Klausur
Grundlagen der Statistik
(WS 2001 / 02)

Bemerkungen:

- Die Klausur besteht aus **5 Aufgaben**, die alle zu lösen sind.
- Lösen Sie die Aufgaben **nur** auf den ausgeteilten Lösungsbögen.
- Vermerken Sie auf **jedem Blatt** Ihren Namen und Ihre Matrikelnummer.
- Schreiben Sie bitte **lesbar**.
- Achten Sie darauf, dass Ihr **Lösungsweg nachvollziehbar** ist.

Aufgabe 1 *(10 Punkte)*

Bewerten Sie jede dieser fünf Aussagen mit *RICHTIG* oder *FALSCH* und **begründen** Sie Ihre Entscheidung. Für (richtige) Bewertungen ohne Begründung gibt es **keine** Punkte.

1. Falls für die Zufallsvariablen X und Y gilt: $Var(X) = Var(Y) = 1$, so folgt daraus, dass ihre Korrelation gleich ihrer Kovarianz ist.
2. Das Konfidenzintervall zum Niveau $\alpha = 0,01$ für den unbekannten Parameter θ sei $[3;7]$. Dann gilt mit einer Wahrscheinlichkeit von 99 %, dass $3 \leq \theta \leq 7$ ist.
3. Die zufällige Variable X sei binomial verteilt $(X \sim B(n, \pi))$. Dann wird $Var(X)$ um so größer, je größer der Parameter π ist ($\pi \in [0;1]$).
4. Beim Durchführen eines statistischen (Hypothesen-) Tests darf man H_0 und H_1 erst nach Realisation der Stichprobe aufstellen.
5. Likelihoodfunktionen geben die Wahrscheinlichkeit dafür an, dass ein (unbekannter) Parameter einen bestimmten Wert annimmt. Dabei ist sie maximal für den wahren Wert des Parameters.

Aufgabe 2 *(20 Punkte)*

Fotografin F. lässt seit drei Jahren ihre Bilder bei einem Fotolabor entwickeln. Von diesem bekommt sie monatlich eine Rechnung Sie erteilt Ihnen den Auftrag, die Labor-Kosten der Jahre 1999 - 2001 statistisch auszuwerten.

Die Jahresabrechnung für das Jahr 2001 (in DM) liegt Ihnen dazu in folgender Form vor:

Monat (i)	1	2	3	4	5	6	7	8	9	10	11	12
Betrag (x_i)	353	209	203	125	126	244	179	215	150	283	237	217

1. Geben Sie das arithmetische Mittel, die empirische Standardabweichung, den Median und den Interquartilsabstand der Daten für das Jahr 2001 an.

   Rechenhilfe : $\sum_{i=1}^{12} x_i^2 = 584649$
2. Fertigen Sie den Boxplot der Monatsbeträge für das Jahr 2001 an. Charakterisieren Sie danach die Form der Verteilung.
3. Ermitteln Sie das arithmetische Mittel und die empirische Standardabweichung *bezogen auf die Anzahl* der entwickelten Bilder für das Jahr 2001.
   Die Entwicklung eines Bildes kostete 0,12 DM. Jeden Monat wurde zusätzlich eine Pauschale von 24,60 DM für die Entwicklung aller Filme erhoben (die bereits im Monatspreis enthalten ist).

Für die Jahre 1999 und 2000 liegen die monatlichen Rechnungsbeträge (in DM) nur in gruppierter Form vor:

Gruppe j von... - unter...	Beobachtungen h_j
50 - 150	8
150 - 200	8
200 - 250	2
250 - 500	4
500 - 800	2

4. Fügen Sie die Daten für das Jahr 2001 hinzu und erstellen Sie das gemeinsame Histogramm über alle drei Jahre. Charakterisieren Sie die Form der Verteilung und vergleichen Sie Ihr Ergebnis mit dem aus Aufgabenteil 2.

Aufgabe 3 *(20 Punkte)*

Assistent A. soll das Skript des Professors "Grundlagen der Statistik" auf Fehler prüfen. Da A. schon lange für den Professor arbeitet weiß er, dass er bei den ihm vorgelegten Texten mit durchschnittlich 1 Fehler je Seite rechnen muss. Man weiß, dass die Fehler unabhängig voneinander auftreten.

