

# Lösungen zur Klausur Statistik II

Dr. Andreas Voß  
Sommersemester 2005

Name: \_\_\_\_\_

Mat.Nr.: \_\_\_\_\_

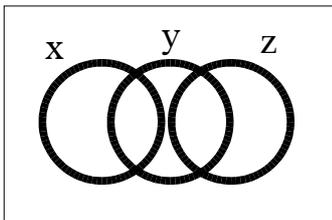
## **Bearbeitungshinweise:**

- Insgesamt können **40 Punkte** erreicht werden. Die Klausur gilt als bestanden, wenn Sie mindestens **20 Punkte** erreichen.
- Verwenden Sie einen **Füller** oder **Kugelschreiber**. Mit Bleistift bearbeitete Klausuren werden nicht akzeptiert!
- Bei den **multiple-choice** Aufgaben sind **Mehrfachnennungen** möglich.
- **Unterschreiben** Sie Ihre Klausur auf der letzten Seite vor dem Abgeben!

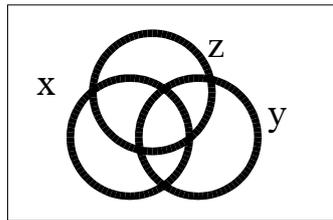
## **1 Bitte ergänzen Sie! (1 Punkt)**

- Die quadrierte multiple Korrelation beschreibt, wieviel Varianz eines Kriteriums durch mehreren Prädiktoren insgesamt aufklärt wird.
- Die quadrierte Semipartialkorrelation beschreibt, wieviel Varianz eines Kriteriums ein neuer Prädiktor zusätzlich zu einem oder mehreren anderen Prädiktoren aufklärt.
- Die Partialkorrelation beschreibt den Zusammenhang zwischen zwei Variablen, der nicht auf eine dritte Variable zurückzuführen ist.
- Die multiple Korrelation beschreibt den Zusammenhang zwischen einem Kriterium und mehreren Prädiktoren.

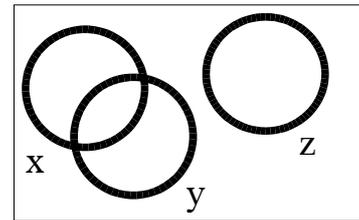
## **2 Folgende Diagramme stellen Zusammenhänge zwischen den Variablen $x$ , $y$ und $z$ dar. Dabei soll $x$ das Kriterium und $y$ und $z$ sollen Prädiktoren sein (2 Punkte).**



(A)



(B)



(C)

- 2.1 In welchem Diagramm gibt es einen Suppressoreffekt?     A
- 2.2. In welchem Fall ist  $r_{y,z}$  am größten?     A
- 2.3. Wie hoch ist die bivariate Korrelation von  $x$  und  $z$  in Diagramm A?      $r = 0$
- 2.4. Besitzt  $z$  in Diagramm B *inkrementelle Validität* gegenüber  $y$ ?     ja

3. Was ist das Ziel einer Multiplen Regression (1 Punkt)

Die Multiple Regression dient der Vorhersage eines Kriteriums durch mehrere Prädiktoren.

4. Was ist der Vorteil der Angabe standardisierter Regressionsgewichte bei der Multiplen Regression? (1 Punkt).

Die standardisierten Koeffizienten sind leichter zu interpretieren, da sie (wie die Korrelation) einen Wertebereich von -1 bis +1 haben.

5. In einer Stichprobe von  $N=78$  wurde eine multiple Korrelation des Kriteriums „Ängstlichkeit“ mit den Prädiktoren „Alter“ und „Optimismus“ von  $R=.27$  gefunden. Ist dieser Zusammenhang statistisch bedeutsam? Hinweis:  $F_{krit} = 3.12$  (2 Punkte)

$$F_{emp} = \frac{R^2 / df_z}{(1 - R^2) / df_N} = \frac{0.27^2 / 2}{(1 - 0.27^2) / 75} = \frac{0.036}{0.012} = 2.95$$

Da  $F_{emp} < F_{krit}$  ist die multiple Korrelation nicht statistisch signifikant!

