

Ergebnisbericht

Die Klausur zum Modul B2 Inferenzstatistik im Sommersemester 2022 fand am 14.01.2022 von 11.00 - 12.00 Uhr in Hörsaal 5, Gebäude 16 der OVGU mit 34 Teilnehmer:innen statt. Die Klausur bestand aus 30 Multiple Choice Aufgaben mit jeweils vier Antwortmöglichkeiten und jeweils genau einer richtigen Antwort. Die Klausurlösung ist diesem Bericht beigelegt.

Bewertungschema

Die Aufteilung der zugelassenen Noten auf die erreichten Prozentpunkte wurde anhand untenstehender Tabelle vorgenommen. Diese trifft folgende Zuordnung der erreichten Prozentpunkte zu den zugelassenen Noten anhand von geschlossenen Prozentpunktintervallen.

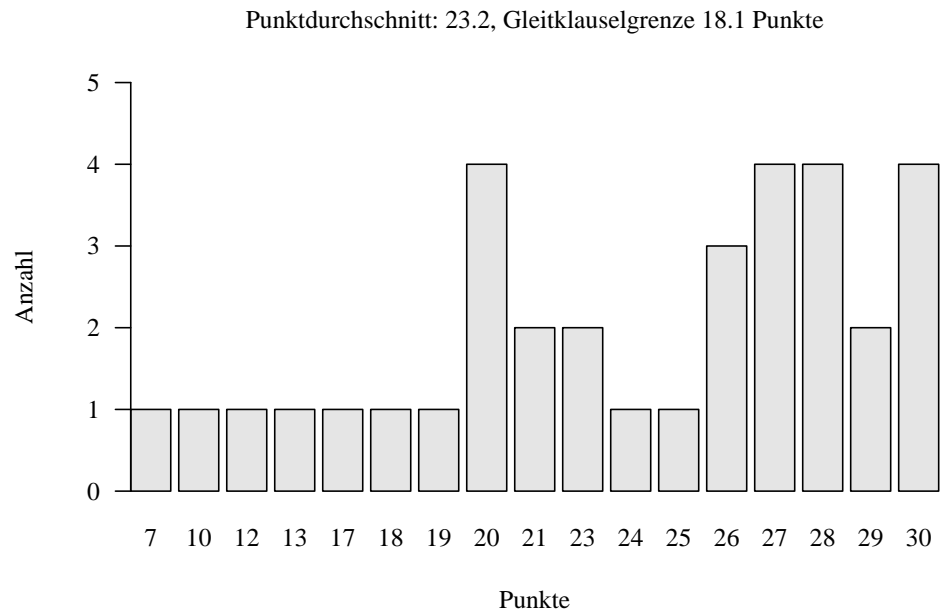
\leq	\geq	Note
100	95	1,0
94	90	1,3
89	85	1,7
84	80	2,0
79	75	2,3
74	70	2,7
69	65	3,0
64	60	3,3
59	55	3,7
54	50	4,0
49	0	5,0

Es ergibt sich folgendes Punktenotenschema, wobei < 15 Punkte mit 5.0 bewertet wurden.

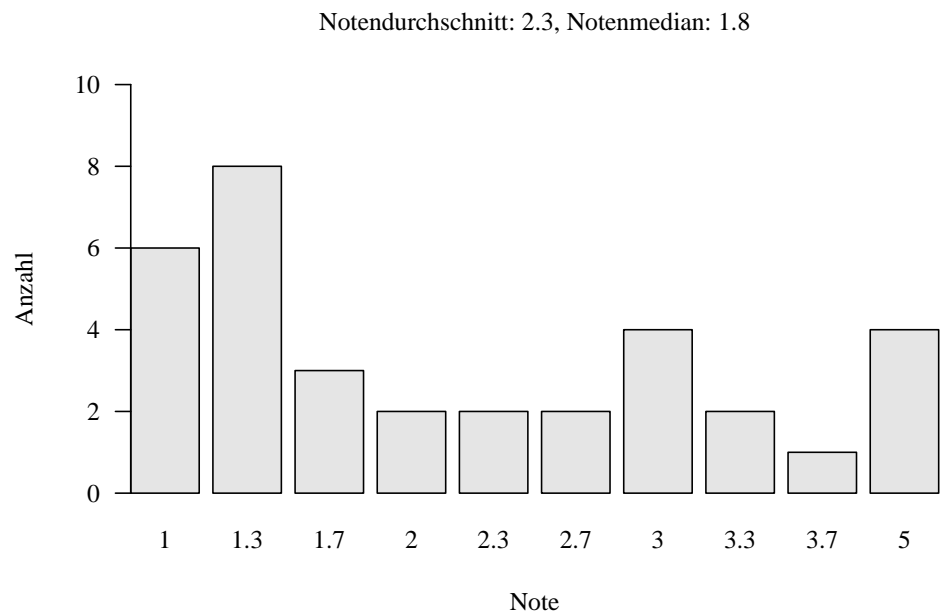
Punkte	Prozent	Note
30	100,0	1,0
29	96,7	1,0
28	93,3	1,3
27	90,0	1,3
26	86,7	1,7
25	83,3	2,0
24	80,0	2,0
23	76,7	2,3
22	73,3	2,7
21	70,0	2,7
20	66,7	3,0
19	63,3	3,3
18	60,0	3,3
17	56,7	3,7
16	53,3	4,0
15	50,0	4,0

