Statistik für Digital Humanities

Vergleich mehrerer Mittelwerte: ANOVA (GLM 1/5)

Dr. Jochen Tiepmar

Institut für Informatik
Computational Humanities
Universität Leipzig

16. Dezember 2019

[Letzte Aktualisierung: 11/02/2020, 11:10]

Agenda

Wiederholung t-Test

- Sind 2 Mittelwerte signifikant verschieden?

2 Arten von t-Tests

- Independent-means t-Test $\rightarrow$ Gruppendesign (verschiedene Probanden) Auch Independent-Measures oder Independent Samples t-Test
- Dependent-means t-Test $\rightarrow$ Messwiederholungsdesign (gleiche Probanden) Auch Matched Pairs oder Paired Samples t-Test

Wiederholung t-Test

- Sind 2 Mittelwerte signifikant verschieden?

2 Arten von t-Tests

- Independent-means t-Test → Gruppendesign (verschiedene Probanden) Auch Independent-Measures oder Independent Samples t-Test
- Dependent-means t-Test → Messwiederholungsdesign (gleiche Probanden) Auch Matched Pairs oder Paired Samples t-Test

Warum nicht mehrere t-Tests durchführen um mehrere Mittelwerte zu vergleichen?

Statistischer Fehler beim Vergleich mehrerer Mittelwerte

		Wirklichkeit	
		H_0 ist wahr	H_1 ist wahr
Entscheidung des Tests	für H_0	Spezifität True Positive Wahrscheinlichkeit: $1 - \alpha$	**Fehler 2. Art** False Negative Wahrscheinlichkeit: β
	für H_1	**Fehler 1. Art** False Positive Wahrscheinlichkeit: α	Sensitivität, Trennschärfe True Negative Wahrscheinlichkeit: $1\text{-}\beta$

Statistischer Fehler beim Vergleich mehrerer Mittelwerte

		Wirklichkeit	
		H_0 ist wahr	H_1 ist wahr
Entscheidung des Tests	für H_0	Spezifität True Positive Wahrscheinlichkeit: $1 - \alpha$	**Fehler 2. Art** False Negative Wahrscheinlichkeit: β
	für H_1	**Fehler 1. Art** False Positive Wahrscheinlichkeit: α	Sensitivität, Trennschärfe True Negative Wahrscheinlichkeit: $1\text{-}\beta$

- Problem: familienbezogene / experimentbezogene Fehlerrate $1 - (1 - \alpha)^k$ mit $k = \textit{Anzahl der Kombinationen}$ α ist die Typ 1 Fehlerwahrscheinlichkeit
- vergößert sich mit jedem zusätzlichen Test
  - 3 Gruppen $\rightarrow$ 3 Tests $\rightarrow$ Wahrscheinlichkeit für Typ 1 Fehler $1 - 0.95^3 = 0.143$
  - 5 Gruppen: $\rightarrow$ 10 Kombinationen $\rightarrow$ Wahrscheinlichkeit für Typ 1 Fehler $1 - 0.95^{10} = 0.401$, also 40%

Agenda

ANOVA Berechnung

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

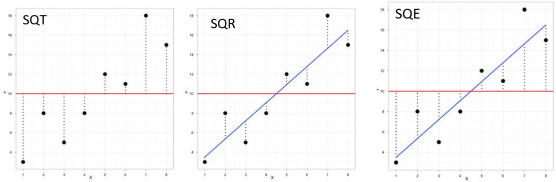
ANOVA Berechnung

- $H_0 =$ Alle Mittelwerte sind gleich / sehr 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-Ratio $F = \frac{MQE}{MQR} = \frac{Systematische\ \ Variation}{Unsystematische\ \ Variation}$
  - $MQE = \frac{SQE}{k-1}$
  - $MQR = \frac{SQR}{n-k}$
  - $k =$ Anzahl der Gruppen

