Mathematik IV f. Elektrotechnik Mathematik III f. Informatik 3. Übungsblatt



Fachbereich Mathematik Prof. Dr. Martin Kiehl Davorin Lešnik, Ph.D. Dipl.-Math. Sebastian Pfaff SoSe 2012 2. Mai 2012

Zum Bearbeiten dieses Aufgabenblattes benötigen Sie den Inhalt der Kapitel 9 und 10 im Skript.

Gruppenübung

Aufgabe G1 (Maximum-Likelihood-Schätzer)

Um die Präzision einer Waage zu überprüfen, wird n-mal das Gewicht eines Kilogramm-Prototyps gemessen. Die entstehende Messreihe soll als Realisierung von unabhängigen, identisch $N(1,\theta)$ -verteilten Zufallsvariablen X_1, \ldots, X_n mit unbekannter Varianz $\theta > 0$ aufgefasst werden.

- (a) Bestimmen Sie den Maximum-Likelihood-Schätzer T_n für $\tau(\theta) = \theta$.
- (b) Ist T_n erwartungstreu für $\tau(\theta) = \theta$?

Lösung:

(a) Für $x_1, ..., x_n \in \mathbb{R}$ ist die Likelihoodfunktion gegeben durch

$$L(\theta; x_1, ..., x_n) = \prod_{i=1}^n f_{\theta}(x_i) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\theta}} \exp(-(x_i - 1)^2 / (2\theta))$$
$$= (2\pi\theta)^{-n/2} \cdot \exp\left(-\frac{1}{2\theta} \sum_{i=1}^n (x_i - 1)^2\right).$$

Logarithmieren ergibt

$$g(\theta) := \ln L(\theta; x_1, \dots, x_n) = -\frac{n}{2} \ln(2\pi\theta) - \frac{1}{2\theta} \sum_{i=1}^{n} (x_i - 1)^2.$$

Diese Funktion gilt es nun zu maximieren, also:

$$\frac{dg}{d\theta} = 0$$

$$\Leftrightarrow \quad -\frac{n}{2} \cdot \frac{1}{\theta} + \frac{1}{2\theta^2} \sum_{i=1}^{n} (x_i - 1)^2 = 0$$

$$\Leftrightarrow \quad \frac{n}{2\theta} = \frac{1}{2\theta^2} \sum_{i=1}^{n} (x_i - 1)^2$$

$$\Leftrightarrow \quad \theta = \frac{1}{n} \sum_{i=1}^{n} (x_i - 1)^2 =: \hat{\theta}.$$

1

Wegen

$$\frac{d^{2}g}{d\theta^{2}}\bigg|_{\theta=\hat{\theta}} = \frac{n}{2\hat{\theta}^{2}} - \frac{1}{\hat{\theta}^{3}} \sum_{i=1}^{n} (x_{i} - 1)^{2}$$
$$= \frac{n}{2\hat{\theta}^{2}} - \frac{1}{\hat{\theta}^{3}} \cdot n\hat{\theta} = -\frac{n}{2\hat{\theta}^{2}} < 0$$

liegt an der Stelle $\hat{\theta}$ tatsächlich ein Maximum vor, d.h.

$$T_n(x_1,...,x_n) = \frac{1}{n} \sum_{i=1}^n (x_i - 1)^2$$

ist der Maximum-Likelihood-Schätzer für θ .

(b) Es gilt, da X_1 normalverteilt ist mit Erwartungswert 1 und Varianz θ :

$$E_{\theta}(T_n(X_1, \dots, X_n)) = \frac{1}{n} E_{\theta} \left(\sum_{i=1}^n (X_i - 1)^2 \right)$$
$$= \frac{1}{n} \cdot n \cdot E_{\theta} \left((X_1 - 1)^2 \right)$$
$$= E_{\theta} \left((X_1 - E_{\theta}(X_1))^2 \right)$$
$$= Var_{\theta}(X_1) = \theta.$$

Also ist T_n erwartungstreu für $\tau(\theta) = \theta$.