6. Für eine multiple Regressionsanalyse ergab sich folgendes Ergebnis. Geben Sie die Regressionsgleichungen in der nicht-standardisierten sowie in der standardisierten Form an! (1 Punkt)

Koeffizientenr						
Modell		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten	T	Signifikanz
		B	Standardfehler	Beta		
1	(Konstante)	273.718	1303.139		.210	.835
	x1	15.131	37.017	.063	.409	.685
	x2	-36.771	16.422	-.377	-2.239	.032
	x3	52.769	21.451	.414	2.460	.019

a. Abhängige Variable: y

nicht standardisiert:  $y = 15.12 \cdot x_1 - 36.77 \cdot x_2 + 52.77 \cdot x_3 + 273.72$   
 standardisiert:  $y = 0.06 \cdot x_1 - 0.38 \cdot x_2 + 0.41 \cdot x_3$

7 **In einer Untersuchung soll die Hypothese geprüft werden, dass sich die Arbeitsleistung zwischen 4 Gruppen einer Firma nicht unterscheidet. Aus jeder Gruppe wird eine Stichprobe erhoben, und die Stichprobenmittelwerte werden paarweise verglichen. Dafür werden 6 t-Test durchgeführt. Einer der Tests wird bei  $\alpha=0.05$  signifikant.**

7.1 Muss damit die Hypothese verworfen werden? Begründen Sie kurz Ihre Antwort. (1 Punkt)

*Nein, wegen der Alpha-Fehler-Kumulierung besteht eine erhöhte Wahrscheinlichkeit dafür, dass ein Test zufällig signifikant wurde, selbst wenn es in der Population keinen Unterschied gibt.*

7.2. Was hätten Sie bei der Datenauswertung anders gemacht? Nennen Sie hier 2 Möglichkeiten! (1 Punkt)

*(1) Alphafehler-Adjustierung (Bonferroni-Korrektur)  
(2) Varianzanalytische Auswertung*

8 **Mit einer einfaktoriellen Varianzanalyse wird die Gedächtnisleistung zwischen 3 Gruppen verglichen. Das Ergebnis ist signifikant. Welche Aussagen können in diesem Fall angenommen werden? (1 Punkt)**

Alle drei Mittelwerte sind gleich.

Der Mittelwert der ersten Gruppe muss sich signifikant von einem der Mittelwerte der anderen Gruppen unterscheiden.

Die Varianz der Mittelwerte ist größer Null.

Die Varianz der Mittelwerte ist kleiner Null.

Mindestens ein Gruppeneffekt muss sich von Null unterscheiden.

Die Summe der Effekte ist ungleich Null.

9 **An einer Untersuchung nehmen 120 Versuchspersonen teil, die in 4 Gruppen aufgeteilt werden.**

9.1 Wie viele Freiheitsgrade hat der entsprechende F-Test? (1 Punkt)

$$df_{\text{between}} = p - 1 = 3$$
$$df_{\text{within}} = N - p = 116$$

9.2 Im Datensatz ergibt sich für  $SS_{\text{total}} = 100$  und für  $SS_{\text{within}} = 85$ . Berechnen Sie den empirischen F-Wert! (2 Punkte)

$$SS_{\text{between}} = SS_{\text{total}} - SS_{\text{within}} = 15$$
$$MS_{\text{between}} = SS_{\text{between}} / df_{\text{between}} = 15 / 3 = 5$$
$$MS_{\text{within}} = SS_{\text{within}} / df_{\text{within}} = 85 / 116 = 0.73$$
$$F = MS_{\text{between}} / MS_{\text{within}} = 5 / 0.73 = 6.82$$

**10 Welche Voraussetzungen besitzt die einfaktorielle Varianzanalyse? (1 Punkt)**

- *Intervallskalierte, normalverteilte abhängige Variable (AV)*
- *Mindestens 20 Elemente pro Gruppe*
- *Ähnlich Gruppengrößen in den Zellen*
- *Varianzhomogenität*

**11 Im Rahmen einer einfaktorielle Varianzanalyse kann der Levene-Test verwendet werden.**

11.1 Wozu dient er? (1 Punkt)

*Zur Überprüfung der Varianzhomogenitätsannahme*

11.2 Welchen Schluss ziehen Sie für die ANOVA, wenn der Levene-Test zu folgendem Ergebnis kommt? (1 Punkt)

F	df1	df2	Signifikanz
.196	2	62	.823

- *Der Levene-Test wurde nicht signifikant, daher ist die Voraussetzung der ANOVA nicht verletzt.*
- *Die Interpretation der Ergebnisse ist somit unproblematisch.*