Wiederholung Regression

- Regression erlaubt Abschätzen von Y für neue Werte aus X
- Regressionsformel $\hat{Y} = (b_0 + b_1 * X_1 + b_2 * X_2 + \dots)$
- Als statistisches Modell hat eine Regressionslinie eine Fitness
  - Residuenquadratsumme, Erklärte Quadratsumme, R^2 = Verhältnis beider
  - F-Test möglich um Modell zu bewerten
  - t-Test möglich um Einflußstärke des Prädiktors zu bewerten

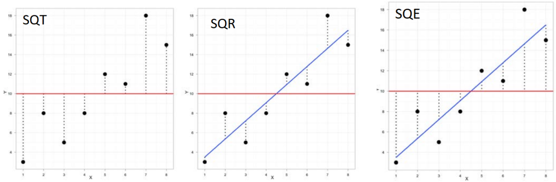
Wiederholung Abweichungsquadrate bei Regression



Abstände von Regression zu Beobachtung sind Residuen (Residuum)

- **Quadratsumme der Abweichungen** $SQT = \sum(y_i - \overline{y})^2$
- **Residuenquadratsumme** $SQR = \sum(y_i - \hat{y}_i)^2$
- **Erklärte Quadratsumme** $SQE = \sum(\overline{y} - \hat{y}_i)^2$
- $R^2 = \frac{SQE}{SQT}$

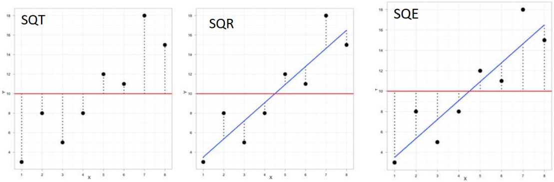
Wiederholung Abweichungsquadrate bei Regression



Abstände von Regression zu Beobachtung sind Residuen (Residuum)

- **Quadratsumme der Abweichungen** $SQT = \sum(y_i - \overline{y})^2$
- **Residuenquadratsumme** $SQR = \sum(y_i - \hat{y}_i)^2$
- **Erklärte Quadratsumme** $SQE = \sum(\overline{y} - \hat{y}_i)^2$
- $R^2 = \frac{SQE}{SQT}$
- MQx = Mittelwert der Quadrate von x
- $MQE = \frac{SQE}{Variablenanzahl}$
- $MQR = \frac{SQR}{Beobachtungen - Regressionskoeffizienten}$

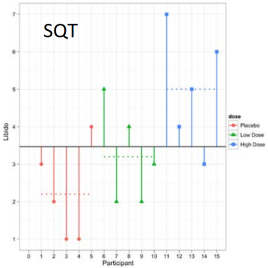
Wiederholung Abweichungsquadrate bei Regression

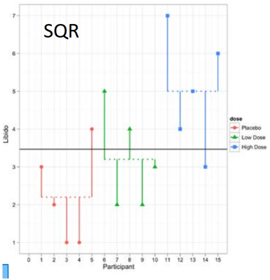


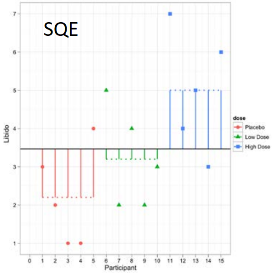
Abstände von Regression zu Beobachtung sind Residuen (Residuum)

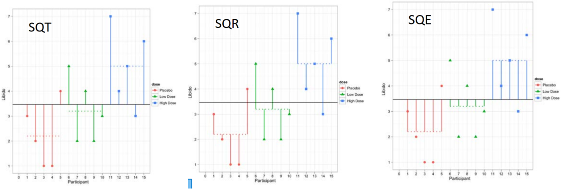
- **Quadratsumme der Abweichungen** $SQT = \sum(y_i - \overline{y})^2$
- **Residuenquadratsumme** $SQR = \sum(y_i - \hat{y}_i)^2$
- **Erklärte Quadratsumme** $SQE = \sum(\overline{y} - \hat{y}_i)^2$
- $R^2 = \frac{SQE}{SQT}$
- MQx = Mittelwert der Quadrate von x
- $MQE = \frac{SQE}{Variablenanzahl}$
- $MQR = \frac{SQR}{Beobachtungen - Regressionskoeffizienten}$
- F-Ratio $F = \frac{MQE}{MQR}$

