
Name, Vorname

--	--	--	--	--	--	--	--

Matrikelnummer



Modulklausur 31821 – Multivariate Verfahren

Datum

Punkte

Note

Termin: 18. März 2016, 9.00 - 11.00 Uhr
Erstprüfer: 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
 - 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

(15 Punkte)

Kennzeichnen Sie die folgenden Aussagen zur kategorialen Regression mit **R** für richtig oder **F** für falsch.

- Bei binären abhängigen Variablen führt das lineare Wahrscheinlichkeitsmodell zu effizienten Schätzern.
- Im Logit-Modell können die Vorzeichen der geschätzten Parameter der erklärenden Variablen nicht interpretiert werden.
- Die Log-odds-ratio ist größer als eins, wenn der geschätzte Koeffizient β positiv ist.
- Der multinomiale Logit-Schätzer erzeugt bei K Kategorien der abhängigen Variable K Vektoren von Parameterschätzern.
- Bei der Likelihood-Quotienten-Statistik wird die Log-Likelihood des vollständigen Modells mit der Log-Likelihood des Nullmodells verglichen. Beim Nullmodell handelt es sich um das Modell, in dem nur die Konstante vorhanden ist und alle anderen Parameter auf 0 gesetzt werden.

Hinweis: Für jede korrekte Kennzeichnung werden 3 Punkte vergeben. Jede falsche Kennzeichnung sowie nicht oder unlesbar gekennzeichnete Felder werden mit 0 Punkten bewertet. Die minimale Punktzahl der Aufgabe beträgt 0 Punkte.

Aufgabe 2

(40 Punkte)

Für ein Unternehmen seien nachfolgende Umsätze (in Mio. Euro) und Werbeausgaben (in Mio. Euro) in den Jahren 2013-2015 bekannt:

Jahr	Umsatz	Werbeausgaben
2013	10	1
2014	11	2
2015	11	4

Untersuchen Sie mit Hilfe der Regressionsanalyse, wie sich die Werbeausgaben auf den Umsatz auswirken.

a) Schätzen Sie zunächst β ! (8 P.)

b) Bestimmen Sie nun die Residuen und die Fehlervarianz. (6 P.)

c) Bestimmen Sie den Determinationskoeffizienten R^2 ! Wieviel Prozent der Varianz wird durch das Modell erklärt? (10 P.)

d) Stellen Sie die ANOVA-Tabelle auf und prüfen Sie, ob $H_0 : \beta_1 = 0$ zum Niveau $\alpha = 5\%$ gilt. (8 P.)

e) Welchen Umsatz erzielt das Unternehmen voraussichtlich, wenn es für die Werbung 7 Mio. Euro ausgibt? Bestimmen Sie einen Schätzer und das Prognoseintervall ($\alpha = 5\%$). (8 P.)

Aufgabe 3

(24 Punkte)

Betrachten Sie den SPSS-Output auf der nächsten Seite. Es wurden 300 bivariate Datenpunkte analysiert.

a) Auf den unteren beiden Grafiken sehen Sie die Daten univariat als Histogramm aufgetragen sowie jeweils eine Normalverteilungskurve. Sprechen die Grafiken für oder gegen eine univariate Normalverteilung der beiden Variablen? Begründen Sie Ihre Antwort.

(2 P.)

b) Gehen Sie von einer bivariaten Normalverteilung der Daten aus. Geben Sie die Maximum-Likelihood-Schätzer für die Verteilungsparameter an (gerundet auf zwei Nachkommastellen)!

(6 P.)

c) Bestimmen Sie zur Erzeugung weiterer Zufallszahlen die Cholesky-Wurzel der Kovarianzmatrix. Nehmen Sie an, dass $z_1 = 0.50$ und $z_2 = -0.2$ zwei Realisierungen von unabhängigen, standard-normalverteilten Zufallszahlen sind. Generieren sie daraus eine bivariate Zufallszahl, die dieselbe Verteilung wie die mit SPSS analysierten Daten hat!

(12 P.)

d) Betrachten Sie wieder den SPSS-Output. Ist die Korrelation zum 5%-Niveau signifikant von 0 verschieden? Begründen Sie Ihre Antwort.

(2 P.)

e) Wenn die Variable 1 einen hohen Wert annimmt, ist dann die Variable 2 eher groß oder eher klein? Begründen Sie Ihre Antwort.

(2 P.)

