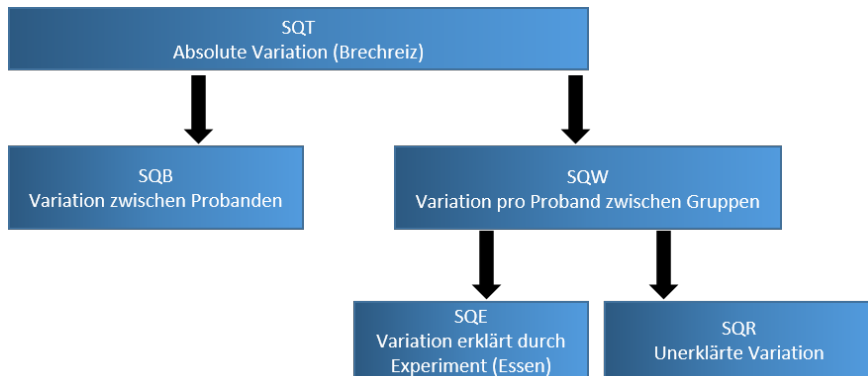
- F-Werte der Tabelle können nicht mehr genutzt werden
- Greenhouse-Geisser- oder Huynh-Feldt- Korrektur oder Durchschnitt beider
- MANOVA oder Multilevel Linear Models benötigen keine Sphärizität

Messwiederholungs-ANOVA

Beispiel Dschungelcamp: Pro Promi Sekunden bis zum Auslösen des Brechreizes bei Verzehr von...

Promi	Maden	Kuhhoden	Rosenkohl	Fischaugen	Mean	s^2
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
3	6	2	3	8	4.75	7.58
4	5	3	1	9	4.50	11.67
5	8	4	5	8	6.25	4.25
6	7	5	6	7	10.3	0.92
7	10	2	7	2	10.3	15.58
8	12	6	8	1	10.3	20.92
Mean	8.13	4.25	4.13	5.75		

Messwiederholungs-ANOVA



Promi	Maden	Kuhhoden	Rosenkohl	Fischaugen	Mean	s^2
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
...

$SQT = 253.89$, $MQT = 8.19$, $SQE = 83.13$, $MQE = 27.71$

Promi	Maden	Kuhhoden	Rosenkohl	Fischaugen	Mean	s^2
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
...

$$SQT = 253.89, MQT = 8.19, SQE = 83.13, MQE = 27.71$$

Jetzt neu:

- **Within-Participant Summenquadrate**

$$\begin{aligned} SQW &= \sum s_p^2(n_p - 1) \text{ mit } p = \text{Proband/Person} \\ &= 9.67 * (4 - 1) + 5.25 * (4 - 1) + \dots = 236.50 \end{aligned}$$

Promi	Maden	Kuhhoden	Rosenkohl	Fischaugen	Mean	s^2
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
...

$$SQT = 253.89, MQT = 8.19, SQE = 83.13, MQE = 27.71$$

Jetzt neu:

- **Within-Participant Summenquadrate**

$$SQW = \sum s_p^2(n_p - 1) \text{ mit } p = \text{Proband/Person}$$
$$= 9.67 * (4 - 1) + 5.25 * (4 - 1) + \dots = 236.50$$

- **Between-Participant Summenquadrate**

$$SQB = SQT - SQW = 253.89 - 236.50 = 17.39$$

Promi	Maden	Kuhhoden	Rosenkohl	Fischaugen	Mean	s^2
1	8	7	1	6	5.50	9.67
2	9	5	2	5	5.25	8.25
...

$$SQT = 253.89, MQT = 8.19, SQE = 83.13, MQE = 27.71$$

Jetzt neu:

