

Einführung in die Statistik für SozialwissenschaftlerInnen

Klausur am 15.02.2005

Name: _____ Vorname: _____

Matrikelnummer: _____

Studienfächer _____

Einige Hinweise zur Organisation:

Ihnen stehen zwei Zeitstunden für die Bearbeitung der Aufgaben zur Verfügung.

Es sind alle Fragen zu beantworten. Sie können innerhalb der ersten 15 Minuten den Raum endgültig verlassen. Anderenfalls bleiben Sie bitte bis zum Ende der Bearbeitungszeit ruhig auf Ihrem Platz.

Bitte nehmen Sie die geheftete Klausur bei der Bearbeitung nicht auseinander.

Versuchen Sie, möglichst prägnant und gut lesbar auf die Fragen zu antworten.

Wenn der vorgegebene Platz nicht ausreichen sollte, können Sie die leeren Rückseiten benutzen.

Viel Erfolg bei der Bearbeitung der Klausur!

Punkte	
Note	

Dr. Anna Brake

Aufgabe 1:

In der Nacht auf den 15. April 1912 sank die als unsinkbar geltende Titanic, nachdem sie auf ihrer Jungfernfahrt im Nordatlantik mit einem Eisberg zusammengestoßen war. Das Unglück forderte 1.503 Todesopfer. Von den insgesamt 1343 Passagieren auf der Titanic waren 337 in der ersten Klasse und 285 in der zweiten Klasse an Bord (die 885 Besatzungsmitglieder werden an dieser Stelle nicht betrachtet). 40% der Passagiere aus der ersten Klasse wurden nach dem Unglück vermisst, von den Passagieren der zweiten Klasse verloren insgesamt 160 ihr Leben und in der dritten Klasse waren dies 541 Passagiere.

- 1 a.) Erstellen Sie auf der Basis der Informationen eine Kreuztabelle, bei der das Merkmal „Passagierklasse“ die Spaltenvariable darstellt. Bestimmen Sie die Zellenhäufigkeiten als Anzahl n und als Zeilenprozente.

n				
%				
n				
%				
n				
%				

- b) Zeigen die Daten einen Zusammenhang zwischen Überlebenschance und Passagierklasse? Beschreiben Sie das Ergebnis.

- c) Welches Assoziationsmaß würden Sie heranziehen, um die Stärke des Zusammenhangs zu bestimmen? Begründung?

Aufgabe 2:

In einer Untersuchung zu statistikbezogenen Einstellungen wurden 109 Soziologie-Studierende, die an einer Vorlesung „Einführung in die Statistik“ teilnahmen, gebeten ihren Erfolg bei der Teilnahme an der abschließenden Klausur einzuschätzen. Nachfolgende Kreuztabelle zeigt das geschlechtsspezifische Antwortverhalten auf das Item „Die Statistik-Klausur wird mir keine Probleme machen.“

Geschlecht * "Die Statistik-Klausur wird mir keine Probleme machen." Kreuztabelle

		keine Probleme mit Statistik-Klausur					Gesamt
		trifft nicht zu	trifft eher nicht zu	weder noch	trifft eher zu	trifft zu	
männlich	Anzahl	1	8	14	11	4	38
	Erwartete Anzahl	4,2	8,4	13,9	9,8	1,7	38,0
	% von Geschlecht	2,6%	21,1%	36,8%	28,9%	10,5%	100,0%
	Standardisierte Residuen	-1,6	-,1	,0	,4	1,7	
weiblich	Anzahl	11	16	26	17	1	71
	Erwartete Anzahl	7,8	15,6	26,1	18,2	3,3	71,0
	% von Geschlecht	15,5%	22,5%	36,6%	23,9%	1,4%	100,0%
	Standardisierte Residuen	1,1	,1	,0	-,3	-1,3	
Gesamt	Anzahl	12	24	40	28	5	109
	Erwartete Anzahl	12,0	24,0	40,0	28,0	5,0	109,0
	% von Geschlecht	11,0%	22,0%	36,7%	25,7%	4,6%	100,0%

Bitte interpretieren Sie unter Verwendung der aufgeführten Zellenwerte das zentrale Ergebnis der Kreuztabelle.

a) Was besagen die Zahlenwerte in der Zeile „Erwartete Anzahl“

b) Welche Information kann der Höhe der standardisierten Residuen entnommen werden?

c) Um die Art des Zusammenhangs näher zu untersuchen, wurde als PRE-Maß Goodman- und Kruskal-Tau berechnet.

