# Tutorium Allgemeines Lineares Modell

BSc Psychologie SoSe 2022

14. Termin: Übungsfragen

Belinda Fleischmann

Welche Aussage bzgl der Kovariationsterme $(x_i - \bar{x})(y_i - \bar{y})$ treffen zu.

1. Häufige Abweichungen der $x_i$ und $y_i$ von ihren Mittelwerten (richtungsgleich oder richtungsungleich)
   $\Rightarrow$ hohe absolute Korrelation
2. Häufige Abweichungen der $x_i$ und $y_i$ von ihren Mittelwerten (richtungsgleich oder richtungsungleich)
   $\Rightarrow$ hohe positive Korrelation
3. Wenige Abweichungen der $x_i$ und $y_i$ von ihren Mittelwerten (richtungsgleich oder richtungsungleich)
   $\Rightarrow$ hohe negative Korrelation
4. Wenige Abweichungen der $x_i$ und $y_i$ von ihren Mittelwerten (richtungsgleich oder richtungsungleich)
   $\Rightarrow$ hohe absolute Korrelation

Es seien $A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ und $B := \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}$. Was ist das Ergebnis des Matrixprodukts $AB$

1. $AB = \begin{pmatrix} 5 & 8 \\ 13 & 20 \end{pmatrix}$
2. $AB = \begin{pmatrix} 10 & 7 \\ 22 & 15 \end{pmatrix}$
3. $AB = \begin{pmatrix} 5 & 7 \\ 13 & 15 \end{pmatrix}$
4. $AB = \begin{pmatrix} 10 & 7 \\ 10 & 24 \end{pmatrix}$

Welche Aussage bzgl des sphärischen Kovarianzmatrixparameters trifft zu?

1. Die Einträge auf der Diagonalen entprechen den Kovarianzen von $n$ univariat normalverteilen Zufallsvariablen.
2. Ein sphärischer Kovarianzmatrixparameter hat die Form $\sigma^2 I_n$.
3. Für alle Elemente des Kovarianzmatrixparametersm, die nicht auf der Diagonalen liegen, gilt $\sigma > 0$.
4. Sphärische Kovarianzmatrixparamter von $n$-variaten Normalverteilungen entsprechen $n$ abhängigen univariaten Normalverteilungen und umgekehrt.

Welche der unten stehenden Modellformulierungen beschreibt **nicht** das Szenario einer einfachen linearen Regression?

1. $y \sim N(X\beta, \sigma^2 I_n)$ mit $X \in \mathbb{R}^{n\times 2}, \beta \in \mathbb{R}^2, \sigma^2 > 0.$
2. $y = X\beta + \varepsilon \Leftrightarrow \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} = \begin{pmatrix} \beta_0 + x_1\beta_1 + \varepsilon_1 \\ \vdots \\ \beta_0 + x_n\beta_1 + \varepsilon_n \end{pmatrix}$
3. $y \sim N(X\beta, \sigma^2 I_n)$ mit $X \in \mathbb{R}^{n\times p}, \beta \in \mathbb{R}^p, \sigma^2 > 0, p = 1$
4. $y = X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 I_n)$ mit $X \in \mathbb{R}^{n\times p}, \beta \in \mathbb{R}^p, \sigma^2 > 0, p = 2$

Es sei $y = X\beta + \varepsilon$ mit $\varepsilon \sim N(0_n, \sigma^2 I_n)$, dann ist ...

1. $\hat{\sigma}^2 := \frac{(y-X\hat{\beta})^T(y-X\hat{\beta})}{n-p}$ ein unverzerrter Schätzer von $\sigma^2 I_n$
2. $\hat{\beta} := (X^T X)^{-1} X^T y$ ein unverzerrter Maximum Likelihood Schätzer von $\beta \in \mathbb{R}^p$
3. $\hat{\sigma}^2 := \frac{(y-X\hat{\beta})^T(y-X\hat{\beta})}{n-p}$ ein Maximum Likelihood Schätzer von $\sigma^2$
4. $\hat{\beta} := (X^T X)^T X^{-1} x$ ein unverzerrter Maximum Likelihood Schätzer von $\beta \in \mathbb{R}^p$

Welche Aussage bzgl der WDF einer nichtzentralen $t$-Zufallsvariablen trifft zu?