Abweichungsquadrate bei mehreren Mittelwerten









- **Quadratsumme der Abweichungen**
  $SQT = \sum(x_i - Grand\ Mean)^2$ //Im Buch SS_T
- **Residuenquadratsumme**
  $SQR = \sum(x_i - \overline{group})^2 = \sum s^2_{group} * (n_{group} - 1)$ //Im Buch SS_R
- **Erklärte Quadratsumme**
  $SQE = \sum n_{group} * (\overline{group} - Grand\ Mean)^2$ //Im Buch SS_M

$SQT = SQR + SQE$

F-Ratio

- F-Ratio $F = \frac{MQE}{MQR} = \frac{Systematische\ \ Variation}{Unsystematische\ \ Variation}$
- $MQE = \frac{SQE}{k-1}$
- $MQR = \frac{SQR}{n-k}$
- $k =$ Anzahl der Gruppen

Interpretation:

F-Ratio

- F-Ratio $F = \frac{MQE}{MQR} = \frac{Systematische\ \ Variation}{Unsystematische\ \ Variation}$
- $MQE = \frac{SQE}{k-1}$
- $MQR = \frac{SQR}{n-k}$
- $k =$ Anzahl der Gruppen

Interpretation:

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

F-Ratio

- F-Ratio $F = \frac{MQE}{MQR} = \frac{Systematische\ \ Variation}{Unsystematische\ \ Variation}$
- $MQE = \frac{SQE}{k-1}$
- $MQR = \frac{SQR}{n-k}$
- $k =$ Anzahl der Gruppen

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 $\rightarrow$

F-Ratio

- F-Ratio $F = \frac{MQE}{MQR} = \frac{Systematische\ \ Variation}{Unsystematische\ \ Variation}$
- $MQE = \frac{SQE}{k-1}$
- $MQR = \frac{SQR}{n-k}$
- $k =$ Anzahl der Gruppen

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 gleich / sehr ähnlich
- $df(Numerator) = k - 1$
- $df(Denominator) = n - k$

Agenda

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G.\ Mean = 30.555$

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G.\ Mean = 30.555$
$SQE = \sum n_{group} * (\overline{group} - G.\ Mean)^2 = 3 * (36.667 - 30.555)^2 + 3 * (41.667 - 30.555)^2 + 3 * (13.333 - 30.555)^2 = 1372.291$
$SQR = \sum (x_i - \overline{group})^2 = 116.667 + 116.667 + 216.667 = 450$

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G.\ Mean = 30.555$
$SQE = \sum n_{group} * (\overline{group} - G.\ Mean)^2 = 3 * (36.667 - 30.555)^2 + 3 * (41.667 - 30.555)^2 + 3 * (13.333 - 30.555)^2 = 1372.291$
$SQR = \sum (x_i - \overline{group})^2 = 116.667 + 116.667 + 216.667 = 450$

$MQE = \frac{SQE}{k-1} = \frac{1372.291}{2} = 686.145$
$MQR = \frac{SQR}{n-k} = \frac{449.998}{6} = 75$

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G. \ Mean = 30.555$
$SQE = \sum n_{group} * (\overline{group} - G. \ Mean)^2 = 3 * (36.667 - 30.555)^2 + 3 * (41.667 - 30.555)^2 + 3 * (13.333 - 30.555)^2 = 1372.291$
$SQR = \sum (x_i - \overline{group})^2 = 116.667 + 116.667 + 216.667 = 450$

