HTW Chur

Hochschule für Technik und Wirtschaft
University of Applied Sciences

Varianzanalyse ANOVA

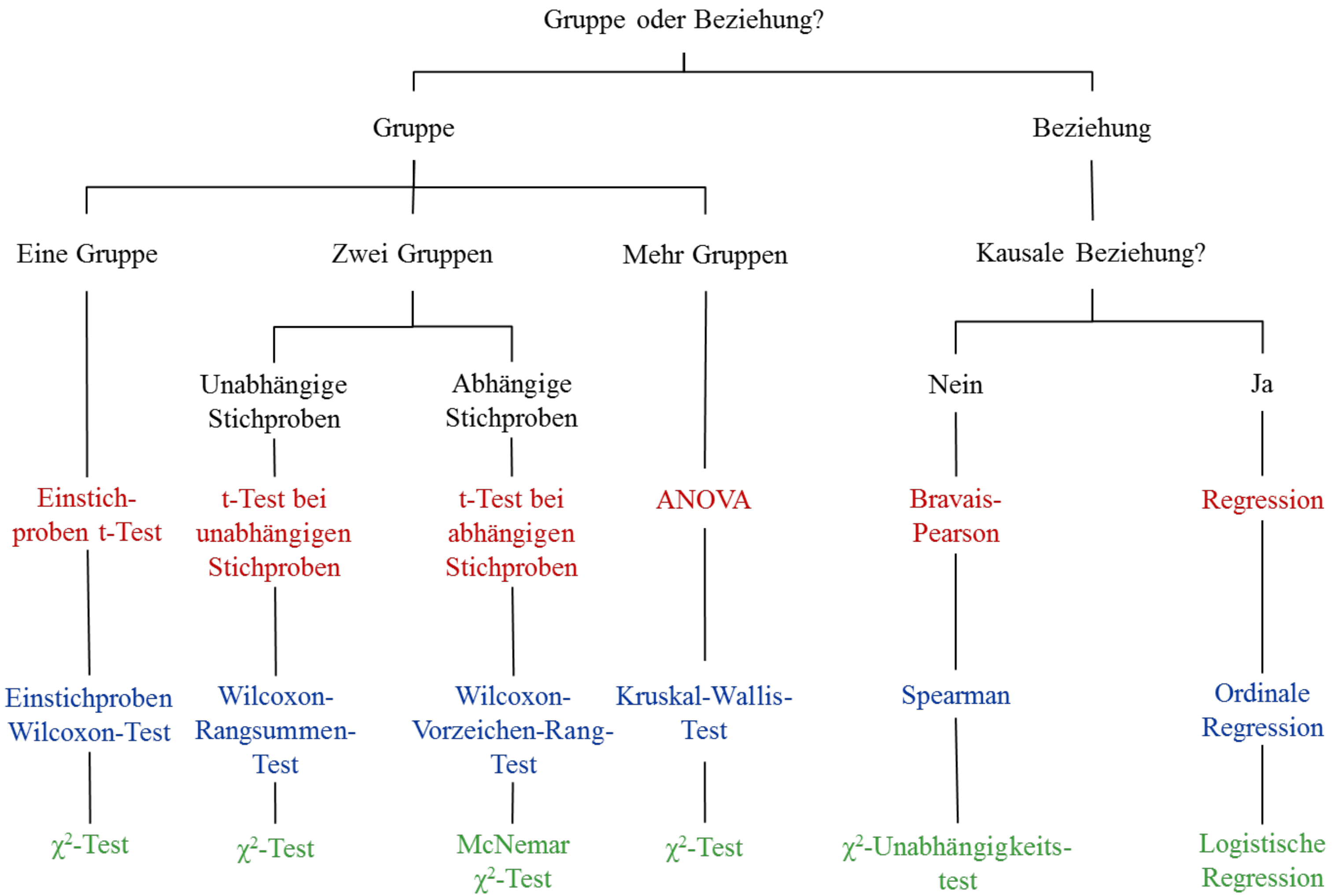
Quantitative Forschungsmethoden

Prof. Dr. Franz Kronthaler

Lernziele

- verstehen der Idee der ANOVA
- wissen wann die ANOVA das angemessene Verfahren ist, um ein Problem zu analysieren
- wissen wie eine ANOVA und eine Kovarianzanalyse ANCOVA durchgeführt wird
- wissen welche Annahmen erfüllt sein müssen, um eine ANOVA oder ANCOVA durchzuführen
- wissen wie die Ergebnisse einer ANOVA und ANCOVA zu interpretieren sind

Einleitung



Variable(n) von Interesse sind:
- metrisch und normal verteilt
- ordinal oder metrisch und nicht-normal verteilt
- nominal

ANOVA

- eine ANOVA ist ein statistisches Verfahren mit welchem Gruppenmittelwerte verglichen werden können
- eine ANOVA analysiert die Beziehung zwischen einer abhängigen Variable und einer oder mehreren unabhängigen Variablen
- die abhängige Variable ist metrisch, die unabhängigen Variablen sind nicht-metrisch beziehungsweise Faktoren mit einer definierten Anzahl an Kategorien

$$\underset{metrisch}{Y} = \underset{nicht-metrisch}{X_1 + X_2 + X_3 + \cdots + X_N}$$

# abhängige Variablen	# unabhängige Variablen	Name
1	1	Einfaktorielle ANOVA
1	2	Zweifaktorielle ANOVA
1	3	Dreifaktorielle ANOVA
…	…	…
>1	1 oder >1	MANOVA