Ergebnisse

Die nachfolgende Abbildung zeigt die absolute Häufigkeitsverteilung der erzielten Punkte.



Die nachfolgende Abbildung zeigt die absolute Häufigkeitsverteilung der erreichten Noten.



OTTO-VON-GUERICKE-UNIVERSITÄT MAGDEBURG

Institut für Psychologie

Abteilung Methodenlehre I: Methoden der experimentellen und neurowissenschaftlichen Psychologie

Prof. Dr. Dirk Ostwald

Klausur Modul B2 Inferenzstatistik

Termin: 14.07.2022

Name, Vorname: _____

Matrikelnummer: _____

Bearbeitungshinweise

- Die Klausur besteht aus **30 Aufgaben**. Sie haben zur Bearbeitung **60 Minuten** Zeit.
- Bei jeder Aufgabe sind jeweils **vier Antwortmöglichkeiten** vorgegeben, es trifft **immer genau eine** Antwort zu. Bitte kreuzen Sie bei jeder Aufgabe die zutreffende Antwort an.
- Für jede richtig gelöste Aufgabe erhalten Sie einen Punkt.

Viel Erfolg!

1. Welche Aussage zur Definition der Ausgleichsgerade trifft zu?
 - a) Die Ausgleichsgerade ist ohne Bezug zu einer Wertemenge definiert.
 - b) **Die Ausgleichsgerade ist eine linear-affine Funktion.**
 - c) Die Ausgleichsgerade minimiert eine Summe von absoluten Abweichungen.
 - d) Die Ausgleichsgerade ist eine quadratische Funktion.

2. $f_{\hat{\beta}} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f_{\hat{\beta}}(x) := \hat{\beta}_0 + \hat{\beta}_1 x$ sei die Ausgleichsgerade eines Datensatzes $\{(x_1, y_1), \dots, (x_n, y_n)\}$. Welche Aussage trifft zu?
 - a) $\hat{\beta}_0$ ist das Stichprobenmittel der $y_i, i = 1, \dots, n$.
 - b) $\hat{\beta}_1$ ist die Stichprobenvarianz der $x_i, i = 1, \dots, n$.
 - c) $\hat{\beta}_0$ ist der Schnittpunkt der Ausgleichsgerade mit der x -Achse.
 - d) $\hat{\beta}_1$ **beschreibt die Steigung der Ausgleichsgerade.**

3. Welche Aussage zum Zusammenhang von Ausgleichsgerade und einfacher linearer Regression trifft zu?
 - a) **Die Ausgleichsgeradenparameter und die ML Parameterschätzer der einfachen linearen Regression sind identisch.**
 - b) Sowohl die Ausgleichsgerade als auch das Modell der einfachen linearen Regression gehen von normalverteilten Fehlervariablen aus.
 - c) Die Ausgleichsgerade erlaubt eine explizite probabilistische Beschreibung der Unsicherheit ihrer Datenpassung.
 - d) Bei der einfachen linearen Regression wird davon ausgegangen, dass die Daten Realisierungen von Bernoullizufallsvariablen sind.

