

**KLAUSUR**  
**Methodenlehre der Statistik I**

Bitte diesen Teil vollständig ausfüllen:

<b>Name:</b> _____	<b>Vorname:</b> _____
<b>Geburtsdatum:</b> _____	<b>Geburtsort:</b> _____
<b>Matrikel-Nr.:</b> _____	<b>Semesterzahl:</b> _____
<b>Anschrift:</b> _____	
<b>Dozent:</b> _____	<b>Hausaufgaben: wann? SS/WS</b> _____
	<b>bei wem?</b> _____
Wann haben Sie folgende Teilklausuren geschrieben?	
Statistik II: SS/WS _____	Note: _____ Dozent: _____
Wirtschaftsstatistik: SS/WS _____	Note: _____ Dozent: _____
Der Leistungsnachweis für Statistik wird als Gesamtschein ausgestellt, wenn alle drei Teilklausuren bestanden sind.	

---

Bitte diesen Teil nicht ausfüllen.

Aufgabe	Punkte
1	
2	
3	
4	
<b>G</b>	

Punkte (Übertrag)	
Bonus (Hausaufg.)	
Gesamtpunktzahl	
Note	

Methodenlehre der Statistik I  
Klausur

**Bitte beachten!:** Lösungen in die vorgehaltenen Zwischenräume und - falls erforderlich - auf die Rückseite des Vorblattes. Bitte dokumentieren Sie Ihre Lösungswege. Die Angabe von Lösungen ohne nachvollziehbare Lösungswege ergibt keine Punkte.

**Aufgabe 1 (14 Punkte):**

Unter 800 Arbeitslosen wurden die Merkmale Alter in Jahren und die Häufigkeit der Arbeitslosigkeit, also das wievielte Mal jemand arbeitslos ist, erhoben. Aus den Ergebnissen wurde folgende Tabelle erstellt:

Häufigkeit \ Alter	1	2	3
15 bis unter 25	90	24	0
25 bis unter 40	124	40	6
40 bis unter 60	286	144	40
60 bis unter 65	20	12	14

- a) Berechnen Sie für alle 800 Arbeitslosen
- das arithmetische Mittel bezüglich des Merkmals „Häufigkeit der Arbeitslosigkeit“
  - die Standardabweichung für eben dieses Merkmal

- b) Berechnen Sie approximativ
- den Median für das Merkmal „Alter“
  - den Anteil der unter 45jährigen
- c) Welcher Anteil (approximativ) der höchstens zum zweiten Mal Arbeitslos ist unter 30 Jahren alt?
- d) Stellen Sie die Dichtefunktion und die Verteilungsfunktion bezüglich des Merkmals „Alter“ tabellarisch dar.

- e) Stellen Sie die Verteilungsfunktion bezüglich des Merkmals „Häufigkeit der Arbeitslosigkeit“ grafisch dar.
- f) In einer weiteren Gruppe von 1200 Arbeitslosen errechnete man approximativ ein Durchschnittsalter von 36,32 Jahren bei einer Standardabweichung von 13,4 Jahren. Welches Durchschnittsalter und welche Standardabweichung der Altersangaben errechnen sich damit für alle 2000 Arbeitslosen approximativ?

## Aufgabe 2 (9 Punkte):

In einem verkehrsreichen Gebiet der Bundesrepublik Deutschland wurden die Merkmale X:=„Bestand an Kraftfahrzeugen in der Jahresmitte, in Mio.“ und Y:=„Straßenverkehrsunfälle mit Personen- oder Sachschaden im Jahr, in Mio.“ für 12 Jahre erhoben. Man errechnete aus den Daten die folgenden Summen:

$$\sum_{i=1}^{12} x_i = 162,8 \quad \sum_{i=1}^{12} y_i = 14,22 \quad \sum_{i=1}^{12} x_i \cdot y_i = 197,074$$
$$\sum_{i=1}^{12} x_i^2 = 2326,3 \quad \sum_{i=1}^{12} y_i^2 = 17,015$$

- a) Berechnen Sie die empirischen Varianzen, die empirische Kovarianz sowie den empirischen Korrelationskoeffizienten der Merkmale X und Y.