- **Within-Participant Summenquadrate**

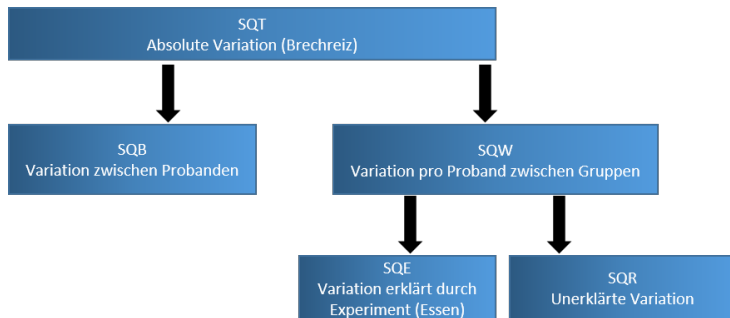
$$SQW = \sum s_p^2 (n_p - 1) \text{ mit } p = \text{Proband/Person}$$

$$= 9.67 * (4 - 1) + 5.25 * (4 - 1) + \dots = 236.50$$

- **Between-Participant Summenquadrate**

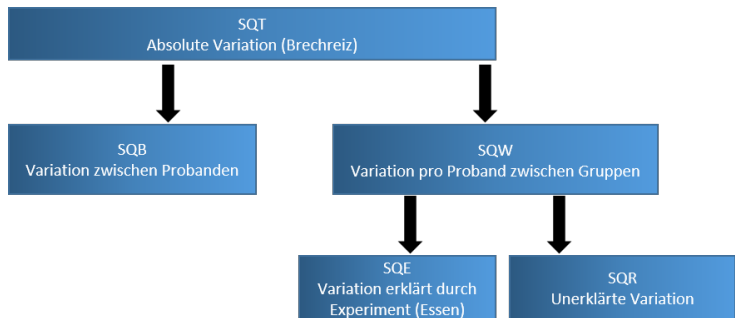
$$SQB = SQT - SQW = 253.89 - 236.50 = 17.39$$

- $SQR = SQW - SQE = 153,37, MQR = \frac{SQR}{(n-1)-(k-1)} = 7.3033$



- **Within-Participant Summenquadrate**

SQW = Individuelle Abweichung zwischen den Gruppen
 $SQW(\text{ahrheit})$



- **Within-Participant Summenquadrate**
 SQW = Individuelle Abweichung zwischen den Gruppen
 $SQW(\text{ahrheit})$
- **Between-Participant Summenquadrate**
 SQB = Varianz erklärt durch individuelle Toleranz/Veranlagung,
Experiment $SQB(\text{ias})$

Genau wie bei unabhängiger ANOVA

- $F = \frac{MQE}{MQR} = \frac{\textit{Systematische Variation}}{\textit{Unsystematische Variation}}$

Interpretation:

Genau wie bei unabhängiger ANOVA

- $F = \frac{MQE}{MQR} = \frac{\text{Systematische Variation}}{\text{Unsystematische Variation}}$

Interpretation:

- Je höher F, desto besser das Modell
- $F < 1 \rightarrow$ Unsystematische Variation ist größer als Systematische Variation
- $F < F_{kr}$ aus Tabelle $\rightarrow H_0$

Genau wie bei unabhängiger ANOVA

- $F = \frac{MQE}{MQR} = \frac{\text{Systematische Variation}}{\text{Unsystematische Variation}}$

Interpretation:

- Je höher F, desto besser das Modell
- $F < 1 \rightarrow$ Unsystematische Variation ist größer als Systematische Variation
- $F < F_{kr}$ aus Tabelle $\rightarrow H_0$ kann nicht abgewiesen werden \rightarrow

Genau wie bei unabhängiger ANOVA

- $F = \frac{MQE}{MQR} = \frac{\text{Systematische Variation}}{\text{Unsystematische Variation}}$

Interpretation:

- Je höher F, desto besser das Modell
- $F < 1 \rightarrow$ Unsystematische Variation ist größer als Systematische Variation
- $F < F_{kr}$ aus Tabelle $\rightarrow H_0$ kann nicht abgewiesen werden \rightarrow Alle Mittelwerte sind statistisch ähnlich
- $df(\text{Numerator}) = k - 1$
- $df(\text{Denominator}) = n - k$

Genau wie bei unabhängiger ANOVA

- $F = \frac{MQE}{MQR} = \frac{\text{Systematische Variation}}{\text{Unsystematische Variation}}$

Interpretation:

