

ANOVA (Einfache)

Vitaly Belik

Institut für Veterinär-Epidemiologie und Biometrie, FU Berlin

24/11/2020

Falls die Antwortvariable (*response*) kontinuierlich ist, und die unabhängigen Variablen (*Faktoren*) kategoriell, macht es Sinn die Unterschiede zwischen den verschiedenen Gruppen von Faktorenniveaus oder Faktorenabstufungen zu betrachten.

Die statistischen Techniken für den Vergleich zwischen den Mittelwerten mehreren Gruppen bezeichnet man als **ANOVA**.

ANOVA kann und soll als GLM (verallgemeinerte *Regression*) betrachtet werden.

- ▶ Einfache ANOVA: eine erklärende (kategoriale) Variable (Faktor). Vergleiche zwischen den Niveaus von diesem Faktor.
- ▶ Zweifache (faktorielle) ANOVA: zwei erklärende (kategoriale) Variable (Faktor). Vergleiche zwischen den Niveaus von diesen Faktoren.
- ▶ Multiple ANOVA – **MANOVA**: mehrere unabhängige Variable.
- ▶ ANCOVA (Analysis of covariances): zusätzlich zur ANOVA-Situation kommt noch eine Kovariate für die abhängige Variable hinzu.
- ▶ Repeated measures (Messwiederholungen): bei den gleichen Objekten wurden die Messungen mehrmals vorgenommen.

ANOVA Tabelle

$$\hat{Y} = \beta_0 + \beta_1 X$$

Um zu beurteilen, wie viel Information die Variable X über die Variable Y enthält, betrachten wir, wie viel von der *Summe der Quadrate* (SS) der Variablen Y durch die Kenntnis der Variablen X erklärt werden könnte.

Table 1: Zerlegung von Quadratsummen im Regressionsmodell mit k Prädiktorvariablen [Dohoo, p. 327]

Source of variation	Sum of squares	Degrees of freedom	Mean square	F-test
Model	$SSM = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$	$dfM = k$	$MSM = \frac{SSM}{dfM}$	$\frac{MSM}{MSE}$
Error	$SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$	$dE = n - (k + 1)$	$MSE = \frac{SSE}{dE}$	
Total	$SST = \sum_{i=1}^n (Y_i - \bar{Y})^2$	$dfT = n - 1$	$MST = \frac{SST}{dfT}$	

$MSE = \sigma^2$ and σ is called *standard error of prediction*.

F-Test zur Beurteilung der Modellgüte

$H_0: \beta_1 = \beta_2 \dots \beta_k = 0, \beta_0 \neq 0$

H_1 : zumindest manche von Koeffizienten $\beta_i \neq 0, i \neq 0$

- ▶ Variablen werden so ausgewählt, dass die F -Statistik maximiert wird
- ▶ F -Test hat eine einfache Bedeutung, wenn unabhängige Variablen in einem kontrollierten Experiment manipuliert werden
- ▶ Vorsicht wird bei Beobachtungsvariablen geboten (beeinflusst von der Anzahl der Variablen, ihren Korrelationen und der Stichprobengröße)

Einfache ANOVA

ANOVA Beispiel

```
df <- read.table('/Users/vitaly/Lehre/WS1920/Beuth/Data files/Viagra.dat', sep = '', header = TRUE)
knitr::kable(df, caption = "Viagra-Datensatz [A. Field]")
```

Table 2: Viagra-Datensatz [A. Field]

person	dose	libido
1	1	3
2	1	2
3	1	1
4	1	1
5	1	4
6	2	5
7	2	2
8	2	4
9	2	2
10	2	3
11	3	7
12	3	4
13	3	5
14	3	3
15	3	6

