



# ANalysis Of VAriance (ANOVA) 2/2

## Wdh: ANOVA - Idee

- ANOVA 1: Zwei Medikamente zur Blutdrucksenkung und Placebo (Faktor  $X$ ). Gibt es einen sign. Unterschied in der Wirkung (kontinuierlich  $Y$ )?

$$Y \sim X + \varepsilon$$

1-weg ANOVA

- ANOVA 2: Zwei Medikamente zur Blutdrucksenkung, Placebo (Faktor  $X_1$ ) und Geschlecht (Faktor  $X_2$ ). Gibt es einen sign. Unterschied in der Wirkung (kontinuierlich  $Y$ ) (evtl. geschlechterspezifisch)?

$$Y \sim X_1 + X_2 + \varepsilon$$

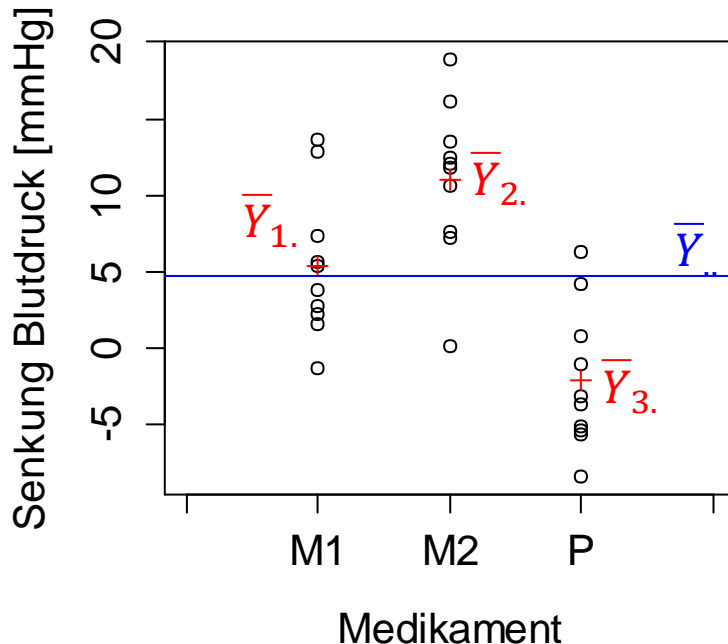
2-weg ANOVA

# Wdh: 1-weg ANOVA

$g$ : Anzahl Gruppen (3)

$p$ : Anzahl Beob. pro Gruppe (10)

Ann:  $p$  in jeder Gruppe gleich



## Streuung zwischen Gruppen:

“Between-Sum-of-Squares” ( $SS_B$ )

RSS der Gruppenmittelwerte (rote Kreuze) um den totalen Mittelwert (blaue Linie)

$$SS_B = p * \sum_{i=1}^g (\bar{Y}_{i.} - \bar{Y}_{..})^2$$

## Streuung innerhalb Gruppen:

“Within-Sum-of-Squares” ( $SS_W$ )

RSS der Einzelbeobachtungen (schwarze Kreise) um die einzelnen Mittelwerte (rote Kreuze)

$$SS_W = \sum_{i=1}^g \sum_{j=1}^p (Y_{ij} - \bar{Y}_{i.})^2$$

$$\text{Teststatistik} \approx \frac{SS_B}{SS_W}$$

# Wdh: 1-weg ANOVA - Modell

- $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma^2) \text{ iid}$

Technische Nebenbedingung:  $\sum_{i=1}^g \alpha_i = 0$

“Analyse der Varianzen”

- $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_g = 0$

- Teststatistik:  $T = \frac{SS_B/(g-1)}{SS_W/(g*(p-1))} = \frac{MS_B}{MS_W}$

“Mean Squares”

- Theorie: Falls  $H_0$  stimmt

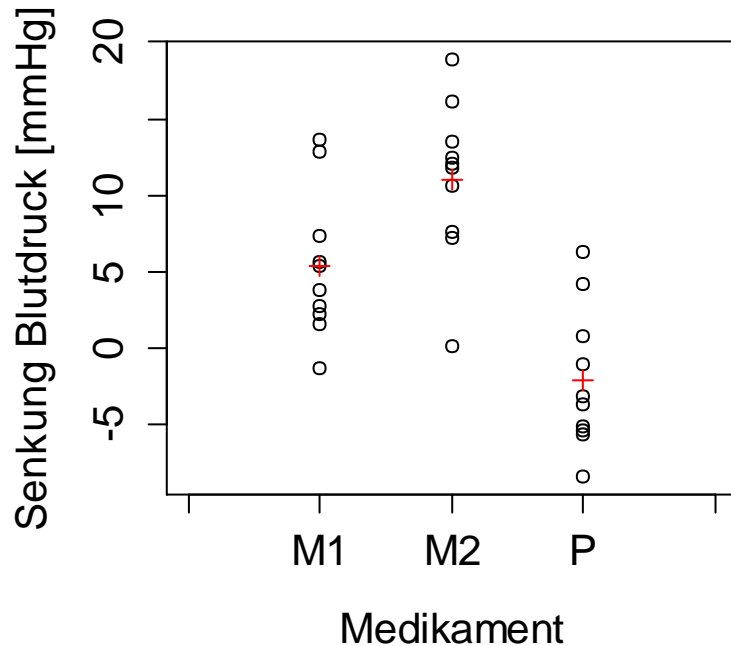
“Degrees of freedom (Df)”

