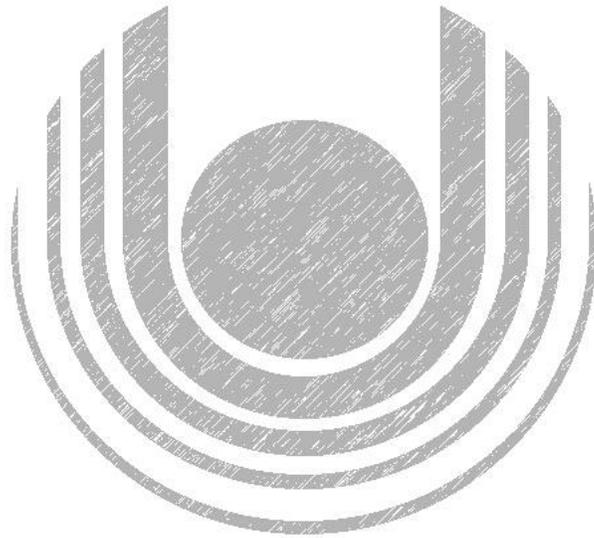


Hinweis zu den zur Verfügung gestellten Klausuren des Lehrstuhls für Angewandte Statistik der FernUniversität in Hagen

Dieser Text ist urheberrechtlich geschützt. Die Ausgabe erfolgt ausschließlich zu persönlichen Übungszwecken der Studierenden. Ausdrücklich untersagt ist die Nutzung des Texts zur Erstellung kommerzieller Angebote sowie die Weitergabe an Dritte. Unerlaubte Vervielfältigung, Verbreitung oder andere Nutzungshandlungen außerhalb der Grenzen des Urheberrechtsgesetzes sind untersagt und werden ggf. rechtlich verfolgt.

<hr/> <p>Name, Vorname</p>	<table border="1"><tr><td style="width: 20px; height: 20px;"></td><td style="width: 20px; height: 20px;"></td></tr></table> <p>Matrikelnummer</p>							



Modulklausur 31821 – Multivariate Verfahren

Datum

Punkte

Note

Termin: 20. März 2020, 11:30 - 13:30 Uhr

Prüfer: Univ.-Prof. Dr. H. Singer

Hinweise zur Bearbeitung der Modulklausur 31821

1. Füllen Sie zunächst den **Kopf des Deckblatts** aus!
2. Es können insgesamt 100 Punkte erreicht werden. Bei Erreichen von 50 Punkten ist die Klausur bestanden. **Bitte kontrollieren Sie sofort, ob Sie ein vollständiges Klausurexemplar erhalten haben.**
3. Zugelassen ist Kurseinheit 1 des Moduls 31821 (Kursnr. 00883) mit farblichen Markierungen, kleinen Aufklebern und/oder textbezogenen Anmerkungen. Nicht zugelassen sind selbst ausgedruckte oder kopierte Kursmaterialien.
Die Verwendung eines Taschenrechners ist dann und nur dann erlaubt, wenn dieser einer der drei folgenden Modellreihen angehört:
 - Casio fx86 oder Casio fx87
 - Texas Instruments TI 30 X II
 - Sharp EL 531

Eventuelle Vorgänger- oder Nachfolgemodelle, die nicht in der oben aufgeführten Liste enthalten sind, sind ebenfalls nicht erlaubt. Die Verwendung anderer Taschenrechnermodelle wird als Täuschungsversuch gewertet und mit der Note „nicht ausreichend“ (5,0) sanktioniert.

Ob ein Taschenrechner einer der drei Modellreihen angehört, können Sie selbst überprüfen, indem Sie die vom Hersteller auf dem Rechner angebrachte Modellbezeichnung mit den oben angegebenen Bezeichnungen vergleichen: Bei vollständiger Übereinstimmung ist das Modell erlaubt. Ist die auf dem Rechner angebrachte Modellbezeichnung umfangreicher, enthält aber eine der oben angegebenen Bezeichnungen vollständig, ist das Modell ebenfalls erlaubt. In allen anderen Fällen ist das Modell nicht erlaubt.

4. Bitte benutzen Sie für Ihre Rechnungen nur die beigelegten Lösungsbögen.
5. Wenn Sie die einzelnen Blätter der Klausur voneinander trennen, **vermerken Sie auf jedem Blatt Ihre Matrikelnummer**. Legen Sie bitte am Ende der Klausur die Blätter wieder zusammen.