ANOVA Beispiel (1)

```
library(gplots)
plotmeans(df$libido ~ df$dose)
```

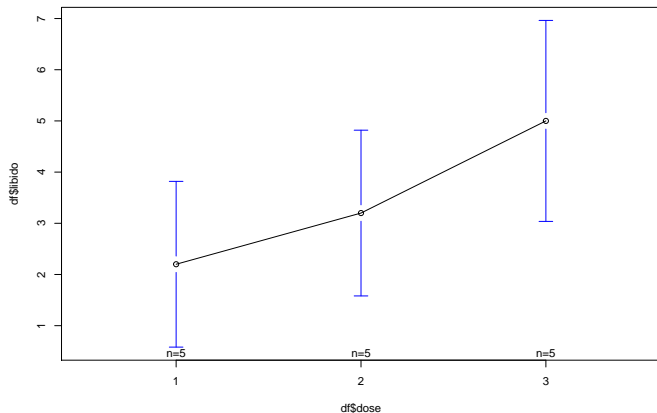


Figure 1: Viagra-Datensatz mit Konfidenzintervallen

ANOVA Beispiel (2)

```
library(pastecs)
by(df$libido, df$dose, stat.desc)
```

```
## df$dose: 1
##   nbr.val  nbr.null  nbr.na      min      max      range
##  5.0000000  0.0000000  0.0000000  1.0000000  4.0000000  3.0000000
##      sum      median      mean  SE.mean CI.mean.0.95      var
##  11.0000000  2.0000000  2.2000000  0.5830952  1.6189318  1.7000000
##   std.dev  coef.var
##   1.3038405  0.5926548
## -----
## df$dose: 2
##   nbr.val  nbr.null  nbr.na      min      max      range
##  5.0000000  0.0000000  0.0000000  2.0000000  5.0000000  3.0000000
##      sum      median      mean  SE.mean CI.mean.0.95      var
##  16.0000000  3.0000000  3.2000000  0.5830952  1.6189318  1.7000000
##   std.dev  coef.var
##   1.3038405  0.4074502
## -----
## df$dose: 3
##   nbr.val  nbr.null  nbr.na      min      max      range
##  5.0000000  0.0000000  0.0000000  3.0000000  7.0000000  4.0000000
##      sum      median      mean  SE.mean CI.mean.0.95      var
##  25.0000000  5.0000000  5.0000000  0.7071068  1.9632432  2.5000000
##   std.dev  coef.var
##   1.5811388  0.3162278
```

ANOVA Beispiel (Levene's test)

```
library(pastecs)
library(car)
leveneTest(df$libido, df$dose, center = median)

## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 2  0.1176  0.89
##      12
```

If the *Levene's test* is non-significant, the data variances are very similar.

Falls nicht signifikant, kann man *Welch's F* oder robuste ANOVA benutzen.

$$\text{libido}_i = \text{dose}_i + \text{error}$$

ANOVA Beispiel (*lm*)

```
df <- read.table('/Users/vitaly/Lehre/WS1920/Beuth/Data files/Viagra.dat', sep = ',', header = TRUE)
df$dose <- as.factor(df$dose)
library(ggplot2)
viagraModel <- lm(libido ~ dose, data = df)
summary(viagraModel)
```

```
##
## Call:
## lm(formula = libido ~ dose, data = df)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
##  -2.0  -1.2  -0.2    0.9    2.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.2000    0.6272   3.508  0.00432 **
## dose2       1.0000    0.8869   1.127  0.28158
## dose3       2.8000    0.8869   3.157  0.00827 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.402 on 12 degrees of freedom
## Multiple R-squared:  0.4604, Adjusted R-squared:  0.3704
## F-statistic: 5.119 on 2 and 12 DF,  p-value: 0.02469
```

ANOVA Beispiel (aov)

```
ViagraModel <- aov(libido ~ dose, data = df)
summary(ViagraModel)
```

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## dose         2  20.13  10.067    5.119 0.0247 *
## Residuals   12  23.60   1.967
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

F-Test prüft, ob Unterschiede zwischen den Gruppen vorliegen.

F-Test sagt uns nicht zwischen welchen Gruppen die Unterschiede vorliegen.