$$T \sim F_{g-1, g*(p-1)}$$

- Damit kann ein Hypothesentest mit den üblichen 6 Schritten durchgeführt werden

# Wdh: 1-weg ANOVA-Tabelle

$g = 3, p = 10$



$$SS_B = 872.3$$

$$SS_W = 642.1$$

$$F = \frac{436.1}{23.8} = 18.34$$

```
> fm <- aov(y ~ g, data = df)
> summary(fm)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
g	2	872.3	436.1	18.34	9.32e-06 ***
Residuals	27	642.1	23.8		

$$g - 1 = 2$$

$$g^*(p-1) = 27$$

$$MS_B = \frac{872.3}{2} = 436.1$$

$$MS_W = \frac{642.1}{27} = 23.8$$

## 2-weg ANOVA: Modell ohne Interaktion

- Oft gibt es mehr als einen Faktor.
- Bsp:
  - Medikament (Faktor  $M$ :  $M_1$ - Medikament,  $M_2$ - Placebo)
  - Geschlecht (Faktor  $G$ :  $G_1$ - Mann,  $G_2$ - Frau)
- Das einfachste Modell ist dann (**ohne Interaktion**):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

Messung von Person  $k$  mit  
Medikament  $i$  und Geschlecht  $j$

Effekt von Medikament  $i$

Effekt von Geschlecht  $j$

Fehlerterm

## 2-weg ANOVA: Modell mit Interaktion

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim N(0, \sigma^2) \text{ iid}$$

Effekt von  
Medikament

Effekt von  
Geschlecht

Evtl. Interaktion:  
Geschlechtsspezifischer  
Effekt von Medikament

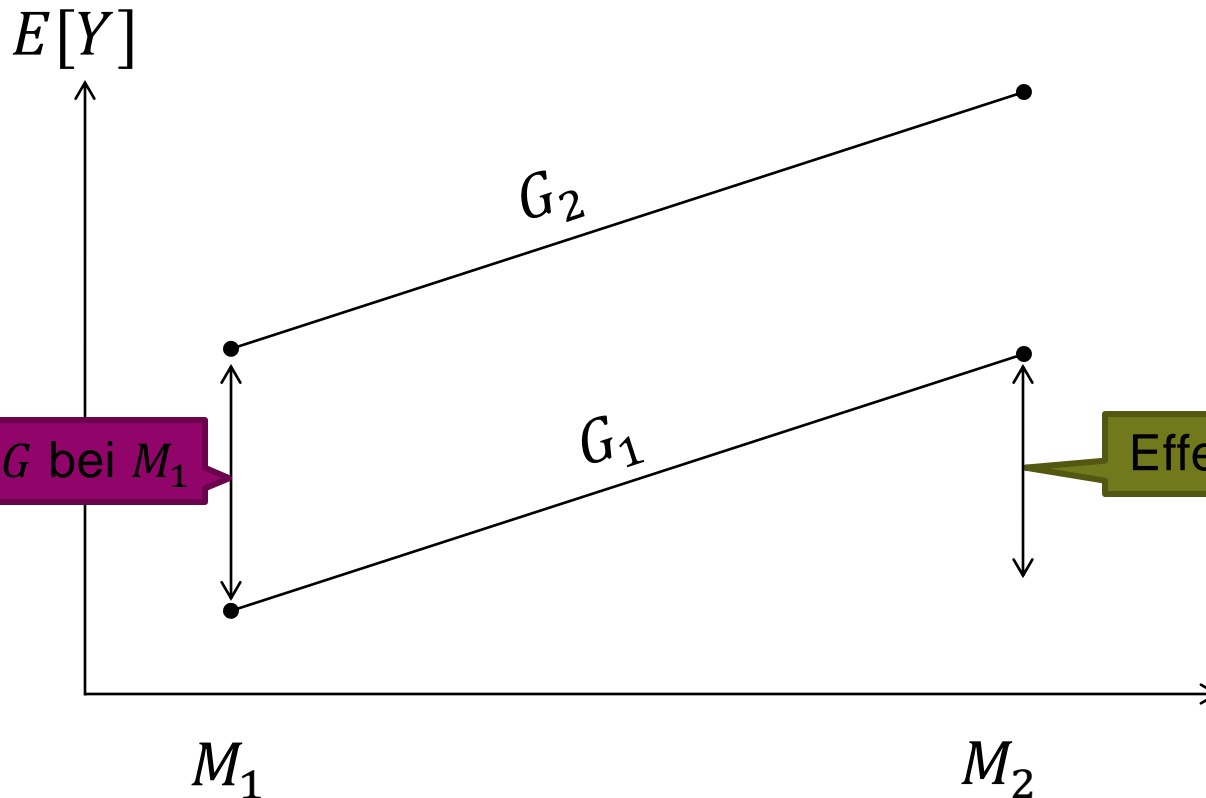
### 3 Nullhypothesen:

- $H_{0,1}: \alpha_i = 0$  für alle  $i \rightarrow$  Kein Medikamenten-Effekt
- $H_{0,2}: \beta_j = 0$  für alle  $j \rightarrow$  Kein Geschlechter-Effekt
- $H_{0,3}: \delta_{ij} = 0$  für alle  $i, j \rightarrow$  Kein Geschlechtsspezifischer Effekt von Medikament (keine Interaktion)

# Modell-Visualisierung: Ohne Interaktion

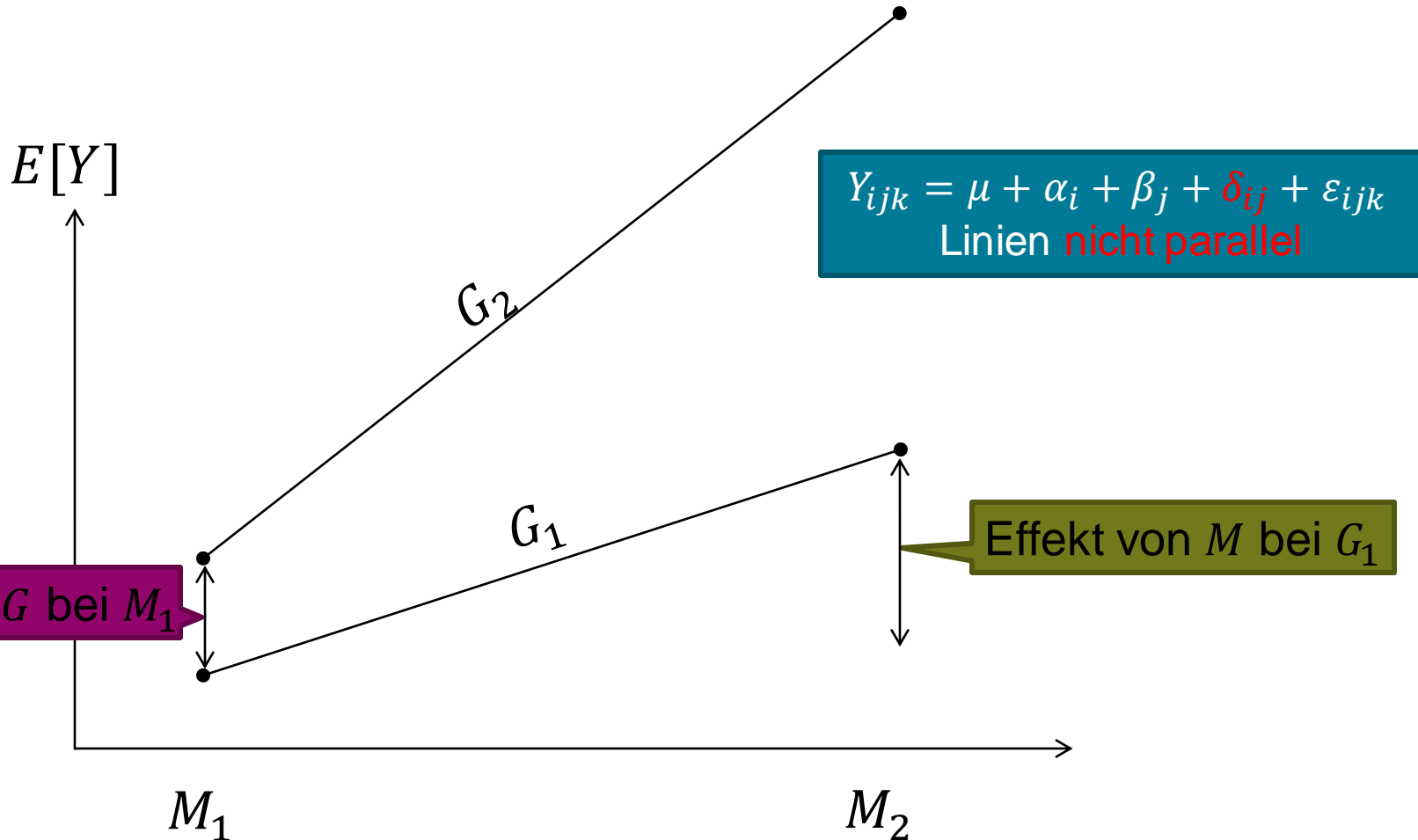
$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

Linien parallel





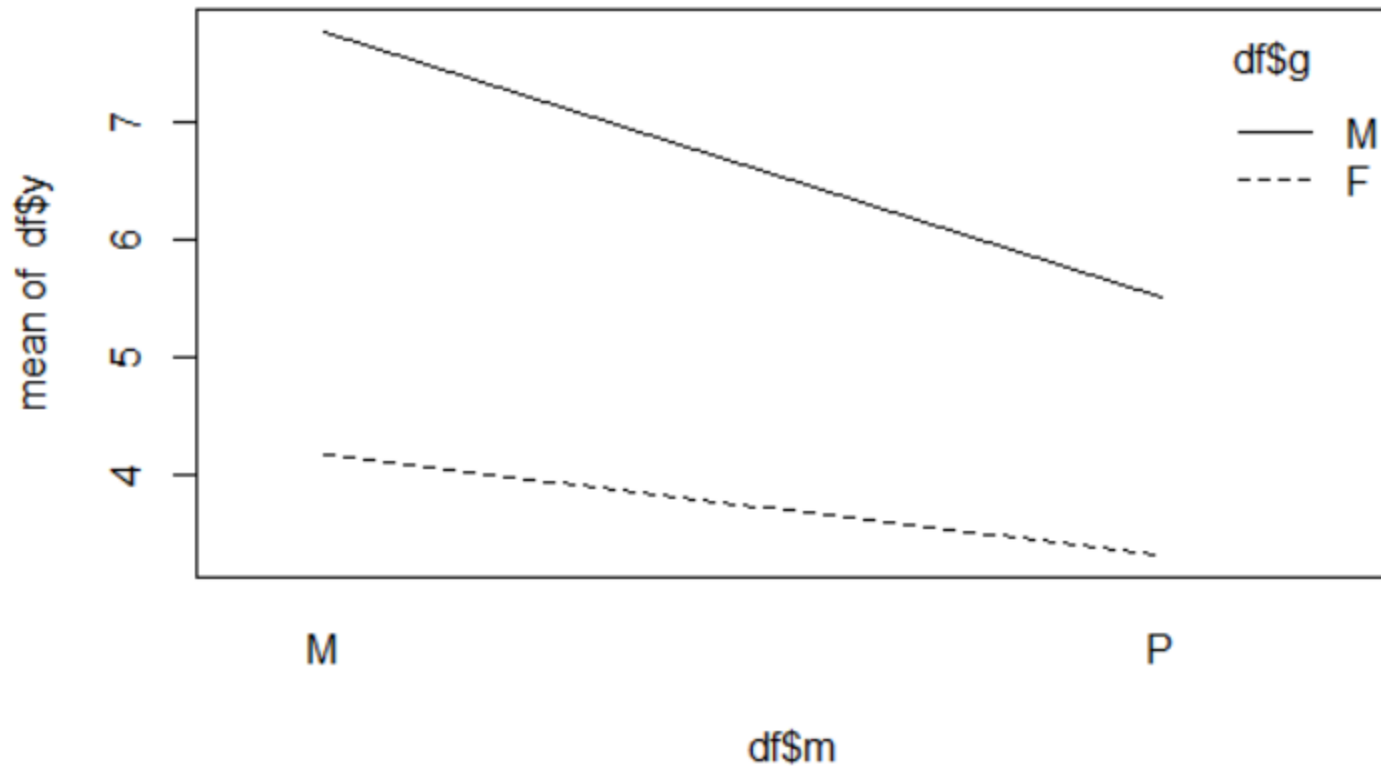
# Modell-Visualisierung: Mit **Interaktion**





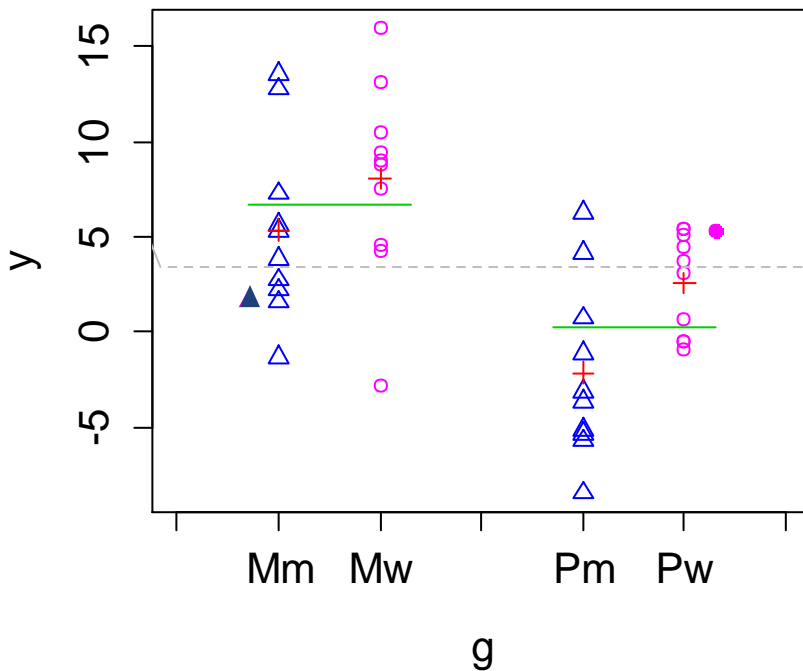
# Interaktionsplot in R

- Funktion 'interaction.plot'



## 2-weg ANOVA: Test 1 / 2

$p$ : Gruppengrösse (10)  
 $g$ : Anz. Geschlechter (2)  
 $m$ : Anz. Medikamente (2)



Keine WW:  $Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$   
 WW:  $Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}$

" $SS_B$ "

$$SS_M = p * g * \sum_{i=1}^m (\bar{Y}_{i..} - \bar{Y}_{...})^2$$

$$SS_G = p * m * \sum_{j=1}^g (\bar{Y}_{.j.} - \bar{Y}_{...})^2$$

$$SS_{MG} = p * \sum_{i=1}^m \sum_{j=1}^g (\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$$

$$SS_{Res} = \sum_{i=1}^m \sum_{j=1}^g \sum_{k=1}^p (Y_{ijk} - \bar{Y}_{ij.})^2$$

" $SS_W$ "

Teststatistiken:  $H_{0,1}: \approx \frac{SS_M}{SS_{Res}}$ ;  $H_{0,2}: \approx \frac{SS_G}{SS_{Res}}$ ;  $H_{0,3}: \approx \frac{SS_{MG}}{SS_{Res}}$