ANOVA

- ANOVA ist die Methode, die typischerweise bei der Analyse von Experimenten zum Einsatz kommt
- typische Anwendungen bzw. Fragestellungen sind
  - Variieren die Verkaufszahlen bei verschiedenen Arten von Werbemassnahmen (z. B. Fernsehen, Zeitung)?
  - Hängen die Verkaufszahlen von der Art der Verpackung ab?
  - Haben unterschiedliche Lehrmethoden (z. B. passive oder aktive Lehrmethoden) einen Einfluss auf den Lernerfolg bei den Studierenden?
  - Variiert der Output in Abhängigkeit der verschiedenen Produktionsmethoden und Arbeitsgruppen?
  - Variiert der Ernteertrag in Abhängigkeit des verwendeten Düngers und der Bodenqualität? Gibt es eine Interaktion zwischen Dünger und Boden?
  - und viele mehr

ANOVA – grundlegende Idee

- die grundlegende Idee der ANOVA ist, dass ein beobachteter Wert mit Hilfe eines linearen Models beschrieben werden kann

  $y_{gk} = \bar{y} + \alpha_g + \varepsilon_{gk}$

  – y_{gk} beobachteter Wert

  – $\bar{y}$ globaler Mittelwert

  – α_g gruppenspezifischer Effekt der Gruppe g

  – ε_{gk} zufälliger Fehler

ANOVA – eine einfache Illustration

- wir nehmen an, dass eine Firma ein Produkt mit vier verschiedenen Produktionsmethoden produziert
- die Frage ist, ob sich die Produktionsmethoden hinsichtlich der Produktivität unterscheiden
- die Firma misst den Output pro Stunde für die vier Produktionsmethoden zu mehreren Zeitpunkten
  - abhängige Variable: Output pro Stunde
  - Faktor: Produktionsmethode
  - Faktorstufen: Produktionsmethoden 1, 2, 3 und 4

ANOVA – eine einfache Illustration

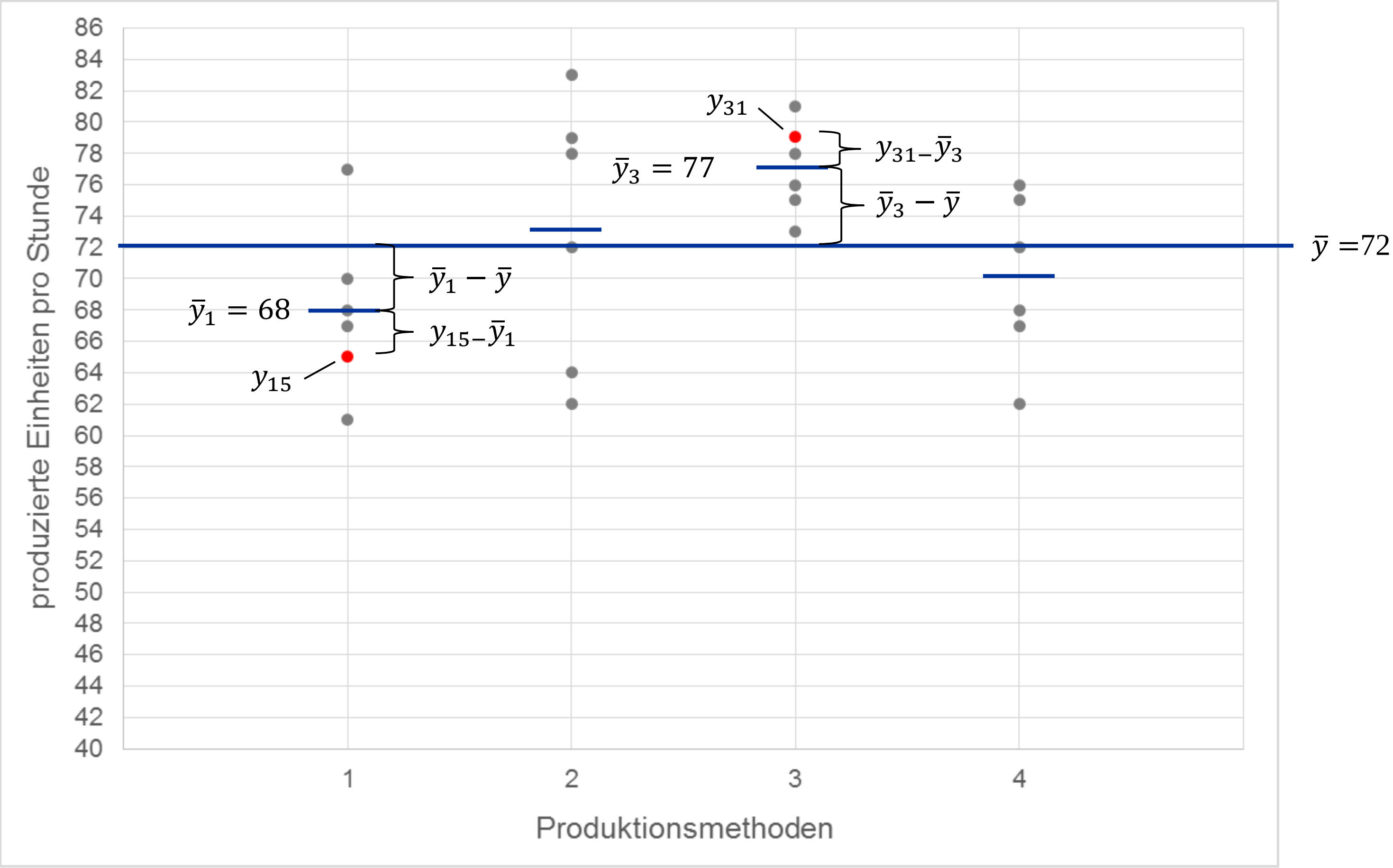
- Output pro Stunde zu den gemessenen Zeiten

	1	2	3	4	5	6	$\bar{y}_g$
Methode 1	68	77	61	67	65	70	$\bar{y}_1 = 68$
Methode 2	79	83	62	64	72	78	$\bar{y}_2 = 73$
Methode 3	79	78	73	76	81	75	$\bar{y}_3 = 77$
Methode 4	75	67	68	62	72	76	$\bar{y}_4 = 70$