$MQE = \frac{SQE}{k-1} = \frac{1372.291}{2} = 686.145$
$MQR = \frac{SQR}{n-k} = \frac{449.998}{6} = 75$
$F = \frac{MQE}{MQR} = \frac{686.145}{75} = 9.149 \ F > F_{kr95}(2,6) = 5.79 \rightarrow H_0$

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G.\ Mean = 30.555$
$SQE = \sum n_{group} * (\overline{group} - G.\ Mean)^2 = 3 * (36.667 - 30.555)^2 + 3 * (41.667 - 30.555)^2 + 3 * (13.333 - 30.555)^2 = 1372.291$
$SQR = \sum (x_i - \overline{group})^2 = 116.667 + 116.667 + 216.667 = 450$

$MQE = \frac{SQE}{k-1} = \frac{1372.291}{2} = 686.145$
$MQR = \frac{SQR}{n-k} = \frac{449.998}{6} = 75$
$F = \frac{MQE}{MQR} = \frac{686.145}{75} = 9.149$ $F > F_{kr95}(2,6) = 5.79 \rightarrow H_0$ kann abgewiesen werden $\rightarrow$

Beispiel

Zeichenlänge des Dokumententitels pro Autor

Autor 1	Autor 2	Autor 3
30	40	10
35	35	25
45	50	5

$\overline{Autor1} = 36.667, \overline{Autor2} = 41.667, \overline{Autor3} = 13.333, G.\ Mean = 30.555$
$SQE = \sum n_{group} * (\overline{group} - G.\ Mean)^2 = 3 * (36.667 - 30.555)^2 + 3 * (41.667 - 30.555)^2 + 3 * (13.333 - 30.555)^2 = 1372.291$
$SQR = \sum (x_i - \overline{group})^2 = 116.667 + 116.667 + 216.667 = 450$

$MQE = \frac{SQE}{k-1} = \frac{1372.291}{2} = 686.145$
$MQR = \frac{SQR}{n-k} = \frac{449.998}{6} = 75$
$F = \frac{MQE}{MQR} = \frac{686.145}{75} = 9.149\ F > F_{kr95}(2,6) = 5.79 \rightarrow H_0$ kann abgewiesen werden $\rightarrow$ Es gibt signifikante Unterschiede in den Mittelwerten