- Je höher F, desto besser das Modell
- $F < 1 \rightarrow$ Unsystematische Variation ist größer als Systematische Variation
- $F < F_{kr}$ aus Tabelle $\rightarrow H_0$ kann nicht abgewiesen werden \rightarrow Alle Mittelwerte sind statistisch ähnlich
- $df(\text{Numerator}) = k - 1$
- $df(\text{Denominator}) = n - k$

Abhängige t-Tests bei Kontrastierung und Post Hoc Analysen verwenden

Mehrfaktorielle Abhängige ANOVA

Drink	Bier		Wein		Wasser	
Probant	Glas	Flasche	Glas	Flasche	Glas	Flasche
1	22	23	44	44	45	55
2	11	2	3	44	55	44

Kurz: Abhängige ANOVA (diese Vorlesung) für jeden Prädiktor und deren Interaktion.

Anschauungsbeispiel im Moodle.

2 Ansätze

- via `ezANOVA(...)`
leichter aber Sphärizität wichtig
- Multilevel Linear Model via
`lme(...)`

Codebeispiele und Daten im Moodle

$$\omega^2 = \frac{\left[\frac{k-1}{n*k} * (MQE - MQR)\right]}{MQR + \frac{MQB - MQR}{k} + \left[\frac{k-1}{n*k} * (MQE - MQR)\right]}$$

Guidelines für ω^2

Kirk, R.E. (1996): *Practical Significance: A concept whose time has come*

- .01 → gering
- .06 → moderat
- .14 → stark
- sehr kontextabhängig

- 1 Mehrfaktorielle ANOVA
- 2 Zweifaktorielle Unabhängige ANOVA
 - Berechnung
 - Beispiel
 - Visualisierung
 - Effektstärke
- 3 ANOVA bei Messwiederholungsdesigns
 - Sphärizität
 - Messwiederholungs-ANOVA
 - Effektstärke
- 4 Gemischte ANOVA am Beispiel
 - Dateneingabe
 - Datenexploration
 - Kontraste erstellen
 - Modell berechnen

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments
- Mehrfaktorielle ANOVA: Experiment durch Veränderung mehrerer unabhängigen Variablen (Gruppenzuordnungen)
- Abhängige ANOVA: Experiment durch Veränderung einer abhängigen Variable

Bekannt:

- (einfaktorielle) ANOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung)
- ANCOVA: Experiment durch Veränderung einer unabhängigen Variable (Gruppenzuordnung) unter Eliminierung einer Variable außerhalb des Experiments
- Mehrfaktorielle ANOVA: Experiment durch Veränderung mehrerer unabhängigen Variablen (Gruppenzuordnungen)
- Abhängige ANOVA: Experiment durch Veränderung einer abhängigen Variable

Jetzt:

- Gemischte ANOVA: Experiment durch Veränderung (mindestens) einer abhängigen und einer unabhängigen Variable

Wir schauen uns jetzt eine dreifaktorielle gemischte ANOVA als Anschauungsbeispiel an.

```
install.packages("ez")  
install.packages("ggplot2")  
install.packages("nlme")  
install.packages("pastecs")  
install.packages("reshape")
```

```
#Initiate packages  
library(ez)  
library(ggplot2)  
library(nlme)  
library(pastecs)  
library(reshape)
```

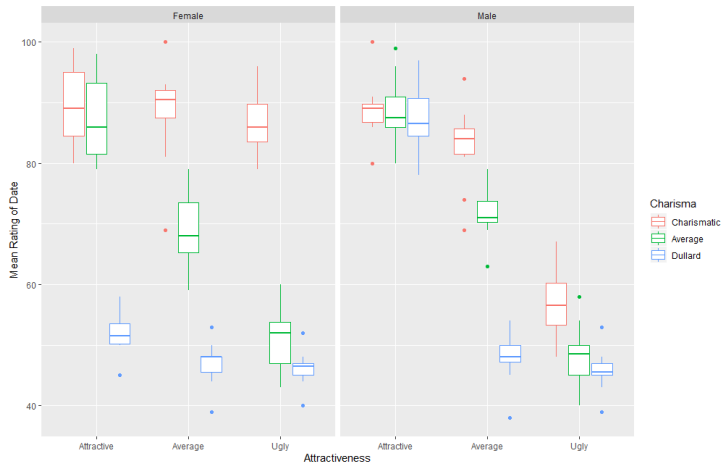
Gender	High Charisma	Low Charisma	Dullard
	Att — Avg — Ug	Att — Avg — Ug	Att — Avg — Ug
Male	86 — 84 — 67	88 — 69 — 50	97 — 48 — 47
	... — ... — — ... — — ... — ...
Female	89 — 91 — 93	88 — 65 — 54	56 — 48 — 52
	... — ... — — ... — — ... — ...