Richtungsmaße

		Wert	Asymptotischer Standardfehler	Näherungsweise Signifikanz
Goodman-und-Kruskal-Tau	Geschlecht abhängig	,078	,039	,078
	keine Probleme mit Statistik-Klausur abhängig	,009	,006	,452

Interpretieren Sie bitte die Ergebnisse.

Aufgabe 3:

Für das in Aufgabe 2 beschriebene Antwortverhalten auf einer Zustimmungsskala von „trifft nicht zu“ (= 1) und „trifft zu“ (= 5) wurden die folgende deskriptiven Statistiken berechnet.

Univariate Statistiken

			Statistik	Standardfehler
keine Probleme mit Statistik-Klausur	Mittelwert		2,92	,100
	95% Konfidenzintervall des Mittelwerts	Untergrenze	2,72	
		Obergrenze	3,12	
	erstes Quartil		2,15	
	Median		2,98	
	drittes Quartil		3,720	
	Standardabweichung		1,050	
	Minimum		1	
	Maximum		5	
	Spannweite		4	
	Interquartilbereich		1,57	
	Schiefe		-,173	,230
	Kurtosis		-,591	,457

Bitte beschreiben Sie unter Verwendung der angemessenen statistischen Kennwerte das Antwortverhalten der 110 Studierenden.

Aufgabe 4:

z-Transformation: Welche Bedingungen müssen erfüllt sein, damit eine z-Transformation der Daten sinnvoll durchgeführt werden kann? Was geschieht mit den ursprünglichen Messwerten, wenn diese z-transformiert werden? Beschreiben Sie einen Anwendungsfall, wann eine solche Transformation der Daten benötigt wird.

Aufgabe 5:

Am Lehrstuhl für Soziologie und empirische Sozialforschung wird erwogen, ein neues Softwarepaket für die statistische Datenanalyse anzuschaffen. Dieses Programmpaket wird als sehr anwenderInnenfreundlich beschrieben, ist aber in der Anschaffung mit sehr hohen Kosten verbunden. Hinzu käme, dass neue Lehrbücher angeschafft und MitarbeiterInnenschulungen durchgeführt werden müssten. Daher soll zunächst überprüft werden, ob die Studierenden mit der neuen Software einen größeren Lernerfolg erzielen. Mit zufällig ausgewählten Studierenden wird ein Übungskurs mit der alten, parallel ein Übungskurs mit der neuen Software durchgeführt. Am Ende wird eine benotete Klausur geschrieben. Formulieren Sie ein dieser Situation angemessenes Hypothesenpaar. Worauf beziehen sich Ihre Hypothesen? Entscheiden und begründen Sie, welchen Test Sie zur Überprüfung heranziehen würden. Wie legen Sie im vorliegenden Fall die Irrtumswahrscheinlichkeit α fest? Warum?

Aufgabe 6:

In einer Untersuchung über die statistikbezogenen Einstellungen von Studierenden wurden diese u.a. gebeten einzuschätzen, inwieweit folgende Aussagen auf sie zutreffen (Antwortskala: „trifft nicht zu“ = 1 und „trifft zu“ = 5):

- Ich glaube, dass mir Statistik Spaß machen wird.
- Die Statistik-Klausur wird mir keine Probleme machen.
- Statistik ist leider ein notwendiges Übel.
- Mit der Statistik wird doch sowieso nur rumgetrickst.

