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

16.07.2013

Nachhol-Klausur - Schätzen und Testen - Wintersemester 2012/13

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 **34** 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.
- In Aufgabe 1 werden mehrfache Antworten bei einer Frage 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) Es gilt:

- ☐ Wenn für A und B gilt $P(A) > 0$, $P(B) > 0$ und $P(A \cap B) = 0$, dann sind A und B abhängig.
- ☐ Für diskrete Zufallsvariablen gilt $P(X = x) = 0$.
- ☐ Für symmetrische Zufallsvariablen X gilt: $F_{-X}(x_\alpha) = 1 - \alpha$, wobei x_α das α-*Quantil* zu X bezeichnet.
- ☐ Für große Werte von n kann die Verteilungsfunktion binomial verteilter Zufallsvariablen durch die Verteilungsfunktion normalverteilter Zufallsvariablen ersetzt werden. Also gilt: $P(X_n \leq k) = \Phi(k)$ für $X_n \sim B(n, p)$.
- ☐ Für stetige Zufallsvariablen gilt: $P(X \leq x) = F_X(X \leq x)$.

b) Es gilt:

- ☐ $E(a \cdot X + c) = a \cdot E(X)$.
- ☐ Ist die Zufallsvariable X exponential verteilt gilt: $E(X) = \frac{1}{\lambda}$ und $Var(X) = \frac{1}{\lambda}$.
- ☐ Für unabhängige Zufallsvariablen X und Y gilt: $E(X \cdot Y) = E(X) \cdot E(Y)$.
- ☐ Quantile stetiger Zufallsvariablen sind immer positiv.
- ☐ X, Y seien beliebige Zufallsvariablen: $Corr(-X, Y) = Corr(X, Y)$.

c) Es gilt:

- ☐ Im Normalverteilungsmodell gilt für den erwartungstreuen Schätzer S^2 der Varianz σ^2: $P(S^2 = \sigma^2) > 0$.
- ☐ Bei stetigen Zufallsvariablen X sind die Likelihood-Funktionen von X Dichtefunktionen von θ, d. h. $\int L(\theta; x) \mathrm{d}\theta = 1$.
- ☐ Likelihood-Funktionen von diskreten Zufallsvariablen sind stetig.
- ☐ Erwartungstreue Schätzer sind bei gleicher Varianz effizienter als Schätzer mit einem $Bias \neq 0$.
- ☐ Im Normalverteilungsmodell gilt für erwartungstreue Schätzer des Erwartungswertes: $E(\mu) = \overline{X}$.

d) Es gilt:

- ☐ Ist $[\hat{g}_1; \hat{g}_2]$ ein Konfidenzintervall für $g(\theta)$ zum Konfidenzniveau $1 - \alpha$, dann sind $\hat{g}_1$ und $\hat{g}_2$ **keine** Zufallsvariablen.
- ☐ Je größer das Signifikanzniveau, umso größer ist das Konfidenzniveau.
- ☐ Lehnt ein Test auf dem Niveau α die Nullhypothese H_0 eines Testproblems nicht ab, so sprechen die Daten nicht signifikant für H_1.
- ☐ Bei kleinerem Signifikanzniveau ist die Fehlerwahrscheinlichkeit 2. Art ebenfalls kleiner.
- ☐ Ergibt der Test auf Korrelation eine signifikante Korrelation, dann ist die Steigung der Regressionsgeraden nicht signifikant.

Aufgabe 2 (6 Punkte)

Bei einem Börsenindex wird angenommen, dass sich die täglichen Kursentwicklungen als unabhängige Zufallsexperimente darstellen lassen mit Wahrscheinlichkeit $p = 0{,}7$ für einen Kursanstieg und $q = 1 - p = 0{,}3$ für einen Kursverlust für die Tagesdaten. Die Zufallsgröße S_n bezeichne die Anzahl von Kursgewinnen in n Tagen.

a) Für die Prognose der nächsten $n = 5$ Tage berechnen Sie die Wahrscheinlichkeit für das Ereignis, dass höchstens ein Tag mit Kursgewinn auftritt. (Geben Sie hierfür ein geeignetes statistisches Modell an.)

b) Wie viele Tage müssen Sie wenigstens beobachten, um mit einer Wahrscheinlichkeit von 99 % oder mehr mindestens einen Kursverlust zu erleben?

c) Es sollen $n = 500$ Handelstage prognostiziert werden. Berechnen Sie **näherungsweise** die Wahrscheinlichkeit dafür, dass die Anzahl der Tage mit Kursverlust mehr als 155 und weniger als 170 beträgt.

   Hinweis: Nutzen Sie den Zentralen Grenzwertsatz (mit Stetigkeitskorrektur).