Moodle: LooksOrPersonality.dat

```
dateData<-read.delim("LooksOrPersonality.dat", header = TRUE)
speedData<-melt(dateData, id = c("participant","gender"), measured = c("att_high",
  "av_high", "ug_high", "att_some", "av_some", "ug_some", "att_none", "av_none",
  "ug_none"))
names(speedData)<-c("participant", "gender", "groups", "dateRating")
speedData$personality<-gl(3, 60, labels = c("Charismatic", "Average", "Dullard"))
speedData$looks<-gl(3,20, 180, labels = c("Attractive", "Average", "Ugly"))
speedData<-speedData[order(speedData$participant),]
```

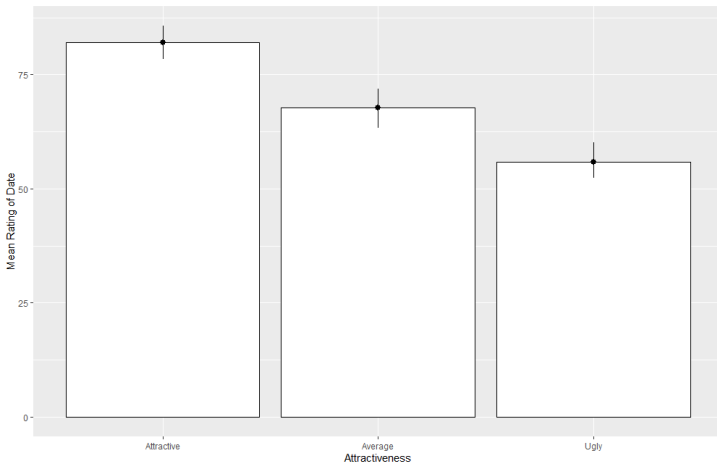
participant	gender	groups	dateRating	personality	looks
P01	Male	att_high	86	Charismatic	Attractive
P01	Male	av_high	84	Charismatic	Average
P01	Male	ug_high	67	Charismatic	Ugly
P01	Male	att_some	88	Average	Attractive
P01	Male	av_some	69	Average	Average
P01	Male	ug_some	50	Average	Ugly
P01	Male	att_none	97	Dullard	Attractive
P01	Male	av_none	48	Dullard	Average
P01	Male	ug_none	47	Dullard	Ugly
P02	Male	att_high	91	Charismatic	Attractive
P02	Male	av_high	83	Charismatic	Average
P02	Male	ug_high	53	Charismatic	Ugly
P02	Male	att_some	83	Average	Attractive
P02	Male	av_some	74	Average	Average
P02	Male	ug_some	48	Average	Ugly

Boxplots



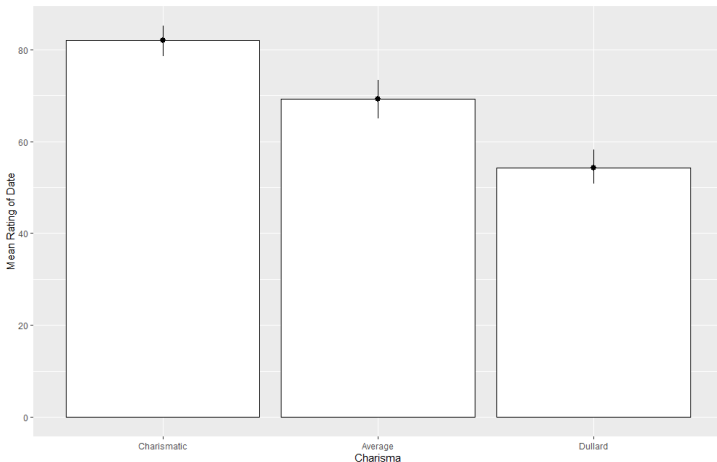
```
dateBoxplot<-ggplot(speedData,aes(looks,dateRating,colour=personality))
dateBoxplot+geom_boxplot()+labs(x="Attractiveness",y="Mean Rating of Date",
  colour="Charisma")+facet_wrap(~gender)
```

