Mathematische und statistische Methoden I

Dr. Malte Persike

persike@uni-mainz.de
lordsofthebortz.de
twitter.com/methodenlehre
tinyurl.com/gplusmethodenlehre

Sprechstunde jederzeit nach Vereinbarung und nach der Vorlesung

Wallstr. 3, 6. Stock, Raum 06-206

WiSe 2011/2012

Fachbereich Sozialwissenschaften
Psychologisches Institut
Johannes Gutenberg Universität Mainz

Annahmen

Fixed vs. Random

Varianzanalyse

Ein- / Mehrfaktorielle ANOVA - Annahmen

- Die Beobachtungen müssen **unabhängig** sein, d.h. sie müssen von verschiedenen Merkmalsträgern stammen
- Die **Fehler** in jeder Treatmentstufe sollten **normalverteilt** sein mit einem erwarteten Mittelwert von 0. $\boxed{\bar{e}_{kj} = 0}$
- Die Fehlervarianzen in jeder der p Treatmentstufen sollen erwartet (nicht numerisch) gleich sein (**Homoskedastizität**). $\boxed{s^2_{e_{1,1}} = s^2_{e_{1,2}} = \ldots = s^2_{e_{q,p}}}$
- Treatmenteffekte und Fehler müssen **additiv** sein, d.h. die Fehler dürfen nicht mit den Erwartungswerten der Treatmentstufen korrelieren.

Annahmen

Fixed vs. Random

Varianzanalyse

Ein- / Mehrfaktorielle ANOVA – Prüfung der Annahmen

- Die **Unabhängigkeit der Beobachtungen** wird i.d.R. „begründet angenommen“ (educated guess).
- Zur Prüfung der **Normalverteilungsannahme** der Fehler (i.e. Zellresiduen) wird der Kolmogoroff-Smirnov Test verwendet
- Zur Prüfung der **Homogenität der Fehlervarianzen** wird zumeist der Bartlett Test, seltener der Levene Test bzw. der F-Test verwendet.
- Die **Unabhängigkeit der Treatmenteffekte und Messfehler** kann über einen Korrelationstest von Zellmittelwerten und Varianzen geprüft werden (selten praktiziert)

Varianzanalyse

Fixed & Random Factors

- **Fixed Factor Modelle**: Die im Experiment realisierten Stufen eines Faktors beschreiben die UV für den Experimentator **vollständig** (kategorial geschlossen).
  ► Das Experiment gibt die Wirkung des Faktors über alle möglichen (bzw. *interessierenden*) Realisierungen an.

- **Beispiele**: Geschlecht, Schulabschluss, Altersgruppen

- **Random Factor Modelle**: Alle im Experiment realisierten Stufen eines Faktors stellen nur eine **zufällige Auswahl** aus vielen möglichen dar
  ► Das Experiment liefert nur eine Probe aus dem gesamten Wirkspektrum des Faktors.

- **Beispiele**: Rauschmitteldosis, Lichtintensität

Annahmen

Fixed vs. Random

Varianzanalyse

Fixed & Random Factors

- **Praktische Konsequenz**: Effekte der Faktoren müssen an anderen Prüfvarianzen (i.e. Nennervarianzen) getestet werden, je nachdem, ob sie fixed oder random sind.
- Dasselbe gilt für die Einzelvergleiche innerhalb von Fixed oder Random Factors.

zu prüfende Varianz	Prüfvarianzen I A fixed B fixed	II A fixed B random	III A random B random
$\widehat{\sigma}^2_A$	$\widehat{\sigma}^2_{Fehler}$	$\widehat{\sigma}^2_{A\times B}$	$\widehat{\sigma}^2_{A\times B}$
$\widehat{\sigma}^2_B$	$\widehat{\sigma}^2_{Fehler}$	$\widehat{\sigma}^2_{Fehler}$	$\widehat{\sigma}^2_{A\times B}$
$\widehat{\sigma}^2_{A\times B}$	$\widehat{\sigma}^2_{Fehler}$	$\widehat{\sigma}^2_{Fehler}$	$\widehat{\sigma}^2_{Fehler}$

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- **Problem**: Ein signifikantes Ergebnis in der ANOVA zeigt nicht an, zwischen welchen Treatmentstufen der Effekt besteht.

- Für die **Prüfung der Mittelwerte** einzelner Faktorstufen gibt es zwei unterschiedliche Verfahrensweisen

  1. **A-Priori Tests** zur Prüfung von Hypothesen, die bereits vor der Untersuchung formuliert worden sind.

  2. **A-Posteriori Tests** (Post-hoc Tests) zur Prüfung von Hypothesen, die nach Ansehen der Daten gebildet wurden.

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- Bei den **Mittelwertevergleichen** im Rahmen der ANOVA können einzelne und Kombinationen von Mittelwerten verglichen werden.

- Es können Fragen beantwortet werden wie z.B.
  1. Sind die Mittelwerte zweier Faktorstufen unterschiedlich?
  2. Ist eine Faktorstufe unterschiedlich zum Mittelwert aller vorhergehenden Faktorstufen?
  3. Sind die letzten beiden Faktorstufen unterschiedlich zu den ersten beiden?

