

## Klausur zur Vorlesung „Statistik I für Statistiker, Mathematiker und Informatiker“ (WS 05/06)

11. Februar 2006, 09.00 - 11.00 Uhr

Hinweise:

- (a) Überprüfen Sie bitte, ob Ihre Angabe vollständig ist. Diese Angabe sollte (inklusive dieser Seite) 4 Seiten umfassen.
- (b) Die Bearbeitungszeit beträgt 120 Minuten. Es können 120 Punkte erreicht werden.
- (c) Als Hilfsmittel sind ausschließlich ein Taschenrechner (nicht programmierbar, ohne Plot-Funktion) sowie die Formelsammlung zugelassen.
- (d) Bei Unterschleif gilt die Klausur als nicht bestanden und es erfolgt eine Meldung an das Prüfungsamt.
- (e) Verwenden Sie für Ihre Notizen und Lösungen ausschließlich die Ihnen zur Verfügung gestellten Papierbögen.
- (f) Alle Ausarbeitungen müssen nachvollziehbar sein. Die Ergebnisse müssen klar ersichtlich sein (unterstreichen oder Antwortsatz). Es erfolgt eine detaillierte Bepunktung des Lösungsweges.
- (g) **Geben Sie am Ende der Klausur alle von Ihnen verwendeten Papierbögen sowie die Angabe ab. Kennzeichnen Sie jedes abgegebene Blatt mit Ihrem Namen und Ihrer Matrikelnummer.**

**Bitte ausfüllen und unterschreiben!!!**

Name (in Druckbuchstaben): .....

Matrikelnummer: ..... Studienfach: .....

Geburtstag: ..... Geburtsort: .....

*Ich bestätige, dass ich obige Hinweise zur Kenntnis genommen habe und sie befolgen werde. Ich bin mit einer Veröffentlichung meines Klausurergebnisses im Internet in der Form <Matrikelnummer><Note> einverstanden. (Falls nicht, den letzten Satz bitte streichen!)*

Unterschrift: .....

**Bitte für die Korrektur freilassen!**

Aufgaben und Punkte					$P_{\max} = 120 + 7$		
Aufgaben	1	2	3	4	5	6	Zusatz-Aufgabe
Punkte	30	26	15	14	23	12	+7
erzielt							
$\sum$ Punkte							

**Aufgabe 1**

3+8+7+5+7=30 Punkte

Von August 2004 bis Juli 2005 haben 25 Statistik-Studenten ihr Studium erfolgreich abgeschlossen. Es ergab sich folgende Häufigkeitsverteilung der Studiendauer (in Semestern):

Semesterzahl	7	8	9	10	11	12	13	14	15	16	17	18
Anzahl d. Studenten	1	1	5	4	4	3	1	3	0	1	1	1

- (a) Wie heißt das statistische Merkmal und wie ist es skaliert? Handelt es sich um ein stetiges oder ein diskretes Merkmal? Begründen Sie Ihre Entscheidungen.
- (b) Für die graphische Darstellung sollen die Semesterzahlen 12 und 13 und die letzten drei Ausprägungen zu jeweils einer Gruppe (links abgeschlossen und rechts offen) zusammengefasst werden. Zeichnen Sie nun ein Histogramm der Daten. Welches Prinzip liegt dieser Darstellung zugrunde?
- (c) Bestimmen Sie den Modus, den Median und das arithmetische Mittel. Charakterisieren Sie anhand dieser Parameter die Form der Verteilung (symmetrisch, rechts-, linkssteil).
- (d) Wieviele Semester höchstens benötigen die 25 Prozent schnellsten Studenten?  
Wieviele Semester mindestens brauchen dagegen die 25 Prozent langsamsten Studenten?
- (e) Zeichnen Sie einen Box-Plot der Daten.

**Aufgabe 2**

5+5+2+14 = 26 Punkte

Im Rahmen einer Studie zu Steuerhinterziehungen wurde bei 150 Akten folgende gemeinsame Häufigkeitsverteilung der Variablen  $X$  : „Schulabschluss“ und  $Y$  : „Entdeckungszeitraum“ betrachtet:

	Y : Entdeckungszeitraum			
X : Schulabschluss	$b_1 =$ bereits bei Versuch entdeckt	$b_2 =$ innerhalb eines Jahres entdeckt	$b_3 =$ nach einem Jahr entdeckt	$h_{i\bullet}$
$a_1 =$ Volksschule	36	30		120
$a_2 =$ höhere Schule			16	
$h_{\bullet j}$	40			150

- (a) Vervollständigen Sie die Kontingenztafel.
- (b) Ist der Anteil der erst nach einem Jahr entdeckten Steuerhinterziehungen bei den Tätern mit höherem Schulabschluss größer als bei den Volksschülern?
- (c) Wie groß ist der Anteil der spätestens bis Ende eines Jahres ertappten Steuersünder?
- (d) Ermitteln Sie für diese Stichprobe, ob und gegebenenfalls inwiefern ein deskriptiver Zusammenhang zwischen dem Schulabschluss und dem Entdeckungszeitraum besteht. Verwenden Sie hierfür eine geeignete Maßzahl mit dem Wertebereich  $[0, 1]$ .