## 2-weg ANOVA: Test 2 / 2

- Sum of Squares:  $SS_M, SS_G, SS_{MG}, SS_{Res}$

- Degrees of Freedom

$$df_M: m - 1; df_G: g - 1; df_{MG}: (m - 1) * (g - 1); df_{Res}: m * g * (p - 1)$$

- Mean Squares:

$$MS_M = \frac{SS_M}{df_M}; MS_G = \frac{SS_G}{df_G}; MS_{MG} = \frac{SS_{MG}}{df_{MG}}; MS_{Res} = \frac{SS_{Res}}{df_{Res}}$$

- Teststatistik und Verteilung unter  $H_{0,1}$ ,  $H_{0,2}$  und  $H_{0,3}$ :

$$\text{Falls } H_{0,1} \text{ stimmt: } T_1 = \frac{MS_M}{MS_{Res}} \sim F_{df_M; df_{Res}}$$

$$\text{Falls } H_{0,2} \text{ stimmt: } T_2 = \frac{MS_G}{MS_{Res}} \sim F_{df_G; df_{Res}}$$

$$\text{Falls } H_{0,3} \text{ stimmt: } T_3 = \frac{MS_{MG}}{MS_{Res}} \sim F_{df_{MG}; df_{Res}}$$



## 2-weg ANOVA: Tabelle

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
m	1	19.62	19.62	28.76	1.03e-05 ***
g	1	66.96	66.96	98.14	1.18e-10 ***
m:g	1	3.95	3.95	5.79	0.023 *
Residuals	28	19.10	0.68		

Medikament hat Effekt

Geschlecht hat Effekt

Effekt vom Medikament hängt vom Geschlecht ab



# Effektstärke: ANOVA & TukeyHSD

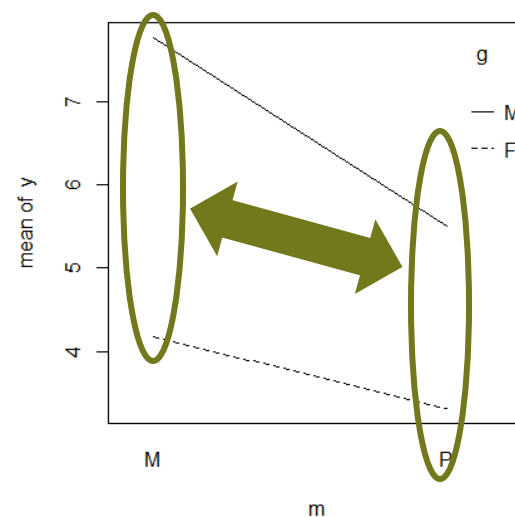
Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: `aov(formula = y ~ m * g, data = df)`

\$m	diff	lwr	upr	p adj
P-M	-1.566005	-2.164193	-0.9678171	1.03e-05

\$g	diff	lwr	upr	p adj
M-F	2.893037	2.294849	3.491225	0

\$`m:g`	diff	lwr	upr	p adj
P:F-M:F	-0.8633291	-1.9909120	0.2642538	0.1808921
M:M-M:F	3.5957128	2.4681299	4.7232956	0.0000000
P:M-M:F	1.3270321	0.1994492	2.4546150	0.0163777
M:M-P:F	4.4590418	3.3314589	5.5866247	0.0000000
P:M-P:F	2.1903611	1.0627783	3.3179440	0.0000683
P:M-M:M	-2.2686807	-3.3962636	-1.1410978	0.0000409



# Effektstärke: ANOVA & TukeyHSD

Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = y ~ m \* g, data = df)

\$m

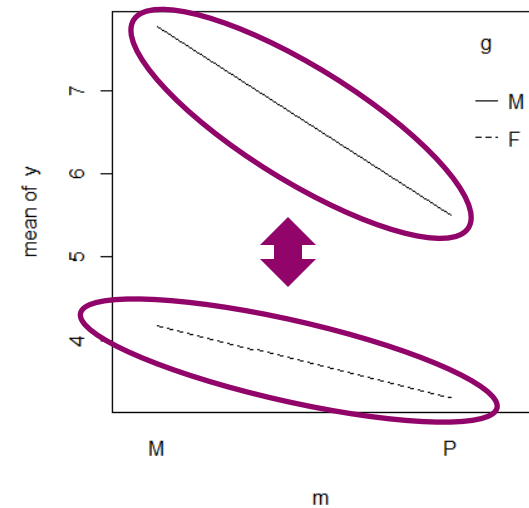
	diff	lwr	upr	p adj
P-M	-1.566005	-2.164193	-0.9678171	1.03e-05

\$g

	diff	lwr	upr	p adj
M-F	2.893037	2.294849	3.491225	0

\$`m:g`

	diff	lwr	upr	p adj
P:F-M:F	-0.8633291	-1.9909120	0.2642538	0.1808921
M:M-M:F	3.5957128	2.4681299	4.7232956	0.0000000
P:M-M:F	1.3270321	0.1994492	2.4546150	0.0163777
M:M-P:F	4.4590418	3.3314589	5.5866247	0.0000000
P:M-P:F	2.1903611	1.0627783	3.3179440	0.0000683
P:M-M:M	-2.2686807	-3.3962636	-1.1410978	0.0000409



# Effektstärke: ANOVA & TukeyHSD

Tukey multiple comparisons of means  
95% family-wise confidence level

Fit: aov(formula = y ~ m \* g, data = df)

\$m

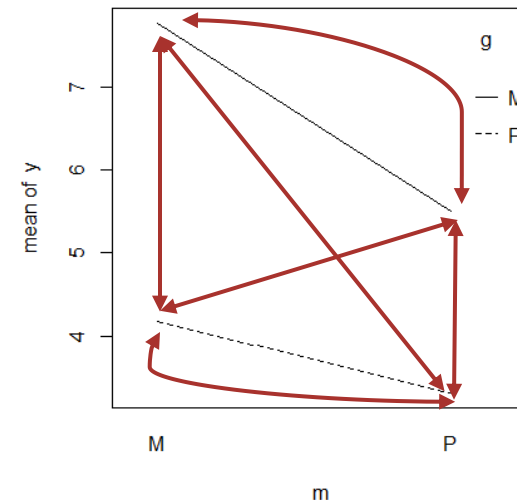
	diff	lwr	upr	p adj
P-M	-1.566005	-2.164193	-0.9678171	1.03e-05

\$g

	diff	lwr	upr	p adj
M-F	2.893037	2.294849	3.491225	0

\$`m:g`

	diff	lwr	upr	p adj
P:F-M:F	-0.8633291	-1.9909120	0.2642538	0.1808921
M:M-M:F	3.5957128	2.4681299	4.7232956	0.0000000