**12 Welche Aussagen zu Einzelvergleichen treffen zu? (1 Punkt)**

- Nach der Berechnung einer Varianzanalyse sollten Einzelvergleiche in jedem Fall erfolgen, d.h. auch wenn der *F*-Test nicht signifikant war.
- Nur wenn *a priori* konkrete Hypothesen formuliert wurden, darf Tukeys HSD-Test verwendet werden.
- Nur wenn *a priori* konkrete Hypothesen formuliert wurden, dürfen Kontraste verwendet werden.
- Der Scheffé-Test sollte gegenüber Tukeys HSD vorgezogen werden, da er ein höhere Power besitzt.
- Kontraste sollten bevorzugt werden, da sie eine höhere Power als Tukeys HSD besitzen.
- Bei der Verwendung von Kontrasten muss darauf geachtet werden, dass die definierten Kontraste voneinander unabhängig sind.

**13 Es wurde untersucht, wie gut Frauen und Männern bestimmte Getränke schmecken. In welchen Fällen gibt es eine Interaktion? (1 Punkt)**

	Getränk	
	A	B
Frauen	10	15
Männer	5	10

	Getränk	
	A	B
Frauen	10	15
Männer	15	10

	Getränk	
	A	B
Frauen	4	5
Männer	5	6

	Getränk	
	A	B
Frauen	8	4
Männer	10	8

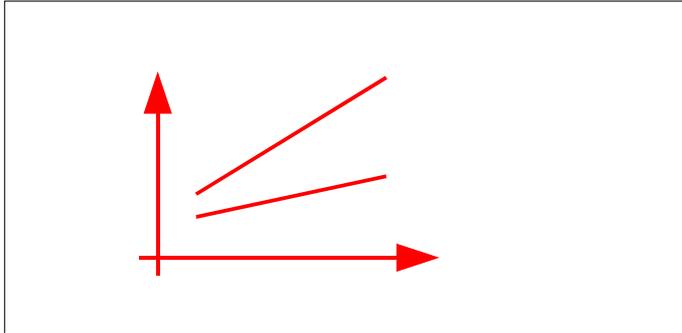
Interaktion ?  
 ja  nein

Interaktion ?  
 ja  nein

Interaktion ?  
 ja  nein

Interaktion ?  
 ja  nein

14 **Skizzieren Sie eine ordinale Interaktion (1 Punkt)**



15 **Ergänzen Sie die fehlenden Zahlen! (1 Punkt)**

**Tests der Zwischensubjekteffekte**

Abhängige Variable: effekt

Quelle	Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz
Korrigiertes Modell	293383.839 <sup>a</sup>	7	41911.977	2.252	.056
Konstanter Term	25826.487	1	25826.487	1.388	.248
sex	63029.225	1	63029.225	<b>3.387</b>	.075
alter	164543.396	3	<b>54847.799</b>	2.948	.048
sex * alter	40676.956	<b>3</b>	13558.985	.729	.543
Fehler	576845.158	31	18607.908		
Gesamt	873178.158	39			
Korrigierte Gesamtvariation	870228.997	38			

a. R-Quadrat = .337 (korrigiertes R-Quadrat = .187)

16 **Welche Aussagen sind richtig? (1 Punkt)**

- Die Annahme von *Zufallseffekten* erlaubt eine Generalisierung der Ergebnisse auch auf nicht untersuchte Faktorstufen.
- Es ist sinnvoll, Geschlecht als *Zufallsfaktor* zu definieren.
- Bei einer Studie zur Auswirkung des Alkoholkonsums können erhobenen Stufen der Alkoholisierung (1 Bier, 2 Bier, 3 Bier) als *Zufallsfaktor* betrachtet werden.
- Bei der Berechnung einer einfaktoriellen ANOVA mit *Zufallseffekten* verändert sich die Art der Berechnung des empirischen *F*-Werts.
- Bei der Berechnung einer mehrfaktoriellen ANOVA mit *Zufallseffekten* verändert sich die Art der Berechnung der empirischen *F*-Werte.
- Wenn man nicht sicher ist, ob ein *fester Effekt* oder ein *Zufallseffekt* vorliegt, sollte man ein Modell mit *gemischten Effekten* berechnen.