4. Für eine Wertemenge $\{(x_1, y_1), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$ und ihre zugehörige Ausgleichsgerade $f_{\hat{\beta}}$ seien \bar{y} und \hat{y}_i für $i = 1, \dots, n$ das Stichprobenmittel der y -Werte und die durch die Ausgleichsgerade erklärten Werte, respektive. Welche Aussage trifft dann zu?
 - a) $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ ist die Total Sum of Squares.
 - b) $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ **ist die Residual Sum of Squares.**
 - c) $\sum_{i=1}^n (y_i - \bar{y})^2$ ist die Explained Sum of Squares.
 - d) Es gilt $SQT = SQE - SQR$.

5. Welche Aussage zum Bestimmtheitsmaß R^2 bei Ausgleichsgerade und Korrelation trifft **nicht** zu?
 - a) R^2 ist die quadrierte Stichprobenkorrelation eines Datensatzes.
 - b) Es gilt $0 \leq R^2 \leq 1$.
 - c) $R^2 = 0$ beschreibt eine denkbar schlechte Erklärung der Daten durch die Ausgleichsgerade.
 - d) R^2 **ergibt sich aus dem Verhältnis von Residual Sum of Squares und Total Sum of Squares.**

6. Es seien

$$A := \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \end{pmatrix}, x := \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \text{ und } b := \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1)$$

Dann gilt

a) $Ax + b = (5, 10)^T$

b) $Ax + b = \begin{pmatrix} 4 \\ 9 \end{pmatrix}$.

c) $Ax + b = \begin{pmatrix} 4 & 0 \\ 8 & 1 \end{pmatrix}$.

d) $Ax + b$ ist nicht definiert.

7. Es sei $X = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$. Dann gilt

a) Die Determinante von X ist -2 .

b) Die Determinante von X ist 0 .

c) Die Determinante von X ist $(1, 0)^T$.

d) Die Determinante von X ist nicht definiert.

8. Welche Aussage zur Inversen einer Matrix trifft zu?

a) Die Inverse einer Matrix ist für Matrizen beliebiger Dimension definiert.

b) Multiplikation einer Matrix mit ihrer Inversen ergibt die Nullmatrix.

c) **Zu manchen quadratischen Matrizen gibt es keine Inverse.**

d) Die Inverse zu einer Matrix A wird mit A^T bezeichnet.

9. Wie lautet die funktionale Form der WDF einer multivariaten Normalverteilung mit Parametern μ, Σ ?

a) $p(v) := (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(v - \mu)^T \Sigma^{-1} (v - \mu)\right)$.

b) $p(v) := (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \ln\left(-\frac{1}{2}(v - \mu)^T \Sigma^{-1} (v - \mu)\right)$.

c) $p(v) := (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}v^T (\Sigma - \mu)^{-1} v\right)$.

d) $p(v) := (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(v - \mu)^T \Sigma^{-1} (v - \mu)\right)$.

10. Welche Aussage zum Theorem zu Unabhängigen normalverteilten Zufallsvariablen trifft **nicht** zu?

a) Das Theorem stellt eine Beziehung zwischen der WDF einer n -variaten Normalverteilung mit sphärischem Kovarianzmatrixparameter und der WDF der gemeinsamen Verteilung von n univariaten normalverteilten Zufallsvariablen her.

b) **Formal besagt das Theorem, dass $N(v; \mu, \sigma^2 I_n) = \sum_{i=1}^n N(v_i; \mu_i, \sigma^2)$.**

c) Als sphärische Kovarianzmatrixparameter bezeichnet man Kovarianzmatrixparameter der Form $\sigma^2 I_n$.

d) Die unabhängigen univariaten normalverteilten Zufallsvariablen des Theorems haben nach Voraussetzung die gleiche Varianz.