P:M-M:F	1.3270321	0.1994492	2.4546150	0.0163777
M:M-P:F	4.4590418	3.3314589	5.5866247	0.0000000
P:M-P:F	2.1903611	1.0627783	3.3179440	0.0000683
P:M-M:M	-2.2686807	-3.3962636	-1.1410978	0.0000409

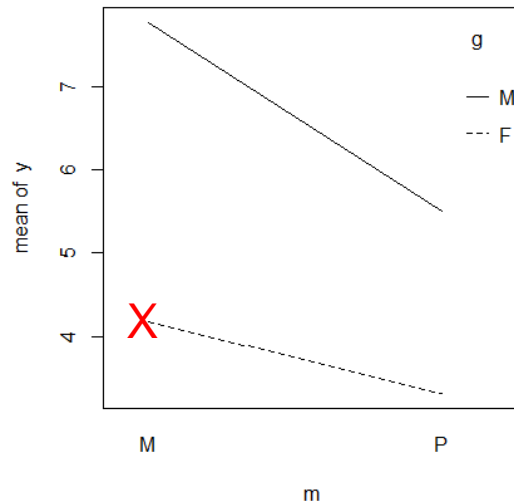






# Interpretation: ANOVA & TukeyHSD vs. Lineare Regression

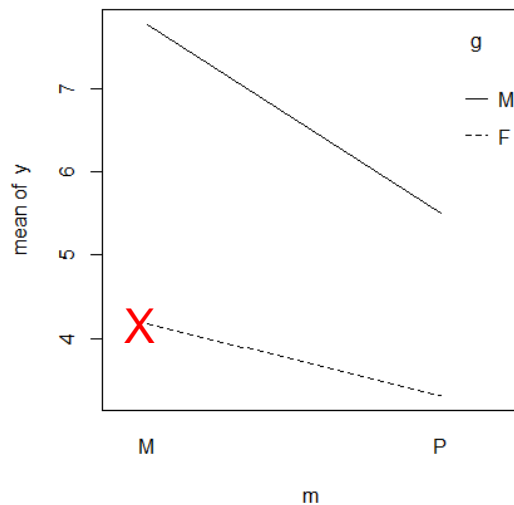
- Methoden technisch gesehen gleichwertig  
ABER: In der Praxis völlig unterschiedliche Interpretation
- ANOVA & TukeyHSD: “Totale Effekte”
- Lineare Regression: Effekte bzgl. **Referenzlevel**



Referenzlevel:  
Medikamentengruppe, Frauen

## Interpretation: Lineare Regression

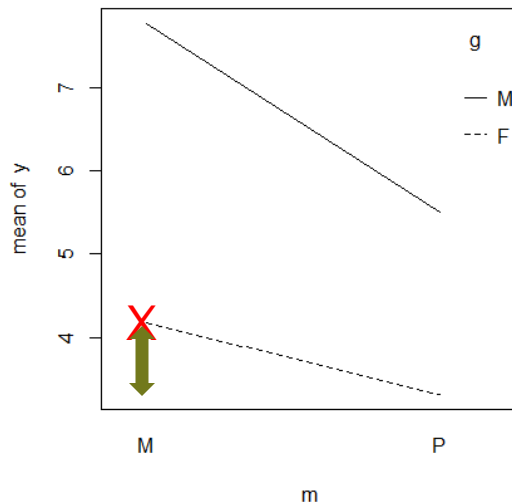
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.1792	0.2920	14.311	2.10e-14
mP	-0.8633	0.4130	-2.090	0.0458
gM	3.5957	0.4130	8.707	1.87e-09
mP : gM	-1.4054	0.5841	-2.406	0.0230



Wie gross ist  $E[Y]$  in Referenzgruppe ?  
(Keine Entsprechung in TukeyHSD)

# Interpretation: Lineare Regression

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.1792	0.2920	14.311	2.10e-14
mP	-0.8633	0.4130	-2.090	0.0458
gM	3.5957	0.4130	8.707	1.87e-09
mP : gM	-1.4054	0.5841	-2.406	0.0230

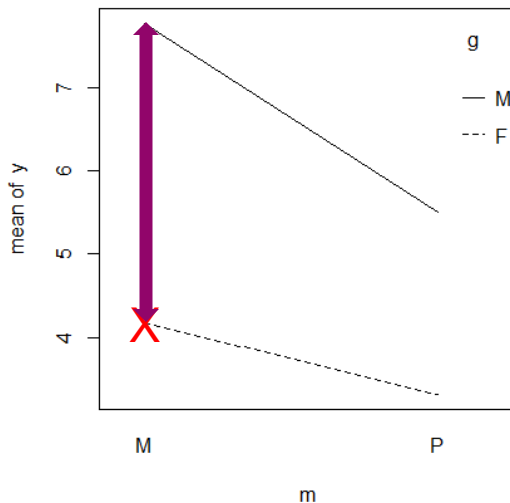


Wie ändert sich  $E[Y]$ , wenn man in der Referenzgruppe “Frauen” von “Medikament” zu “Placebo” wechselt?

(Entspricht P:F-M:F in TukeyHSD; VI & p-Wert wegen Korrektur für multiples Testen in TukeyHSD anders)

# Interpretation: Lineare Regression

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.1792	0.2920	14.311	2.10e-14
mP	-0.8633	0.4130	-2.090	0.0458
gM	3.5957	0.4130	8.707	1.87e-09
mP : gM	-1.4054	0.5841	-2.406	0.0230

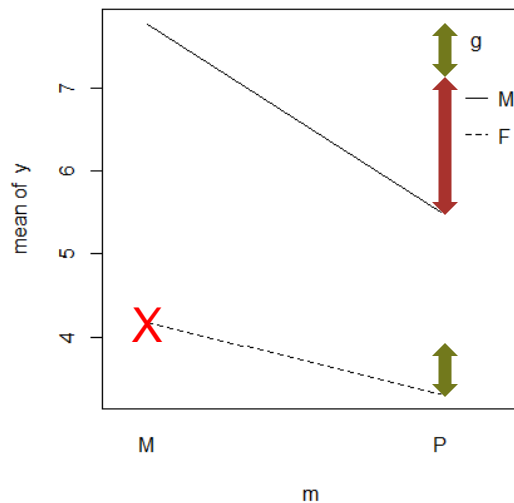


Wie ändert sich  $E[Y]$ , wenn man in der Referenzgruppe “Medikament” von “Frauen” zu “Männer” wechselt?

(Entspricht M:M-M:F in TukeyHSD; VI & p-Wert wegen Korrektur für multiples Testen in TukeyHSD anders)

# Interpretation: Lineare Regression

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	4.1792	0.2920	14.311	2.10e-14
mP	-0.8633	0.4130	-2.090	0.0458
gM	3.5957	0.4130	8.707	1.87e-09
mP : gM	-1.4054	0.5841	-2.406	0.0230

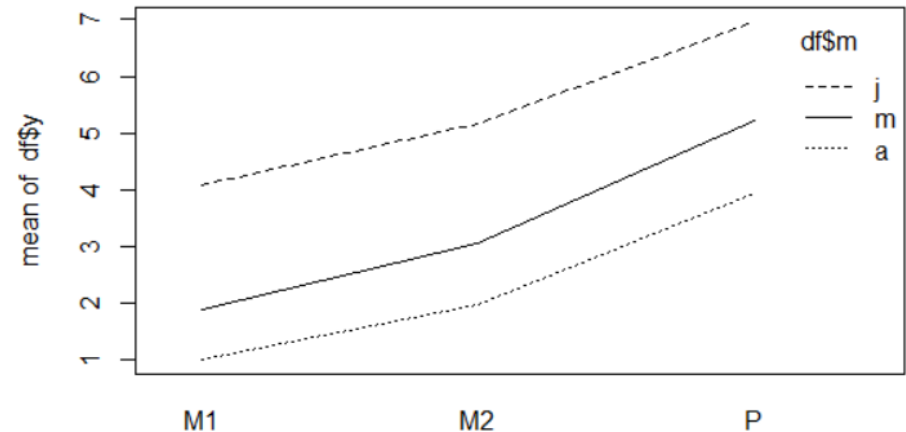


Um wieviel ist der Medikamenten-Effekt bei Männern anders als bei Frauen?

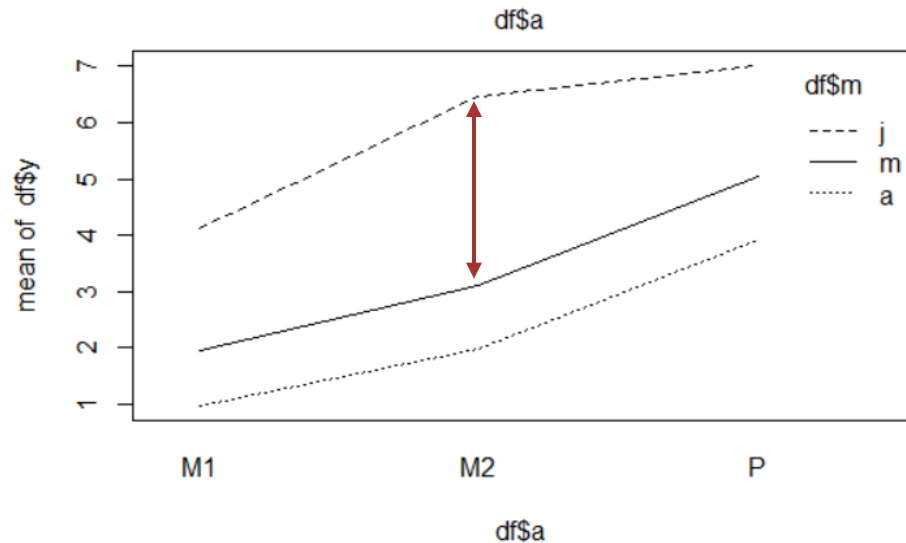
(Entspricht  $(P:M-M:M - P:F-M:F)$ ;  
kein entsprechendes VI oder p-Wert  
in TukeyHSD)

# Mehr als zwei Faktorstufen (Bsp: Empfinden nach Schmerzmittel)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
a	2	1436.2	718.1	670.657	<2e-16
m	2	1486.2	743.1	693.971	<2e-16
a:m	4	6.7	1.7	1.567	0.181
Residuals	891	954.1	1.1		



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
a	2	1347.8	673.9	705.24	<2e-16