Aufgabe G2 (Konfidenzintervalle)

Bei der Größenmessung in einer Gruppe von 9 Personen ergaben sich folgende Körpergrößen [in cm]:

Diese Messwerte werden als Realisationen der Zufallsvariablen X_1, \dots, X_9 angenommen, die unabhängig und identisch $N(\mu, \sigma^2)$ -verteilt seien.

- (a) Geben Sie ein Konfidenzintervall zum Konfidenz
niveau 0.99 für den Erwartungswert μ an, falls die Standardabweichung bekannt ist und $\sigma=2.4$ [cm] beträgt.
- (b) Welches Konfidenzintervall ergibt sich in (a) für dasselbe Konfidenzniveau, falls die Standardabweichung als unbekannt angenommen wird?
- (c) Ermitteln Sie im letzteren Fall ein Konfidenzintervall zum Konfidenzniveau 0.9 für die Varianz σ^2 .
- (d) Formulieren Sie in eigenen Worten, was Ihre Resultate aus den ersten drei Aufgabenteilen besagen.

Lösung

(a) Das arithmetische Mittel der Messwerte ist $\bar{X}_{(9)}=184.8$ und das 0.995-Quantil der N(0,1)-Verteilung $u_{0.995}\approx 2.5758$. Daher lautet das gesuchte Konfidenzintervall

$$\left[\bar{X}_{(9)} - u_{0.995} \frac{\sigma}{\sqrt{9}}, \bar{X}_{(9)} + u_{0.995} \frac{\sigma}{\sqrt{9}}\right] \approx [182.7393, 186.8607].$$

(b) Die Stichprobenvarianz beträgt $S_{(9)}^2=1.725$ und das 0.995-Quantil der t_8 -Verteilung $t_{8;0.995}\approx 3.3554$. Daher lautet das gesuchte Konfidenzintervall

$$\left[\bar{X}_{(9)} - t_{8;0.995} \sqrt{\frac{S_{(9)}^2}{9}}, \bar{X}_{(9)} + t_{8;0.995} \sqrt{\frac{S_{(9)}^2}{9}}\right] \approx [183.3310, 186.2690].$$

(c) Das 0.95-Quantil und das 0.05-Quantil der χ^2_8 -Verteilung sind $\chi^2_{8;0.95} \approx 15.507$ und $\chi^2_{8;0.05} \approx 2.733$. Daher lautet das gesuchte Konfidenzintervall

$$\left[\frac{(n-1)S_{(9)}^2}{\chi_{8;0.95}}, \frac{(n-1)S_{(9)}^2}{\chi_{8;0.05}}\right] \approx [0.8899, 5.0494].$$

- (d) (a): Bei bekannter Varianz $\sigma = 2.4$ liegt der Erwartungswert der Körpergrößen mit einer Wahrscheinlichkeit von mindestens 99% zwischen 182.7393 und 186.8607.
 - (b): Bei unbekannter Varianz liegt der Erwartungswert der Körpergrößen mit einer Wahrscheinlichkeit von mindestens 99% zwischen 183.3310 und 186.2690.
 - (c): Bei unbekanntem Erwartungswert liegt die Varianz der Körpergrößen mit einer Wahrscheinlichkeit von mindestens 90 % zwischen 0.8899 und 5.0494.

Aufgabe G3 (Testverfahren)

Eine bestimmte Weizensorte wird auf 9 vergleichbaren, gleich großen Versuchsflächen angebaut. Aus Erfahrung weiß man, dass die Erträge der einzelnen Versuchsflächen als eine Stichprobe unabhängiger, identisch $N(\mu, 3.24)$ -verteilter Zufallsvariablen angesehen werden können. Es ergibt sich ein arithmetisches Mittel von 105.0 [dz].

- (a) Überprüfen Sie mit einem geeigneten Testverfahren die Nullhypothese H_0 : $\mu=106.0$ auf dem Signifikanzniveau $\alpha=0.1$.
- (b) Welche Entscheidung würde sich auf dem Niveau $\alpha = 0.05$ ergeben?
- (c) Überprüfen Sie mit einem geeigneten Testverfahren die Nullhypothese $H_0: \mu \geq 106.0$ auf dem Niveau $\alpha = 0.01$.
- (d) Überprüfen Sie mit einem geeigneten Testverfahren die Nullhypothese $H_0: \mu \leq 106.0$ auf dem Niveau $\alpha = 0.01$.