11. Welche Aussage zur Definition des Allgemeinen Linearen Modells $y = X\beta + \varepsilon$ in generativer Form trifft zu?
- y ist ein nicht beobachtbarer Zufallsvektor.
 - $X\beta$ ist der Erwartungswert von y .
 - Die Komponenten von ε sind identisch, aber nicht unabhängig verteilt.
 - Es gilt $\varepsilon \sim N(1_n, \sigma^2 I_n)$.
12. Welche Aussage zum Szenario von n unabhängig und identisch normalverteilten Zufallsvariablen in ALM Matrixschreibweise trifft **nicht** zu?
- Für die Designmatrix gilt $X := 1_n$.
 - Der Betaparametervektor entspricht dem Erwartungswertparameter der Zufallsvariablen.
 - Das Modell hat genau einen Parameter.**
 - Das Modell kann als $y \sim N(X\beta, \sigma^2 I_n)$ geschrieben werden.
13. Welche Aussage zum Szenario der einfachen linearen Regression in ALM Matrixschreibweise trifft **nicht** zu?
- Für den Betaparametervektor gilt $\beta \in \mathbb{R}^3$.**
 - Eine Spalte der Designmatrix besteht nur aus Einsen.
 - Eine Spalte der Designmatrix entspricht den Werten der unabhängigen Variable.
 - Die Betaparameter können als **y -Achsenabschnitt** und Steigung der Gerade $X\beta$ interpretiert werden.
14. Welche Aussage zur Frequentistischen Verteilung des Betaparameterschätzers des ALMs trifft zu?
- $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$.
 - $\hat{\beta} \sim N(\beta, \sigma^2 I_n)$.
 - $\hat{\beta} \sim N(y, \sigma^2 (X^T X)^{-1})$.
 - $\hat{\beta} \sim \chi^2(n - p)$.
15. Welche Aussage zur Definition der T-Teststatistik trifft zu?
- Der Zähler der T-Teststatistik hat die Form $c^T \beta - c^T \beta_0$.
 - Der Zähler der T-Teststatistik hat die Form $c^T \hat{\beta} - c^T \beta_0$.**
 - Der Nenner der T-Teststatistik hat die Form $\sqrt{\hat{\sigma}^2 c^T (X^T X)^{-1} c}$.
 - Der Nenner der T-Teststatistik hat die Form $\sqrt{\sigma^2 c^T (X^T X)^{-1} c}$.

16. Welche Aussage zur F-Statistik trifft **nicht** zu?
- a) Die F-Statistik ist eng mit einem Likelihood-Quotienten verwandt.
 - b) In der F-Statistik kommen Residualquadratsummen eines vollständigen und eines reduzierten Modells vor.
 - c) **Die F-Statistik ist immer nichtzentral t -verteilt.**
 - d) Die F-Statistik misst eine standardisierte Residualquadratsummenreduktion.
17. Wie viele Faktoren mit jeweils wie vielen Leveln hat ein $2 \times 3 \times 2$ Studiendesign?
- a) **3 Faktoren mit 2, 3 und 2 Leveln, respektive.**
 - b) 12 Faktoren mit jeweils $2 \times 3 \times 2$ Leveln.
 - c) 7 Faktoren mit jeweils $2 \times 3 \times 2$ Leveln.
 - d) 3 Faktoren mit jeweils 12 Leveln.
18. Wie viele experimentelle Bedingungen hat ein $2 \times 3 \times 2$ faktorielles Studiendesign mit Crossed Design?
- a) 5.
 - b) 7.
 - c) **12.**
 - d) 64.
19. Welche Aussage zum Theorem zur Parameterschätzung im Einstichproben-T-Test Modell trifft **nicht** zu?
- a) Der Betaparameterschätzer und das Stichprobenmittel der y_1, \dots, y_n sind identisch.
 - b) Der Betaparameterschätzer ergibt sich zu $\hat{\beta} = \frac{1}{n} \sum_{i=1}^n y_i$.
 - c) Der Varianzparameterschätzer und die Stichprobenvarianz der y_1, \dots, y_n sind identisch.
 - d) **Der Varianzparameterschätzer ergibt sich zu $\hat{\sigma}^2 = \sum_{i=1}^n (y_i - \bar{y})$.**
20. Welche Aussage zum Theorem zur T-Teststatistik des Einstichproben-T-Tests trifft zu?
- a) Der Kontrastgewichtsvektor hat die Form $c = (1, -1)^T$.
 - b) Die T-Teststatistik hängt nicht von der Stichprobengröße n ab.
 - c) **Die T-Teststatistik ist im Allgemeinen nichtzentral t -verteilt.**
 - d) Der Nichtzentralitätsparameter der Verteilung der T-Teststatistik ist gegeben durch $\mu - \mu_0$.
21. Welche Aussage zum Anwendungsszenario eines Zweistichproben-T-Tests trifft **nicht** zu?
- a) Man geht von zwei Gruppen (Stichproben) randomisierter experimenteller Einheiten aus.
 - b) Man nimmt an, dass die Datenvariablen jeder Gruppe jeweils unabhängig und identisch normalverteilt sind.
 - c) **Man setzt die Varianzparameter beider Gruppen als bekannt voraus.**
 - d) Man beabsichtigt, die Unsicherheit beim inferentiellen Vergleich der Gruppenerwartungswertparameter zu quantifizieren.