ANOVA Beispiel (aov 1)

```
library(car)
qqPlot(lm(libido ~ dose, data = df), simulate = TRUE, main = 'Q-Q Plot', labels = FALSE)
```

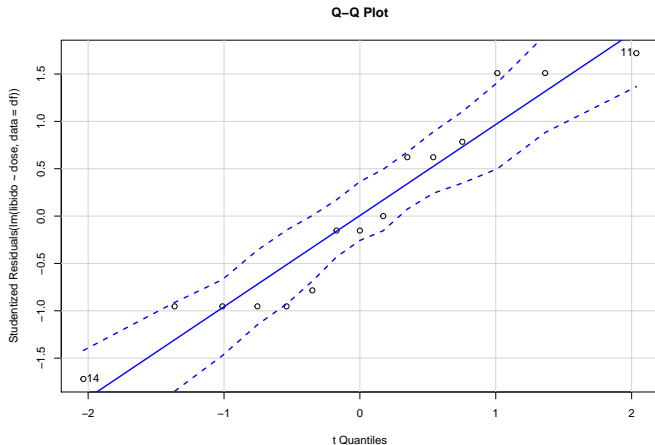


Figure 2: Q-Q Plot.

ANOVA Beispiel: Bartlett-Test

```
bartlett.test(libido ~ dose, data = df)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data: libido by dose  
## Bartlett's K-squared = 0.1853, df = 2, p-value = 0.9115
```

ANOVA Beispiel: Ausreißer

```
outlierTest(ViagraModel)
```

```
## No Studentized residuals with Bonferroni p < 0.05  
## Largest |rstudent|:  
##   rstudent unadjusted p-value Bonferroni p  
## 11  1.71959          0.11348          NA
```


ANOVA Beispiel (aov 3)

```
par(mfrow=c(2,2))  
plot(ViagraModel)
```

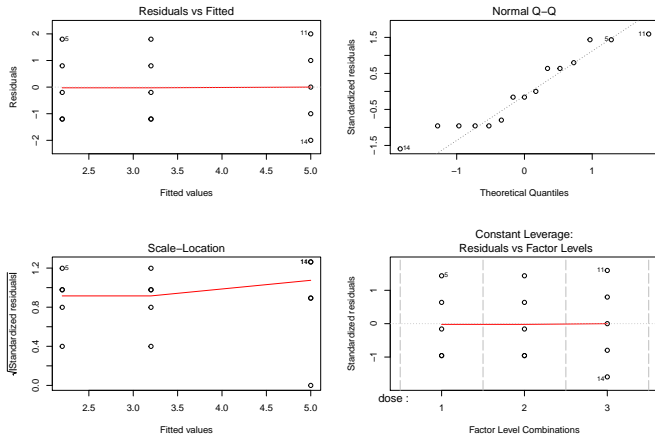


Figure 3: Ausgabe von aov-Funktion

Welch's F-Test berücksichtigt die vorhandenen Unterschiede in den Gruppenvarianzen.

```
oneway.test(libido ~ dose, data = df)
```

```
##  
## One-way analysis of means (not assuming equal variances)  
##  
## data: libido and dose  
## F = 4.3205, num df = 2.0000, denom df = 7.9434, p-value = 0.05374
```

In unserem Beispiel gibt's keine Unterschiede in den Gruppenvarianzen!

Robuste ANOVA

```
viagraWide <- unstack(df, libido ~ dose)
viagraWide1 <- data.frame(data = viagraWide)
colnames(viagraWide1) = c('Placebo', 'Low.Dose', 'High.Dose')
knitr::kable(viagraWide1, caption = "Viagra-Datensatz im breiten Format [A. Field]")
```

Table 3: Viagra-Datensatz im breiten Format [A. Field]