m	2	2034.8	1017.4	1064.69	<2e-16
a:m	4	79.9	20.0	20.91	<2e-16
Residuals	891	851.4	1.0		



# Residuenanalyse bei ANOVA

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + \varepsilon_{ijk}, \quad \varepsilon_{ijk} \sim N(0, \sigma^2) \text{ iid}$$

1. Daten in jeder Gruppe normalverteilt
2. Gleiche Varianz in Gruppen
3. Unabhängige Fehler  $\varepsilon_{ij}$

In R: Funktion “plot” wie bei Linearer Regression

Vorteil: “Balanciertes Experiment” (gleiche Anzahl pro Gruppe):  
ANOVA ist robuster gegen Abweichungen obiger Annahmen

## Randomized Block Design:

### Verallgemeinerung des gepaarten t-Test

- Gepaarter t-Test:  
Pro Person Medikament & Placebo  
Reihenfolge Medi / Placebo pro Person zufällig
- Randomized Block Design:  
Pro Person **mehrere** Medikamente & Placebo  
Reihenfolge Medi / Placebo pro Person zufällig
- Auswertung: **2-weg ANOVA**  
 $Y \sim \text{Medikament} + \text{Person}$
- **Blockfaktor** (hier “Person”): Nicht von Interesse; nur um Streuung zu reduzieren
- Konvention: Keine Interaktion mit Blockfaktor



# Bsp: Randomized Block Design

## Medikament gegen Juckreiz

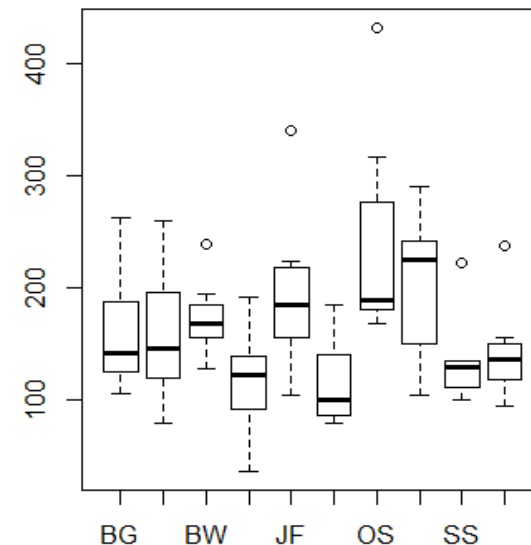
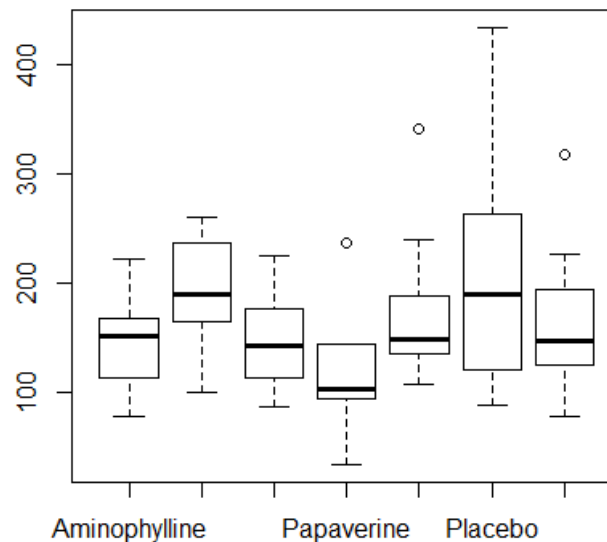
- 10 freiwillige Männer zw. 20 und 30
- Eine Behandlung pro Tag: Medikamentengabe; anschliessend Anwendung von Mittel, das starken Juckreiz auslöst (Juckbohne)
- Zielgrösse: Dauer des Juckreizes (in Sekunden)
- 5 Medikamente, 1 Placebo, einmal keine Behandlung
- Jede Person bekam jede Behandlung einmal; Reihenfolge zufällig





# Bsp: Juckreiz

	keine Behandlung	Placebo	Papaverine	Morphine	Aminophylline	Pentobarbital	Tripelennamine
BG	174	263	105	199	141	108	141
JF	224	213	103	143	168	341	184
BS	260	231	145	113	78	159	125
SI	255	291	103	225	164	135	227
BW	165	168	144	176	127	239	194
TS	237	121	94	144	114	136	155
GM	191	137	35	87	96	140	121
SS	100	102	133	120	222	134	129
MU	115	89	83	100	165	185	79
OS	189	433	237	173	168	188	317

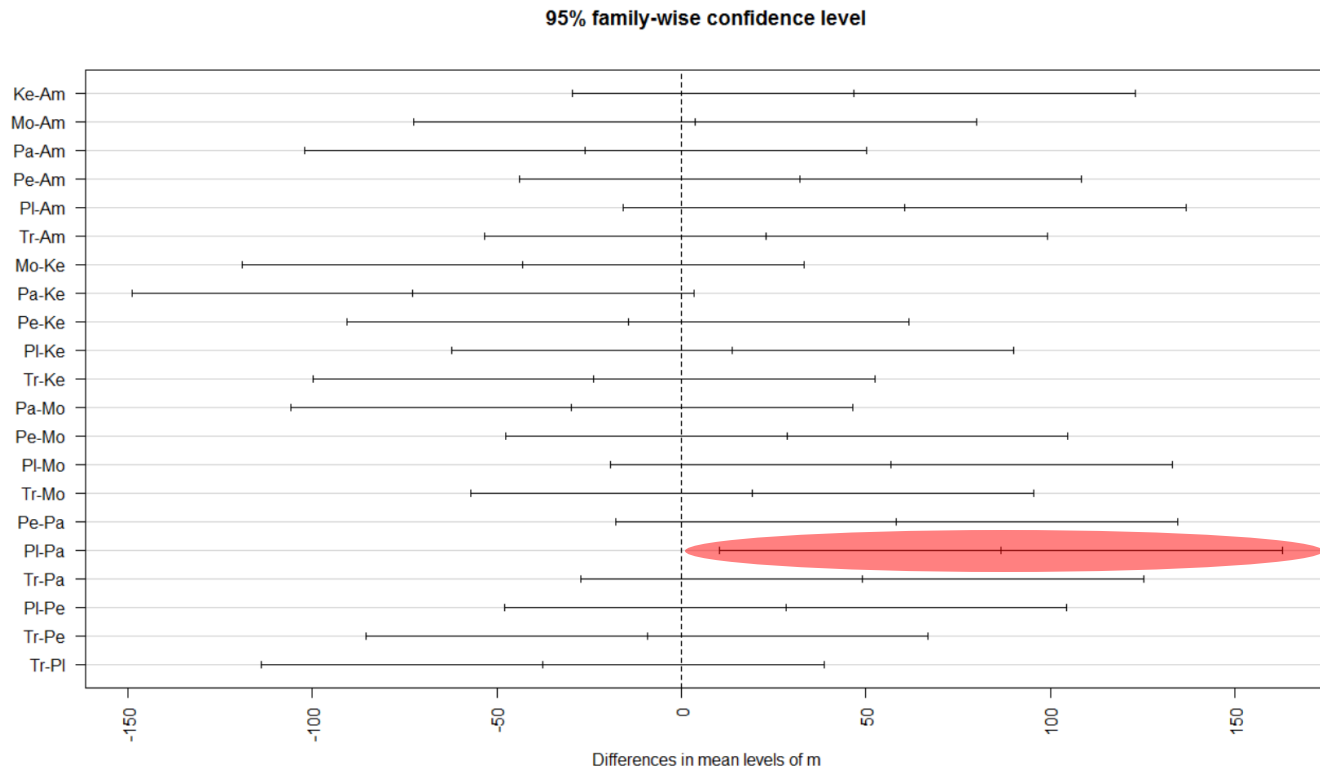


# Bsp: Juckreiz

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
m	6	53013	8835	2.855	0.01730
pers	9	103280	11476	3.708	0.00112
Residuals	54	167130	3095		

Es gibt sign. Unterschiede  
bzgl. Behandlungserfolg

Einzig zulässige Aussage:  
**Papavarine ist sign.  
wirksamer als Placebo.**





# Unbalanciertes Design

- **Bis jetzt “balanciertes” Design:**  
Alle Zellen enthalten die gleiche Anzahl Stichproben
- Mit einem balancierten Design kann man den Effekt eines Faktors bestimmen und dabei die **übrigen Faktoren ignorieren.**
- Bei einem unbalancierten Design gilt diese bequeme Regel nicht mehr.
- Faustregel: Balanciertes Design bevorzugen