**Aufgabe 3**

15 Punkte

Die Anzahl der Erkältungen, an denen man jährlich erkrankt, sei poissonverteilt mit Parameter  $\lambda = 5$ . Nun ist ein neues Wundermittel auf dem Markt, das den Verteilungsparameter für 75 Prozent der Bevölkerung auf  $\lambda = 3$  reduziert. (Bei den übrigen 25 Prozent wirkt es nicht.) Sie testen das Mittel ein Jahr lang und sind in diesem Zeitraum zweimal erkältet. Wie hoch ist die Wahrscheinlichkeit, dass das Mittel bei Ihnen wirkt?

**Aufgabe 4** $3+3+3+5 = 14$  Punkte

Bei einem bestimmten Flugzeugtyp versagt jeder Flugzeugmotor bei einem Flug unabhängig von den anderen Motoren mit einer Wahrscheinlichkeit von 20 Prozent. Ein Flugzeug kann sich in der Luft halten (und auch sicher landen), wenn mindestens zwei Motoren funktionsfähig sind. Sei

$X$  : „Anzahl der funktionierenden Motoren eines dreimotorigen Flugzeugs“ und

$Y$  : „Anzahl der funktionierenden Motoren eines viermotorigen Flugzeugs“.

- Wie sind  $X$  und  $Y$  jeweils verteilt? Bestimmen Sie den Verteilungstyp und den/die zugehörigen Parameter.
- Berechnen Sie  $E(X)$  und  $E(Y)$ . Interpretieren Sie kurz Ihre Ergebnisse.
- Berechnen Sie  $Var(X)$  und  $Var(Y)$ .
- Wie groß ist die Wahrscheinlichkeit eines Absturzes für die beiden Flugzeugtypen? Für welchen Flugzeugtyp ist die Wahrscheinlichkeit eines Absturzes geringer?

**Aufgabe 5** $4+6+6+7=23$  Punkte

Intelligenztests sind in der Regel so konstruiert, dass die IQ-Punkte angenähert einer Normalverteilung folgen. Bei einem bestimmten Intelligenztest sind die Parameter  $\mu = 100$  und  $\sigma^2 = 100$ .

- Wie groß ist die Wahrscheinlichkeit dafür, dass eine zufällig ausgewählte Person nach diesem Test einen IQ von höchstens 75 besitzt?
- Welche untere Grenze wird von 33 Prozent der Personen nicht unterschritten?
- Geben Sie die Grenzen eines um den Erwartungswert symmetrischen Intervalls an, in das der Intelligenzquotient bei 98 Prozent der Personen fällt.
- Rohfassungen von Intelligenztests weisen oft noch nicht die gewünschten Parameterwerte auf. Sie müssen dann noch geeicht werden, d.h. die Punktezahlen sind einer geeigneten Lineartransformation zu unterwerfen. Bei einer dieser Rohfassungen haben 5 Prozent weniger als 104 und 50 Prozent mehr als 120 IQ-Punkte erreicht. Bestimmen Sie die Parameter der zugehörigen Normalverteilung der IQs.

**Aufgabe 6**

12 Punkte

Ein Taxiunternehmer hat fünf Taxen. Die Tages-Nettoeinnahmen (d.h. nach Entlohnung der Fahrer) pro Taxi seien Zufallsvariablen  $X_1, \dots, X_5$  mit folgenden Spezifikationen:

$$\begin{aligned} E(X_i) &= 150 \text{ Euro}, \quad i = 1, \dots, 5, \\ E(X_i^2) &= 25000 \text{ Euro}^2, \quad i = 1, \dots, 5. \end{aligned}$$

Außerdem ist bekannt, dass die Korrelation  $\rho_{X_i X_j} = 0.4$  für  $i \neq j$  ist.

Berechnen Sie Erwartungswert und Standardabweichung der Tages-Nettoeinnahmen des Taxiunternehmers.

**Zusatz-Aufgabe**

+7 Zusatz-Punkte

Seien  $X$  und  $Y$  stochastisch unabhängige, exponentialverteilte Zufallsvariablen zum Parameter  $\lambda > 0$ . Bestimmen Sie die Verteilung von  $Z := \min(X, Y)$ .