17 Geben Sie die Formel zur Berechnung der Quadratsumme  $SS_{treatment}$  für eine Messwiederholte ANOVA an. (1 Punkt)

$$SS_{treatment} = \sum_{j=1}^p N \cdot (\bar{y}_{\cdot j} - \bar{y}_{\cdot\cdot})^2$$

18 Bei der der Varianzanalyse mit Messwiederholung wird im Vergleich zur Varianzanalyse ohne Messwiederholung eine zusätzliche Annahme gemacht.

18.1 Welche Annahme ist dies? Was sagt sie aus? (1 Punkt)

*Die Sphärizitätsannahme (Zirkularitätsannahme): Die Varianzen und Kovarianzen unter den einzelnen Faktorstufen müssen homogen sein.*

18.2 Wie kann eine Verletzung dieser Annahme ausgeglichen werden? (1 Punkt)

*Greenhouse-Geisser-Korrektur der Freiheitsgrade des F-Tests*

19 Sie erwarten, dass die Arbeitsmotivation von Frauen und von Männern durch eine Personalentwicklungsmaßnahme verbessert wird. Die Motivation wird vor und nach dem Training mit einem Motivationsfragebogen (Werte 0 bis 20) erhoben. Eine Varianzanalyse mit Messwiederholung kommt zu folgendem Ergebnis:

	sex	Mittelwert	Standardabweichung	N
motiv1	1.00	12.3000	5.79367	10
	2.00	9.7000	1.94651	10
	Gesamt	11.0000	4.41290	20
motiv2	1.00	9.9000	2.42441	10
	2.00	8.4000	1.57762	10
	Gesamt	9.1500	2.13431	20

Quelle		Quadratsumme vom Typ III	df	Mittel der Quadrate	F	Signifikanz
mzp	Sphärizität angenommen	34.225	1	34.225	5.744	.028
	Greenhouse-Geisser	34.225	1.000	34.225	5.744	.028
	Huynh-Feldt	34.225	1.000	34.225	5.744	.028
	Untergrenze	34.225	1.000	34.225	5.744	.028
mzp * sex	Sphärizität angenommen	3.025	1	3.025	.508	.485
	Greenhouse-Geisser	3.025	1.000	3.025	.508	.485
	Huynh-Feldt	3.025	1.000	3.025	.508	.485
	Untergrenze	3.025	1.000	3.025	.508	.485
Fehler(mzp)	Sphärizität angenommen	107.250	18	5.958		
	Greenhouse-Geisser	107.250	18.000	5.958		
	Huynh-Feldt	107.250	18.000	5.958		
	Untergrenze	107.250	18.000	5.958		

19.1 Wie viele Probanden nahmen an der Untersuchung teil? (0.5 Punkte) 20

19.2 Ist die inhaltliche Hypothese bestätigt? (0.5 Punkte) nein

20 Welche Vorteile bietet eine Varianzanalyse mit Messwiederholung? (1 Punkt)

- Es werden weniger Versuchspersonen benötigt, da dieselben Vpn mehrmals getestet werden.*
- Höhere Teststärke (Power), da die Fehlervarianz verringert wird.*

21 ***In welchem Fall sollte eine Varianzanalyse mit Messwiederholung verwendet werden? (1 Punkt)***

- Beim Vergleich der Gedächtnisleitung bei Männern und Frauen im Allgemeinen.
- Um zu überprüfen, ob in Beziehungen Frauen generell extravertierter sind als ihre Partner.
- Um zu überprüfen, ob die Ängstlichkeit von Klienten nach einer Therapie geringer ist als vorher.
- Um zu überprüfen, ob zwei Aufgaben gleich schwierig sind.
- Um zu testen, ob die Studierenden von zwei unterschiedlichen Universitäten sich in ihren Statistikkenntnissen unterscheiden.
- Um zu testen, ob sich die Aggressivität von Kindern im Laufe des ersten Schuljahres ändert.