ANOVA in R

```
titlelength<-c(30,35,45,40,35,50,10,25,5)
group<-gl(3,3,labels=c("autor1","autor2","autor3"))
df<-data.frame(group,titlelength)

anovamodel<-aov(titlelength~group, data=df)
summary(anovamodel)
```

```
            Df Sum Sq Mean Sq F value Pr(>F)
group        2   1372   686.1   9.148 0.0151 *
Residuals    6    450    75.0
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$Pr(> F) < 0.05 \rightarrow H_0$ kann abgewiesen werden

Agenda

Robustheit von ANOVA

- ANOVA ist grundsätzlich erstmal paramatrisch
- Wird oft als robust angesehen…
- …aber Robustheit in vielen Fällen nicht untersucht und viele Hinweise dagegen
- Bei gleichen Gruppengrößen eher robust.
- Ausführlicher im Begleitmaterial "Andy Field - Robustheit von Anova" (Moodle)

Lokalisierung der Unterschiede

Wie finden wir jetzt heraus, welche Gruppen sich unterscheiden ohne die familienbezogene Fehlerrate zu stark ansteigen zu lassen?

Lokalisierung der Unterschiede

Wie finden wir jetzt heraus, welche Gruppen sich unterscheiden ohne die familienbezogene Fehlerrate zu stark ansteigen zu lassen?

- Geplante Kontrastierung: Spezifische Hypothese vorhanden
- Post Hoc Tests: Keine spezifische Hypothese vorhanden

Agenda

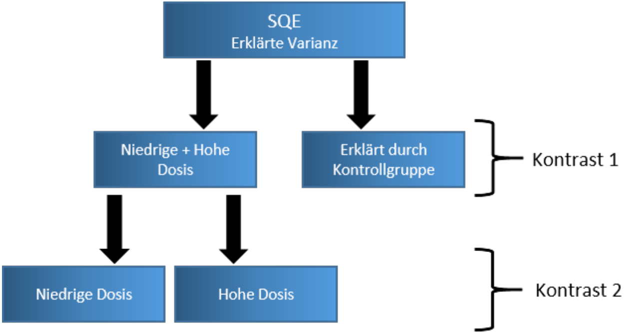
Geplante Kontrastierung

	Viagra	Libido
1	1	3
2	1	2
...		

	Viagra	Libido
6	2	5
7	2	2
...		

	Viagra	Libido
11	3	7
12	3	4
...		

Viagra: 1 = Kontrollgruppe, 2 = niedrige Dosierung, 3 = hohe Dosierung



Geplante Kontrastierung

- Wiederholter Vergleich zweier Variationen
- Hierarchisches Vorgehen
- Gezieltes Annähern anhand der Hypothese , die vorher festgelegt werden muss!

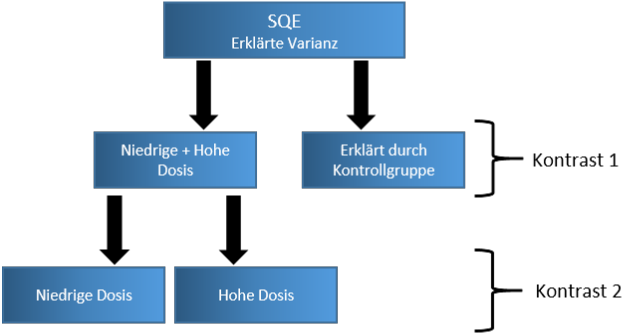
Geplante Kontrastierung

- Wiederholter Vergleich zweier Variationen
- Hierarchisches Vorgehen
- Gezieltes Annähern anhand der Hypothese , die vorher festgelegt werden muss!

3 Regeln:

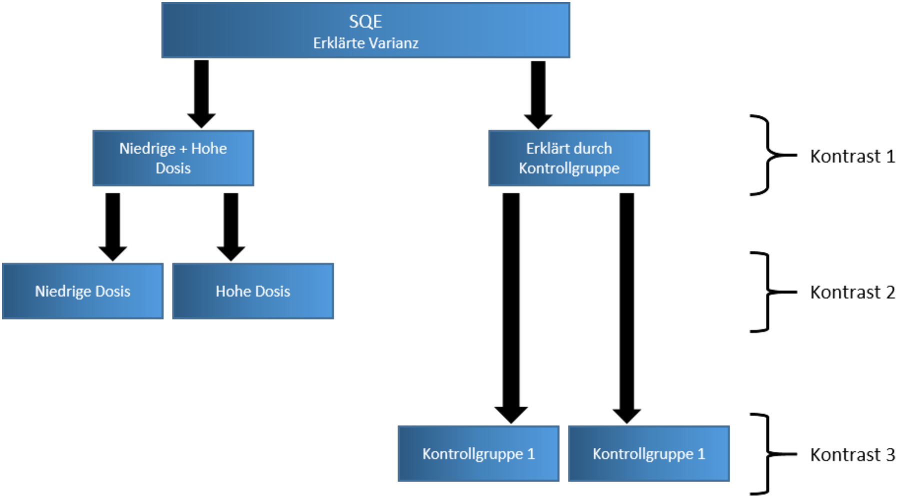
- Variation der Kontrollgruppen sollten gegen Variation der Experimentgruppen kontrastiert werden
- Jeder Kontrast darf nur 2 Variationen vergleichen
- Jede Gruppe darf nur Teil eines Kontrasts sein

Kontrastierung



- Auf übergeordneten Ebenen Vergleichswerte für zusammengelegte Gruppen neu berechnen
