

# Empirische Forschungsmethoden

## Nachklausur vom 10. Januar 2007

**Name:**

**Matrikelnummer:**

Erlaubte Hilfsmittel: Außer Schreibwerkzeug diesmal auch Taschenrechner (wird aber nicht nötig sein); Antworten bitte im Anschluss an die Fragen, Rückseiten können für Antworten und Entwürfe genutzt werden; Entwürfe bitte anschließend durchstreichen!

Diese Klausur soll — nicht — als Freiversuch gewertet werden (gilt nur für Studierende nach alter Ordnung, die anderen können das später entscheiden),

.....  
Unterschrift

Für die Teilaufgaben wird es folgende Punktzahlen geben; die Summe wird durch 2 dividiert, hinzu gezählt werden die Punkte aus den beiden besten Übungsaufgaben. Aus der so entstehenden Gesamtpunktzahl von maximal 80 wird die Modulnote berechnet: 80 Punkte geben eine 1.0. Die Mindestpunktzahl für eine 4.0 wird höchstens 40 betragen.

Aufgabe	Maximal erreichbar	Sie haben erreicht	
1.1	4	12	
1.2	4		
1.3	2		
1.4	2		
2	7	7	
3.1	5	14	
3.2	5		
3.3	4		
4.1	8	16	
4.2	4		
4.3	4		
5.1	5	19	
5.2	4		
5.3	4		
5.4	6		
6.1	3	12	
6.2	7		
6.3	2		
Summe	80	80	

## 1. Grundlagen

1.1 Bitte schauen Sie sich die Verteilung der Studierenden auf die Studiengänge an, wie sie aus der letzten Spalte der Tabelle auf Seite 5 ersichtlich ist. Könnten Sie auf der Basis dieser Stichprobe sinnvolle Aussagen über den Unterschied zwischen den Bachelor- und den Master- Studierenden im Informationsmanagement in der Verteilung auf die Regionen machen? Begründen Sie Ihre Beurteilung!

1.2 Welche (möglicherweise bessere) Alternative hätte es für die Stichprobenziehung in diesem Fall gegeben?

1.3 Welches Skalenniveau und welche Wertebereiche haben die beiden Variablen von Aufgabe 4?

1.4 Entstammen die Daten, die in Aufgabe 3 benutzt werden, einer Stichprobe oder einer Grundgesamtheit? Begründen Sie Ihre Einschätzung (es kommt vor allem darauf an, wie überzeugend Sie Ihre Einschätzung begründen können !!)

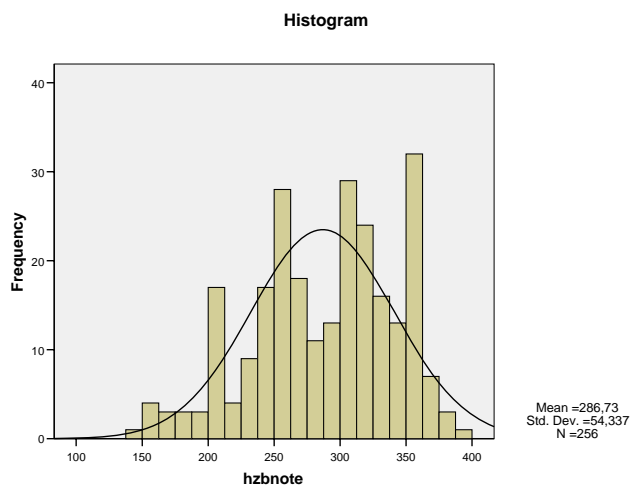
## 2. Skalenniveaus

In den Aufgaben dieser Klausur werden verschiedene Variable benutzt. Bitte geben Sie hier schon zusammenfassend an, welches Skalenniveau nach Ihrer Einschätzung die einzelnen Variablen haben und begründen Sie Ihre Einschätzung. Sie können auch einer Variablen verschiedene Skalenniveaus zuweisen, müssen dann aber auch das begründen.

	Dichotom	Nominal	Ordinal	Intervall	Ratio	Begründung
Wechselkurs zwischen Euro und Schweizer Franken						
ECTS-Punktzahl						
Die Regionen, wie sie in Aufgabe 4 definiert sind						
Abiturnote						
Semesterzahl						
Studiengänge, wie sie in Aufgabe 4 definiert sind						
Regressionskoeffizienten, wie sie in Aufg. 5 vorkommen						

## 3. Univariate Statistik und Mittelwertvergleiche

Die Abiturnoten der Studierenden gegenwärtig im Bachelorstudiengang Informationsmanagement eingeschriebenen Studierenden (Wintersemester 2006/7) folgen der im Histogramm beschriebenen Verteilung. Parameter der Verteilung finden Sie unten ebenfalls.