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- Sind die Mittelwerte zweier Faktorstufen unterschiedlich?

$$\boxed{D = \bar{A}_j - \bar{A}_k \neq 0}$$

- Ist eine Faktorstufe unterschiedlich zum Mittelwert aller vorhergehenden Faktorstufen?

$$\boxed{D = \bar{A}_p - \frac{1}{p-1} \cdot \left(\bar{A}_1 + \bar{A}_2 + \ldots + \bar{A}_{p-1} \right) \neq 0}$$

- Sind die letzten beiden Faktorstufen unterschiedlich zu den ersten beiden?

$$\boxed{D = \frac{1}{2} \cdot \left(\bar{A}_1 + \bar{A}_2 \right) - \frac{1}{2} \cdot \left(\bar{A}_{p-1} + \bar{A}_p \right) \neq 0}$$

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- **Frage:** Wann ist ein so berechnetes D verschieden genug von Null, damit der Mittelwerteunterschied als statistisch signifikant bewertet wird?

- **Lösung:** Aus dem D muss eine **Prüfgröße** mit bekannter Häufigkeitsverteilung konstruiert werden, um die Auftretenswahrscheinlichkeit des gemessenen D zu ermitteln, wenn in Wahrheit kein Unterschied besteht.

- Zur Berechnung einer **F-verteilten Prüfgröße** wird die (theoretische) **Varianz des D** benötigt, die man erhielte, wenn man dasselbe Experiment sehr oft mit immer neuen Stichproben durchführte

- Die Berechnung der Varianz ist dabei von der **Art des Mittelwertevergleichs** abhängig.

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- Kontraste können verwendet werden, um beliebige **Zellen innerhalb der Mittelwertetabelle**, jedoch keine Stufenmittelwerte zu vergleichen
- Es muss eine **a-priori Hypothesenbildung** stattgefunden haben.

	A_1	A_2	A_3
B_1	$\overline{A_1B_1}$	$\overline{A_2B_1}$	$\overline{A_3B_1}$
B_2	$\overline{A_1B_2}$	$\overline{A_2B_2}$	$\overline{A_3B_2}$

verschieden?

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

Prinzip des Testens

Beobachtung im Experiment: Stufen- oder Zellmittelwerte

Frage: Sind die Mittelwerte in Wahrheit gleich?
Ist ihre Differenz D=0?

(1) Festlegung eines Signifikanzniveaus α

↓

(2) Berechnung der Prüfgröße: F
deren Häufigkeitsverteilung theoretisch bekannt ist (F-Verteilung)

↓

(3) Berechnung der Wahrscheinlichkeit für diese Prüfgröße: $p(F)$

↓

(4) Rückschluss: $p(F) = p(D = 0)$

↓

(5) Vergleich von p mit α und Treffen der Signifikanzaussage

Aber: Bei dieser Aussage irrt man sich mit einer Wahrscheinlichkeit von $\alpha \cdot 100\%$

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- Zur statistischen Prüfung der Mittelwertsdifferenz lässt sich nun eine **F-verteilte Prüfgröße** konstruieren

$$\boxed{F = \frac{D^2}{\hat{\sigma}_D^2}} \quad \text{mit} \quad \begin{array}{l} df_{Zähler} = 1 \\ df_{Nenner} = df_{Fehler} \end{array}$$

- Aus der F-Verteilung lässt sich nun die **Wahrscheinlichkeit** $p(F)$ ermitteln, mit der der beobachtete Wert der Prüfgröße auftritt – unter der Annahme, dass es in Wahrheit keinen Unterschied zwischen den Stufenmittelwerten gibt ($D=0$).

- Ist $p(F)$ zu klein, liegt ein statistisch signifikanter Unterschied zwischen den verglichenen Mittelwerten vor, die Treatmentstufen haben verschieden große Effekte.

Varianzanalyse

Einfaktorielle ANOVA – Mittelwertevergleiche

- Die **Populationsvarianz** einer paarweisen Mittelwertedifferenz D für Kontraste in der ANOVA ist prinzipiell:

$$\boxed{\hat{\sigma}_D^2 = Var\left(D\right) = \frac{2}{N}\hat{\sigma}^2_{Fehler}}$$

mit N = Anzahl Personen, die in das D eingehen, wobei sich für spezifische Kontraste andere Werte ergeben

- Es wird die in der ANOVA bereits berechnete **Fehlervarianz** zugrunde gelegt.

- **Hinweis:** Ein Mittelwert hat natürlich keine Varianz. $\hat{\sigma}_D^2$ ist die theoretisch zu erwartende Streuung bei vielen Wiederholungen der ANOVA mit neuen Stichproben.

