Prof. Dr. Rainer Schwabe
Otto-von-Guericke-Universität Magdeburg
Institut für Mathematische Stochastik

13.02.2014

Klausur - Schätzen und Testen - Wintersemester 2013/14

Name: ____________________________, Vorname: ____________________________

Matr.-Nr. ________________________

Hinweise: (Bitte lesen Sie diese Hinweise genau durch.)

- Bitte tragen Sie als erstes Ihre persönlichen Daten auf dieser Seite **und** auf dem Deckblatt des ausgegebenen Materials, in das Sie Ihre Lösungen schreiben, ein.
- Zugelassene Hilfsmittel: Ein beidseitig mit der Hand beschriebenes DIN-A4-Blatt (Markierungen sind erlaubt), ein (vom WiWi-Prüfungsamt erlaubter) Taschenrechner, der bereit gestellte (aktualisierte) Formelzettel mit den Tabellen (ohne Ergänzungen; Markierungen sind erlaubt) sowie ein Geodreieck (oder Lineal und Winkelmesser).
- Die Klausur besteht aus **6** Aufgaben. Insgesamt sind maximal **35** Punkte erreichbar.
- Sie haben die Klausur bestanden, wenn Sie insgesamt **mindestens 14** Punkte erreicht haben.
- Die Lösung einer Aufgabe (außer Aufgabe 1) bedeutet nicht einfach die Angabe eines Ergebnisses, sondern erfordert Begründungen, die den Lösungsweg nachvollziehbar machen.
- Interpretieren Sie die statistischen Ergebnisse und geben Sie dort, wo sinnvoll, einen Antwortsatz an.
- Bitte markieren Sie deutlich in den Lösungen, wo eine Aufgabe beginnt und wo sie endet. Mehrfache Antworten bei einer Frage werden als falsch beantwortet bewertet. Kennzeichnen Sie daher Korrekturen oder Streichungen deutlich.

Aufgabe 1 (4 Punkte)
In jeder Teilaufgabe ist **genau eine** Antwort richtig.
Für eine korrekte Antwort erhalten Sie einen Punkt. Für eine falsche Antwort oder eine nicht beantwortete Frage erhalten Sie weder einen Punkt noch wird Ihnen etwas abgezogen. Mehrfachantworten werden als falsch gewertet.

a) Sei X eine Zufallsvariable. Dann gilt:

- ☐ Die Verteilungsfunktion $F_X(x)$ gibt an, mit welcher Wahrscheinlichkeit Beobachtungen größer oder gleich x auftreten.
- ☐ $P(a \leq X \leq b) = F_X(b) - F_X(a)$ für $a < b$.
- ☐ Für stetige Zufallsvariablen X gilt $F_X(x_\alpha) = \alpha$ für das α-Quantil x_α zu X.
- ☐ Quantile stetiger Zufallsvariablen sind immer positiv.
- ☐ Für symmetrische Zufallsvariablen X gilt: $F_X(x_\alpha) = 1 - \alpha$, wobei x_α das α-*Quantil* bezeichnet.

b) Es gilt:

- ☐ Für unabhängige Ereignisse A und B gilt: $P(A \cup B) = P(A) + P(B)$.
- ☐ Für bedingte Wahrscheinlichkeiten gilt: $P(A|B) = P(B|A)$.
- ☐ Für beliebige Zufallsvariablen X, Y gilt: $Var(X + Y) = Var(X) + Var(Y)$.
- ☐ Für unabhängige Zufallsvariablen X, Y gilt: $Var(X-Y) = Var(X)+Var(Y)$.
- ☐ X, Y seien beliebige Zufallsvariablen. Dann gilt: $Corr(X, Y) = -Corr(Y, X)$.

c) Es gilt:

- ☐ $\hat{g}$ sei ein Schätzer für $g(\theta)$. $\hat{g}$ ist erwartungstreu, wenn der Bias von $\hat{g}$ gleich Null ist.
- ☐ Im Normalverteilungsmodell ist der Maximum-Likelihood-Schätzer $\hat{\sigma}^2_{\mathrm{ML}} = \frac{n-1}{n} S^2$ erwartungstreu.
- ☐ Die Likelihood-Funktionen ist eine Dichtefunktion.
- ☐ $MSE(\hat{g}; \theta) = Var_\theta(\hat{g}) - B(\hat{g}; \theta)^2$.
- ☐ Likelihood-Funktionen von diskreten Zufallsvariablen sind nie stetig.