Balken Looks



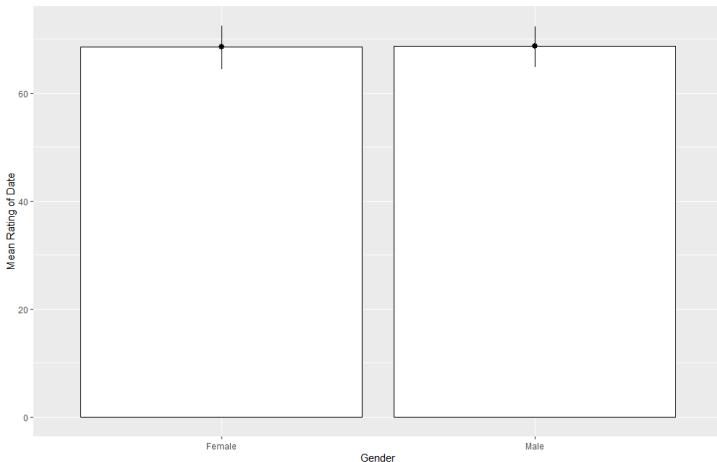
```
looksBar <- ggplot(speedData, aes(looks, dateRating))
looksBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,
  geom = "pointrange") + labs(x = "Attractiveness", y = "Mean Rating of Date")
```

Balken Charisma



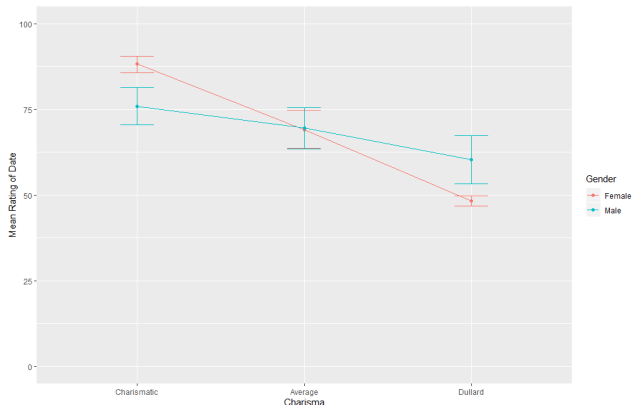
```
charismaBar <- ggplot(speedData, aes(personality, dateRating))  
charismaBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",  
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,  
  geom = "pointrange") + labs(x = "Charisma", y = "Mean Rating of Date")
```

Balken Gender



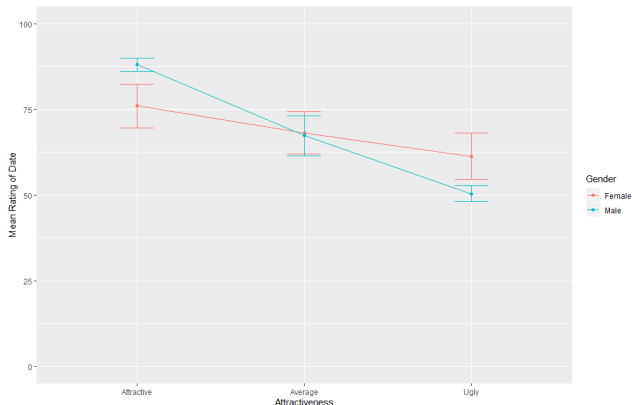
```
genderBar <- ggplot(speedData, aes(gender, dateRating))
genderBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,
  geom = "pointrange") + labs(x = "Gender", y = "Mean Rating of Date")
```