22 ***Nennen Sie drei Möglichkeiten, mit Störvariablen umzugehen (1 Punkt)***

- (1) *Die Störvariable wird über alle Bedingungen konstant gehalten.*
- (2) *Aufnahme der Störvariablen als Faktor im Versuchsplan.*
- (3) *Die Störvariable wird als Kovariate verwendet.*

23 ***Welche Aussagen sind richtig? (1 Punkt)***

- Eine Kovarianzanalyse ist im Prinzip eine Varianzanalyse mit den Regressionsresiduen.
- Der Vorteil einer Kovarianzanalyse liegt darin, dass durch die Berücksichtigung der Kovariate die Quadratsumme  $SS_{\text{within}}$  vergrößert wird.
- Es ist möglich, dass eine Kovarianzanalyse ein signifikantes Ergebnis hat, auch wenn die entsprechende Varianzanalyse ohne Kovariate nicht signifikant wurde.
- Die Kovarianzanalyse hat eine höhere Power, weil sich die Zahl der Freiheitsgrade reduziert.
- In der Kovarianzanalyse werden die Quadratsummen vor der Berechnung des  $F$ -Bruchs „adjustiert.“
- Signifikanz im  $F$ -Test der Kovarianzanalyse zeigt an, dass die Kovariate mit der AV korreliert.

24 ***Das Ergebnis einer Varianzanalyse ist statistisch signifikant. Es zeigt sich bei einer anschließenden Berechnung eine Effektstärke von  $\hat{\omega}^2 = 0.20$ .***

24.1 Was sagt dieser Wert aus? (1 Punkt)

*Die Treatments-Varianz beträgt 20% der Gesamtvarianz oder:  
Es werden 20% der Merkmalsvarianz aufgeklärt*

24.2 Ist hier von einem *kleinen*, einem *mittleren* oder einem *großen* Effekt auszugehen? (1 Punkt)

*Dies ist ein großer Effekt nach Cohen.*

- 25 **Nennen Sie 5 Möglichkeiten, die Teststärke einer empirischen Untersuchung zu erhöhen. (2 Punkte)**
- (1) *Höheres Alpha-Niveau wählen*
  - (2) *Einseitig testen*
  - (3) *Stichprobe vergrößern*
  - (4) *Abhängige Stichproben verwenden*
  - (5) *Teststärkeres Verfahren auswählen*
- 26 **Bei einer statistischen Untersuchung kann es dazu kommen, dass das Ergebnis keine eindeutige Unterscheidung zwischen der  $H_0$  und der  $H_1$  zulässt. Wann ist dies der Fall? (1 Punkt)**
- Wenn der Test nicht signifikant wurde ( $p > .05$ ) und die Power gering ist (z.B.  $1 - \beta < .90$ )*
- 27 **Welche Informationen werden bei einer a priori-Teststärkenanalyse festgelegt und was wird berechnet? (1 Punkt)**
- Festlegen:  $\alpha$ ,  $\beta$ , Effekt*  
*Berechnen:  $N$*
- 28 **Was ist das Ziel einer explorativen Faktorenanalyse? (1 Punkt)**
- Zusammenfassen vieler Items zu wenigen Faktoren.*  
*Oder: Konstruktion von Fragebögen: Bildung von Subskalen.*
- 29 **Erläutern Sie kurz die beiden Begriffe „Eigenwert“ und „Kommunalität“. (2 Punkte)**
- *Die Kommunalität ( $h^2$ ) einer Variablen ist die insgesamt durch alle Faktoren aufgeklärte Varianz dieser Variablen.*
  - *Der Eigenwert ( $\lambda$ ) eines Faktors gibt an, wieviel Varianz dieser Faktor an allen Variablen aufklärt.*
- 30 **Welche beiden Arten der Achsen-Rotation gibt bei der Faktorenanalyse? Was ist der grundlegende Unterschied? (1 Punkt)**
- *Bei der orthogonalen Rotation bleiben die Faktoren unabhängig, d.h. sie stehen senkrecht aufeinander.*
  - *Bei der obliquen Rotation Zusammenhänge zwischen den Faktoren erlaubt.*

Unterschrift: \_\_\_\_\_