

Nachhol-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 stetige Zufallsvariable. Dann gilt:

- Die Verteilungsfunktion $F_X(x)$ gibt an, mit welcher Wahrscheinlichkeit Beobachtungen kleiner oder gleich x auftreten.
- $P(X = \mu) = 0,5$.
- Die Dichtefunktion nimmt nur Werte zwischen 0 und 1 an.
- Die Verteilungsfunktion $F_X(x)$ ist immer streng monoton wachsend.
- $F_X(b) = \sum_{i=-\infty}^b f_X(i)$.

b) Es gilt:

- X, Y seien beliebige Zufallsvariablen: $E(XY) = E(X)E(Y)$.
- Unkorrelierte Zufallsvariablen sind unabhängig.
- Für Zufallsvariablen X_1, \dots, X_n entspricht der Erwartungswert dem Median.
- $Var(X) = E(X^2) - (E(X))^2$.
- Ist die Zufallsvariable X normalverteilt, gilt $E(X) = 0$ und $Var(X) = 1$.

c) Es gilt:

- Erwartungstreue Schätzer sind bei gleicher Varianz nicht effizienter als Schätzer mit einem Bias $\neq 0$.
- Sei $\hat{\mu}_1$ ein erwartungstreuer Schätzer und $\hat{\mu}_2$ kein erwartungstreuer Schätzer für μ . Dann gilt $Var(\hat{\mu}_2) > Var(\hat{\mu}_1)$.
- Schätzer sind keine Zufallsvariablen.
- Maximum-Likelihood-Schätzer haben einen Bias = 0.
- $\hat{\theta}$ maximiert die Likelihood-Funktion $L(\theta; x)$ bei Beobachtung x genau dann, wenn $\hat{\theta}$ die Log-Likelihood-Funktion $\ln(L(\theta; x))$ maximiert.

d) Es gilt:

- Je größer das Konfidenzniveau $1 - \alpha$, umso kleiner wird das Konfidenzintervall.
- Der Fehler 2. Art ist immer größer als das Signifikanzniveau α .
- Ist der p -Wert kleiner als α , dann wird die Nullhypothese abgelehnt.
- Lehnt ein Test auf dem Niveau α die Nullhypothese H_0 eines Testproblems nicht ab, so sprechen die Daten signifikant für H_1 .
- Ergibt der Test auf Korrelation eine signifikante Korrelation, dann ist die Steigung der Regressionsgeraden nicht signifikant.

Aufgabe 2 (10 Punkte)

Ihre Firma stellt Zahnräder in großen Mengen in 3 Schichten her. Abhängig von der Schicht werden 15, 10 bzw. 25 Zahnräder aus der Produktion entnommen und gründlich untersucht. Man geht davon aus, dass in jeder Schicht gleich zuverlässig produziert wird. Sie nutzen als Ihren Standardschätzer für den Ausschussanteil der Produktion den Schätzer

$$\hat{p} = \frac{1}{50}(X_1 + X_2 + X_3),$$

wobei X_i den beobachteten Ausschuss in der Schicht i beschreibt.

- a) Wie sind X_1 , X_2 und X_3 verteilt? Geben Sie geeignete stochastische Modelle an.
 - b) Überprüfen Sie, ob der Schätzer erwartungstreu für den Ausschussanteil ist.
 - c) Eine vorhergehende Untersuchung hat ergeben, dass der Ausschussanteil 10% ist.
 - i) Mit welcher Wahrscheinlichkeit ist in der Stichprobe von 10 Zahnrädern höchstens ein Ausschussstück?
 - ii) Ermitteln Sie **näherungsweise** die Wahrscheinlichkeit dafür, dass von 1000 in dieser Schicht produzierten Zahnrädern höchstens 110 Ausschuss sind.
Hinweis: Nutzen Sie den Zentralen Grenzwertsatz (mit Stetigkeitskorrektur).
 - d) Wochen später treten in der ersten Schicht Unregelmäßigkeiten auf. Bei einer erneuten Stichprobe von 50 Zahnrädern werden 9 Ausschussstücke festgestellt. Testen Sie zum Niveau $\alpha = 10\%$, ob der Anteil der Ausschussstücke größer als 10% ist und bestimmen Sie ein approximatives Konfidenzintervall zum Konfidenzniveau 0,90 für den Anteil des Ausschusses unter den produzierten Zahnrädern.
-

Aufgabe 3 (5 Punkte)

Sie sind der verantwortliche Produktionsmanager für zwei Fertigungsstätten zur Fußcremeherstellung. Für einen Vergleich der beiden Standorte wurden bei der ersten Produktionsstätte 10 Tage und bei der zweiten Produktionsstätte 10 andere Tage zufällig ausgewählt und die Produktionsmengen (x_i) bzw. (y_i) (in 1000 Packungen) ermittelt:

x_i	83,432	65,208	79,173	77,991	70,414	59,359	67,140	79,704	88,762	90,111
y_i	78,906	65,392	82,141	86,302	81,153	77,651	67,833	78,712	74,901	78,134

Aus diesen Daten berechnen sich folgende Maßzahlen:

$$\bar{x} = 76,129, s_x = 10,263, \bar{y} = 77,113, s_y = 6,338, \bar{d} = -0,983, s_d = 9,864.$$

- Testen Sie zum Niveau $\alpha = 0,1$, ob die beiden Varianzen der Fertigungsstätten in der täglichen Produktionsmenge signifikant verschieden sind und geben Sie dafür Ihr Modell an.
- Sind die durchschnittlichen täglichen Produktionsmengen gleich? Geben Sie Ihr Modell an und testen Sie zum Niveau $\alpha = 0,1$. Hinweis: Sie dürfen hier $\sigma_x^2 = \sigma_y^2$ annehmen.