Interaktion Gender Charisma



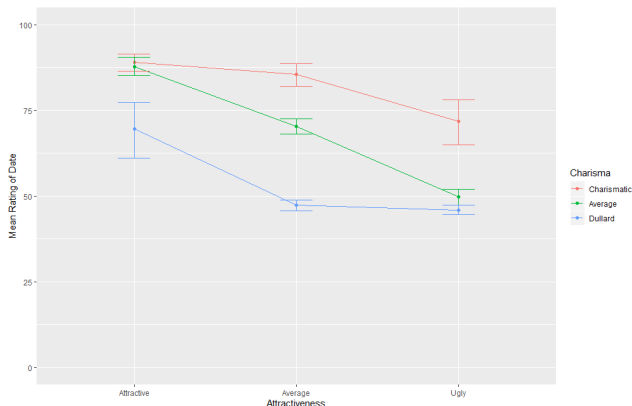
```
genderCharisma <- ggplot(speedData, aes(personality, dateRating,  
  colour = gender))  
genderCharisma + stat_summary(fun.y = mean, geom = "point") +  
  stat_summary(fun.y = mean, geom = "line", aes(group= gender)) +  
  stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2) +  
  labs(x = "Charisma", y = "Mean Rating of Date", colour = "Gender") +  
  scale_y_continuous(limits = c(0,100))
```

Interaktion Gender Looks



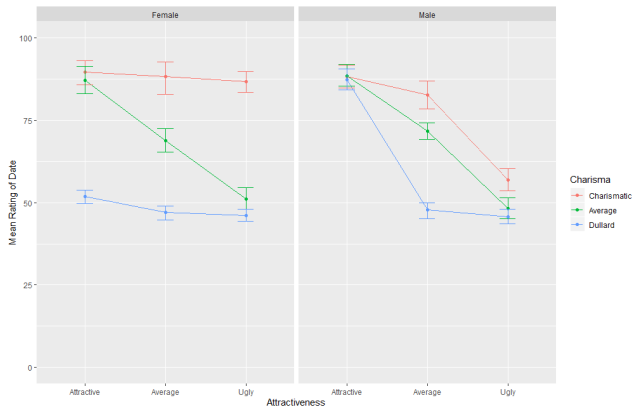
```
genderLooks <- ggplot(speedData, aes(looks, dateRating, colour = gender))
genderLooks + stat_summary(fun.y = mean, geom = "point")
  + stat_summary(fun.y = mean, geom = "line", aes(group= gender))
  + stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2)
  + labs(x = "Attractiveness", y = "Mean Rating of Date", colour = "Gender") +
  scale_y_continuous(limits = c(0,100))
```

Interaktion Looks Charisma



```
looksCharisma <- ggplot(speedData, aes(looks, dateRating, colour = personality))
looksCharisma + stat_summary(fun.y = mean, geom = "point")
  + stat_summary(fun.y = mean, geom = "line", aes(group= personality))
  + stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2)
  + labs(x = "Attractiveness", y = "Mean Rating of Date", colour = "Charisma")
  + scale_y_continuous(limits = c(0,100))
```

Interaktion Gender Looks Charisma



```
looksCharismaGender <- ggplot(speedData, aes(looks, dateRating, colour = personality))
looksCharismaGender + stat_summary(fun.y = mean, geom = "point")
  + stat_summary(fun.y = mean, geom = "line", aes(group= personality))
  + stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2)
  + labs(x = "Attractiveness", y = "Mean Rating of Date", colour = "Charisma")
  + scale_y_continuous(limits = c(0,100)) + facet_wrap(~gender)
```

```
by(speedData$dateRating, list(speedData$looks, speedData$personality,  
  speedData$gender), stat.desc, basic = FALSE)
```

```
: Attractive
```

```
: Charismatic
```

```
: Female
```

median	mean	SE.mean	CI.mean.0.95	var	std.dev	coef.var
89.00000000	89.60000000	2.09867683	4.74753683	44.04444444	6.63659886	0.07406918

```
-----  
: Average
```

```
: Charismatic
```

```
: Female
```

