Thema der Stunde

I. Einführung in die Varianzanalyse: Übersicht und Grundprinzip

II. Kennenlernen am Beispiel einer 2 – faktoriellen ANOVA

Varianzanalyse

Ziele:

- Erklärung einer metrischen Meßgröße X durch ein additives Komponentenmodell
  - a) Abschätzung der Wirkungsweise der Komponenten und ihres Zusammenwirkens
  - b) Zerlegung der Variation der Meßgröße X anteilig auf die Komponenten
- Statistische Absicherung aller im Modell durchgeführten Testungen auf einem Signifikanzniveau

Kern: Zerlegung von X in additive Komponenten

Das Komponentenmodell

<u>Beispiel:</u> 2 Faktoren A (Blutalkohol) und B (Geschlecht).
Abhängige Variable *X*: Punkte im Fahrtest.

			A	
Person Nr		0PM	05PM	1PM
1	w	17	11	8
2	w	20	11	16
3	w	16	13	12
4	w	16	14	13
B 5	w	18	13	14
6	m	21	20	7
7	m	23	19	8
8	m	24	18	7
9	m	20	16	5
10	m	23	22	1
	Summe	198	157	91
	Mittel	19.8	15.7	9.1

Jede Messung X kann genau einer Person unter einer bestimmten Stufe von A und einer bestimmten Stufe von B zugeordnet werden.

Das Komponentenmodell

Beispiel: 2 – faktorielle ANOVA

$$X_{ijm} = \mu + \alpha_i + \beta_j + \alpha_i\beta_j + \varepsilon_{ijm}$$

Hierin ist:

X_{ijm}	Meßwert der $m-$ ten Person auf Variable X in der $i-$ ten Stufe von A und der $j-$ ten Stufe von B.
α_i	Effekt der $i-$ te Stufe des Faktors A.
β_j	Effekt der $j-$ te Stufe des Faktors B.
$\alpha_i\beta_j$	Interaktion der $i-$ ten Stufe des Faktors A mit der $j-$ ten Stufe des Faktors B.
ε_{ijm}	Meßfehler der $m-$ ten Person auf Variable X in der $i-$ ten Stufe von A und der $j-$ ten Stufe von B.

Jede Messung X_{ijm} wird vollständig durch das Komponentenmodell erklärt. Die Frage ist, Gültigkeit der Additivitätsannahme vorausgesetzt, wie groß der Meßfehler ist.

Die Quadratsummenzerlegung

Beispiel: 2 – faktorielle ANOVA

$$QS_{tot} = QS_A + QS_B + QS_{A\times B} + QS_{Fehler}$$

Abschätzung der aufgeklärten Varianz:

$$\eta_A^2 = \frac{QS_A}{QS_{tot}} \qquad \eta_B^2 = \frac{QS_B}{QS_{tot}}$$

Fehler:

$$\eta_\varepsilon^2 = 1 - \eta_{effekt}^2$$

$$\eta_{A\times B}^2 = \frac{QS_{A\times B}}{QS_{tot}} \qquad \eta_{effekt}^2 = \eta_A^2 + \eta_B^2 + \eta_{A\times B}^2$$

Die Quadratsummenzerlegung erlaubt die Abschätzung der Anteile der einzelnen Komponenten und eine Bewertung der Güte des Gesamtmodells

Die unabhängigen Schätzungen der Fehlervarianz

1. Schätzung aus der Variation innerhalb Zellen:

$$E\left[\hat{\sigma}^2_{Fehler}\right] = \frac{QS_{Fehler}}{df_{Fehler}} = \sigma^2_{\varepsilon}$$

2. Schätzung aus der Variation zwischen Zellen

$$E\left[\hat{\sigma}^2_{A}\right] = n \cdot q \cdot \sigma^2_{\alpha} + \sigma^2_{\varepsilon}$$

Hat der Faktor A keinen Effekt, so kann man die Fehlervarianz auch aus der Treatmentvarianz von A schätzen.

Der *F* - Test

In der Nullhypothese fordern wir, daß ein Treatmentfaktor (hier im Beispiel A) keinen Effekt habe:

$$\sigma_\alpha^2 = 0$$

Dann erwarten wir:

$$F = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_{Fehler}^2} = \frac{n \cdot q \cdot \sigma_\alpha^2 + \sigma_\varepsilon^2}{\sigma_\varepsilon^2} = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2} = 1$$

Quotienten von Varianzen sind *F*- verteilt. Der Erwartungswert unter der Nullhypothese für den *F*- Quotienten ist 1.

Der F- Test gibt die Auftretenswahrscheinlichkeit für den gefundenen oder einen extremeren F- Bruch an unter der Annahme, das der wahre F- Quotient 1 beträgt.

Die F - Ergebnistabelle

Die Ergebnisse der F- Testung aller Effekte werden in einer Tabelle zusammengefasst:

QdV	QS	df	$\widehat{\sigma}^2$	F	p
A	$582.8\overline{6}$	2	$291.4\overline{3}$	60.505	$4.22 \cdot 10^{-10}$
B	$16.1\overline{3}$	1	$16.1\overline{3}$	3.3495	0.07967
$A \times B$	$272.8\overline{6}$	2	$136.4\overline{3}$	28.325	$4.8 \cdot 10^{-7}$
Fehler	115.6	24	$4.81\overline{6}$		
Total	$987.4\overline{6}$	29			

Die Freiheitsgrade (*df*) für Zählervarianz und Nennervarianz ergeben sich sinnvoll aus den Tabellierungen der Zellmittelwerte.

Die F- Test Ergebnistabelle fasst die Hauptresultate der Signifikanzprüfung durch die ANOVA zusammen.

Die η^2- Ergebnistabelle

Die η^2- Tabelle fasst die Varianzaufklärung zusammen:

$$
\begin{array}{llcl}
\text{Faktor A} & \frac{582.8\overline{6}}{987.4\overline{6}} \cdot 100\% & = & 59.43\% \\
\text{Faktor B} & \frac{16.1\overline{3}}{987.4\overline{6}} \cdot 100\% & = & 1.634\% \\
\text{Interaktion } A \times B & \frac{272.8\overline{6}}{987.4\overline{6}} \cdot 100\% & = & 27.63\% \\
\hline
\text{Total} & & = & 88.69\%
\end{array}
$$

Anteil der Fehlervarianz: 100%- 88.69% = 11.31%

Die η^2- Tabelle gibt einen Überblick über die Modellgüte und den Anteil der einzelnen Faktoren.

Kontraste

- Sind Faktoren oder deren Interaktionen signifikant, beantwortet die Analyse der Kontraste die Frage, welche Faktufenmittelwerte sich von welchen anderen signifikant unterscheiden

- In der ANOVA können alle Mittelwertsvergleiche auf einem Signifikanzniveau abgesichert werden.

Die Kontrastanalyse liefert die wichtigsten Informationen zu der Interpretation der Analyse und der Prüfung der inhaltlichen Hypothesen.

Voraussetzungen

1) Normalverteilung der Fehlerkomponenten in den einzelnen Stichproben;

2) Varianzhomogenität der Fehlerkomponenten: Die Fehlervarianzen unter den einzelnen Faktorstufen müssen gleich sein;

3) Treatmenteffekte und Fehlereffekte müssen additiv sein. (Fehlervarianz darf nicht mit dem Mittelwerten der Faktorstufen korrelieren)

Die Voraussetzungen können jeweils mit geeigneten Testverfahren überprüft werden. Zu beachten bei der Wahl des α- Niveaus ist, daß es für die Beibehaltung der H0 keine Sicherheit gibt.