- $\bar{y} = \frac{1}{n}\sum y_{gk} = 72$
- wir beobachten drei Dinge
  - jede Beobachtung unterscheidet sich i.d.R vom globalen Mittelwert
  - der Gruppenmittelwert unterscheidet sich i.d.R vom globalen Mittelwert
  - jede Beobachtung unterscheidet sich i.d.R vom Gruppenmittelwert

ANOVA – eine einfache Illustration

- Output pro Stunde zu den gemessenen Zeiten



ANOVA – eine einfache Illustration

- die Abweichung einer Beobachtung vom globalen Mittelwert
$$y_{gk} - \bar{y}$$
kann in eine durch die Faktorstufe erklärte Abweichung
$$\bar{y}_g - \bar{y}$$
und eine durch die Faktorstufe nicht erklärte Abweichung eingeteilt werden
$$y_{gk} - \bar{y}_g$$
- oder für alle Punkte

$$SS_{insgesamt} = SS_{erklärt} + SS_{nicht\ erklärt}$$

$$\sum_{g=1}^{G}\sum_{k=1}^{K}\left(y_{gk} - \bar{y}\right)^2 = \sum_{g=1}^{G} K\left(\bar{y}_g - \bar{y}\right)^2 + \sum_{g=1}^{G}\sum_{k=1}^{K}\left(y_{gk} - \bar{y}_g\right)^2$$

ANOVA – eine einfache Illustration

- wenn der Anteil der erklärten Varianz gross ist, ist der Anteil der unerklärten Varianz klein
- wenn der Anteil der erklärten Varianz gross ist, erklärt der Faktor viel vom Ergebnis
- wenn der Anteil der nicht-erklärten Varianz gross ist, erklärt der Faktor wenig vom Ergebnis

Einfaktorielle ANOVA

- die Einfaktorielle ANOVA analysiert die Beziehung zwischen zwei Variablen

Faktor A (nicht-metrisch) → abhängige Variable Y (metrisch)

Faktor A hat G Stufen (Kategorien), wir können G Gruppen bilden

Gruppen

Gruppe 1
Beobachtungen mit Wert 1 für Faktor A

Gruppe 2
Beobachtungen mit Wert 2 für Faktor A

Gruppe 3
Beobachtungen mit Wert 3 für Faktor A

...

Gruppe G
Beobachtungen mit Wert G für Faktor A

Einfaktorielle ANOVA

- die Einfaktorielle ANOVA testet die folgende Hypothese
  - H_0: Alle Gruppenmittelwerte sind gleich.
  - H_A: Wenigstens zwei der Gruppenmittelwerte sind nicht gleich.

    oder
  - H_0: $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \cdots = \bar{y}_G$
  - H_A: $\bar{y}_i \neq \bar{y}_j$ für wenigstens ein Paar von ij
- die Teststatistik ist die F-Statistik

$$F = \frac{MSS_{erklärt}}{MSS_{nicht\ erklärt}}$$

  - $MSS_{erklärt}$ ist die durchschnittliche erklärte Quadratsumme
  - $MSS_{nicht\ erklärt}$ ist die durchschnittliche nicht erklärte Quadratsumme
  - die Testverteilung ist die F-Verteilung mit $Fg_1 = G - 1$ und $Fg_2 = G \times (K - 1)$

Einfaktorielle ANOVA

– die Quadratsumme steigt mit der Anzahl an Beobachtungen an, dies korrigieren wir indem wir die durchschnittliche Quadratsumme berechnen

$$MSS_{ingesamt} = \frac{SS_{insgesamt}}{G \times K - 1}$$

$$MSS_{erklärt} = \frac{SS_{erklärt}}{G - 1}$$

$$MSS_{nicht\ erklärt} = \frac{SS_{nicht\ erklärt}}{G \times (K - 1)}$$

– der F-Test der ANOVA ist ein Omnibus-Test
  - er testet, ob es zwischen den verschiedenen Gruppen irgendeinen Unterschied gibt
  - er testet nicht, zwischen welchen Gruppen ein Unterschied besteht

Einfaktorielle ANOVA

- um zu testen zwischen welchen Gruppen ein signifikanter Unterschied besteht müssen die Gruppen verglichen werden
  - im Prinzip könnten wir einfach alle Gruppen paarweise verglichen (t-Test bei unabhängigen Stichproben), aber wir würden den α-Fehler inflationieren
  - ein Beispiel
    - ein Bergsteiger nutzt ein Seil mit 20 Knoten
    - jeder Knoten hat eine Fehlerwahrscheinlichkeit von 0.05 (5%)
    - alle Knoten zusammen haben dann eine Fehlerwahrscheinlichkeit von $1 - (1 - 0.05)^{20} = 0.64$
    - das Risiko dass das Seil reist ist 64%!
    - wenn wir wollen, dass das Risiko bei 5% bleibt darf das Risiko für jeden Knoten näherungsweise nicht mehr als $5\%/20 = 0.25\%$ sein
  - entsprechend müssen wir beim Test sicherstellen, dass der α-Fehler nicht inflationiert wird
  - wir machen dies über die sogenannten Post-hoc-Tests

Einfaktorielle ANOVA

- eine Möglichkeit ist der Test nach Tukey
  - er vergleicht jede Gruppe zu allen andern Gruppen
  - er korrigiert für den α-Fehler
  - H_0: Es gibt keinen Unterschied zwischen Gruppe i und j
  - H_A: Es gibt einen Unterschied zwischen Gruppe i und j

  oder
  - H_0: $\bar{y}_i = \bar{y}_j$
  - H_A: $\bar{y}_i \neq \bar{y}_j$

Annahmen der ANOVA

- bei Durchführung der ANOVA sollten folgende Annahmen erfüllt sein
  - Unabhängigkeit der Beobachtungen
  - gleiche Varianzen
  - abhängige Variable ist normalverteilt