1. Die zufällige Variable X sei die Anzahl der Tippfehler auf einer beliebigen Seite des Skripts. Geben Sie ihre Verteilung und Parameter an.
2. Stellen Sie die Likelihood-Funktion für den Erwartungswert der Anzahl Fehler pro Seite auf und leiten sie allgemein für eine Stichprobe von n Seiten den ML-Schätzer her.
3. Zeigen Sie, dass dieser ML-Schätzer erwartungstreu ist.
4. Berechnen Sie die Wahrscheinlichkeit, dass auf einer beliebigen Seite die A. liest mindestens ein Tippfehler enthalten ist.
5. Im vorliegenden Skript wurden folgende Fehlerzahlen gefunden:

Fehler pro Seite	0	1	2	3	4
Anzahl der Seiten	20	10	15	40	15

   Berechnen Sie den ML-Schätzwert.
6. Interpretieren Sie Ihr Ergebnis.
7. Wie groß ist die Wahrscheinlichkeit, dass im neuen Skript zur Vorlesung "Bayesianische Statistik", welches der Professor gerade schreibt und das 100 Seiten umfassen soll, kein Fehler enthalten ist. Welche Zufallsvariable haben Sie verwendet.

Aufgabe 4 *(20 Punkte)*

Wissenschaftler W. ist stolz, dass er nun schon seit mehreren Jahren sein Idealgewicht von $\mu^* = \mathbf{85\,kg}$ halten kann. Als er am 1. Februar von einer längeren Forschungsreise aus Island zurückkehrt, vermutet er, dass die dortigen Ernährungsgewohnheiten Einfluss auf sein Gewicht hatten.

Zu diesem Zweck misst er an den folgenden 5 Tagen sein Gewicht. Dabei erhält er folgende Werte:

Datum	Gewicht in kg
2. 2.	89
3. 2.	88
4. 2.	88
5. 2.	89
6. 2.	90

Hinweis: Nehmen Sie an, dass Ws täglich ermitteltes Gewicht (G_i) zufällig um sein wahres Gewicht (μ_W) schwankt:

$$G_i = \mu_w + \varepsilon_i \text{ mit } \varepsilon_i \overset{i.i.d.}{\sim} N\left(0\,;\sigma^2\right)$$

1. Schätzen Sie Ws mittleres Gewicht und die Varianz der Messungen aus den gegebenen Daten.

Mit Hilfe eines statistischen Tests will W entscheiden, ob er von einer signifikanten Veränderung seines Gewichts ausgehen muss. Der Test soll dabei ein Niveau von $\alpha = \mathbf{0{,}01}$ haben.

2. Stellen Sie die Hypothesen zu diesem Test auf und begründen Sie die Hypothesenwahl.
3. Führen Sie den Test durch. Ermitteln Sie dazu die Prüfgröße und deren Verteilung. Fällen Sie anschließend die Testentscheidung. Interpretieren Sie Ihre Testentscheidung.
4. Vergleichen Sie die Konstruktion des Annahmebereichs des oben durchgeführten Tests mit der Konstruktion eines Konfidenzintervalls für Ws Gewicht. Welche Gemeinsamkeiten, welche Unterschiede gibt es?

Aufgabe 5 *(20 Punkte)*

An der Gauß-Universität muss ein Student der Informatik eine Klausur in Mathematik und eine im Fach Statistik schreiben. Die Resultate des letzten Semester ergaben, dass 63% der Studenten die Statistik-Klausur bestanden haben, 58% die Mathematik-Klausur und 68% der Studenten wenigstens eine der beiden Klausuren bestanden haben.

1. Wie groß ist die Wahrscheinlichkeit, dass ein zufällig befragter Student in beiden Klausuren erfolgreich war und wie groß , dass er in keiner der beiden Klausuren erfolgreich war.
2. Wie groß ist die Wahrscheinlichkeit, dass ein befragter Student, der in der Statistik-Klausur keinen Erfolg hatte, die Mathematik-Klausur erfolgreich bestanden hat?
3. In diesem Semester haben 73% der zur Statistik-Klausur angemeldeten Studenten die Übung zur Statistik besucht. Aus Erfahrung weiß man, dass ein Student der die Übung besucht hat mit einer Wahrscheinlichkeit von 81% die Klausur besteht, ein Student, der die Übung nicht besucht hat jedoch nur mit 31%iger Wahrscheinlichkeit erfolgreich ist. Mit welchem Anteil an bestandenen Statistik-Klausuren kann die Gauß -Universität in diesem Semester rechnen?
4. Haben Sie in diesem Semester die Übung besucht?