d) Es gilt:

- ☐ Die Nullhypothese wird abgelehnt, wenn der p-Wert größer als das Signifikanzniveau α ist.
- ☐ Die Nullhypothese $H_0 : \theta = \theta_0$ gegen die Alternative $H_1 : \theta \neq \theta_0$ wird zum Niveau α abgelehnt, wenn θ_0 nicht im Konfidenzintervall für θ zum Niveau $1 - \alpha$ liegt.
- ☐ Konfidenzintervalle sind länger bei niedrigerem Konfidenzniveau $1 - \alpha$.
- ☐ Lehnt ein Test auf dem Niveau α die Nullhypothese H_0 eines Testproblems nicht ab, so sprechen die Daten signifikant für H_0.
- ☐ Wird bei einem Test auf Korrelation die Nullhypothese $\rho_{xy} \geq 0$ abgelehnt, dann sind die Daten signifikant unkorreliert.

Aufgabe 2 (8 Punkte)

Ein großes Internetcafe hat Plätze an 60 PCs. Umfangreiche Untersuchungen haben gezeigt, dass innerhalb der Kernzeit durchschnittlich 70% aller PCs belegt sind und jeder PC mit einer Wahrscheinlichkeit von 5% ausfallen kann. Belegung und Ausfall eines PCs sind unabhängig von der Belegung und dem Ausfall der übrigen PCs.

a) Mit welcher Wahrscheinlichkeit sind zu einem bestimmten Zeitpunkt in der Kernzeit alle PCs funktionstüchtig? Geben Sie ein geeignetes stochastisches Modell an.

b) Ermitteln Sie **näherungsweise** die Wahrscheinlichkeit dafür, dass zu einem bestimmten Zeitpunkt innerhalb der Kernzeit mindestens 17, aber höchstens 20 PCs frei sind.

c) Bei einem zufälligen Besuch des Internetcafes in der Kernzeit waren von den 60 PCs 45 belegt. Testen Sie zum Niveau $\alpha = 10\%$, ob der Anteil der Belegung der PCs in der Kernzeit größer als 70% ist und bestimmen Sie ein approximatives Konfidenzintervall zum Konfidenzniveau 0,90 für den Anteil der belegten PCs in der Kernzeit.

Aufgabe 3 (5 Punkte)

Zwei Bäckereien im Stadtteil Cracau konkurrieren um den Namen „Cracauer Brötchenkönig“. Dazu wurden an 10 verschiedenen Sonntagen die folgenden Verkaufszahlen notiert, wobei x_i die Anzahl verkaufter Brötchen der Bäckerei „A“ und y_i die Anzahl verkaufter Brötchen der Bäckerei „B“ am Sonntag i bezeichne:

Sonntag i	1	2	3	4	5	6	7	8	9	10
x_i	315	283	280	312	312	282	306	298	315	282
y_i	312	290	287	301	305	285	298	303	310	286

Aus dieser Stichprobe ergeben sich die folgenden Maßzahlen:

$$\bar{x} = 298{,}5,\ s_x = 15{,}25,\ \bar{y} = 297{,}7,\ s_y = 10{,}11,\ s_{x-y} = 6{,}75,\ s = 12{,}94.$$

a) Geben Sie ein geeignetes Modell an und testen Sie zum Niveau $\alpha = 0{,}05$, ob zwischen den beiden Bäckereien in Hinblick auf die Verkaufszahlen der Brötchen ein signifikanter Unterschied besteht.

b) Berechnen Sie ein 90%-Konfidenzintervall für den Erwartungswert der Sonntagsverkaufszahlen der Brötchen der Bäckerei „A“.