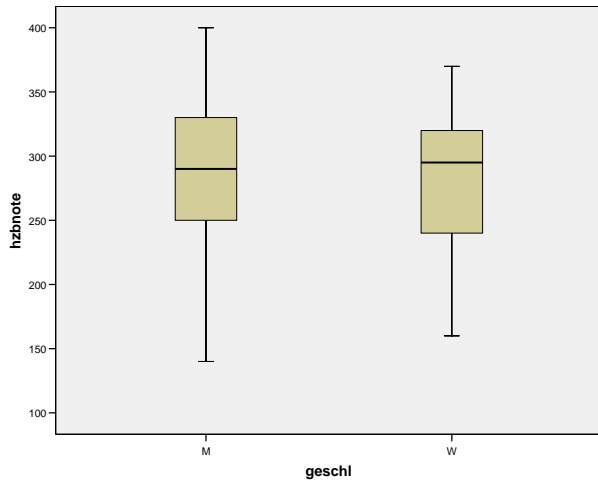


**Statistics**

hzbnote		
N	Valid	256
	Missing	4
Mean		286.73
Median		290.00
Mode		320
Std. Deviation		54.337
Skewness		-.394
Std. Error of Skewness		.152
Kurtosis		-.559
Std. Error of Kurtosis		.303
Range		260
Percentiles	25	250.00
	50	290.00
	75	330.00

3.1 Erläutern Sie, was Sie am Histogramm und an den Parametern erkennen können (beachten Sie, dass die Abiturnoten mit 100 multipliziert sind!). Woran mag es liegen, dass die Werte 1.5, 2.0, 2.5, 3.0 und 3.5 jeweils viel häufiger sind als die jeweils benachbarten Werte?

3.2 Die folgende Abbildung zeigt vereinfacht die Verteilung der Abiturnoten für männliche und weibliche Studierende des Bachelorstudiengangs Informationsmanagement. Was entnehmen Sie dieser graphischen Darstellung? Zur Erinnerung: die „Boxes“ umfassen die mittleren beiden Quartile, die „Whiskers“ sind höchstens anderthalb mal so lang wie die „Boxes“.



3.3 Eine Auswertung der Abiturnoten aller Studierenden des Bachelor-Studiengangs Informationsmanagement ergibt folgende Mittelwerte für männlichen und die weiblichen Studierenden:

**Group Statistics**

geschl		N	Mean	Std. Deviation	Std. Error Mean
hzbnote	M	184	289.82	53.283	3.928
	W	72	278.85	56.555	6.665

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower		Upper
hzbnote	Equal variances assumed	.684	.409	1.456	254	.147	10.973	7.537	-3.869	25.816
	Equal variances not assumed			1.418	123.125	.159	10.973	7.736	-4.340	26.287

Kann man aus dieser Stichprobe schließen, dass die Studentinnen bessere Abiturnoten haben als ihre männlichen Kommilitonen? (Beachten Sie wieder, dass die Abiturnoten mit 100 multipliziert sind!) Um wie viel im Mittel? Was können Sie mit der unteren Tabelle anfangen? Was sagt Ihnen die Zahl 0.147?

### 4. Bivariate Statistik

Eine Stichprobe von 500 Studierenden am Campus Koblenz verteilt sich folgendermaßen auf die Studiengänge und die Herkunftsregionen (region, in der die Hochschulzugangsberechtigung erworben wurde, mit „nachbarn“ sind die Landkreise außerhalb von Rheinland-Pfalz gemeint, die an den RegBez Koblenz grenzen)::

