Multiples Testen

-Kontrasttests-

Dr. Martin Scharpenberg

MSc Medical Biometry/Biostatistics

WiSe 2019/2020

Kontrasttests im ANOVA-Modell

- k Gruppen mit Zielvariablen

$$Y_{ij}, \qquad j = 1, \ldots, k \text{ (Gruppe)}, \quad i = 1, \ldots, n_j \text{ (Individuum)}$$

- **Modellannahmen:** $Y_{ij} \sim N(\mu_j, \sigma^2)$ unabhängige Beobachtungen.
- Beachte die Varianzinhomogenität (σ^2 in allen Gruppen gleich)
- Varianzanalyse testet nur die "globale Nullhypothese":

$$H_0 : \mu_1 = \cdots = \mu_k \ .$$

- Oft Frage, welche Gruppen sich nun voneinander unterscheiden.
- Frage kann mit sogenannten *Kontrasten* beantwortet werden.

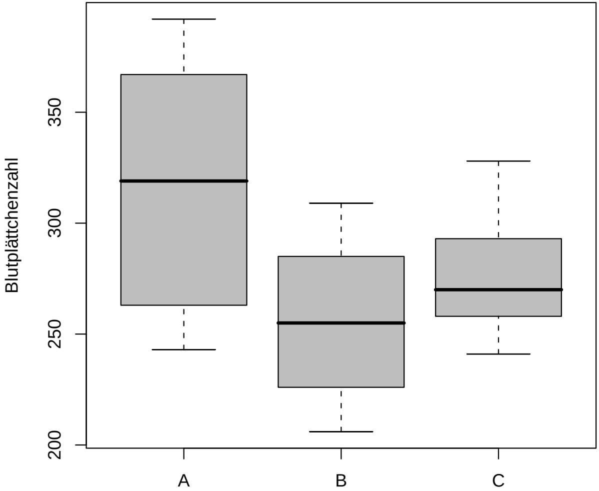
Beispiel: Drei Beatmungsmethoden (I)

22 Patienten mit künstlicher Beatmung wurden drei Beatmungsgruppen per Zufall zugeteilt *(randomisiert)*

- **Gruppe A:** 50% Stickstoff und 50% Sauerstoffgemisch für 24 Stunden.
- **Gruppe B:** 50% Stickstoff und 50% Sauerstoffgemisch nur wärend der Operation.
- **Gruppe C:** kein Stickstoff, 35-50% Sauerstoff für 24 Stunden.

Endpunkt: Die Wirkung der Beatmung wird durch die *Blutplättchenzahl* nach 24 stündiger Beatmung beurteilt.

Beispiel: Drei Beatmungsmethoden (II)



Beispiel: Kontraste für paarweise Vergleiche

Wir interessieren uns nun für die folgenden Nullhypothesen:

$$H_0^{12}: \mu_1 = \mu_2, \quad H_0^{13}: \mu_1 = \mu_3, \quad H_0^{23}: \mu_2 = \mu_3$$

oder etwas umgeschrieben

$$H_0^{12}: \mu_1 - \mu_2 = 0, \qquad H_0^{13}: \mu_1 - \mu_3 = 0$$

$$\text{und} \qquad H_0^{23}: \mu_2 - \mu_3 = 0$$

Paarvergleiche für Blutplättchenzahl

```
proc glm data=redf;
class Group;
model Foliate = Group;
estimate '1 - 2' group 1 -1  0;
estimate '1 - 3' group 1  0 -1;
estimate '2 - 3' group 0  1 -1;
run;
```

Parameter	Estimate	Standard Error	t Value	Pr > |t|
1 - 2	**60.1805556**	**22.2159400**	**2.71**	**0.0139**
1 - 3	38.6250000	26.0644258	1.48	0.1548
2 - 3	-21.5555556	25.5014129	-0.85	0.4085

Wir können H_0^{12} aber weder H_0^{23} noch H_0^{13} verwerfen.

D.h. nur zwischen den Beatmungsmethoden A und B kann ein Unterschied nachgewiesen werden.

Paarvergleiche für Blutplättchenzahl

```
proc glm data=redf;
class Group;
model Foliate = Group;
estimate '1 - 2' group 1 -1  0;
estimate '1 - 3' group 1  0 -1;
estimate '2 - 3' group 0  1 -1;
run;
```

Parameter	Estimate	Standard Error	t Value	Pr > |t|
1 - 2	**60.1805556**	**22.2159400**	**2.71**	**0.0139**
1 - 3	38.6250000	26.0644258	1.48	0.1548
2 - 3	-21.5555556	25.5014129	-0.85	0.4085

Wie berechnet man die t-Teststatistiken?

$$\texttt{t Value} = \texttt{Estimate}/\texttt{Standard Error}$$

Wie berechnet man Estimate und Standard Error?

Beispiel: Kontrasttest für A versus B (H_0^{12})

- **Schätzer für Differenz zwischen A und B**

$$\texttt{Estimate "1 - 2"} = (\bar{y}_1 - \bar{y}_2)$$

- **Standardfehler der Differenz zwischen A und B**

$$\texttt{Standard Error "1 - 2"} = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

wobei $\hat{\sigma}^2 = \sum_{j=1}^{3}(n_j - 1) \cdot s_j^2/(N - 3)$

- **Unterschied zum t-Test:** Beim t-Test schätzt man den Standardfehler durch

$$\hat{\sigma}_{12}^2 = [(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2]/(n_1 + n_2 - 2)$$

Berechnung des p-Wertes

- Die Teststatistik für H_0^{12} ist

$$T_{12} = (\bar{y}_1 - \bar{y}_2) / \left(\hat{\sigma} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)$$

- Der p-Wert für H_0^{12} ist damit

$$P(|T_{12}^*| \geq |t|)$$

  wobei t der beobachtete Wert von T_{12} (also `t Value`) und T_{12}^* die Teststatistik einer unabhängigen Studie unter der Nullhypothese H_0^{12} sind.
- Was ist die Verteilung von T_{12}^* (d.h. von T_{12} unter H_0^{12})?