Die Auswertung der Daten ergab folgende Korrelationsmatrix.

Korrelationen: Kendall-Tau-b

		dass mir Statistik Spaß machen wird	keine Probleme mit Statistik-Klausur	Statistik: ein notwendiges Übel	Statistik: wird nur rumgetrickst
dass mir Statistik Spaß machen wird	Korrelationskoeffizient	1,000	,431**	-,482**	,023
	Sig. (2-seitig)	.	,000	,000	,778
	N	109	109	108	108
keine Probleme mit Statistik-Klausur	Korrelationskoeffizient	,431**	1,000	-,367**	-,135
	Sig. (2-seitig)	,000	.	,000	,097
	N	109	110	109	109
Statistik: ein notwendiges Übel	Korrelationskoeffizient	-,482**	-,367**	1,000	,049
	Sig. (2-seitig)	,000	,000	.	,543
	N	108	109	109	109
Statistik: wird nur rumgetrickst	Korrelationskoeffizient	,023	-,135	,049	1,000
	Sig. (2-seitig)	,778	,097	,543	.
	N	108	109	109	109

** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).

Interpretieren Sie die zentralen Ergebnisse der Korrelationsanalyse.

Aufgabe 7:

Verfügen männliche und weibliche Studierende in unterschiedlichem Maß über Geld zur freien Verfügung? An einer Zufallsauswahl von 96 Studierenden wurde diese Frage untersucht. Der t-Test zeigt folgende Ergebnisse.

Kolmogorov-Smirnov-Anpassungstest

		Geld zur freien Verfügung
N		96
Parameter der Normalverteilung ^a	Mittelwert	260,8125
	Standardabweichung	58,98623
Extremste Differenzen	Absolut	,081
	Positiv	,065
	Negativ	-,081
Kolmogorov-Smirnov-Z		,791
Asymptotische Signifikanz (2-seitig)		,560

a. Die zu testende Verteilung ist eine Normalverteilung.

Gruppenstatistiken

	Geschlecht	N	Mittelwert	Standardabweichung	Standardfehler des Mittelwertes
Geld zur freien Verfügung	männlich	35	275,1429	51,82388	8,75983
	weiblich	60	252,6667	62,14981	8,02351

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit	
		F	Signifikanz
Geld zur freien Verfügung	Varianzen sind gleich	2,465	,120
	Varianzen sind nicht gleich		

Test bei unabhängigen Stichproben

		T-Test für die Mittelwertgleichheit		
		T	df	Sig. (2-seitig)
Geld zur freien Verfügung	Varianzen sind gleich	1,804	93	,075
	Varianzen sind nicht gleich	1,892	81,801	,062

Test bei unabhängigen Stichproben

		T-Test für die Mittelwertgleichheit			
		Mittlere Differenz	Standardfehler der Differenz	95% Konfidenzintervall der Differenz	
				Untere	Obere
Geld zur freien Verfügung	Varianzen sind gleich	22,4762	12,46084	-2,26857	47,22095
	Varianzen sind nicht gleich	22,4762	11,87903	-1,15585	46,10823

Interpretieren Sie diese Ergebnisse.

This image shows a full page of white paper with horizontal blue or grey ruling lines. The lines are evenly spaced and run across the width of the page. There is no handwriting or other markings on the paper.

a) Was genau besagt die in diesem t-Test ausgewiesene „Sig. (2-seitig)“?

Aufgabe 8:

In einem Abschlussbericht finden sich die nachfolgenden Aussagen. Bitte kommentieren Sie diese.

Aussage 1: „Angesichts dieses statistisch höchstsignifikanten Ergebnisses muss an dieser Stelle unbedingt etwas passieren.“

Aussage 2: „Da die Nullhypothese nicht mit der erforderlichen Irrtumswahrscheinlichkeit verworfen werden konnte, ist nachgewiesen, dass sich Männer und Frauen in dieser Frage nicht unterscheiden.“