Annahmen der ANOVA

- Unabhängigkeit der Beobachtungen
  - Beobachtungen zwischen und innerhalb den Gruppen stehen zueinander nicht in Beziehung
    - keine Lern- oder Beobachtungseffekte
    - kein Selektionbias bzgl. der Gruppen
  - Annahme erfordert eine sorgfältige Planung des Experiments
- gleiche Varianzen
  - die Varianzen für alle Gruppen sollten gleich sein
    - kann anhand der Gruppenvarianzen geprüft werden
    - kann mit Hilfe des Tests von Levene auf Gleichheit der Varianzen getestet werden
  - die Annahme ist insbesondere dann wichtig, wenn die Anzahl der Beobachtungen pro Gruppe unterschiedlich ist

Annahmen der ANOVA

- abhängige Variable ist normalverteilt
  - kann geprüft/ getestet werden mit
    - dem Histogramm
    - dem Quantil-Quantil-Plot (Q-Q-Plot)
    - dem Shapiro-Wilk-Test
  - im Falle einer grossen Stichprobe ist die ANOVA relativ robust gegenüber einer Verletzung der Normalverteilungsannahme

Stichprobengrösse bei der Einfaktoriellen ANOVA

- die benötigte Stichprobengrösse hängt von der Anzahl der Faktorstufen ab
- die empfohlene Minimalbesetzung pro Faktorstufe ist 20 Beobachtungen
- mit einem Faktor und 5 Faktorstufen sollten ungefähr 100 Beobachtungen (gleich mässig verteilt auf die Faktorstufen) zur Verfügung stehen
- der Forscher sollte versuchen ungefähr die gleiche Anzahl an Beobachtungen pro Faktorstufe zu erheben

Beispiel Einfaktorielle ANOVA

- um die Einfaktorielle ANOVA mit R zu berechnen nutzen wir das Beispiel auf Seite 7
  - der Datensatz ist production.Rdata (der Originaldatensatz ist von Universität Kassel und wird zu Lehrzwecken verwendet)
  - der Datensatz enthält Informationen zum produzierten Output pro Stunde mit Hilfe von unterschiedlichen Produktionsmethoden
- # lade den Datensatz “production.RData”

Beispiel Einfaktorielle ANOVA

- Schritt 1:
  - Haben die verschiedenen Produktionsmethoden eine unterschiedliche Produktivität?
  - H_0: $\bar{y}_1 = \bar{y}_2 = \bar{y}_3 = \bar{y}_4$
  - H_A: $\bar{y}_i \neq \bar{y}_j$ für wenigstens ein Paar von ij
  - $\alpha = 0.1$
- Schritt 2:
  - bevor wir die ANOVA berechnen müssen wir die Daten vorbereiten, die Annahmen prüfen und sollten die deskriptiven Statistiken betrachten
    - ❖ teile R mit, dass die Variable method eine Faktorvariable ist

```
# nutze die Benutzeroberfläche des R Commanders oder den folgenden Befehl
production$f_method<-as.factor(production$method)
```

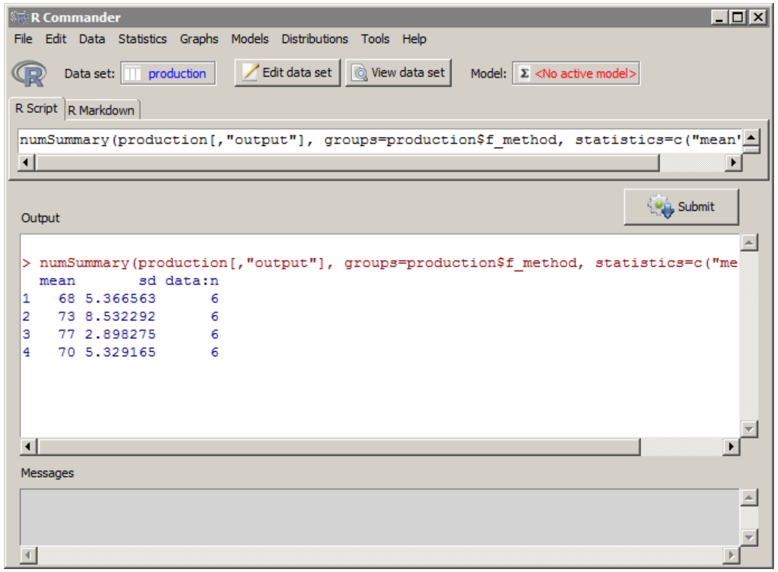
Beispiel Einfaktorielle ANOVA

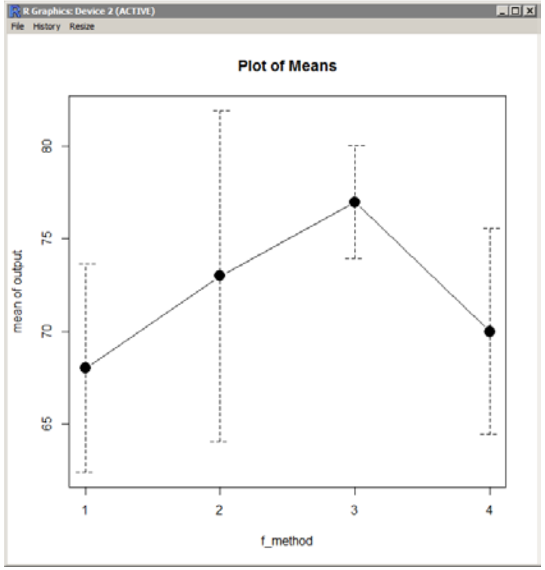
- betrachte die deskriptiven Statistiken