Wir wünschen Ihnen viel Erfolg!

Aufgabe 1

(20 Punkte)

Die gemeinsame Verteilung von X und Y sei durch die folgende Kontingenztafel der Auftretenswahrscheinlichkeiten gegeben:

	$Y = 1$	$Y = 2$	$Y = 3$
$X = 1$	0.25	0.15	0.10
$X = 2$	0.10	0.15	0.25

Man bestimme:

- a) den Erwartungswert, sowie die Varianz von X und Y . (5 P.)
- b) die Kovarianz, sowie die Korrelation von X und Y . (5 P.)
- c) die Varianz von $X + Y$. (5 P.)
- d) die bedingten Verteilungen von $X|Y = y$ und $Y|X = x$. (5 P.)

Aufgabe 2

(20 Punkte)

Gegeben sei eine bivariate Zufallsvariable $\mathbf{x} = [X, Y]' \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Diese soll auf den Erwartungswert $\boldsymbol{\mu}_0 = [5, 5]'$ zum Signifikanzniveau $\alpha = 5\%$ getestet werden. Die entsprechenden Mittelwerte seien $\bar{\mathbf{x}} = [6, 6.5]'$. Die wahre Kovarianzmatrix $\boldsymbol{\Sigma}$ ist bekannt und numerisch gleich der empirischen Kovarianzmatrix $\mathbf{S} = \begin{bmatrix} 7.5 & 2.5 \\ 2.5 & 6 \end{bmatrix}$. Zudem sei $N = 34$.

- a) Wie lautet die Nullhypothese? (3 P.)
- b) Wie lautet die dazugehörige Teststatistik? Geben Sie die allgemeine Formel an und rechnen Sie diese explizit aus. (11 P.)
- c) Wie lautet der kritische Wert in diesem Fall? (4 P.)
- d) Wird die Hypothese abgelehnt? (2 P.)

Aufgabe 3

(30 Punkte)

Betrachten Sie das folgende multiple lineare Regressionsmodell

$$Y_i = \beta_1 + \beta_2 X_{i2} + \epsilon_i, \quad i = 1, \dots, N.$$

Nehmen Sie an, dass $\epsilon_i \sim N(0, \sigma^2)$ und unabhängig ist. Folgende Werte wurden nun beobachtet:

Y_i	-5	5	-5	-12.5	17.5
X_{i2}	5	10	0	15	-5

- a) Beschreiben Sie das Modell in Matrix-Form und führen Sie eine kleinste-Quadrate-Schätzung für β_1 und β_2 durch. (14 P.)
- b) Berechnen Sie den Schätzer für die Fehlervarianz $\hat{\sigma}^2$. (8 P.)
- c) Stellen Sie die geschätzte Regressionsgleichung auf und geben Sie entsprechenden Standardabweichungen für $\hat{\beta}_1$ und $\hat{\beta}_2$ an. (8 P.)

Aufgabe 4

(30 Punkte)

Bei einem häufig benutzten Werkstoff, der auf drei verschiedene Weisen hergestellt werden kann, vermutet man einen unterschiedlichen Gehalt an einer krebserregenden Substanz. Von dem Werkstoff wurden für jede der drei Herstellungsmethoden (A, B, C) vier Proben (x_1, x_2, x_3, x_4) je 100g entnommen und folgende Werte für den Gehalt an dieser speziellen krebserregenden Substanz in mg pro Methode gemessen:

	A	B	C
x_1	61	62	65
x_2	58	59	62
x_3	60	61	63
x_4	60	61	62

a) Schätzen Sie den Effekt der Herstellungsmethode auf den Gehalt an der krebserregenden Substanz und interpretieren Sie die Unterschiede. (12 P.)

b) Gehen Sie davon aus, dass der Gehalt an der krebserregenden Substanz approximativ normalverteilt ist. Prüfen Sie zum Signifikanzniveau $\alpha = 0.05$, ob sich die drei Herstellungsmethoden hinsichtlich des Gehalts an der krebserregenden Substanz unterscheiden. Für eine bessere Überschaubarkeit erstellen Sie eine entsprechende ANOVA-Tabelle und tragen Sie die Zwischenergebnisse ein. (18 P.)