Varianzanalyse

A-Priori Tests – Kontraste I

- Oft sollen in einer zweifaktoriellen ANOVA **beliebige Zellmittelwerte** paarweise miteinander verglichen werden.
- Hat a-priori eine Hypothesenbildung stattgefunden, kann auch dies über **Kontraste** erreicht werden

	A_1	A_2	A_3
B_1	$\overline{A_1B_1}$	$\overline{A_2B_1}$	$\overline{A_3B_1}$
B_2	$\overline{A_1B_2}$	$\overline{A_2B_2}$	$\overline{A_3B_2}$

verschieden?

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

A-Priori Tests – Kontraste I

- In diesem Fall verläuft die **Berechnung der Prüfgröße** exakt wie soeben für den allgemeinen Fall gesehen

Die Prüfgröße ist
$$\boxed{F = \frac{D^2}{\hat{\sigma}_D^2}}$$

mit
$$\boxed{\hat{\sigma}_D^2 = \frac{2}{n} \cdot \hat{\sigma}_{Fehler}^2}$$

- Für die Prüfgröße F gilt: $df_{Zähler} = 1$
$$df_{Nenner} = df_{Fehler}$$

Varianzanalyse

A-Priori Tests – Kontraste II

- Oft sollen auch **Stufenmittelwerte** statt der Zellmittelwerte verglichen werden.
- Hier gehen in die Mittelwertsdifferenz D mehr Personen als nur die n Personen einer Zelle ein

	A_1	A_2	A_3	
B_1	$\overline{A_1B_1}$	$\overline{A_2B_1}$	$\overline{A_3B_1}$	$\bar{B}_1$
B_2	$\overline{A_1B_2}$	$\overline{A_2B_2}$	$\overline{A_3B_2}$	$\bar{B}_2$
	$\bar{A}_1$	$\bar{A}_2$	$\bar{A}_3$	verschieden?

verschieden?

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

A-Priori Tests – Kontraste II

- Sei k die Anzahl der Stufen im anderen Faktor (d.h. der Faktor, zwischen dessen Stufen der Kontrast *nicht* berechnet wird)

  Die Prüfgröße ist $$\boxed{F = \frac{D^2}{\hat{\sigma}_D^2}}$$

  mit $$\boxed{\hat{\sigma}_D^2 = \frac{2}{k \cdot n} \cdot \hat{\sigma}^2_{Fehler}}$$

  weil ja nun $k \cdot n$ Personen in das D eingegangen sind

- Für die Prüfgröße F gilt wieder: $df_{Zähler} = 1$

  $$df_{Nenner} = df_{Fehler}$$

Varianzanalyse

Einfaktorielle ANOVA – A-Posteriori Tests

- Gerade bei umfangreicheren Versuchsdesigns und unklaren Hypothesen werden während der Auswertung Effekte entdeckt, für die zuvor **keine Hypothesen** bestanden.

- In solchen Fällen ist es trotzdem sinnvoll zu prüfen, ob sich Signifikanzen ergeben, um gezielt Fragestellungen für weitere Untersuchungen zu entwickeln

- **Achtung**: Eine im Nachhinein aufgestellte Hypothese mit einem a-posteriori Test zu prüfen und zu belegen, hat faktisch **keine Aussagekraft**
(ist jedoch in der empirischen Forschung durchaus verbreitet)

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

A-Posteriori Tests – Scheffé Test

- Mit dem Scheffé Test können **beliebige Zellen** miteinander verglichen werden.
- Dieser Test ist ein **a-posteriori Test** und kommt eher spät (i.e. bei größeren Unterschieden) zu einer Signifikanzaussage.

	A_1	A_2	A_3
B_1	$\overline{A_1B_1}$	$\overline{A_2B_1}$	$\overline{A_3B_1}$
B_2	$\overline{A_1B_2}$	$\overline{A_2B_2}$	$\overline{A_3B_2}$

verschieden?

Einführung
A-Priori Kontraste
Scheffé Test

Varianzanalyse

A-Posteriori Tests – Scheffé Test

- Für a-posteriori Vergleiche **beliebiger Zellen** (aber nicht der Stufenmittelwerte) einer ANOVA mit $p \cdot q$ Zellen und n Personen je Zelle gilt im Scheffé Test

$$\boxed{F_{corr} = \frac{1}{df_{Zähler}} \cdot F = \frac{1}{df_{Zähler}} \cdot \frac{D^2}{\hat{\sigma}_D^2}}$$

wieder mit:

$$\boxed{\hat{\sigma}_D^2 = \frac{1}{n} \cdot 2 \cdot \hat{\sigma}_{Fehler}^2}$$

- Für die Prüfgröße F_{corr} gilt $df_{Zähler} = p \cdot q - 1$

$$df_{Nenner} = df_{Fehler}$$

Einführung

A-Priori Kontraste

Scheffé Test

Varianzanalyse

Einfaktorielle ANOVA – Weitere A-Posteriori Tests

- Der Scheffé Test ist der konservativste unter den üblichen a-posteriori Tests, d.h. er kommt erst bei größeren Mittelwertsunterschiede zu einer signifikanten Entscheidung.
- Er ist zudem robust gegenüber Verletzungen der Voraussetzungen der ANOVA
- Andere, progressivere Tests sind der Duncan-Test, der Newman-Keuls Test oder der Tukey Test