  # nutze die Benutzeroberfläche des R Commanders oder die folgenden Befehle

  numSummary(production[,"output"], groups=production$f_method, statistics=c("mean", "sd"))

  with(production, plotMeans(output, f_method, error.bars="conf.int", level=0.95))





Beispiel Einfaktorielle ANOVA

- prüfe die Annahme – Unabhängigkeit der Beobachtungen
  - ✓ ist ein Problem des Designs des Experiments
  - ✓ kann nicht mit Hilfe der Daten getestet werden
- prüfe die Annahme – gleiche Varianzen

```
# nutze die Benutzeroberfläche des R Commanders oder folgende Befehle
numSummary(production[,"output"], groups=production$f_method,
statistics=c("sd"))
leveneTest(output ~ f_method, data=production, center="mean")
```

- prüfe die Annahme – abhängige Variable ist normalverteilt

```
# nutze die Benutzeroberfläche des R Commanders oder folgende Befehle
attach(production)
hist(output, scale="frequency", breaks="Sturges", col="purple")
qqPlot(output, dist="norm", id.method="y", id.n=2,
labels=rownames(production))
shapiro.test(output)
detach(production)
```

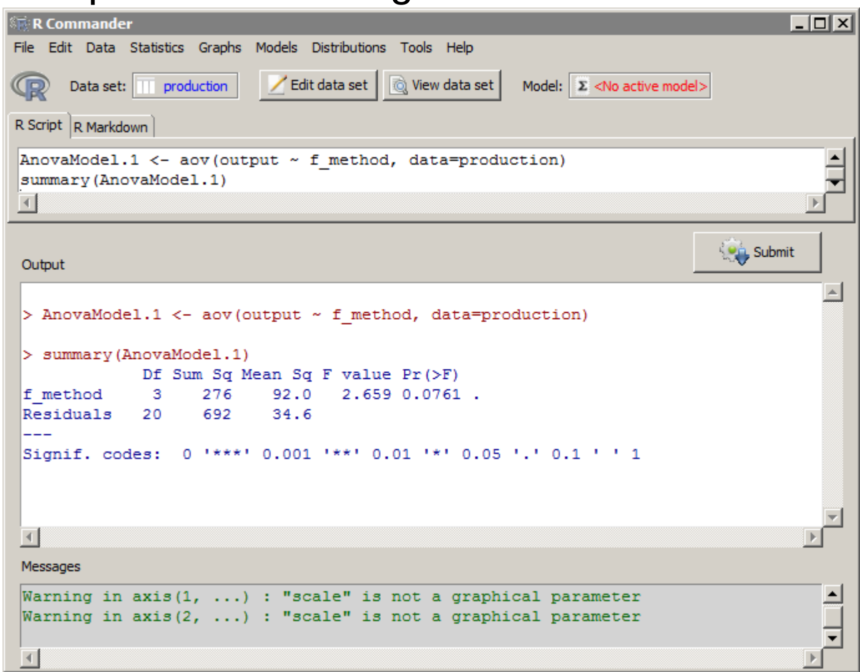
Beispiel Einfaktorielle ANOVA

- Stichprobengrösse
  - ✓ wir haben nur sechs Beobachtungen pro Faktorstufe
  - ✓ Stichprobe ist eher klein
  - ✓ wenn möglich sollte die Stichprobengrösse erhöht werden
- rechne die ANOVA und den Post-hoc-Test

```
# nutze die Benutzeroberfläche des R Commanders oder die folgenden Befehle
AnovaModel.1 <- aov(output ~ f_method, data=production)
summary(AnovaModel.1)
TukeyHSD(AnovaModel.1)
```

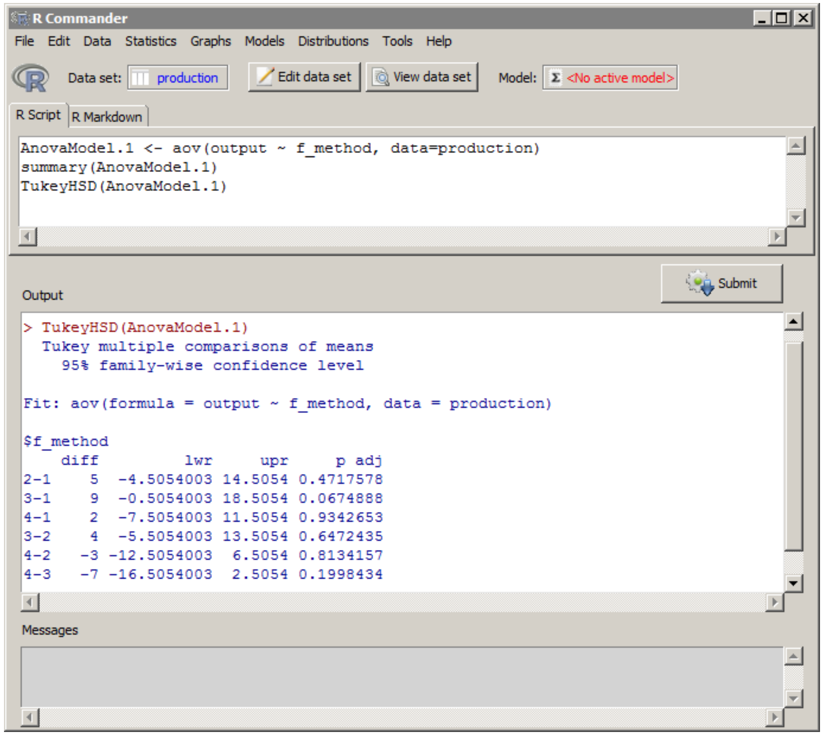
Beispiel Einfaktorielle ANOVA

- Schritt 3:
  - Interpretation des Ergebnisses der ANOVA



Beispiel Einfaktorielle ANOVA

- Interpretation des Post-hoc-Tests «TukeyHSD»