Lösung:

Geeignetes Testverfahren: Gauß-Test

Begründung: Die Varianz σ^2 wird als bekannt vorausgesetzt, und man sucht ein Testverfahren zum Vergleich von μ mit dem vorgegebenen Zahlenwert ($\mu_0 = 106.0$).

(a) Nullhypothese: H_0 : $\mu = 106.0$, Alternative H_1 : $\mu \neq 106.0$ Testgröße:

$$T(X_1,...,X_n) = \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}$$

Realisierung der Testgröße:

$$T(x_1,...,x_9) = \frac{\bar{x}_9 - \mu_0}{\sqrt{\sigma^2/9}} = \frac{105.0 - 106.0}{\sqrt{3.24/9}}$$

= $-\frac{5}{3} \approx -1.667$.

Kritischer Bereich: Wir lehnen H_0 ab, falls gilt:

$$|T(x_1,\ldots,x_n)| > u_{1-\alpha/2}$$

Entscheidung: Wegen $|T(x_1,...,x_9)| = \frac{5}{3} > 1.645 = u_{0.95}$ wird die Nullhypothese **abgelehnt**.

(b) Da sich gegenüber a) lediglich das Niveau α ändert, bleiben Nullhyothese, Testgröße, deren Realisierung und der kritische Bereich identisch. Dagegen ändert sich die Entscheidung, weil nun mit einem anderen Quantil verglichen wird. Wegen $|T(x_1,\ldots,x_9)|=\frac{5}{3}<1.96=u_{0.975}$ wird die Nullhypothese nicht abgelehnt.

(c) Nullhypothese: $H_0: \mu \ge 106.0$, Alternative $H_1: \mu < 106.0$ Testgröße:

$$T(X_1,\ldots,X_n) = \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}$$

Realisierung der Testgröße:

$$T(x_1,...,x_9) = \frac{\bar{x}_9 - \mu_0}{\sqrt{\sigma^2/9}} = \frac{105.0 - 106.0}{\sqrt{3.24/9}}$$

= $-\frac{5}{3} \approx -1.667$.

Kritischer Bereich: Wir lehnen H_0 ab, falls gilt:

$$T(x_1,\ldots,x_n) < u_\alpha$$

Entscheidung: Wegen $T(x_1,...,x_9)=-\frac{5}{3}>-2.326=u_{0.01}$ wird die Nullhypothese nicht abgelehnt

(d) Nullhypothese: $H_0: \mu \leq 106.0$, Alternative $H_1: \mu > 106.0$ Testgröße:

$$T(X_1,...,X_n) = \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}$$

Realisierung der Testgröße:

$$T(x_1,...,x_9) = \frac{\bar{x}_9 - \mu_0}{\sqrt{\sigma^2/9}} = \frac{105.0 - 106.0}{\sqrt{3.24/9}}$$

= $-\frac{5}{3} \approx -1.667$.

Kritischer Bereich: Wir lehnen H_0 ab, falls gilt:

$$T(x_1,\ldots,x_n)>u_{1-\alpha}$$

Entscheidung: Wegen $T(x_1,...,x_9) = -\frac{5}{3} < 2.326 = u_{0.99}$ wird die Nullhypothese **nicht abgelehnt**.

Hausübung

Aufgabe H1 (Schätzverfahren)

Die Zufallsvariablen X_1, \dots, X_n seien unabhängig und identisch $R(\theta - 1, \theta + 1)$ -verteilt, mit $\theta \in \mathbb{R}$ unbekannt.

- (a) Zeigen Sie, dass das arithmetische Mittel $T_n(X_1,...,X_n) = \overline{X}_{(n)}$ ein erwartungstreuer Schätzer für $\tau(\theta) = \theta$ ist.
- (b) Berechnen Sie die Varianz des Schätzers T_n . $Hinweis: X \sim R(a,b) \Rightarrow Var(X) = \frac{(b-a)^2}{12}$
- (c) Ist die Schätzerfolge T_1, T_2, \dots konsistent für $\tau(\theta) = \theta$?

