

Statistik für Digital Humanities

Gemischte ANOVA (GLM 5/5)

Dr. Jochen Tiepmar

Institut für Informatik
Abteilung Computational Humanities
Universität Leipzig

27. Januar 2020

[Letzte Aktualisierung: 27/01/2020, 12:01]

1 Gemischte ANOVA

2 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)
```

1 Gemischte ANOVA

2 Beispiel

- Dateneingabe
- Datenexploration
- Kontraste erstellen
- Modell berechnen

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),]
```

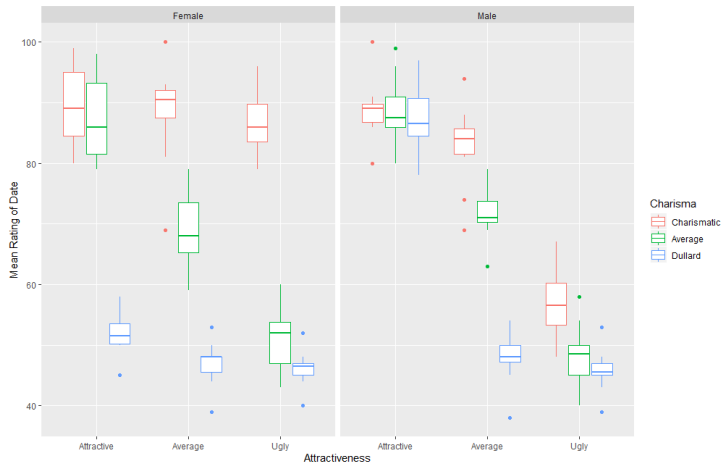

participant	gender	groups	dateRating	personality	looks
P01	Male	att_high	86	Charismatic	Attractive
P01	Male	av_high	84	Charismatic	Average
P01	Male	ug_high	67	Charismatic	Ugly
P01	Male	att_some	88	Average	Attractive
P01	Male	av_some	69	Average	Average
P01	Male	ug_some	50	Average	Ugly
P01	Male	att_none	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

1 Gemischte ANOVA

2 Beispiel

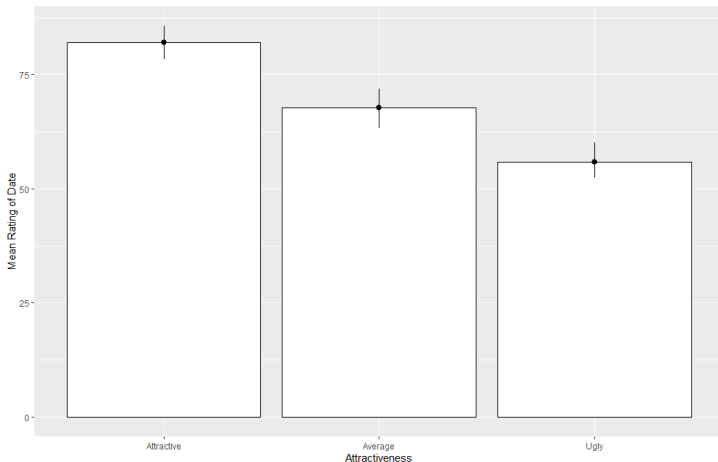
- Dateneingabe
- **Datenexploration**
- Kontraste erstellen
- Modell berechnen

Boxplots



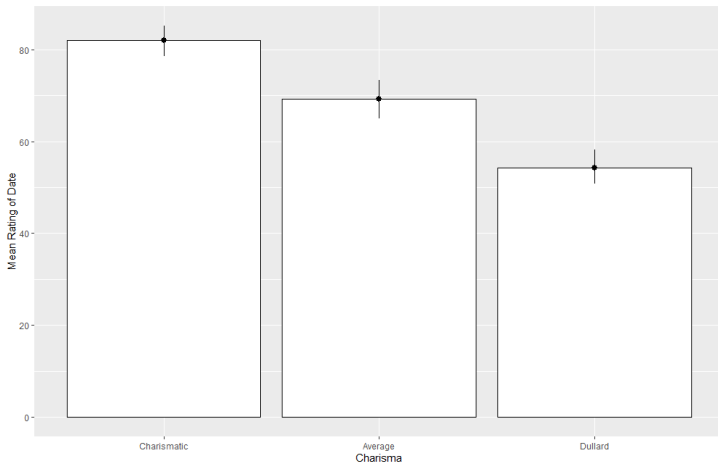
```
dateBoxplot<-ggplot(speedData,aes(looks,dateRating,colour=personality))
dateBoxplot+geom_boxplot()+labs(x="Attractiveness",y="Mean Rating of Date",
  colour="Charisma")+facet_wrap(~gender)
```