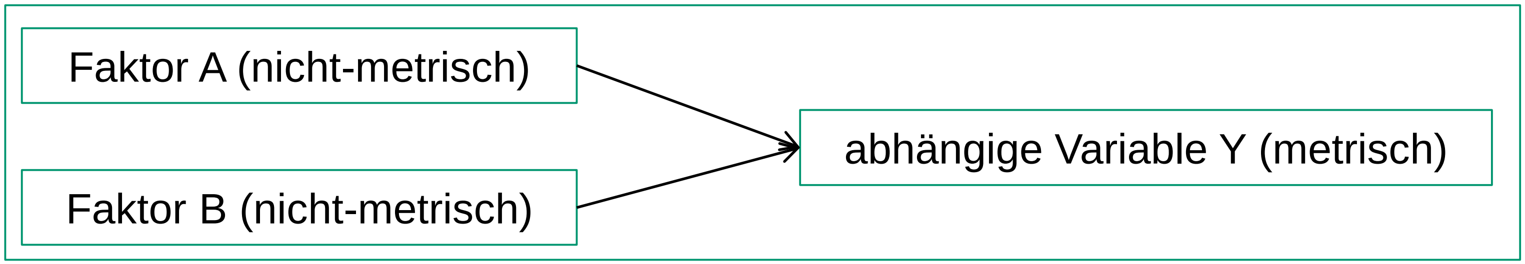


Anwendungen

- Berechne $SS_{insgesamt}$, $MSS_{insgesamt}$, $SS_{erklärt}$, $MSS_{erklärt}$, $SS_{nicht\ erklärt}$, $MSS_{nicht\ erklärt}$ und die F-Statistik für das Beispiel von Seite 7 und 8 mit der Hand. Entnehme den kritischen Werte der F-Verteilungstabelle und vergleiche den kalkulierten F-Wert mit dem kritischen Wert. Wird H_0 verworfen?
- Nutze den Datensatz production.RData. Teste mit der Hilfe der ANOVA, ob es einen signifikanten Unterschied in der Produktivität der verschiedenen Arbeitsgruppen gibt. Führe alle relevanten Schritte durch.

Zweifaktorielle ANOVA

- die Zweifaktorielle ANOVA analysiert die Beziehung zwischen einer abhängigen und zwei unabhängigen Variablen (Faktoren)



- Gruppen werden aus einer Kombination der unabhängigen Variablen gebildet
  - Faktor A hat G Gruppen
  - Faktor B hat H Gruppen
  - zusammen gibt es $G \times H$ Gruppen

Zweifaktorielle ANOVA

- die Modellgleichung ist

  $y_{ghk} = \bar{y} + \alpha_g + \beta_h + \varepsilon_{ghk}$

  – y_{ghk} beobachteter Wert

  – $\bar{y}$ globaler Mittelwert

  – α_g gruppenspezifischer Effekt der Gruppe g

  – β_h gruppenspezifischer Effekt der Gruppe h

  – ε_{ghk} zufälliger Fehler

Zweifaktorielle ANOVA

- die Zweifaktorielle ANOVA nutzt wiederrum das Aufteilen der Varianz in einen erklärten und einen nicht erklärten Anteil
- es gibt drei mögliche Quellen der erklärten Varianz

- Quadratsumme insgesamt $SS_{insgesamt}$
  - erklärte Quadratsumme $SS_{erklärt}$
    - Quadratsumme Faktor A $SS_{erklärt_A}$
    - Quadratsumme Faktor B $SS_{erklärt_B}$
    - erklärte Quadratsumme über die Interaktion zwischen Faktor A und B $SS_{erklärt_{AxB}}$
  - nicht erklärte Quadratsumme

$$SS_{insgesamt} = SS_{erklärt_A} + SS_{erklärt_B} + SS_{erklärt_{AxB}} + SS_{nicht\ erklärt}$$

Zweifaktorielle ANOVA

- die Zweifaktorielle ANOVA testet die folgenden drei Hypothesen
  - 1. H_0: Alle Gruppenmittelwerte des ersten Faktors sind gleich.

       H_A: Wenigstens zwei Gruppenmittelwerte des ersten Faktors sind nicht gleich.
  - 2. H_0: Alle Gruppenmittelwerte des zweiten Faktors sind gleich.

       H_A: Wenigstens zwei Gruppenmittelwerte des zweiten Faktors sind nicht gleich.
  - 3. H_0: Es gibt keine Interaktion zwischen den beiden Faktoren.

       H_A: Es gibt eine Interaktion zwischen den beiden Faktoren.

Zweifaktorielle ANOVA

- die Teststatistik ist die F-Statistik
  - für Faktor A

$$F_A = \frac{MSS_{erklärt_A}}{MSS_{nicht\ erklärt}} \text{ mit } MSS_{erklärt_A} = \frac{SS_{erklärt_A}}{G-1} \text{ und } MSS_{nicht\ erklärt} = \frac{SS_{nicht\ erklärt}}{G \times H \times (K-1)}$$

  - für Faktor B

$$F_B = \frac{MSS_{erklärt_B}}{MSS_{nicht\ erklärt}} \text{ mit } MSS_{erklärt_B} = \frac{SS_{erklärt_B}}{H-1}$$

  - für den Effekt der Interaktion

$$F_{A \times B} = \frac{MSS_{erklärt_{A \times B}}}{MSS_{nicht\ erklärt}} \text{ mit } MSS_{erklärt_{A \times B}} = \frac{SS_{erklärt_{A \times B}}}{(G-1)(H-1)}$$

  - der F-Test der ANOVA ist ein Omnibus-Test
    - er testet nicht zwischen welchen Gruppen es signifikante Unterschiede gibt
    - hierfür benötigen wir wiederrum einen Post-hoc-Test

Stichprobengrösse bei der Zweifaktoriellen ANOVA

- die Stichprobengrösse hängt von der Anzahl der Gruppen (Zellen) ab
- es gibt $G \times H$ Gruppen
- nehmen wir an, wir haben einen Faktor A mit 4 Faktorstufen und einen Faktor B mit 3 Faktorstufen, dann haben wir $G \times H = 4 \times 3 = 12$ Gruppen
- pro Gruppe werden im Minimum 20 Beobachtungen empfohlen
- d. h., wir brauchen in etwa 240 Beobachtungen
- eine näherungsweise gleiche Anzahl an Beobachtungen pro Gruppe ist von Vorteil

- !die diskutierten Annahmen auf den Seiten 17ff. gelten für die ANOVA im Allgemeinen!