stggrp Studiengang \* gebiet Region Crosstabulation

			gebiet Region					Total
			1.00 Koblenz-Umland	2.00 Nachbarkreise außerhalb Rheinland-Pfalz	3.00 Rest Rheinland-Pfalz	4.00 Rest Deutschland	5.00 Ausland	
stggrp Studien gang	2000.00 Magister	Count	12	0	0	2	1	15
		% within stggrp Studiengang	80.0%	.0%	.0%	13.3%	6.7%	100.0%
	6000.00 Promotion	Count	4	0	1	5	0	10
		% within stggrp Studiengang	40.0%	.0%	10.0%	50.0%	.0%	100.0%
	11052.00 Pädagogik (Diplom)	Count	24	5	11	9	2	51
		% within stggrp Studiengang	47.1%	9.8%	21.6%	17.6%	3.9%	100.0%
	11079.00 Informatik (Diplom)	Count	30	2	5	3	2	42
		% within stggrp Studiengang	71.4%	4.8%	11.9%	7.1%	4.8%	100.0%
	11845.00 Computervisualistik (Diplom)	Count	18	3	7	21	0	49
		% within stggrp Studiengang	36.7%	6.1%	14.3%	42.9%	.0%	100.0%
	20000.00 Lehramt Grund- und Hauptschule	Count	81	31	58	22	3	195
		% within stggrp Studiengang	41.5%	15.9%	29.7%	11.3%	1.5%	100.0%
	23000.00 Lehramt realschule	Count	82	6	23	5	5	121
		% within stggrp Studiengang	67.8%	5.0%	19.0%	4.1%	4.1%	100.0%
	82079.00 Informatik (BSc)	Count	0	1	1	1	1	4
		% within stggrp Studiengang	.0%	25.0%	25.0%	25.0%	25.0%	100.0%
	82714.00 Informationsmanagement (BSc)	Count	14	1	7	7	1	30
		% within stggrp Studiengang	46.7%	3.3%	23.3%	23.3%	3.3%	100.0%
	82845.00 Computervisualistik (BSc)	Count	1	1	1	1	0	4
		% within stggrp Studiengang	25.0%	25.0%	25.0%	25.0%	.0%	100.0%
	88714.00 Informationsmanagement (MSc)	Count	9	0	1	3	1	14
		% within stggrp Studiengang	64.3%	.0%	7.1%	21.4%	7.1%	100.0%
	99999.00 Sonstige	Count	6	4	6	26	5	47
		% within stggrp Studiengang	12.8%	8.5%	12.8%	55.3%	10.6%	100.0%
Total		Count	281	54	121	105	21	582
		% within stggrp Studiengang	48.3%	9.3%	20.8%	18.0%	3.6%	100.0%

4.1 Formulieren Sie an Hand dieser Kreuztabelle einige Sätze, die den Zusammenhang zwischen den beiden Variablen beschreiben! Welche Hypothese hätten Sie schon vor Betrachten dieser Tabelle gehabt bzw. welche Hypothese können Sie aus ihr ableiten?

4.2 Für diesen Zusammenhang ist Phi 0.541 und Cramers' V 0.271. Erklären Sie den Unterschied dieser beiden eng verwandten Maße. Sind diese beiden Maßzahlen hier hilfreich?

4.3 Für diesen Zusammenhang ist Lambda 0.047 und der Uncertainty Coefficient 0.087. Erklären Sie den Unterschied auch dieser beiden eng verwandten Maße . Sind diese beiden Maßzahlen hier hilfreich?

## 5. Multivariate Statistik / Multiple Regression

Die bisher erreichten ECTS-Punktzahlen (Summevonbonus) hängen vermutlich einerseits von der Semesterzahl (stgsem) ab, andererseits aber vielleicht auch von der Leistungsfähigkeit, wie man sie zum Beispiel mit der Abiturnote (hzbnote) messen könnte.

Die Regression ergibt folgende Koeffizienten:

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.847 <sup>a</sup>	.717	.715	30.725

a. Predictors: (Constant), hzbnote, stgsem

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.087	10.693		.756	.450
	stgsem	17.227	.686	.840	25.097	.000
	hzbnote	-.088	.035	-.083	-2.488	.013

a. Dependent Variable: Summevonbonus

5.1 Welche Bedeutung haben die beiden Zahlen 17.227 und -0.088?

5.2 Welche (ziemlich fiktive) Bedeutung hat die Zahl 8.087?

5.3 Wie können Sie sich die unterschiedlichen Vorzeichen der beiden Koeffizienten der beiden unabhängigen Variablen erklären?