  - bspw. Mittelwerte für Niedrige + Hohe Dosis
- Kontrastierung mit Multipler Regression mittels Dummy Variablen

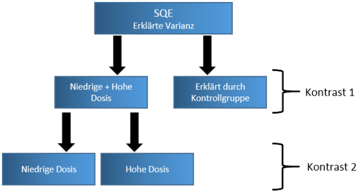
Kontraste bei 4 Gruppen (Bsp: 2 Kontrollgruppen)



Berechnung Orthogonaler Kontraste

Schritt 1: Tabelle aufstellen
Schritt 2:
Schritt 3:
Schritt 4:
(Schritt 5:)

Gruppe	Kontrast 1	Kontrast 2
Placebo		
niedrig		
hoch		



Berechnung Orthogonaler Kontraste

Schritt 1: Tabelle aufstellen

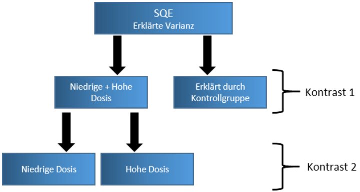
Schritt 2: Unbeteiligte Gruppen $\rightarrow 0$

Schritt 3: Linke Gruppe $+$, Rechte Gruppe $-$

Schritt 4:

(Schritt 5:)

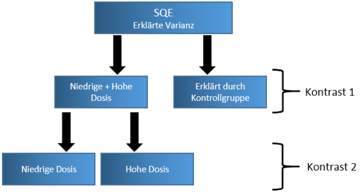
Gruppe	Kontrast 1	Kontrast 2
Placebo	-	0
niedrig	$+$	-
hoch	$+$	$+$



Berechnung Orthogonaler Kontraste

Schritt 1: Tabelle aufstellen
Schritt 2: Unbeteiligte Gruppen $\rightarrow$ 0
Schritt 3: Linke Gruppe $-$, Rechte Gruppe $+$
Schritt 4: Gegenüberliegende Gruppenzahl eintragen
(Schritt 5:)

Gruppe	Kontrast 1	Kontrast 2
Placebo	-2	0
niedrig	+1	-1
hoch	+1	+1



Berechnung Orthogonaler Kontraste

Schritt 1: Tabelle aufstellen
Schritt 2: Unbeteiligte Gruppen $\rightarrow$ 0
Schritt 3: Linke Gruppe $-$, Rechte Gruppe $+$
Schritt 4: Gegenüberliegende Gruppenzahl eintragen
(Schritt 5: Orthogonalität prüfen)

Gruppe	Kontrast 1	Kontrast 2	Kontrast 1 * Kontrast 1
Placebo	-2	0	0
niedrig	+1	-1	-1
hoch	+1	+1	+1
SUM	0	0	0

Berechnung Orthogonaler Kontraste

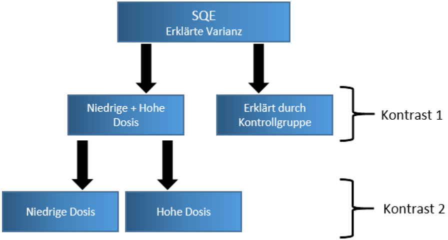
	Viagra	Libido
1	1	3
2	1	2
...		
6	2	5
7	2	2
...		
11	3	7
12	3	4
...		

$\rightarrow$

	Viagra	Libido	Kontr1	Kontr2
1	1	3	-2	0
2	1	2	-2	0
...				
6	2	5	1	-1
7	2	2	1	-1
...				
11	3	7	1	1
12	3	4	1	1
...				

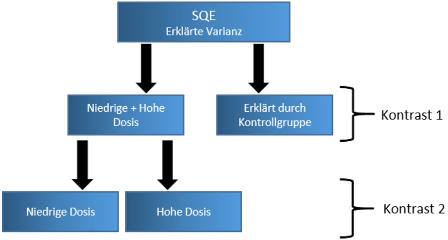
Kontrastierung über Multiple Regression

- Regressionsformel $\hat{Y} = (b_0 + b_1 * Kontrast1 + b_2 * Kontrast2 + \ldots)$



Kontrastierung über Multiple Regression

- Regressionsformel $\hat{Y} = (b_0 + b_1 * Kontrast1 + b_2 * Kontrast2 + \dots)$



- $b_0 = Grand\ \ Mean$
- $b_i = \frac{\overline{Links} - \overline{Rechts}}{Gruppenzahl_{Kontrast\ \ i}}$
- $b_1 = \frac{\overline{Niedrig + Hoch} - \overline{Kontrollgruppe}}{3}$, $b_2 = \frac{\overline{Niedrig} - \overline{Hoch}}{2}$

Kontrastierung über Multiple Regression