Beispiel Zweifaktorielle ANOVA

- um die Zweifaktorielle ANOVA mit dem R Commander zu berechnen nutzen wir das selbe Beispiel und erweitern es
  - der Datensatz ist production.RData
  - der Datensatz enthält Informationen zum produzierten Output pro Stunde für die verschiedenen Produktionsmethoden und Arbeitsgruppen
    - abhängige Variable: Output pro Stunde
    - Faktor A: Produktionsmethode, 4 Faktorstufen
    - Faktor B: Arbeitsgruppen, 3 Faktorstufen
  - # lade den Datensatz “production.RData”

Beispiel Zweifaktorielle ANOVA

Beobachtung	k	g (Produktionsmethode)	h (Arbeitsgruppe)	Y_{ghk} (Output)
1 2	1 2	1 1	1 1	68 77
3 4	1 2	1 1	2 2	61 67
5 6	1 2	1 1	3 3	65 70
7 8	1 2	2 2	1 1	79 83
9 10	1 2	2 2	2 2	62 64
11 12	1 2	2 2	3 3	72 78
13 14	1 2	3 3	1 1	79 78
15 16	1 2	3 3	2 2	73 76
...	...	...	...	...

Beispiel Zweifaktorielle ANOVA

– Schritt 1:

- Unterscheiden sich die verschiedenen Produktionsmethoden und Arbeitsgruppen bezüglich der Produktivität?
- H_0: Alle Gruppenmittelwerte des Faktors Produktionsmethoden sind gleich.

  H_A: Wenigstens zwei Gruppenmittelwerte des Faktors Produktionsmethoden sind nicht gleich.

  H_0: Alle Gruppenmittelwerte des Faktors Arbeitsgruppen sind gleich.

  H_A: Wenigstens zwei Gruppenmittelwerte des Faktors Arbeitsgruppen sind nicht gleich.

  H_0: Es gibt keine Interaktion zwischen den Produktionsmethoden und den Arbeitsgruppen.

  H_A: Es gibt eine Interaktion zwischen den Produktionsmethoden und den Arbeitsgruppen.
- $\alpha = 0.05$

Beispiel Zweifaktorielle ANOVA

- Schritt 2:
  - bevor wir die ANOVA berechnen müssen wir die Daten vorbereiten, die Annahmen prüfen und sollten die deskriptiven Statistiken betrachten
    - ❖ teile R mit, dass die Variablen method und group Faktorvariablen sind

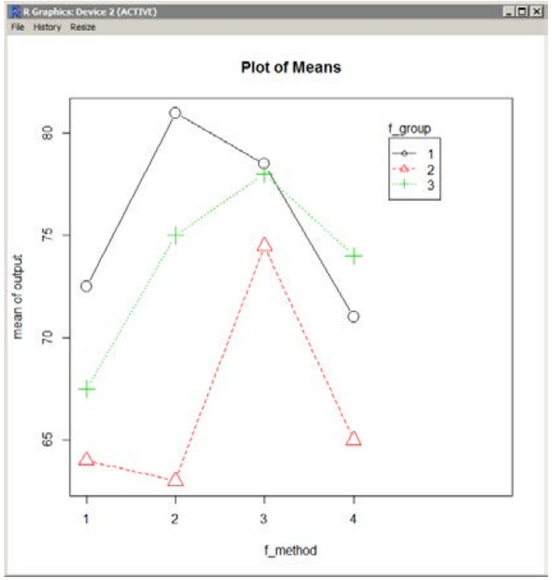
```
# nutze die Benutzeroberfläche des R Commanders oder die folgenden Befehle
production$f_method<-as.factor(production$method)
production$f_group<-as.factor(production$group)
```

Beispiel Zweifaktorielle ANOVA

❖ betrachte die deskriptiven Statistiken

nutze am Besten die folgenden Befehle

```
with(production, (tapply(output, list(f_group, f_method), mean, na.rm=TRUE)))
with(production, (tapply(output, list(f_group, f_method), sd, na.rm=TRUE)))
with(production, (tapply(output, list(f_group, f_method), function(x) sum(!is.na(x)))))
with(production, plotMeans(output, f_method, f_group, error.bars="none"))
```