Placebo	Low.Dose	High.Dose
3	5	7
2	2	4
1	4	5
1	2	3
4	3	6

Robuste ANOVA (1)

```
library(WRS2)
t1way(libido ~ dose, data = df, tr = 0.1)

## Call:
## t1way(formula = libido ~ dose, data = df, tr = 0.1)
##
## Test statistic: F = 4.3205
## Degrees of freedom 1: 2
## Degrees of freedom 2: 7.94
## p-value: 0.05374
##
## Explanatory measure of effect size: 0.71
mediway(libido ~ dose, data = df) # median
```

```
## Call:
## mediway(formula = libido ~ dose, data = df)
##
## Test statistic F: 4.7829
## Critical value: 5.473
## p-value: 0.07
```

```
t1waybt(libido ~ dose, data = df)
```

```
## Call:
## t1waybt(formula = libido ~ dose, data = df)
##
## Effective number of bootstrap samples was 384.
##
## Test statistic: 3
## p-value: 0.09115
## Variance explained 0.623
## Effect size 0.789
```

Alle Tests sind nicht signifikant und somit hat die Dosis keine Aswirkung auf Libido.

Kontraste

Kontraste

```
summary.lm(viagraModel)
```

```
##
## Call:
## lm(formula = libido ~ dose, data = df)
##
## Residuals:
##   Min     1Q   Median     3Q      Max
##  -2.0  -1.2  -0.2   0.9   2.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.2000     0.6272   3.508  0.00432 **
## dose2        1.0000     0.8869   1.127  0.28158
## dose3        2.8000     0.8869   3.157  0.00827 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.402 on 12 degrees of freedom
## Multiple R-squared:  0.4604, Adjusted R-squared:  0.3704
## F-statistic: 5.119 on 2 and 12 DF, p-value: 0.02469
```

```
contrasts(df$dose)
```

```
##    2 3
```

```
## 1 0 0
```

```
## 2 1 0
```

```
## 3 0 1
```

Geplante Kontraste

Sei gegeben die Variablen oder deren Statistiken $\theta_1, \theta_2, \dots, \theta_k$ und die Konstanten a_1, a_2, \dots, a_k .

Die lineare Kombination $\sum_i a_i \theta_i$ heißt *Kontrast*.

Zwei Kontraste $\sum_i a_i \theta_i$ und $\sum_i b_i \theta_i$ heißen *orthogonal* wenn $\sum_i a_i b_i = 0$.

Geplante Kontraste (1)

Contrast	Description
<code>contr.helmert</code>	Contrasts the second level with the first, the third level with the average of the first two, the fourth level with the average of the first three, and so on.
<code>contr.poly</code>	Contrasts are used for trend analysis (linear, quadratic, cubic, and so on) based on orthogonal polynomials. Use for ordered factors with equally spaced levels.
<code>contr.sum</code>	Contrasts are constrained to sum to zero. Also called <i>deviation contrasts</i> , they compare the mean of each level to the overall mean across levels.
<code>contr.treatment</code>	Contrasts each level with the baseline level (first level by default). Also called <i>dummy coding</i> .
<code>contr.SAS</code>	Similar to <code>contr.treatment</code> , but the baseline level is the last level. This produces coefficients similar to contrasts used in most SAS procedures.

Figure 4: Verschiedene eingebaute Kontraste [Quelle: A. Kabacoff]

```
contrasts(df$dose) <- contr.helmert(3)
df$dose

## [1] 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3
## attr(,"contrasts")
## [1,] [,2]
## 1 -1 -1
## 2 1 -1
## 3 0 2
## Levels: 1 2 3
```

Geplante Kontraste (3)

```
contrasts(df$dose) <- cbind(c(-2,1,1),c(0,-1,1))
df$dose
```

```
## [1] 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3
## attr(,"contrasts")
## [1] [,2]
## 1 -2 0
## 2 1 -1
## 3 1 1
## Levels: 1 2 3
```

- ▶ Bei den geplanten Kontrasten vergleichen wir die Gruppen mit dem positiven Vorzeichen gegen die Gruppen mit dem negativen.
- ▶ Im Kontrast 1 vergleichen wir Placebo-Gruppe und die zwei Versuchsgruppen
- ▶ Im Kontrast 2 vergleichen wir niedrige Dosis-Gruppe und die hohe Dosis-Gruppe