Lösung:

(a) Zunächt stellen wir fest, dass

$$E_{\theta}(X_1) = \frac{(\theta - 1) + (\theta + 1)}{2} = \theta$$

gilt. Nun können wir die Rechenregeln für den Erwartungswert anwenden, d.h.

$$E_{\theta}(T_{n}(X_{1},...,X_{n})) = E_{\theta}\left(\frac{1}{n}(X_{1}+...+X_{n})\right) = \frac{1}{n}E_{\theta}(X_{1}+...+X_{n})$$

$$= \frac{1}{n}\left(E_{\theta}(X_{1})+...+E_{\theta}(X_{n})\right) = \frac{1}{n}\cdot n\cdot E_{\theta}(X_{1})$$

$$= E_{\theta}(X_{1}) = \theta.$$

(b) Für die Varianz erhalten wir wegen $Var_{\theta}(X_1) = \frac{(\theta + 1 - (\theta - 1))^2}{12} = \frac{1}{3}$ analog

$$Var_{\theta}(T_n(X_1, \dots, X_n)) = Var_{\theta}\left(\frac{1}{n}(X_1 + \dots + X_n)\right) = \left(\frac{1}{n}\right)^2 Var_{\theta}(X_1 + \dots + X_n)$$

$$= \left(\frac{1}{n}\right)^2 \left(Var_{\theta}(X_1) + \dots + Var_{\theta}(X_n)\right) = \left(\frac{1}{n}\right)^2 \cdot n \cdot Var_{\theta}(X_1)$$

$$= \frac{Var_{\theta}(X_1)}{n} = \frac{1}{3n}.$$

(c) Aus der Eigenschaft

$$\lim_{n\to\infty} Var_{\theta}(T_n(X_1,\ldots,X_n)) = \lim_{n\to\infty} \frac{1}{3n} = 0$$

folgt zusammen mit der Erwartungstreue des Schätzers aus Satz 9.1.5 die Konsistenz der Schätzerfolge.

Aufgabe H2 (Konfidenzintervalle für Defektwahrscheinlichkeiten)

Man ist an einem Konfidenzintervall für die Defektwahrscheinlichkeit $\theta \in (0,1)$ eines Produktionsprozesses interessiert. Um die Anzahl defekter Produkte in einer Stichprobe vom Umfang n zu zählen verwenden wir unabhängig identisch $B(1,\theta)$ -verteilte Zufallsvariablen $X_1, \ldots X_n$, wobei $X_i=1$ für $i=1,\ldots n$, falls das i-te Produkt defekt ist. Die Anzahl defekter Produkte in der Stichprobe, also die Summe $Y=X_1+\ldots X_n$, ist dann aufgrund der Unabhängigkeitsannahme $B(n,\theta)$ -verteilt.

(a) Zeigen Sie, dass gilt

$$P_{\theta}\left(-u_{1-\frac{\alpha}{2}} \leq \frac{Y - n\theta}{\sqrt{n\theta(1-\theta)}} \leq u_{1-\frac{\alpha}{2}}\right) = P_{\theta}\left((Y - n\theta)^2 \leq u_{1-\frac{\alpha}{2}}^2 \cdot n\theta(1-\theta)\right) \approx 1 - \alpha.$$

(b) Folgern Sie daraus, dass θ mit Wahrscheinlichkeit $1-\alpha$ im Konfidenzintervall

$$I(X_1, \dots X_n) = \left[\frac{1}{n + u_{1 - \frac{\alpha}{2}}^2} \left(Y + \frac{u_{1 - \frac{\alpha}{2}}^2}{2} - u_{1 - \frac{\alpha}{2}} \sqrt{Y(1 - \frac{Y}{n}) + \frac{u_{1 - \frac{\alpha}{2}}^2}{4}} \right),$$

$$\frac{1}{n + u_{1 - \frac{\alpha}{2}}^2} \left(Y + \frac{u_{1 - \frac{\alpha}{2}}^2}{2} + u_{1 - \frac{\alpha}{2}} \sqrt{Y(1 - \frac{Y}{n}) + \frac{u_{1 - \frac{\alpha}{2}}^2}{4}} \right) \right]$$

liegt.