Beispiel Zweifaktorielle ANOVA

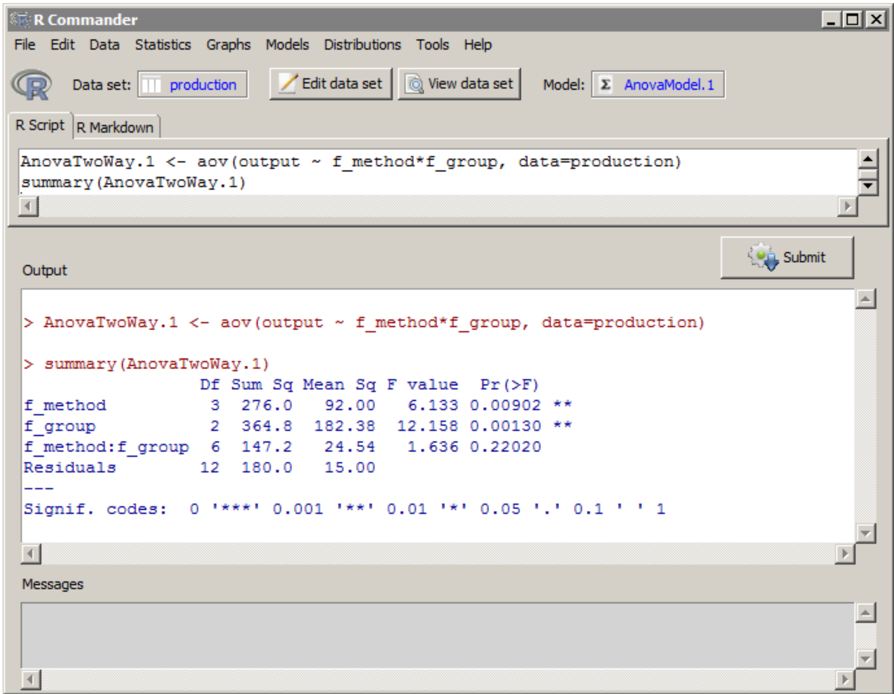
- ❖ prüfe die Annahmen
  - ✓ Unabhängigkeit der Beobachtungen
  - ✓ gleiche Varianzen
  - ✓ abhängige Variable ist normal verteilt
- ❖ Stichprobengrösse
  - ✓ nur zwei Beobachtungen pro Gruppe
  - ✓ Stichprobe ist viel zu klein
- berechne die ANOVA

```
# nutze die Benutzeroberfläche des R Commanders oder die folgenden Befehle
AnovaTwoWay.1 <- aov(output ~ f_method*f_group, data=production)
summary(AnovaTwoWay.1)
TukeyHSD(AnovaTwoWay.1)
```

Beispiel Zweifaktorielle ANOVA

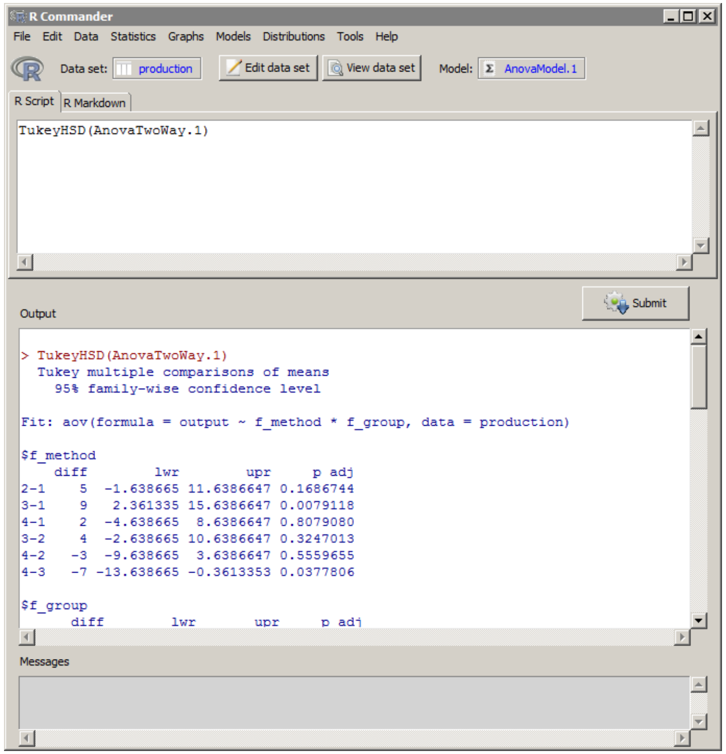
– Schritt 3:

- Interpretation des Ergebnisses der ANOVA



Beispiel Zweifaktorielle ANOVA

- Interpretation des Post-hoc-Tests «TukeyHSD»



Anwendungen

- Nutze den Datensatz gssft.RData (aus dem Lehrbuch von Norusis 2008). Analysiere mit Hilfe der ANOVA, ob die Ausbildungsstufe und das Geschlecht einen Einfluss darauf haben wie viele Stunden pro Woche gearbeitet werden. Führe alle relevanten Schritte durch.
- Nutze den Datensatz sales.RData (aus dem Lehrbuch von Backhaus et al. 2011). Prüfe mit Hilfe der ANOVA, ob die Platzierung und die Verpackungsart einen Einfluss auf die Verkaufszahlen haben. Führe alle relevanten Schritte durch.

Eine sehr kurze Einführung in die ANCOVA

- ANCOVA ist eine Erweiterung der ANOVA, es können auch metrische unabhängige Variablen aufgenommen werden

$$\begin{array}{ccc} Y & = & X_1 + X_2 + X_3 + \cdots + X_N \\ metrisch & & nicht - metrisch, \mathrm{metrisch} \end{array}$$

  – nicht-metrische Variablen sind Faktoren

  – metrische Variablen werden als “Kovariate” bezeichnet

  – Kovariate sind notwendig, wenn die abhängige Variable nicht nur von nicht-metrischen Variablen sondern auch von metrischen Variablen beeinflusst wird

  – Kovariate werden genutzt um für externe Einflüsse zu kontrollieren, die nicht über das Experiment aufgefangen werden

  – eine effektive Kovariate ist eine Variable, welche hoch korreliert mit der abhängigen Variable, aber nicht korreliert mit den unabhängigen Variablen

Zusammenfassende Fragen

- Wann nutzen wir eine ANOVA?
- Welche Annahmen sollen erfüllt sein, wenn eine ANOVA durchgeführt wird?
- Wann sollen Kovariate in der Analyse berücksichtigt werden?
- Welche Hypothesen werden bei der ANOVA geprüft?
- Wie werden die Ergebnisse interpretiert?