Probeklausur zur Vorlesung Statistik II für Studierende der Soziologie und Nebenfachstudierende im Sommersemester 2012

Prof. Dr. H. Küchenhoff, J. Brandt, G. Schollmeyer, G. Walter

Aufgabe 1

Betrachten Sie den SPSS-Output zur linearen Regression über den Betrag in Schweizer Franken, den eine Person jährlich für Kulturveranstaltungen ausgibt, in Abhängigkeit vom Lebensalter, dem Geschlecht, welches als binäre Variable aufgenommen wurde (Ausprägung 0 bedeutet weiblich, Ausprägung 1 männlich), und dem Jahreseinkommen in Schweizer Franken. Der Stichprobenumfang beträgt $n = 699$ Personen.

Koeffizienten[a]

		Nicht standardisierte Koeffizienten		Standardisierte Koeffizienten		
Modell		B	Standardfehler	Beta	T	Signifikanz
1	(Konstante)	149.985	10.640		14.097	.000
	Alter	1.231	.136	.332	9.066	.000
	Geschlecht	-8.123	3.909	-.078	-2.078	.038
	Jahresgehalt in Tausend Schweizer Franken	.329	.152	.084	2.161	.031

a. Abhängige Variable: Ausgaben für Kultur

(a) Interpretieren Sie die Parameterschätzwerte aus Spalte B.

(b) Welche der Variablen haben einen signifikanten ($\alpha = 0.05$) Einfluss?

(c) Welche Hypothesen werden bei der Überprüfung der Signifikanz der Variable Alter gegeneinander getestet?

(d) Wie kann bei dem Test aus (c) das Signifikanzniveau von 5% interpretiert werden?

(e) Berechnen Sie ein Konfidenzintervall zum Sicherheitsgrad von 95% für den Parameter zur Variable Jahresgehalt.

Aufgabe 2

Man interessiert sich für die Essgewohnheiten von Studenten, welche einen eigenen Haushalt führen. Die Grundgesamtheit setzt sich zu 56 Prozent aus Frauen zusammen. Es zeigt sich, dass sich Frauen mit einer Wahrscheinlichkeit von 0.45 und Männer mit einer Wahrscheinlichkeit von 0.61 regelmäßig von Fast-Food ernähren.

(a) Zeichnen Sie (unter Angabe der unbedingten und bedingten Wahrscheinlichkeiten sowie der zugehörigen Bezugspopulation) den zugrundeliegenden Wahrscheinlichkeitsbaum!

(b) Wie hoch ist die Wahrscheinlichkeit, dass sich ein zufällig ausgewählter Student nicht regelmäßig von Fast-Food ernährt?

(c) Wie hoch ist die Wahrscheinlichkeit, dass ein Student, der sich nicht regelmäßig von Fast-Food ernährt, weiblich ist?

Es wird nun eine Stichprobe von 200 Personen gezogen und der Anteil für den Fast-Food Konsum getrennt für Männer und Frauen berechnet. In dieser Stichprobe ernähren sich 43 der 100 befragten Frauen und 64 der 100 befragten Männer regelmäßig von Fast-Food.

(d) Kann man mit einer Irrtumswahrscheinlichkeit von 0.05 nachweisen, dass der Anteil für regelmäßigen Fast-Food-Konsum unter den Männern höher ist als unter den Frauen?

Aufgabe 3

Es seien zwei unabhängige Zufallsvariablen X und Y gegeben, für die gilt: $P(X = 0) = \frac{1}{2}$, $P(X = 1) = \frac{1}{2}$, $P(Y = 0) = \frac{1}{4}$, $P(Y = 1) = \frac{3}{4}$.

(a) Geben Sie die gemeinsame Verteilung von X und Y an (z.B. in einer Kontingenztafel).

(b) Welche Ausprägungen hat die aus X und Y konstruierte Zufallsvariable $Z = X \cdot Y$?

(c) Welchem Verteilungsmodell entspricht die Verteilung von Z?

(d) Berechnen Sie den Erwartungswert und die Varianz der Zufallsvariable Z.

Aufgabe 4

Der Weg von Universitätsabsolventen zu ihrer ersten Arbeitsstelle soll mit Hilfe eines (homogenen) Markov-Modells analysiert werden. Dabei werden die folgenden Zustände und Übergangswahrscheinlichkeiten zwischen den Zuständen betrachtet:

a_1: in Universitätsausbildung

a_2: arbeitslos gemeldet

a_3: in Praktikum oder Fortbildungsmaßnahme

a_4: an erster Arbeitsstelle

		i+1			
		a_1	a_2	a_3	a_4
	a_1	0	0.2	0.4	
i	a_2	0	0.2	0.6	
	a_3	0	0.3	0	
	a_4				1

(a) Vervollständigen Sie die Matrix der Übergangswahrscheinlichkeiten.

(b) Stellen Sie den zugehörigen Übergangsgraphen auf.

(c) Berechnen Sie $P(A_{i+2,4} \mid A_{i,1})$, d.h. die Wahrscheinlichkeit, in zwei Schritten von Zustand a_1 zum Zustand a_4 zu gelangen.

(d) Diskutieren Sie kurz die Markov-Eigenschaft für dieses Beispiel.

Aufgabe 5

In einer Studie werden Drogensüchtige nach einem Entzug beobachtet, bis sie zum ersten Mal rückfällig werden. Skizzieren und interpretieren Sie zwei mögliche Hazardraten in dem entsprechenden Überlebenszeitmodell.