# Unbalanciertes Design: Beispiel

Laufzeit für eine gewisse Strecke in Sekunden; ungepaart  
In Rot: Anzahl Beobachtungen pro Zelle

	<i>Energy Drink A</i>	<i>Energy Drink B</i>
<b>Men</b>	40.6, 49.7, 42.1, 42.2, 39.0, 44.2, 44.1, 43.1, 44.7, 46.3 <b>10</b>	49.7, 48.1, 49.7, 52.0, 51.5, 49.9, 55.6, 53.0, 53.5, 51.1 <b>10</b>
<b>Women</b>	55.7, 61.0, 58.0, 54.1, 51.9, 54.2, 54.4, 55.4, 55.4, 56.1 <b>10</b>	62.0, 60.3, 59.9, 61.2, 66.2, 56.5, 59.7, 63.0, 58.4, 61.7, 61.4, 62.6, 56.8, 55.2, 66.1, 60.6, 58.9, 59.1, 56.8, 62.5, 58.5, 61.3, 62.2, 62.5, 60.8, 57.1, 61.6, 65.9, 58.6, 60.6, 56.1, 53.6, 62.4, 62.2, 59.2, 62.9, 57.0, 58.5, 60.9, 63.4, <b>40</b>

# Unbalancierte Daten: Beispiel

- Wenn wir die Geschlechterstruktur ignorieren und den Effekt von Energy Drink schätzen, bekommen wir den falschen Effekt.
- Warum ? Es gibt mehr Frauen mit Energy Drink B.
- Die Variablen “Geschlecht” und “Energy Drink” sind korreliert
- Wenn wir von Drink A zu Drink B wechseln, wechseln wir automatisch zu einer **anderen Geschlechterstruktur**
- In einem balancierten Design wäre das nicht passiert.
- Parameter können nicht nacheinander sondern müssen **gleichzeitig geschätzt** werden.
- Analog: **Quadratsumme** (SS) kann nicht mehr einfach den einzelnen erklärenden Variablen zugewiesen werden

# Unbalancierte Daten: Modell Vergleich

- “Lösung” für das Problem der Quadratsummen: Modell Vergleiche
- Zur Wdh:  $SS_B$  ist die Reduktion der Sum of Squares, wenn der Faktor B dem Modell hinzugefügt wird.
- Balanciertes Design:  $SS_B$  ist immer gleich, egal ob A schon im Modell ist oder nicht
- Unbalanciertes Design:  $SS_B$  ist unterschiedlich, je nachdem welche anderen Variablen schon im Modell sind
- Notation:  $SS(B \mid 1, A)$  ist **Reduktion der Sum of Squares** wenn wir das Modell  $(1, A, B)$  mit dem Modell  $(1, A)$  vergleichen.

# Unbalancierte Daten

- Mögliche Variablen:  $(1, A, B, AB)$
- Mögliche Vergleiche:
  - $SS(A \mid 1)$
  - $SS(B \mid 1, A)$
  - $SS(A \mid 1, B)$
  - $SS(AB \mid 1, A, B)$
  - $SS(A \mid 1, B, AB)$
  - ...
- $SS_E$  oder  $MS_E$  werden mit dem Modell berechnet, das alle Terme enthält



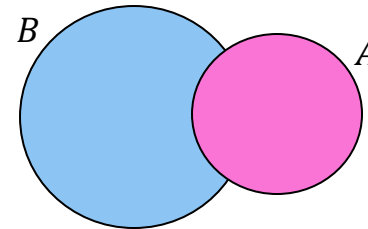


# Sum of Squares: "Typ 1" & "Typ 3"

## ■ Typ 1: Sequentiell

Modell wird sequentiell aufgebaut

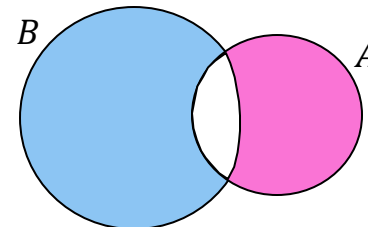
- $SS(A \mid 1)$
- $SS(B \mid 1, A)$
- Abhängig von der Reihenfolge !!!
- R: `aov`, `summary`



## ■ Typ 3: Adjustiert für alle übrigen Faktoren

Alle anderen Faktoren bleiben im Modell

- $SS(A \mid 1, B)$
- $SS(B \mid 1, A)$
- R: `drop1`
- Wird üblicherweise bevorzugt.





# drop1 in R

- Bsp: `fm <- aov(y ~ drink)`

Mit nur einem erklärenden Faktor sind Typ1 und Typ3 identisch