5.4 Was bedeutet die Zahl 0.013 am Ende der letzten Zeile der Tabelle?

## 6. Multivariate Statistik / Faktorenanalyse

In einer Studie für das Ministerium des Innern und für Sport des Landes Rheinland-Pfalz und den Kriminalpräventiven Rat der Verbandsgemeinde Weilerbach wurde den Jugendlichen der Verbandsgemeinde in einem Fragebogen unter anderem eine Fragebatterie vorgelegt, in der sie angeben sollten, mit welchen Aktivitäten sie wie häufig ihre Freizeit gestalten (von „sehr häufig“ bis „nie“). In den beiden Tabellen finden Sie die Ergebnisse zweier Faktorenanalysen, einmal wurden drei, ein anderes Mal wurden vier Faktoren extrahiert. Faktorladungen unter 0.2 wurden unterdrückt.

**Rotated Component Matrix<sup>a</sup>**

	Component		
	1	2	3
f1106 Disco besuchen	.805		
f1105 Cafes, Kneipen besuchen	.718	.220	
f1119 Musik hören	.461		.231
f1102 Bücher lesen	-.450		.203
f1101 Freunde treffen, Unternehmungen	.445	.330	
f1115 Malen, nähen	-.421		
f1118 Kino gehen	.412	.329	.293
f1108 Besuch von Jugendtreffs	.381	.210	
f1111 Sport privat		.763	
f1112 Am Roller arbeiten		.611	
f1110 Sport im Fitnessclub		.538	
f1109 Sport im Verein		.534	
f1117 Schwimmen		.501	.269
f1104 selbst Musik machen		.247	
f1107 Video-, Computerspiele	-.276	.250	.679
f1114 Computer allgemein	.200		.678
f1103 Fernsehen, Videos			.636
f1116 Nichts tun	.276	-.203	.401
f1113 Mitarbeit in Organisationen			-.265

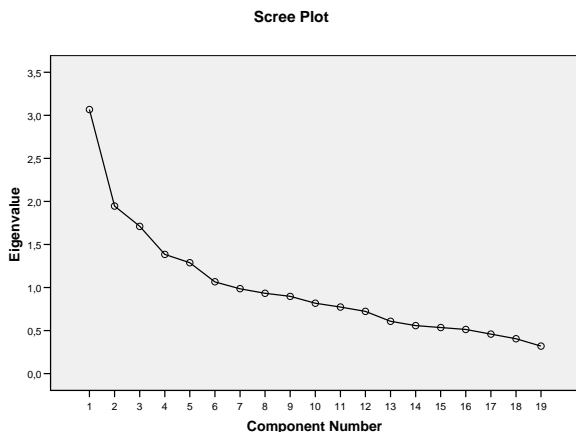
Extraction Method: Principal Component Analysis.  
 Rotation Method: Equamax with Kaiser Normalization.  
 a. Rotation converged in 5 iterations.

**Rotated Component Matrix<sup>a</sup>**

	Component			
	1	2	3	4
f1106 Disco besuchen	.709			-.379
f1105 Cafes, Kneipen besuchen	.689			-.270
f1119 Musik hören	.575			
f1118 Kino gehen	.511	.246	.225	
f1108 Besuch von Jugendtreffs	.461			
f1101 Freunde treffen, Unternehmungen	.432	.302		
f1104 selbst Musik machen	.286			.222
f1111 Sport privat		.712		.224
f1109 Sport im Verein		.625		
f1112 Am Roller arbeiten		.615		
f1110 Sport im Fitnessclub		.555		
f1107 Video-, Computerspiele		.273	.755	
f1103 Fernsehen, Videos			.653	
f1114 Computer allgemein	.291		.652	
f1113 Mitarbeit in Organisationen	.270		-.377	
f1116 Nichts tun	.298	-.254	.353	
f1115 Malen, nähen				.760
f1102 Bücher lesen				.642
f1117 Schwimmen	.317	.384		.384

Extraction Method: Principal Component Analysis.  
 Rotation Method: Equamax with Kaiser Normalization.  
 a. Rotation converged in 7 iterations.

6.1 Beurteilen Sie die Unterschiede zwischen den beiden Ladungsmatrizen. Welche der beiden Lösungen lässt sich besser interpretieren? Begründen Sie Ihre Entscheidung und ziehen Sie dazu auch das Scree-Plot heran.



6.2 Geben Sie den drei bzw. vier Faktoren Namen (die Namen sollten Eigenschaften oder Gewohnheiten der befragten Jugendlichen benennen):

Faktor	Lösung 1	Lösung 2
1		
2		
3		
4		

6.3 Wie beurteilen Sie die Einfachstruktur? Sind die Faktoren sauber genug getrennt?