(c) Ein Hersteller von Elektrogeräten möchte eine grössere Lieferung Transistoren auf ihre Qualität testen. Dazu überprüft er 400 zufällig ausgewählte Transistoren, von denen 12 nicht den Qualitätsanforderungen genügen. Berechnen Sie ein Konfidenzintervall für die Ausschusswahrscheinlichkeit zum Niveau 0.95.

Lösung:

- (a) Nach dem zentralen Grenzwertsatz ist $\frac{Y-n\theta}{\sqrt{n\theta(1-\theta)}}$ N(0,1)-verteilt. Damit folgt die Aussage mit der Definition des Quantils, also mit $\Phi(-u_{1-\frac{\alpha}{2}})=\frac{\alpha}{2}$, siehe auch Skript, Abschnitt 8.8.
- (b) Die Frage, die es zu beantworten gilt, ist welche θ die geforderte Ungleichung

$$(Y - n\theta)^2 \le u_{1 - \frac{\alpha}{2}}^2 \cdot n\theta(1 - \theta)$$

erfüllen. Die Ungleichung kann geschrieben werden als

$$(Y - n\theta)^2 - u_{1 - \frac{\alpha}{2}}^2 \cdot n\theta(1 - \theta) \le 0$$

$$Y^2 - 2n\theta Y + n^2\theta^2 - u_{1 - \frac{\alpha}{2}}^2 \cdot n\theta + u_{1 - \frac{\alpha}{2}}^2 n\theta^2 \le 0$$

$$(n^2 + nu_{1 - \frac{\alpha}{2}}^2)\theta^2 - (u_{1 - \frac{\alpha}{2}}^2 \cdot n + 2nY)\theta + Y^2 \le 0.$$

Die Nullstellen dieses Polynoms 2. Ordnung sind,

$$\begin{split} \theta_{1,2} &= \frac{1}{2(n^2 + nu_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 \cdot n + 2\,n\,Y) \pm \sqrt{(u_{1-\frac{\alpha}{2}}^2 \cdot n + 2\,n\,Y)^2 - 4(n^2 + nu_{1-\frac{\alpha}{2}}^2)Y^2} \right) \\ &= \frac{1}{2(n + u_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 + 2\,Y) \pm \sqrt{(u_{1-\frac{\alpha}{2}}^2 + 2\,Y)^2 - 4(1 + \frac{1}{n}u_{1-\frac{\alpha}{2}}^2)Y^2} \right) \\ &= \frac{1}{2(n + u_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 + 2\,Y) \pm \sqrt{(u_{1-\frac{\alpha}{2}}^4 + 4\,Yu_{1-\frac{\alpha}{2}}^2 + 4Y^2 - 4Y^2 - \frac{4}{n}u_{1-\frac{\alpha}{2}}^2Y^2} \right) \\ &= \frac{1}{2(n + u_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 + 2\,Y) \pm \sqrt{4u_{1-\frac{\alpha}{2}}^2 (\frac{u_{1-\frac{\alpha}{2}}^2}{4} + Y - \frac{1}{n}Y^2)} \right) \\ &= \frac{1}{2(n + u_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 + 2\,Y) \pm 2u_{1-\frac{\alpha}{2}} \sqrt{\frac{u_{1-\frac{\alpha}{2}}^2}{4} + Y - \frac{1}{n}Y^2} \right) \\ &= \frac{1}{(n + u_{1-\frac{\alpha}{2}}^2)} \left((u_{1-\frac{\alpha}{2}}^2 + 2\,Y) \pm u_{1-\frac{\alpha}{2}} \sqrt{\frac{u_{1-\frac{\alpha}{2}}^2}{4} + Y - \frac{1}{n}Y^2}} \right). \end{split}$$

Das Polynom ist kleinergleich Null genau zwischen diesen Nullstellen und somit folgt das angegebene Intervall.