```
> fm <- aov(y~drink)
> summary(fm)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
drink	1	1146	1145.9	41.69	1.33e-08 ***
Residuals	68	1869	27.5		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> drop1(fm)
```

SS: Vgl. Modellvorhersagen von  $y \sim 1$  und  $y \sim \text{drink}$

Single term deletions

Model:

$y \sim \text{drink}$

	Df	Sum of Sq	RSS	AIC
<none>			1869.0	233.93
drink	1	1145.9	3014.9	265.40

RSSdrink: Residuenquadratsumme;  
vgl. Beobachtungen und Modellvorhersagen von  $y \sim \text{drink}$

RSS1: Residuenquadratsumme;  
vgl. Beobachtungen und Modellvorhersagen von  $y \sim 1$

$y \sim \text{drink}$   
drink wurde "gedroppt":  $y \sim 1$

# Typ 1: Reihenfolge hat Einfluss falls unbalanciert

- Typ 3 (drop1): Reihenfolge egal, daher **bevorzugt**
- Typ 1 (summary): Reihenfolge hat Einfluss auf Ergebnis

```
> fit <- aov(y~gender+drink)
> summary(fit)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
gender	1	2024.0	2024.0	253.17	< 2e-16 ***
drink	1	455.2	455.2	56.94	1.58e-10 ***
Residuals	67	535.6	8.0		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> fit2 <- aov(y~drink+gender)
> summary(fit2) ## Residuenquadratsummen & p-Werte sind unterschiedlich !!!
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
drink	1	1145.9	1146	143.3	<2e-16 ***
gender	1	1333.4	1333	166.8	<2e-16 ***
Residuals	67	535.6	8		