Geplante Kontraste (4)

```
ViagraModel3 <- aov(libido ~ dose, data = df)
summary.lm(ViagraModel3)
```

```
##
## Call:
## aov(formula = libido ~ dose, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0    -1.2    -0.2     0.9     2.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.4667     0.3621   9.574 5.72e-07 ***
## dose1         0.6333     0.2560   2.474 0.0293 *
## dose2         0.9000     0.4435   2.029 0.0652 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.402 on 12 degrees of freedom
## Multiple R-squared:  0.4604, Adjusted R-squared:  0.3704
## F-statistic: 5.119 on 2 and 12 DF,  p-value: 0.02469
```

Geplante Kontraste (5)

- ▶ Weil wir davon ausgehen, dass Libido durch die Gabe des Medikaments steigt, können wir den zweiseitigen Test durch den einseitigen ersetzen und die entsprechenden p -Werte halbieren.
- ▶ Ohne die Hypothese, dass die Wirkung mit der Dosis steigt, wäre unser Ergebnis nicht signifikant
- ▶ Das zeigt, dass es wichtig ist, die Hypothese zu haben, bevor wir die Daten sammeln.

Post hoc-Tests

- ▶ *F*-Test sagt uns nur, dass es die Unterschiede vorhanden sind, aber nicht zwischen welchen Gruppen.
- ▶ Um die paarweisen Vergleiche durchzuführen, brauchen wir *post hoc*-Tests.

Bonferroni und Benjamini-Hochberg *post hoc*-Tests

```
pairwise.t.test(df$libido, df$dose, p.adjust.method = "bonferroni")
```

```
##  
## Pairwise comparisons using t tests with pooled SD  
##  
## data: df$libido and df$dose  
##  
## 1 2  
## 2 0.845 -  
## 3 0.025 0.196  
##  
## P value adjustment method: bonferroni
```

```
pairwise.t.test(df$libido, df$dose, p.adjust.method = "BH")
```

```
##  
## Pairwise comparisons using t tests with pooled SD  
##  
## data: df$libido and df$dose  
##  
## 1 2  
## 2 0.282 -  
## 3 0.025 0.098  
##  
## P value adjustment method: BH
```

Tukey HSD paarweise Vergleiche

```
TukeyHSD(ViagraModel3)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = libido ~ dose, data = df)
##
## $dose
##      diff      lwr      upr      p adj
## 2-1  1.0 -1.3662412  3.366241 0.5162761
## 3-1  2.8  0.4337588  5.166241 0.0209244
## 3-2  1.8 -0.5662412  4.166241 0.1474576
```


Tukey HSD paarweise Vergleiche (1)

```
par(las=2)
par(mar=c(5,8,4,2))
plot(TukeyHSD(ViagraModel3), cex = 2, cex.main=2, cex.lab=2, cex.axis=2)
```

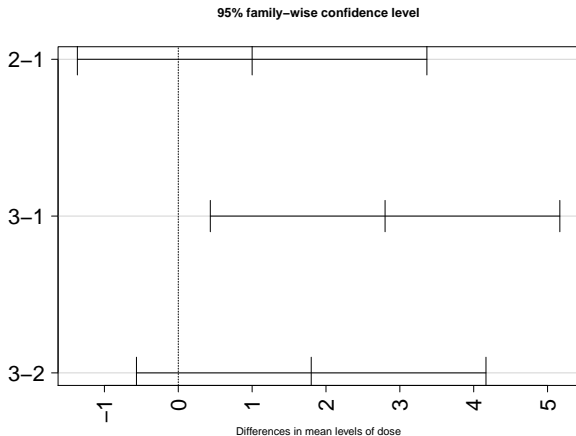


Figure 5: Tukey HSD paarweise Vergleiche.