Aufgabe 3 (6 Punkte)

Bei einer Umfrage zum Thema „Zufriedenheit am Arbeitsplatz" äußerten 493 von 611 befragten jüngeren Angestellten (≤ 40 Jahre) und 474 von 629 befragten älteren Angestellten (> 40 Jahre), dass sie am Arbeitsplatz zufrieden seien.

a) Testen Sie zum Niveau $\alpha = 10\%$ die Behauptung, dass anteilsmäßig gleich viele junge Angestellte und ältere Angestellte an ihrem Arbeitsplatz zufrieden sind.

b) Bestimmen Sie ein approximatives Konfidenzintervall zum Konfidenzniveau 0,90 für die Anteilsdifferenz der beiden Angestelltengruppen bezüglich der Zufriedenheit am Arbeitsplatz.

Aufgabe 4 (6 Punkte)
Die ausfallfreie Arbeitszeit X (in Zeiteinheiten) eines Maschinentyps lässt sich durch die Verteilungsfunktion

$$F_X(x) = \begin{cases} 1 - e^{-ax^2} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0, \end{cases}$$

mit dem Parameter $a > 0$ beschreiben.

a) Bei Beobachtungen der ausfallfreien Arbeitszeit von 10 unabhängigen Maschinen dieses Typs ergab sich die konkrete Stichprobe

$$9;\ 8;\ 10;\ 11;\ 10;\ 8;\ 9;\ 10;\ 10;\ 9$$

(in Zeiteinheiten). Schätzen Sie mit Hilfe der Maximum-Likelihood-Methode den Parameter a.
Hinweis: Für die Dichte $f_X(x)$ gilt $f_X(x) = F_X'(x)$ für $x \neq 0$, $f_X(0) = 2a$.

b) Der Parameter a sei nun bekannt mit $a = 0{,}01$.

(i) Berechnen Sie die Wahrscheinlichkeit, dass eine Maschine dieses Typs mindestens 5 Zeiteinheiten ohne Ausfall arbeitet.

(ii) Bis zu welchem Zeitpunkt x arbeitet eine Maschine dieses Typs mit genau der Wahrscheinlichkeit 0,75 ohne Ausfall?.

Aufgabe 5 (5 Punkte)
Nach der Fusion zweier Medikamentenhersteller sollen die Vertriebskosten reduziert werden. Der Vertrieb hat zwei Abteilungen, von denen je eine aus einer der fusionierten Firmen stammt. Es wurden für 10 Gebiete die Leistungen der beiden Vertriebsabteilungen analysiert. Man nimmt an, dass die beobachteten Werte normalverteilt sind.

Gebiet	1	2	3	4	5	6	7	8	9	10
Abteilung A: x_i	370	396	390	372	210	415	330	430	298	525
Abteilung B: y_i	428	430	369	385	239	418	313	464	289	547

Hiermit ergibt sich: $\overline{x}$=373,6, s_x=83,6556, $\overline{y}$=388,2, s_y=90,2574, s_{x-y}=25,5135

a) Ist eine der beiden Vertriebsabteilungen signifikant besser? Geben Sie Ihr Modell an und testen Sie zum Signifikanzniveau $\alpha = 5\%$.

b) Berechnen Sie ein 95%- Konfidenzintervall für die erwartete Differenz der Leistungen der beiden Vertriebsabteilungen.

Aufgabe 6 (7 Punkte)
Eine Unternehmensabteilung ist ausschließlich mit der Herstellung eines einzigen Produktes beschäftigt. Wir untersuchen den Zusammenhang zwischen der Produktionsmenge (Anzahl x_i der produzierten Stücke) und den Gesamtkosten (y_i in Euro) anhand der Daten von 10 Produktionsperioden dieser Abteilung:

Periode i	1	2	3	4	5	6	7	8	9	10
Menge x_i	9	12	14	12	12	13	10	11	12	15
Kosten y_i	1216	1300	1356	1288	1276	1292	1260	1244	1288	1360

Demnach ergibt sich:

$$\overline{x}=12\ , \quad \overline{y}=1288\ , \quad \sum_{i=1}^{10}(x_i-\overline{x})^2=28\ , \quad \sum_{i=1}^{10}(y_i-\overline{y})^2=18016\ ,$$

$$\sum_{i=1}^{10}(x_i-\overline{x})(y_i-\overline{y})=672\ , \quad s_r=15{,}36 \text{ (Schätzung für } \sigma).$$

a) Bestimmen Sie den Korrelationskoeffizienten zwischen Produktionsmenge und Gesamtkosten.

b) Passen Sie eine Regressionsgerade für die Gesamtkosten in Abhängigkeit von der Produktionsmenge an und interpretieren Sie die Parameter, soweit dies sinnvoll ist.

c) Testen Sie zum Signifikanzniveau $\alpha = 0{,}05$, ob die Steigung der Gesamtkosten pro Stück signifikant größer als 23 Euro/Stück ist.

d) Geben Sie einen Vorhersagewert für die Gesamtkosten bei einer Produktion von 15 Stück dieses Produktes in einer zukünftigen Produktionsperiode an.