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> |
```

# Typ 3: Reihenfolge hat keinen Einfluss falls unbalanciert

- Typ 3 (drop1): Reihenfolge egal, daher **bevorzugt**
- Typ 1 (summary): Reihenfolge hat Einfluss auf Ergebnis

```
> drop1(fit, test = "F")
Single term deletions
```

Model:

y ~ gender + drink

	Df	Sum of Sq	RSS	AIC	F value	Pr(>F)
<none>			535.64	148.45		
gender	1	1333.41	1869.05	233.93	166.789	< 2.2e-16 ***
drink	1	455.25	990.89	189.51	56.944	1.583e-10 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

```
> drop1(fit2, test = "F")
```

Single term deletions

Model:

y ~ drink + gender

	Df	Sum of Sq	RSS	AIC	F value	Pr(>F)
<none>			535.64	148.45		
drink	1	455.25	990.89	189.51	56.944	1.583e-10 ***
gender	1	1333.41	1869.05	233.93	166.789	< 2.2e-16 ***

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Take home message:

## Balanciert vs. unbalanciert

- ANOVA hat eine “Schwäche”, falls Daten unbalanciert
  - p-Werte in `summary()` hängen von Reihenfolge ab
  - p-Werte in `drop1()` nicht → bevorzugen
- `lm()` hat keine vergleichbare Schwäche
  - output von `summary()` ist bzgl. Reihenfolge stabil