median	mean	SE.mean	CI.mean.0.95	var	std.dev	coef.var
90.50000000	88.40000000	2.63396617	5.95844544	69.37777778	8.32933237	0.09422322

```
-----  
: Average
```

```
: Dullard
```

```
: Male
```

median	mean	SE.mean	CI.mean.0.95	var	std.dev	coef.var
48.00000000	47.80000000	1.32329555	2.99350251	17.51111111	4.18462795	0.08754452

```
-----  
: Ugly
```

```
: Dullard
```

```
: Male
```

median	mean	SE.mean	CI.mean.0.95	var	std.dev	coef.var
45.50000000	45.80000000	1.13333333	2.56377812	12.84444444	3.58391468	0.07825141

```
*Und so weiter*
```

Wir bauen orthogonale Kontraste analog zu vorher

- *ugly* und *dullard* als Kontrollgruppen
- *att vs avg* und *high vs low* als Untersuchungseinheit

Wir bauen orthogonale Kontraste analog zu vorher

- *ugly* und *dullard* als Kontrollgruppen
- *att vs avg* und *high vs low* als Untersuchungseinheit

Gruppe	Kontr 1	Kontr 2
Attractive	1	-1
Average	1	1
Ugly	-2	0

Gruppe	Kontr 1	Kontr 2
Charismatic	1	-1
Average	1	1
Dullard	-2	0

```
SomevsNone<-c(1, 1, -2)
```

```
HivsAv<-c(1, -1, 0)
```

```
contrasts(speedData$personality)<-cbind(SomevsNone, HivsAv)
```

```
AttractivevsUgly<-c(1, 1, -2)
```

```
AttractvsAv<-c(1, -1, 0)
```

```
contrasts(speedData$looks)<-cbind(AttractivevsUgly, AttractvsAv)
```

Modell berechnen (als ANOVA)

```
options(digits = 3)
speedModel<-ezANOVA(data = speedData, dv = .(dateRating), wid = .(participant),
  between = .(gender), within = .(looks, personality), type = 3, detailed = TRUE)
speedModel
options(digits = 7)
```

	Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ges
1	(Intercept)	1	18	846249.8	760	2.00e+04	7.01e-29	*	9.94e-01
2	gender	1	18	0.2	760	4.74e-03	9.46e-01		4.07e-05
3	looks	2	36	20779.6	883	4.24e+02	9.59e-26	*	8.09e-01
5	personality	2	36	23233.6	1274	3.28e+02	7.69e-24	*	8.26e-01
4	gender:looks	2	36	3944.1	883	8.04e+01	5.23e-14	*	4.45e-01
6	gender:personality	2	36	4420.1	1274	6.24e+01	1.97e-12	*	4.74e-01
7	looks:personality	4	72	4055.3	1993	3.66e+01	1.10e-16	*	4.52e-01
8	gender:looks:personality	4	72	2669.7	1993	2.41e+01	1.11e-12	*	3.52e-01

\$'Mauchly's Test for Sphericity'

	Effect	W	p	p<.05
3	looks	0.960	0.708	
4	gender:looks	0.960	0.708	
5	personality	0.929	0.536	
6	gender:personality	0.929	0.536	
7	looks:personality	0.613	0.534	
8	gender:looks:personality	0.613	0.534	

\$'Sphericity Corrections'

	Effect	GGe	p[GG]	p[GG]<.05	HFe	p[HF]	p[HF]<.05
3	looks	0.962	7.62e-25	*	1.074	9.59e-26	*
4	gender:looks	0.962	1.49e-13	*	1.074	5.23e-14	*
5	personality	0.934	2.06e-22	*	1.038	7.69e-24	*
6	gender:personality	0.934	9.44e-12	*	1.038	1.97e-12	*
7	looks:personality	0.799	9.00e-14	*	0.992	1.43e-16	*
8	gender:looks:personality	0.799	1.47e-10	*	0.992	1.34e-12	*