Aufgabe 6

Formulieren Sie den Hauptsatz der Statistik in Ihren eigenen Worten. Gehen Sie dabei auch explizit auf die Voraussetzungen des Hauptsatzes ein.

Aufgabe 7

Betrachten Sie den Einstichproben-Gauß-Test für den Erwartungswert μ bei der Nullhypothese $H_0 : \mu \leq \mu_0$ gegen die Alternative $H_1 : \mu > \mu_0$ bei normalverteilten $X_1, \ldots, X_n$. Wie verändert sich der Annahmebereich, wenn

- der Stichprobenumfang n steigt.
- das Signifikanzniveau α abnimmt.

Die jeweils anderen Größen werden als konstant angenommen, d.h. Sie führen eine ceteris-paribus-Analyse durch.

Aufgabe 8

Eine Lady behauptet, sie könne, wenn sie einen Tee probiert, der einen Zusatz Milch enthält, unterscheiden, ob zuerst die Milch oder zuerst der Tee eingegossen wurde. Um dies nachzuprüfen, werden 50 Tassen bereitgestellt, in die zufällig (d.h. mit einer Wahrscheinlichkeit von jeweils 0.5) zuerst Tee und dann Milch oder zuerst Milch und dann Tee eingegossen wird. Die Lady wird nun gebeten, aus diesen Tassen zu kosten und sie zu klassifizieren. Hierbei gibt sie 30 richtige und 20 falsche Beurteilungen ab.

Kann aus diesem Ergebnis mit einer Irrtumswahrscheinlichkeit von 5% auf die besonderen Fähigkeiten der Lady geschlossen werden, d.h. kann man davon ausgehen, dass die Lady mehr richtige Beurteilungen abgibt, als wenn sie nur raten würde?

Aufgabe 9

Eine Firma hat ein neues Getränk entwickelt und möchte wissen ob das Getränk in Deutschland positiv aufgenommen wird. Es soll eine Stichprobe erhoben werden, die der Firma einen Eindruck über die Beliebtheit des Getränks gibt. Jeder Befragte muss deshalb nach einer Trinkprobe angeben: „mir schmeckt das Getränk" oder „mir schmeckt das Getränk nicht".

(a) In der Stichprobe wurden $n = 500$ Personen befragt, von denen 200 angaben, dass ihnen das Getränk schmecke. Berechnen Sie ein Konfidenzintervall zum Sicherheitsniveau $\gamma = 95\%$ für den Anteil der Personen in Deutschland, denen das Getränk schmeckt, und interpretieren Sie Ihr Ergebnis.

(b) Dem Chef der Firma ist das angegebene Intervall nicht genau genug (d.h. die Länge des Intervalls ist ihm zu groß). Wie groß hätte der Stichprobenumfang sein müssen, um bei gleichem Sicherheitsniveau $\gamma = 0.95$ eine maximale Genauigkeit von 0.005 zu erhalten?

Beiblatt zur Probeklausur

- Quantile der Standardnormalverteilung

$z_{0.9000} = 1.2816$
$z_{0.9500} = 1.6449$
$z_{0.9750} = 1.9600$
$z_{0.9900} = 2.3263$
$z_{0.9950} = 2.5758$
$z_{0.9990} = 3.0902$
$z_{0.9995} = 3.2905$

- Quantile der t-Verteilung

$t_{0.9000}(10) = 1.3722$	$t_{0.9000}(15) = 1.3406$	$t_{0.9000}(693) = 1.2828$	$t_{0.9000}(694) = 1.2828$
$t_{0.9500}(10) = 1.8125$	$t_{0.9500}(15) = 1.7531$	$t_{0.9500}(693) = 1.6471$	$t_{0.9500}(694) = 1.6471$
$t_{0.9750}(10) = 2.2281$	$t_{0.9750}(15) = 2.1314$	$t_{0.9750}(693) = 1.9634$	$t_{0.9750}(694) = 1.9634$
$t_{0.9900}(10) = 2.7638$	$t_{0.9900}(15) = 2.6025$	$t_{0.9900}(693) = 2.3317$	$t_{0.9900}(694) = 2.3317$
$t_{0.9950}(10) = 3.1693$	$t_{0.9950}(15) = 2.9467$	$t_{0.9950}(693) = 2.5829$	$t_{0.9950}(694) = 2.5829$
$t_{0.9990}(10) = 4.1437$	$t_{0.9990}(15) = 3.7328$	$t_{0.9990}(693) = 3.1020$	$t_{0.9990}(694) = 3.1020$
$t_{0.9995}(10) = 4.5869$	$t_{0.9995}(15) = 4.0728$	$t_{0.9995}(693) = 3.3046$	$t_{0.9995}(694) = 3.3046$

- Quantile der χ^2-Verteilung

$q_{0.9000}(1) = 2.7055$	$q_{0.9000}(2) = 4.6052$
$q_{0.9500}(1) = 3.8415$	$q_{0.9500}(2) = 5.9915$
$q_{0.9750}(1) = 5.0239$	$q_{0.9750}(2) = 7.3778$
$q_{0.9900}(1) = 6.6349$	$q_{0.9900}(2) = 9.2103$
$q_{0.9950}(1) = 7.8794$	$q_{0.9950}(2) = 10.5966$
$q_{0.9990}(1) = 10.8276$	$q_{0.9990}(2) = 13.8155$
$q_{0.9995}(1) = 12.1157$	$q_{0.9995}(2) = 15.2018$