Tukey HSD paarweise Vergleiche (2)

```
library(multcomp)
par(mar=c(5,4,6,2))
tuk <- glht(ViagraModel3, linfct = mcp(dose = "Tukey"))
summary(tuk)

##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = libido ~ dose, data = df)
##
## Linear Hypotheses:
##           Estimate Std. Error t value Pr(>|t|)
## 2 - 1 == 0  1.0000    0.8869   1.127  0.5163
## 3 - 1 == 0  2.8000    0.8869   3.157  0.0209 *
## 3 - 2 == 0  1.8000    0.8869   2.029  0.1475
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

Tukey HSD paarweise Vergleiche (3)

```
confint(tuk)

##
## Simultaneous Confidence Intervals
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: aov(formula = libido ~ dose, data = df)
##
## Quantile = 2.6658
## 95% family-wise confidence level
##
## Linear Hypotheses:
##      Estimate lwr      upr
## 2 - 1 == 0  1.0000 -1.3644  3.3644
## 3 - 1 == 0  2.8000  0.4356  5.1644
## 3 - 2 == 0  1.8000 -0.5644  4.1644
```

Für signifikante Unterschiede dürfen die Konfidenzintervalle die Null nicht durchkreuzen!

Tukey HSD paarweise Vergleiche (4)

```
par(mar=c(5,8,4,2))  
plot(cld(tuk, level = .05, col = "lightblue"), cex = 2, cex.main=2, cex.lab=2, cex.axis=2)
```

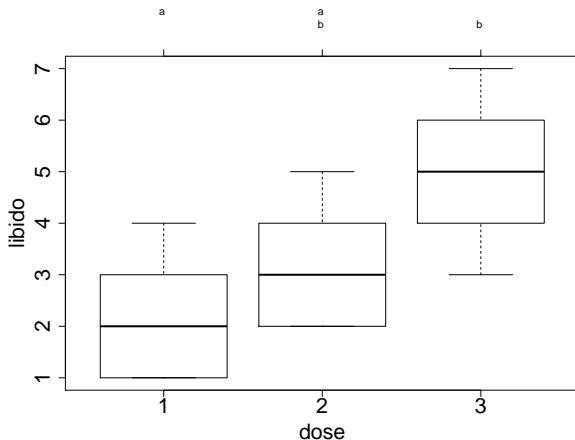


Figure 6: Tukey HSD paarweise Vergleiche. Gruppen, die die gleichen Buchstaben oben zu stehen haben, sind nicht signifikant unterschiedlich.

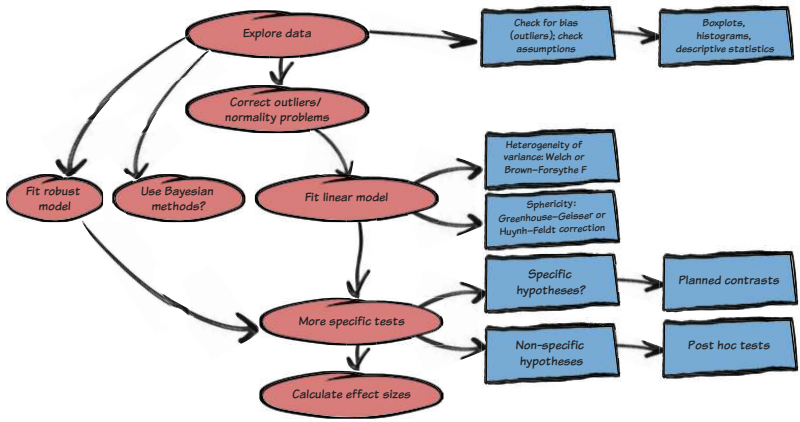


Figure 7: Vergleich mehrerer Mittelwerte [Quelle: A. Field]

$$R^2 = \frac{SSM}{SST}$$

Effektgröße ist gegeben durch R (0.68)

Anstatt r^2 benutzt man in ANOVA η^2 .

$$\omega^2 = \frac{SSM - (df_M)MSE}{SST + MSE}$$

$$\omega = 0.60$$

- ▶ $\omega^2 = 0.01$ (small effects)
- ▶ $\omega^2 = 0.06$ (medium effects)
- ▶ $\omega^2 = 0.14$ (large effects)