```
libido<-c(3,2,1,1,4,5,2,4,2,3,7,4,5,3,6)
dose<-gl(3,5, labels = c("Placebo", "Low Dose", "High Dose"))
viagraDataContrast<-data.frame(dose, libido)
viagraData<-data.frame(dose, libido)
contrast1<-c(-2,1,1)
contrast2<-c(0,-1,1)
contrasts(viagraDataContrast$dose)<-cbind(contrast1, contrast2)
viagraPlannedContrast<-aov(libido~dose, data = viagraDataContrast)
viagraPlanned<-aov(libido~dose, data = viagraData)

summary.lm(viagraPlannedContrast)

                 Estimate Std. Error t value Pr(>|t|)
    (Intercept)    3.4667     0.3621   9.574 5.72e-07 ***
    dosecontrast1  0.6333     0.2560   2.474   0.0293 *    //t-Wert f"ur Kontrast1
    dosecontrast2  0.9000     0.4435   2.029   0.0652 .    //t-Wert f"ur Kontrast2

summary.lm(viagraPlanned)                                  //Das hier ist falsch
    Coefficients:
                 Estimate Std. Error t value Pr(>|t|)
    (Intercept)    2.2000     0.6272   3.508  0.00432 **
    doseLow Dose   1.0000     0.8869   1.127  0.28158
    doseHigh Dose  2.8000     0.8869   3.157  0.00827 **
```

Agenda

Post Hoc Tests

Grundidee:

- Mache paarweise Tests
- Angepasste Signifikanz verhindert, dass der kummulierte Typ 1 Fehler im Toleranzbereich bleibt

Post Hoc Tests

Grundidee:

- Mache paarweise Tests
- Angepasste Signifikanz verhindert, dass der kummulierte Typ 1 Fehler im Toleranzbereich bleibt
- (naiv) Bonferroni Korrektur: $p_{kr} = \frac{\alpha}{Anzahl\ der\ Vergleiche}$

Post Hoc Tests

Grundidee:

- Mache paarweise Tests
- Angepasste Signifikanz verhindert, dass der kummulierte Typ 1 Fehler im Toleranzbereich bleibt
- (naiv) Bonferroni Korrektur: $p_{kr} = \frac{\alpha}{Anzahl\ der\ Vergleiche}$
- Pro Test steigt die Wahrscheinlichkeit von Typ 2 Fehlern (Vorhandener Effekt nicht festgestellt)

Post Hoc Tests

Verbesserung durch Abstufung: (Hochberg:1988, Holm:1979, Hommel:1988)
Grundidee Holm:1979 : Bereits abgearbeitete Vergleiche werden nicht mehr beachtet

- Berechne p für jeden Vergleich
- Sortiere p <u>absteigend</u> und vergebe Rang j
- $p_{kr} = \frac{\alpha}{j}$
- From j=6 TO j=1:
  if $p < p_{kr} \rightarrow$ Unterschied ist signifikant
- Stoppe sobald $p > p_{kr}$
- Alle weiteren Unterschiede sind nicht signifikant

Post Hoc Tests

Verbesserung durch Aufstufung: (Benjamini & Hochberg:1995)
Grundidee : Lieber Typ 2 Fehler als Typ 1 Fehler optimieren

- Berechne p für jeden Vergleich