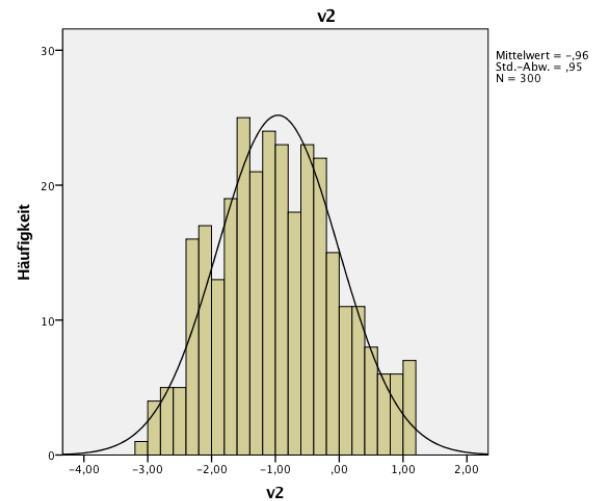
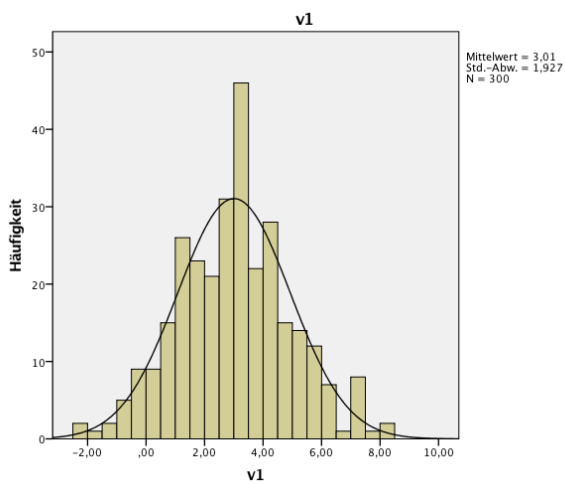
Deskriptive Statistiken

	Mittelwert	Standardabweichung	N
v1	3,0063	1,92708	300
v2	-,9590	,95034	300

Korrelationen

		v1	v2
v1	Korrelation nach Pearson	1	,560**
	Signifikanz (2-seitig)		,000
	Quadratsummen und Kreuzprodukte	1110,378	306,752
	Kovarianz	3,714	1,026
	N	300	300
v2	Korrelation nach Pearson	,560**	1
	Signifikanz (2-seitig)	,000	
	Quadratsummen und Kreuzprodukte	306,752	270,044
	Kovarianz	1,026	,903
	N	300	300

** . Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.



Aufgabe 4

(21 Punkte)

Drei verschiedene Unterrichtsarten sollen untersucht werden. Dazu werden 24 Personen zufällig auf drei Gruppen aufgeteilt. Am Ende des Kurses wird eine Abschlussprüfung durchgeführt und die Punkte jedes Teilnehmers dokumentiert. Es soll untersucht werden, ob sich die einzelnen Gruppen signifikant unterscheiden.

- a) Unterscheiden sich die Gruppen zum 10%-Niveau von einander? Berücksichtigen Sie bei Ihrer Antwort die SPSS-Outputs unten auf dieser Seite und auf der nächsten Seite. (9 P.)
- b) Berechnen Sie das Bonferroni-Konfidenzintervall für den Unterschied zwischen den Gruppen 2 und 3 zum Niveau $\alpha_r = 0.01$. (10 P.)
- c) Welcher Anteil der Varianz des Testergebnisses lässt sich durch die unterschiedlichen Unterrichtsarten erklären? (2 P.)

ONEWAY deskriptive Statistiken

Testergebnis

	N	Mittelwert	Standardabweichung	Standardfehler	95%-Konfidenzintervall für den Mittelwert		Minimum	Maximum	Varianz zwischen den Komponenten
					Untergrenze	Obergrenze			
1,00	8	77,2500	9,77972	3,45765	69,0740	85,4260	60,00	90,00	
2,00	8	79,7500	10,63350	3,75951	70,8602	88,6398	57,00	90,00	
3,00	8	66,1250	11,08973	3,92081	56,8538	75,3962	54,00	86,00	
Gesamt	24	74,3750	11,72812	2,39399	69,4226	79,3274	54,00	90,00	
Modell									
Feste Effekte			10,51501	2,14637	69,9114	78,8386			
Zufallseffekte				4,18766	56,3570	92,3930			38,78869

Einfaktorielle ANOVA

Testergebnis

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	841,750	2	420,875	3,807	,039
Innerhalb der Gruppen	2321,875	21	110,565		
Gesamt	3163,625	23			

T-Test

Gruppenstatistiken

	Gruppe	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Testergebnis	1,00	8	77,2500	9,77972	3,45765
	2,00	8	79,7500	10,63350	3,75951

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
								Untere	Obere	
Testergebnis	Varianzen sind gleich	,023	,883	-4,89	14	,632	-2,50000	5,10777	-13,45507	8,45507
	Varianzen sind nicht gleich			-4,89	13,903	,632	-2,50000	5,10777	-13,46224	8,46224

T-Test

Gruppenstatistiken

	Gruppe	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Testergebnis	1,00	8	77,2500	9,77972	3,45765
	3,00	8	66,1250	11,08973	3,92081

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
								Untere	Obere	
Testergebnis	Varianzen sind gleich	,156	,699	2,128	14	,052	11,12500	5,22763	-,08715	22,33715
	Varianzen sind nicht gleich			2,128	13,784	,052	11,12500	5,22763	-,10362	22,35362

T-Test

Gruppenstatistiken

	Gruppe	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Testergebnis	2,00	8	79,7500	10,63350	3,75951
	3,00	8	66,1250	11,08973	3,92081

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
								Untere	Obere	
Testergebnis	Varianzen sind gleich	,232	,637	2,508	14	,025	13,62500	5,43201	1,97450	25,27550
	Varianzen sind nicht gleich			2,508	13,975	,025	13,62500	5,43201	1,97258	25,27742

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte

LÖSUNGSBOGEN

--	--	--	--	--	--	--

Punkte