1. Je größer Stichprobengröße $n$ desto höher ist der Wert des Nichtzentralitätsparameters $\delta$.
2. Je größer Stichprobengröße $n$ desto geringer ist der Wert des Nichtzentralitätsparameters $\delta$.
3. Für $\delta = 0$ gilt, dass die WDF um $t = 0$ herum symmetrisch ist.
4. Keine der Antworten ist zutreffend.

# 7

Welche Aussage trifft **nicht** zu?

1. Im Nenner der T-Teststatistik steht der Term $\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}$, welche die die geschätzte stichprobenumfangskalierte Datenvariabilität quantifiziert.
2. Im Nenner der T-Statistik steht $\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}$
3. Die T-Statisik ist eine Funktion der Parameterschätzer und des Nullhypothesenparameters $\beta_0$ und quantifiziert somit das Verhältnis von Signal zu Rauschen.
4. Die T-Teststatistik ist nicht-zentral $t$-verteilt und zentral $t$-verteil, wenn $\delta = 0$.

# 8

Welche Aussage bzgl eines faktoriellen Studiendesigns ist **nicht** zutreffend?

1. Die Werte der kontinuierlichen unabhängigen Variable, werden oft Level genannt.
2. In einem zweifaktoriellen experimentellen Design mit 16 Bedingungen und acht Leveln in Faktor B, muss Faktor A vier Level haben.
3. Intuitiv spricht man vom Vorliegen einer Interaktion zweier Faktoren mit jeweils zwei Leveln, wenn der Unterschied der Gruppenmittelwerte von Faktor A zwischen Level 1 und 2 unterschiedlich für Level 1 und Level 2 von Faktor B ausgeprägt ist.
4. Wenn in einem zweifaktoriellen experimentellem Design keine Haupteffekte vorliegen, ist es unwahrscheinlich, dass eine Interkation der Faktoren besteht.

Das praktische Vorgehen bei bei Durchführung eines zweiseitgen Level-$\alpha_0$-Einstichproben-T-Tests beinhaltet...

1. ...die Annahme, dass ein Datensatz $v_1, ...v_n$ eine Realisierung von $y_i \sim N(\mu, \sigma^2)$ u.i.v. für $i = 1, ..., n$ mit unbekannten Parametern $\mu$ und $\sigma^2 > 0$ ist.
2. ...die Berechnung eines Signikanzniveaus, auch p-Wert genannt, bei welchem man die Nullhypothese basierend auf einem vorliegendem Wert der Teststatistik ablehnen würde.
3. ...die Berechnung eines Nichtzentralitätsparameters $\delta$, um eine Aussage darüber zu treffen, wie stark die Verteilung der T-Teststatistik von einer zentralen $t$-Verteilung abweicht.
4. ... die Entscheidung über Annahme oder Ablehnung einer Nullhypothese, basierend darauf, wie wahrscheinlich die Alternativhypothese in $\alpha_0 \cdot 100$ von $100$ Fällen zutrifft.

Für $i = 1, ..., p$ und $j = 1, ..., n_i$ sei $y_{ij}$ die $j$te Datenvariable in der $i$ten Gruppe eines EVA Szenarios. Weiterhin seien

$\bar{y} := \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{n_i} y_{ij}$ das *Gesamtstichprobenmittel*

$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij}$ das $i$*te Stichprobenmittel*

Welche Aussage bzgl der Quadratsummenzerlegung bei einfaktorieller Varianzanalyse ist nicht zutreffend?

1. SQT $:= \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$ ist die Total Sum of Squares.
2. SQB $:= \sum_{i=1}^{p} n_i (\bar{y}_i - \bar{y})^2$ ist die Within Sum of Squares.
3. SQT $=$ SQB $+$ SQW
4. Die Between Sum of Squares quantifiziert die Variabilität der Daten zwischen Gruppen, definiert als die Summe der quadrierten Abstände zwischen allen gruppenspezifischen Stichprobenmitteln zum Gesamtstichprobenmittel. Mit anderen Worten, quantifiziert die SQB, inwieweit die gruppenspezifischen Stichprobenmittel voneinander abweichen.