Verteilung der t-Teststatistik T_{12} unter H_0^{12}

- **Schätzer von $\mu_1 - \mu_2$:**

$$(\bar{y}_1 - \bar{y}_2) \underset{H_0^{12}:\mu_1=\mu_2}{\sim} N\left(0, \sigma^2(\frac{1}{n_1} + \frac{1}{n_2})\right)$$

- **Schätzer für Standardfehler von $\bar{y}_1 - \bar{y}_2$:**

$$s_{12} = \hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- **Teststatistik für H_0^{12}:**

$$T_{12} = (\bar{y}_1 - \bar{y}_2) / \left(\hat{\sigma}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) \underset{H_{12}}{\sim} t_{N-3}$$

Kontrasttests im Allgemeinen

- **Definition von Konrasten beim Vergleich von k Gruppen**

  Ein Kontrast ist ein k-dimensionaler Vektor $c = (c_1, \dots, c_k)$ mit der Eigenschaft

  $$c_1 + \ldots + c_k = 0$$

- **Definition der Nullhypothese eines Kontrasts von k Gruppen**

  Die Nullhypothese eines Kontrasts $c = (c_1, \dots, c_k)$ ist

  $$H_0^c : c^T \cdot \mu = c_1 \cdot \mu_1 + \ldots + c_k \cdot \mu_k = 0.$$

- **Beispiel: Paarvergleich bei drei Gruppen**

  Die Nullhypothese des Paarvergleichs $H_0^{12} : \mu_1 - \mu_2 = 0$ ist identisch zur Nullhypothese H_0^c des Kontrasts $c = (1, -1, 0)$.

Kontrasttests im Allgemeinen

- **Schätzer von $c^T \cdot \mu$:**

$$c^T \cdot \hat{\mu} = \sum_{j=1}^{k} c_j \cdot \bar{y}_j \underset{H_0^c}{\sim} N\left(0, \sigma^2 \sum_{j=1}^{k} \frac{c_j^2}{n_j}\right)$$

- **Schätzer für Standardfehler von $c^T \cdot \hat{\mu}$:**

$$s_c = \hat{\sigma}\sqrt{\sum_{j=1}^{k} \frac{c_j^2}{n_j}} \quad \text{wobei} \quad \hat{\sigma}^2 = \sum_{j=1}^{k} (n_j - 1) s_j^2 / (N - k)$$

- **Teststatistik für H_0^c:**

$$T = (c^T \cdot \hat{\mu})/s_c \underset{H_0^c}{\sim} t_{N-k}$$

Beispiel: Vergleich von drei Gruppen

- **Gruppe A:** 50% Stickstoff und 50% Sauerstoffgemisch für 24 Stunden.
- **Gruppe B:** 50% Stickstoff und 50% Sauerstoffgemisch nur wärend der Operation.
- **Gruppe C:** kein Stickstoff, 35-50% Sauerstoff für 24 Stunden.

Wir könnten die folgenden Fragen stellen:

- Gibt es Unterschiede zwischen den beiden Methoden mit Stickstoff (Gruppen A versus B)?
- Gibt es einen Unterschied zwischen der Methode ohne Stickstoff (Gruppe C) und den Methoden mit Stickstoff (Gruppe C versus Gruppen A und B zusammen)? D.h.: Gibt es einen Effekt von Stickstoff?

Helmert-Kontraste (für drei Gruppen)

Man nennt die folgenden Kontraste *Helmert-Kontraste*

$$
\begin{aligned}
c^{(1)} &= (\ \ 1, -1, 0) &\leftrightarrow\quad & H_0^{c^{(1)}} : \mu_1 - \mu_2 = 0 \\
c^{(2)} &= (-0.5, -0.5, 1) &\leftrightarrow\quad & H_0^{c^{(2)}} : \mu_3 - (\mu_1 + \mu_2)/2 = 0
\end{aligned}
$$

```
proc glm data=redf;
 class Group;
 model Foliate = Group;
 estimate '1 - 2' group 1 -1  0;
 estimate '3 - (1+2)/2' group -0.5  -0.5 1;
run;
```

Parameter	Estimate	Standard Error	t Value	Pr > |t|
1 - 2	60.1805556	22.2159400	2.71	0.0139
3 - (1+2)/2	-8.5347222	23.2691035	-0.37	0.7178

Wir können $H_0^{c^{(2)}}$ nicht verwerfen, d.h. wir können keinen Effekt von Stickstoff nachweisen.

Multipler Fehler bei Gruppenvergleichen

- Vergleichen k unabh. Gruppen paarweise durch Kontrasttests.
- Signifikanzniveau bei jedem Kontrasttest $\alpha = 0.05$.
- Es gilt: $H_0 : \mu_1 = \cdots = \mu_k$.
- Kontrasttests sind **nicht** statistisch unabhängig.
- Daher multipler Fehler nicht wie bei unabhängigen Tests.

Multipler Fehler: Wie groß ist die Wahrscheinlichkeit, dass mindestens einer der $h = k(k-1)/2$ Kontrasttests signifikant ist?

Zahl der Gruppen k	2	3	4	5
Zahl der Hypothsesen h	1	3	6	10
Multiple Fehlerwahr.	0.05	0.12	0.20	0.28
zum Vergl. $1-(1-0.05)^h$	0.05	0.14	0.27	0.40

Übung: Verifizieren Sie die multiplen Fehlerwahrscheinlichkeiten durch eine Simulation (in R oder SAS).