Balken Looks



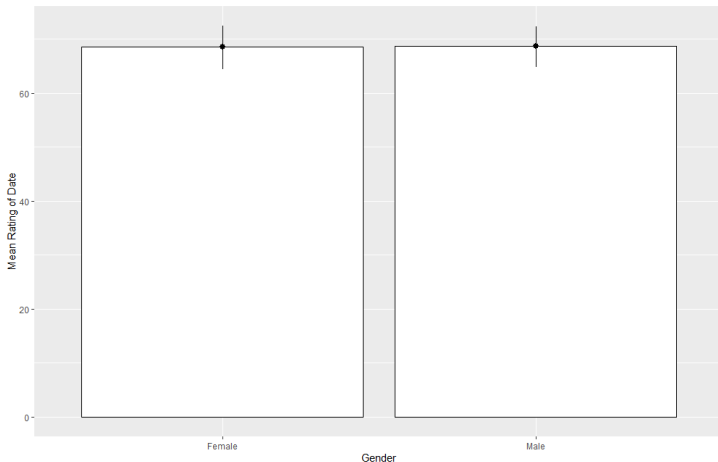
```
looksBar <- ggplot(speedData, aes(looks, dateRating))
looksBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,
  geom = "pointrange") + labs(x = "Attractiveness", y = "Mean Rating of Date")
```

Balken Charisma



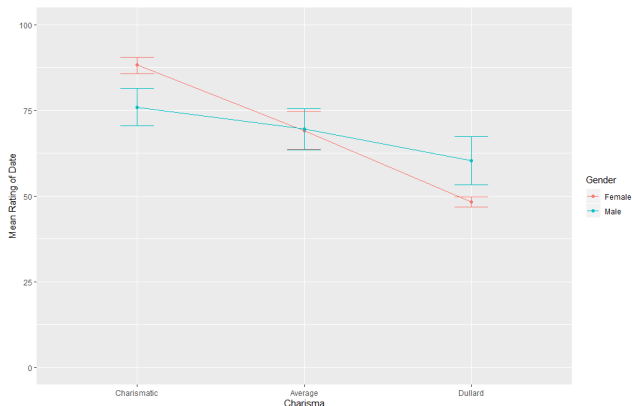
```
charismaBar <- ggplot(speedData, aes(personality, dateRating))  
charismaBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",  
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,  
  geom = "pointrange") + labs(x = "Charisma", y = "Mean Rating of Date")
```

Balken Gender



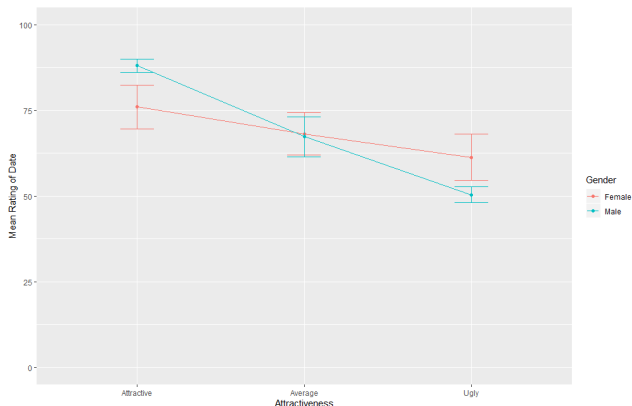
```
genderBar <- ggplot(speedData, aes(gender, dateRating))
genderBar + stat_summary(fun.y = mean, geom = "bar", fill = "White",
  colour = "Black") + stat_summary(fun.data = mean_cl_boot,
  geom = "pointrange") + labs(x = "Gender", y = "Mean Rating of Date")
```