22. Welche Aussage zum Anwendungsszenario einer einfaktoriellen Varianzanalyse (EVA) trifft **nicht** zu?
- Man geht von zwei oder **mehr** Gruppen (Stichproben) randomisierter experimenteller Einheiten aus.
 - Man nimmt an, dass die Datenvariablen jeder Gruppe jeweils unabhängig und identisch normalverteilt sind.
 - Man setzt die Varianzparameter aller Gruppen als identisch voraus.
 - Man beabsichtigt Interaktionen der Faktoren des Designs inferentiell zu überprüfen.**
23. Welche Aussage zur Betaparameterschätzung im EVA Modell in Effektdarstellung mit Referenzgruppe trifft zu?
- Der Referenzgruppeneffektparameter wird durch das Stichprobenmittel aller Datenpunkte geschätzt.
 - Der Referenzgruppeneffektparameter wird durch die Stichprobenvarianz aller Datenpunkte geschätzt.
 - Die Effektparameter der Nicht-Referenzgruppen werden durch die Stichprobenmittel der entsprechenden Gruppen geschätzt.
 - Die Effektparameter der Nicht-Referenzgruppen werden durch die Differenzen zwischen ihren Stichprobenmitteln und dem Stichprobenmittel der Referenzgruppe geschätzt.**
24. Welcher Zusammenhang besteht zwischen dem Effektstärkenmaß η^2 und der F-Teststatistik F der EVA?
- $\eta^2 = F$.
 - $\eta^2 = \frac{F(p-1)}{F(p-1)+(n-p)}$.
 - $\eta^2 = \frac{F(p-1)}{(n-p)}$.
 - η^2 kann nicht aus der F-Teststatistik F errechnet werden.
25. Aus wie vielen Datenpunkten besteht ein Datensatz eines 2×2 Zweifaktoriellen Varianzanalyse (ZVA) Designs mit 10 Datenpunkten pro Zelle?
- 4.
 - 14.
 - 40.**
 - 200
26. Welchen Wert hat der Erwartungswertparameter μ_{12} in einem additiven 2×2 ZVA Modell mit Referenzgruppe für $\mu_0 = 0$, $\alpha_2 = 2$ und $\beta_2 = 2$?
- 0.
 - 2.**
 - 4.
 - μ_{12} kann aus diesen Angaben nicht berechnet werden.

27. Gegeben seien die Matrizen

$$X_1 := \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad X_2 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \quad X_3 := \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad X_4 := \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \quad (2)$$

Welche dieser Matrizen ist eine mögliche Designmatrix des Modells einer 2×2 ZVA mit Interaktion mit 2 Datenpunkten pro Zelle?

- X_1 .
 - X_2 .
 - X_3 .
 - X_4 .
28. Was ist **kein** typisches Ziel einer Multiplen Regressionsanalyse?
- Die Quantifizierung von Unsicherheit beim inferentiellen Vergleich zweier Gruppenerwartungswertparameter.**
 - Die Quantifizierung des Erklärungspotentials der Variation unabhängiger Variablen für die Variation einer abhängigen Variablen.
 - Die Quantifizierung des Einflusses einzelner unabhängiger Variablen auf die abhängige Variable im Kontext anderer unabhängiger Variablen.
 - Die Prädiktion von Werten einer abhängigen Variable aus Werten von unabhängigen Variablen nach Parameterschätzung.
29. Welchen Wert hat die Betaparameterschätzerkomponente des ersten kontinuierlichen Regressors in einem multiplen Regressionmodell mit Interzeptprädiktor und zwei kontinuierlichen Prädiktoren, wenn die partielle Korrelation der Daten mit dem ersten kontinuierlichen Regressor gegeben den zweiten kontinuierlichen Regressor gleich Null und die Standardabweichung der Daten gleich Eins ist?
- 1.
 - 0.
 - $\frac{1}{2}$.
 - 1.
30. Welcher Kontrastgewichtsvektor kann zum Testen der Nullhypothese zweier gleicher Regressoreffekte im Rahmen eines T-Tests bei einer Multiplen Regressionsanalyse mit drei Regressoren benutzt werden?
- $c = (1, -1)$.
 - $c = (1, 0, -1)$.
 - $c = (1, 0, 1)$.
 - $c = (1, -1, 1, -1)$.