- Sortiere p aufsteigend und vergebe Rang j
- $p_{kr} = \frac{j}{k} * \alpha$
- From j=6 TO j=1:
  if $p > p_{kr} \rightarrow$ Unterschied ist nicht signifikant
- Stoppe sobald $p < p_{kr}$
- Alle weiteren Unterschiede sind signifikant

Beispiel

Daten Siehe Superhero.dat im Moodle
* = signifikant

	p
NT,Super	.000
Super,Hulk	.0014
Spider,Super	.0127
NT,Spider	.0252
NT,Hulk	.1704
Spider,Hulk	.3431

Beispiel

Daten Siehe Superhero.dat im Moodle
* = signifikant

	p	Bonferroni $\frac{\alpha}{k}$
NT,Super	.000	.0083*
Super,Hulk	.0014	.0083*
Spider,Super	.0127	.0083
NT,Spider	.0252	.0083
NT,Hulk	.1704	.0083
Spider,Hulk	.3431	.0083

Beispiel

Daten Siehe Superhero.dat im Moodle

* = signifikant

	p	Bonferroni $\frac{\alpha}{k}$	j_1	Holm $\frac{\alpha}{j_1}$
NT,Super	.000	.0083*	6	.0083*
Super,Hulk	.0014	.0083*	5	.0100*
Spider,Super	.0127	.0083	4	.0125
NT,Spider	.0252	.0083	3	.0167
NT,Hulk	.1704	.0083	2	.0250
Spider,Hulk	.3431	.0083	1	.0500

Beispiel

Daten Siehe Superhero.dat im Moodle

* $=$ signifikant

	p	Bonferroni $\frac{\alpha}{k}$	j_1	Holm $\frac{\alpha}{j_1}$	j_2	B. & H. $\frac{j}{k} * \alpha$
NT,Super	.000	.0083*	6	.0083*	1	.0083*
Super,Hulk	.0014	.0083*	5	.0100*	2	.0167*
Spider,Super	.0127	.0083	4	.0125	3	.0250*
NT,Spider	.0252	.0083	3	.0167	4	.0333*
NT,Hulk	.1704	.0083	2	.0250	5	.0417
Spider,Hulk	.3431	.0083	1	.0500	6	.0500

Effektstärke und ANOVA

- Determinationskoeffizient $R^2 = \frac{SQE}{SQT}$
- Weniger Bias: $Omega^2$ $\omega^2 = \frac{SQE-(k*MSR)}{SQT+MSR}$ mit k = Anzahl der Gruppen

Effektstärke und ANOVA

- Determinationskoeffizient $R^2 = \frac{SQE}{SQT}$
- Weniger Bias: *Omega*^{2} $\omega^2 = \frac{SQE-(k*MSR)}{SQT+MSR}$ mit k = Anzahl der Gruppen
- Guidelines für ω^2 (Kirk, R.E.(1996) Practical Significance: A concept whose time has come, *Educational and Psychological Measurements*, 56(5), 746-759.)
    - .01 $\rightarrow$ gering
    - .06 $\rightarrow$ moderat
    - .14 $\rightarrow$ stark
    - sehr kontextabhängig

Zusammenfassung

- ANOVA vergleicht mehrere Mittelwerte als Omnibus Test
- F-Ratio
- Robustheit von ANOVA
- Familienbezogene Fehler bei paarweisen Vergleichen
- Lokalisierung von Unterschieden
    - Kontrastierung
    - Post Hoc Tests
- Determinationskoeffizient R^2 und quadriertes Omega ω^2 als Effektstärke
- Übersprungen: Weitere Kontrastierungen (Nicht-Orthogonal, Polynomial)