Aufgabe 4 (6 Punkte)
Ein Künstler erhält den Auftrag, ein Mosaik auf dem Marktplatz vor dem Rathaus seiner Gemeinde auszulegen. Die quadratischen Mosaiksteine bezieht er von einer Firma in mehreren Großlieferungen. Es kann vorausgesetzt werden, dass die Lieferungen unabhängig voneinander sind. Leider geraten manchmal runde Mosaiksteine in diese Lieferungen. Sei das Merkmal X die Anzahl der runden Steine pro Lieferung. Gehen Sie davon aus, dass X Poisson-verteilt ist mit dem Parameter λ, es gilt also

$$P(X = x) = \frac{\lambda^x}{x!}\, e^{-\lambda}$$

mit $x \in \{0, 1, 2, \dots\}$, $\lambda > 0$ und $E(X) = \lambda, Var(X) = \lambda$. Bei einer zufälligen Stichprobe von sechs Lieferungen betrug die Anzahl der runden Steine für die jeweilige Lieferung :

$$x_1 = 3, \quad x_2 = 3, \quad x_3 = 7, \quad x_4 = 4, \quad x_5 = 6, \quad x_6 = 1$$

a) Zeigen Sie, dass der Maximum-Likelihood-Schätzer für den unbekannten Parameter λ der Poisson-Verteilung gegeben ist durch $\hat{\lambda}_{ML} = \bar{x}$.

b) Bestimmen Sie die ML-Schätzung für den Erwartungswert und die Varianz von X.

Die Herstellerfirma der Mosaiksteinchen gibt zu, dass für gewöhnlich ein gewisser Anteil an runden Steinen in jeder Lieferung enthalten ist. Laut Herstellerangaben ist der Anteil runder Steine Poisson-verteilt mit einem Parameter λ_0, und $P(X = 0)$, also die Wahrscheinlichkeit, eine Lieferung ohne runde Steine zu erhalten beträgt 6,72%.

c) Bestimmen Sie den Parameter λ_0, der vom Hersteller angegeben wird.

Aufgabe 5 (8 Punkte)
Ein Filialunternehmen will den Einfluss der Größe der Ladenverkaufsfläche auf den Jahresumsatz untersuchen. In einem bestimmten Jahr lieferten 5 Filialen folgende Daten:

Filiale i	Fläche x_i in m^2	Jahresumsatz y_i in Mio €
1	300	3
2	700	4
3	1 000	5
4	1 200	7
5	1 800	11

Demnach ergibt sich:

$$\overline{x} = 1\,000 \;, \quad \overline{y} = 6 \;, \quad \sum_{i=1}^{5}(x_i - \overline{x})^2 = 1\,260\,000 \;, \quad \sum_{i=1}^{5}(y_i - \overline{y})^2 = 40 \;,$$
$$\sum_{i=1}^{5}(x_i - \overline{x})(y_i - \overline{y}) = 6\,900 \;, \quad s_r = 0{,}859 \text{ (Schätzung für } \sigma).$$

a) Bestimmen Sie aufgrund der Daten eine Regressionsgerade für den Jahresumsatz in Abhängigkeit von der Ladenverkaufsfläche. Zeichnen Sie die Datenpunkte und die geschätzte Regressionsgerade in ein Streudiagramm mit **geeigneter** Skalierung der Achsen.

b) Interpretieren Sie – soweit sinnvoll – die geschätzten Parameter der Regressionsgeraden.

c) Geben Sie den Vorhersagewert für einen zukünftigen Jahresumsatz für eine Filiale mit einer Fläche von 1 000 m^2 an, und bestimmen Sie für diesen Jahresumsatz einen Vorhersagebereich zum 90%-Niveau.

Aufgabe 6 (4 Punkte)
Bei Folien, die aus einer Titanlegierung hergestellt werden, soll geprüft werden, ob die Zugfestigkeit an allen Stellen (Ecke, Mitte, Kante) dieselbe ist. Es wurden jeweils vier Folien untersucht. Die Ergebnisse sind in der folgenden Tabelle dagestellt.

1. Gruppe (Ecke)	137	142	128	137
2. Gruppe (Mitte)	140	139	117	137
3. Gruppe (Kante)	142	140	133	141

Aus diesen Daten errechnet sich:

$$\hat{\mu} = 136{,}083 \text{ (Gesamtmittelwert)},$$
$$\sum_{i=1}^{3} n_i(\hat{\mu}_i - \hat{\mu})^2 = 66{,}17 \;, \quad \sum_{i=1}^{3}\sum_{j=1}^{n_i}(y_{ij} - \hat{\mu}_i)^2 = 508{,}75.$$

a) Erstellen Sie die zugehörige Varianzanalysetafel.

b) Unterscheidet sich die Zugfestigkeit signifikant zum 5%-Niveau an den drei Stellen?