	Effect	DFn	DFd	SSn	SSd	F	p	p<.05	ges
1	(Intercept)	1	18	846249.8	760	2.00e+04	7.01e-29	*	9.94e-01
2	gender	1	18	0.2	760	4.74e-03	9.46e-01		4.07e-05
3	looks	2	36	20779.6	883	4.24e+02	9.59e-26	*	8.09e-01
5	personality	2	36	23233.6	1274	3.28e+02	7.69e-24	*	8.26e-01
4	gender:looks	2	36	3944.1	883	8.04e+01	5.23e-14	*	4.45e-01
6	gender:personality	2	36	4420.1	1274	6.24e+01	1.97e-12	*	4.74e-01
7	looks:personality	4	72	4055.3	1993	3.66e+01	1.10e-16	*	4.52e-01
8	gender:looks:personality	4	72	2669.7	1993	2.41e+01	1.11e-12	*	3.52e-01

- Mauchly's Test überall nicht signifikant, also Sphärizität gegeben
- Bei *gender* Effekt nicht signifikant → Bei Ignorieren von *personality* und *looks* kein signifikanter Unterschied
- Signifikanter Effekt bei *looks* → Bei Ignorieren von *personality* und *gender* signifikanter Unterschied bei *looks*
- Signifikanter Effekt bei *gender:looks* → Effekt bei *looks* verschieden je nach *gender*
- Signifikanter Effekt bei *gender:looks:personality* → Der signifikante Effekt bei *looks:personality* ist verschieden je nach *gender*
 - Analog bei anderen (und umgekehrten) Kombinationen

- Mehrfaktorielle ANOVA: Variation mehrerer Variablen, Mehrfache Gruppierung
- 2-Faktorielle unabhängige ANOVA: 2 Gruppierungen (Alkohol und Geschlecht) mit verschiedenen Probanden
- Interaktionsgrad als Messwert
- Berechnung analog zu einfaktorieller ANOVA aber für jede Gruppierung einzeln sowie für Interaktion
 - F interpretieren
 - Kontrastierung, Post Hoc Tests
 - Effektstärke
- Einfache Effektanalysen bereits mit Visualisierung möglich
- Robust: Siehe Wilcox, 2005

- Repeated Measures ANOVA: Experiment durch Veränderung einer abhängigen Variable
- Sphärizität
 - Korrektur möglich mit Greenhouse-Geisser oder Huynh-Feldt
 - Alternativ MANOVA oder Multilevel Lineare Modelle
- Within-Participant Summenquadrat SQW
- Between-Participant Summenquadrat SQB
- ANOVA durchführen
 - F interpretieren
 - Kontrastierung, Post Hoc Tests
 - Effektstärke
- Robust: Wilcox (2005)

- Gemischte ANOVA: Experiment durch Veränderung (mindestens) einer abhängigen und einer unabhängigen Variable
- Schritte
 - 1 Dateneingabe und -Exploration
 - 2 Kontraste erstellen und Modell berechnen
 - 3 Auswerten und ggfalls weitere Auswertung mittels paarweiser t-Tests (gezielt, Kontrastierung) oder per Post Hoc Test (explorativ)
- Übersprungen: Robust (Siehe Wilcox, 2005), Gemischtes Design als Lineares Modell/Regression (Siehe Begleitmaterial im Moodle)
- Lineares Modell/Regression flexibler einsetzbar, erlaubt genauere Analysen der Interaktion und kommt ohne Bedingung der Sphärizität aus
- Im Moodle finden sich ein paar teils sehr komplexe Beispielanalysen als Anschauungsbeispiele

Bayer, J & Häusler, J & Bader, M (2016): *A New Diagnostic for Cyclic Wh-Movement: Discourse Particles in German Questions*

- Verwendung des Wortes *denn* in Fragesätzen (Wh-Clause)
- Linguistisch anspruchsvoll "Our data show that in wh-questions, the DiP *denn* can only occur in a clausal complement X if a long (i.e., transclausal) wh-dependency in the sense of cyclic movement connects X with the interrogative Force head of the root clause."
- Umfassende Analyse des Wortes *denn* und ähnlichen Worten (*wohl, schon*)
- Mixed-Effects Analyse → ANOVA über Satzakzeptanz