- b) Unterstellen Sie die lineare Regression  $y_i = a + bx_i$  für  $i=1, \dots, 12$ .  
Berechnen Sie

- b1) die empirischen Regressionskoeffizienten a und b nach der Methode der kleinsten Quadrate und interpretieren Sie kurz das Ergebnis für b.

b2) den Anteil der durch diese Regression „erklärten“ Streuung an der Gesamtstreuung der Y-Daten.

c) Mit welcher Schadenshöhe ist nach vorliegender Regression in einem Jahr zu rechnen, in dessen Mitte der Kraftfahrzeugbestand 14 Mio. beträgt und als durchschnittliche Schadenshöhe 5.000 DM/Schadensfall angenommen wird?

### **Aufgabe 3 (7 Punkte):**

In einem Unternehmen wird auf zwei verschiedenen Produktionsbändern I und II derselbe Fernsehrohrentyp produziert. Band I liefert 25% und Band II die restlichen Fernsehrohren der Gesamtproduktion dieses Typs. Es kann nur Fehler A auftreten. Bei Band I beträgt der Anteil der Röhren mit Fehler A 8% und bei Band II 5%..

Aus der Gesamtproduktion des Unternehmens dieses Fernsehrohrentyps wird eine Röhre zufällig ausgewählt. Wie groß ist die Wahrscheinlichkeit, dass eine solche Fernsehrohre

a) Fehler A aufweist und von Band II stammt?

b) einwandfrei ist?

c) von Band I stammt, wenn sie einwandfrei ist?

d) Sind die Ereignisse „Die Fernsehöhre stammt von Band I und ist einwandfrei“ und „Die Fernsehöhre stammt von Band II und ist einwandfrei“ stochastisch unabhängig voneinander? (Mit Begründung!)

#### **Aufgabe 4 (10 Punkte):**

4.1 Aus einer Mischung sehr, sehr vieler gleichartiger und einzeln greifbarer weißer, grüner und blauer Kügelchen eines Reinigungspulvers werden zufällig 25 entnommen.

- a) Das Mischungsverhältnis von weiß : grün : blau sei  $7 : 1 : 2$ , bzw.  $70\% : 10\% : 20\%$ . Wie ist die Zufallsvariable  $X :=$  „Anzahl der grünen unter den entnommenen Kügelchen“ verteilt (Wertebereich, Verteilungstyp, Parameter)?

- b) Wie groß müsste der Anteil der weißen Kügelchen mindestens sein, damit die Wahrscheinlichkeit dafür, nur weiße in der Auswahl zu erhalten, mindestens  $0,001$  beträgt?

4.2 Nun werden aus dem Reinigungspulver mit dem Mischungsverhältnis  $7 : 1 : 2$  für weiß : grün : blau nacheinander einzelne Kügelchen zufällig entnommen.

- a) Wie groß ist die Wahrscheinlichkeit dafür, als 15tes erstmals ein grünes Kügelchen zu erhalten?

- b) Mit wie vielen andersfarbenen Kügelchen müssen Sie im Mittel rechnen, bevor Sie das erste blaue erhalten?

4.3 Aus 20 K ugelchen, die genau dem Mischungsverh altnis 7 : 1 : 2 f ur wei  : gr n : blau entsprechen, w hlen Sie zuf allig drei aus. Wie gro  ist die Wahrscheinlichkeit daf ur, dass

a) alle gr n sind?

b) keines blau ist?

c) Wie viele wei e K ugelchen erwarten Sie in einer derartigen Auswahl und welche Varianz besitzt die Anzahl der wei en K ugelchen in einer derartigen Auswahl?