Aufgabe 4 (5 Punkte)

Die Zufallsvariable X beschreibe die Elastizität von Gummibändern einer Firma (maximaler Streckfaktor, bis zu dem das Gummiband nicht zerreißt). X besitze die folgende Dichtefunktion:

$$f_X(x) = \begin{cases} \alpha\beta x^{\beta-1} e^{-\alpha x^\beta} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0, \end{cases}$$

mit Parametern $\alpha, \beta > 0$. Die folgenden Werte wurden in einer Qualitätskontrolle beobachtet: 2,33 1,67 1,83 2,17 1,93. Es sei $\beta = 3$ bekannt.

- Geben Sie den Maximum-Likelihood-Schätzer für α an und berechnen Sie die ML-Schätzung für α aus den Daten.
- Geben Sie die Verteilungsfunktion für X für allgemeines $\alpha > 0$ an. Berechnen Sie für den speziellen Wert $\alpha = 0,2$
 - $P(X \leq 3,0)$,
 - $P(X > 2,0)$ und
 - $P(X \leq 3,0 | X > 2,0)$.

$$\text{Hinweis: } \int_a^b g'(x) e^{g(x)} dx = e^{g(x)} \Big|_a^b.$$

Aufgabe 5 (7 Punkte)

Ein Kinderpsychologe vermutet, dass sich häufiges Fernsehen negativ auf das Schlafverhalten von Kindern auswirkt. Um dieser Frage nachzugehen, wurde bei neun zufällig ausgewählten Kindern gleichen Alters die Dauer (Y) der Tiefschlafphasen einer Nacht in Stunden gemessen. Außerdem wurde ebenfalls in Stunden angegeben, wie lange das Kind am Tag ferngesehen (X) hat. Es ergeben sich folgende Beobachtungen:

Kind i	1	2	3	4	5	6	7	8	9
x_i	0,3	2,2	0,5	0,7	1,0	1,8	3,0	0,2	2,3
y_i	5,8	4,4	6,5	5,8	5,6	5,0	4,8	6,0	6,1

Daraus ergibt sich:

$$\bar{x} = 1,33, \quad \bar{y} = 5,56, \quad \sum_{i=1}^9 (x_i - \bar{x})^2 = 8,24, \quad \sum_{i=1}^9 (y_i - \bar{y})^2 = 3,72,$$
$$\sum_{i=1}^9 (x_i - \bar{x})(y_i - \bar{y}) = -3,71, \quad s_r = 0,54 \text{ (Schätzung für } \sigma).$$

- Bestimmen Sie aufgrund der Daten eine Regressionsgerade für die Tiefschlafdauer in Abhängigkeit von der Fernsehzeit. Zeichnen Sie die Datenpunkte und die geschätzte Regressionsgerade in ein Streudiagramm mit **geeigneter** Skalierung der Achsen.
 - Interpretieren Sie – soweit sinnvoll – die geschätzten Parameter der Regressionsgeraden.
 - Testen Sie zum Signifikanzniveau $\alpha = 0,05$, ob die Steigung der Tiefschlafdauer pro Stunde Fernsehzeit signifikant kleiner als $-0,30$ ist.
 - Geben Sie den Vorhersagewert für eine künftige Tiefschlafdauer bei einer Fernsehzeit von 2,2 Stunden an, und bestimmen Sie für diese Tiefschlafdauer einen 90%-Vorhersagebereich.
-

Aufgabe 6 (4 Punkte)

In einer Untersuchung sollte überprüft werden, ob die durchschnittlichen Gewinne pro Einzelhandelsunternehmen in den 3 Bezirken einer Stadt sich unterscheiden. Aus jedem Bezirk wurden jeweils 4 Unternehmen zufällig ausgewählt und der zugehörige Jahresgewinn (in 100 000 €) festgestellt. Die Ergebnisse sind in der folgenden Tabelle dargestellt.

1. Bezirk	1	0	1	2
2. Bezirk	3	4	3	2
3. Bezirk	2	0	3	3

Aus diesen Daten errechnet sich:

$$\hat{\mu} = 2,0 \text{ (Gesamtmittelwert)}, \quad \hat{\mu}_1 = 1,0, \quad \hat{\mu}_2 = 3,0, \quad \hat{\mu}_3 = 2,0,$$
$$\sum_{i=1}^3 n_i (\hat{\mu}_i - \hat{\mu})^2 = 8,00, \quad \sum_{i=1}^3 \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu}_i)^2 = 10,00.$$

- Erstellen Sie die zugehörige Varianzanalysetafel.
- Gibt es einen signifikanten Unterschied zwischen den drei Bezirken hinsichtlich des erwarteten Gewinns pro Unternehmen zum 5%-Niveau?