Interaktion Gender Charisma



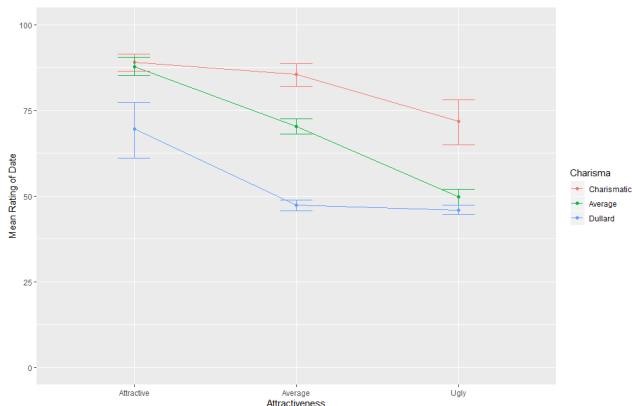
```
genderCharisma <- ggplot(speedData, aes(personality, dateRating,  
  colour = gender))  
genderCharisma + stat_summary(fun.y = mean, geom = "point") +  
  stat_summary(fun.y = mean, geom = "line", aes(group= gender)) +  
  stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2) +  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

Interaktion Gender Looks



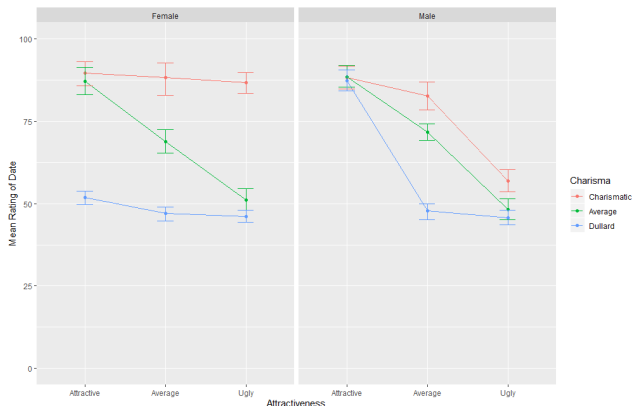
```
genderLooks <- ggplot(speedData, aes(looks, dateRating, colour = gender))
genderLooks + stat_summary(fun.y = mean, geom = "point")
  + stat_summary(fun.y = mean, geom = "line", aes(group= gender))
  + stat_summary(fun.data = mean_cl_boot, geom = "errorbar", width = 0.2)
  + labs(x = "Attractiveness", 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*
```

1 Gemischte ANOVA

2 Beispiel

- Dateneingabe
- Datenexploration
- **Kontraste erstellen**
- Modell berechnen

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)
```

1 Gemischte ANOVA

2 Beispiel

- Dateneingabe
- Datenexploration
- Kontraste erstellen
- **Modell berechnen**

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	

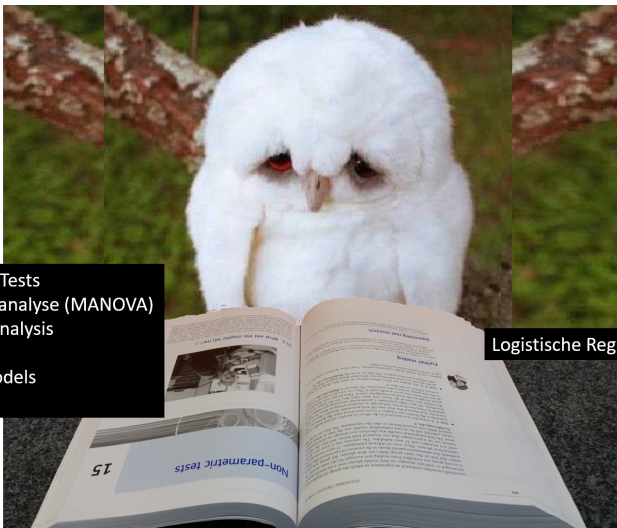
\$'Sphericity Corrections'

	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

- 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 (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



Nichtparametrische Tests
Multivariate Varianzanalyse (MANOVA)
Exploratory Factor Analysis
Kategorische Daten
Multilevel Linear Models

Logistische Regression