(c) Hier ist n=400, Y=12 und $\alpha=0.05$. Wir lesen aus der Tabelle $u_{0.975}=1.96$. Einsetzen ergibt das Intervall

Damit liegt die Ausschusswahrscheinlichkeit mit einer Wahrscheinlichkeit von 95% zwischen 1.72% und 5.17 %.

Aufgabe H3 (Testverfahren)

Um die Genauigkeit eines neu entwickelten Gerätes zur Messung von Weglängen im Gelände zu kontrollieren, wurde eine bestimmte Strecke von genau $1000 \ m$ zehnmal vermessen. Es ergaben sich folgende Meßwerte (in m):

Es wird angenommen, dass die Messwerte eine Realisierung unabhängiger $N(\mu, \sigma^2)$ - verteilter Zufallsvariablen sind.

- (a) Überprüfen Sie zum Niveau $\alpha = 0.05$ die Hypothese, dass das Gerät mindestens die korrekte Entfernung als Erwartungswert hat.
- (b) Überprüfen Sie unter der Voraussetzung, dass $\sigma^2 = 4$ gilt, zum Niveau $\alpha = 0.05$ die Hypothese, dass das Gerät die korrekte Entfernung als Erwartungswert hat.

Hinweis: Es gilt $\bar{X}_{(10)} = 999.3$ und $S_{(10)}^2 = 3.9$.

Lösung:

(a) Geeignetes Testverfahren: t-Test

Begründung: Die Varianz σ^2 ist unbekannt, und man sucht ein Testverfahren zum Vergleich von μ mit dem vorgegebenen Zahlenwert ($\mu_0 = 1000$).

Nullhypothese: $H_0: \mu \ge 1000$, Alternative $H_1: \mu < 1000$

Testgröße:

$$T(X_1,...,X_n) = \frac{\bar{X}_{(n)} - \mu_0}{\sqrt{S_{(n)}^2/n}}$$

Realisierung der Testgröße:

$$\bar{X}_{(10)} = \frac{1}{10} \sum_{i=1}^{10} x_i = 999.3,$$

$$S_{(10)}^2 = \frac{1}{9} \left(\sum_{i=1}^{10} x_i^2 - 10 \cdot \bar{X}_{(10)}^2 \right) = 3.9$$

Damit folgt:

$$T(x_1,...,x_{10}) = \frac{\bar{X}_{(10)} - \mu_0}{\sqrt{\frac{S_{(10)}^2}{10}}} = \frac{999.3 - 1000}{\sqrt{\frac{3.9}{10}}} \approx -1.1209.$$

Kritischer Bereich: Wir lehnen H_0 ab, falls gilt:

$$T(x_1, \ldots, x_n) < t_{n-1:\alpha}$$

Entscheidung: Wegen $T(x_1,...,x_{10}) \approx -1.1209 > -1.8331 = t_{9;0.05}$ wird die Nullhypothese **nicht** abgelehnt.

(b) Geeignetes Testverfahren: Gauß-Test

Begründung: Entgegen der Aufgabenstellung in a) ist die Varianz σ^2 nunmehr bekannt.

Nullhypothese: $H_0: \mu = 1000$, Alternative $H_1: \mu \neq 1000$

Testgröße:

$$T(X_1,\ldots,X_{(n)}) = \frac{\bar{X}_n - \mu_0}{\sqrt{\sigma^2/n}}$$

Realisierung der Testgröße:

$$T(x_1,...,x_{10}) = \frac{\bar{X}_{(10)} - \mu_0}{\sqrt{\sigma^2/10}} = \frac{999, 3 - 1000}{\sqrt{4/10}}$$

= $-\frac{7}{2\sqrt{10}} \approx -1.1068$.

Kritischer Bereich: Wir lehnen H_0 ab, falls gilt:

$$|T(x_1,...,x_n)| > u_{1-\alpha/2}$$

Entscheidung: Wegen $|T(x_1,...,x_9)| \approx 1.1068 < 1.96 = u_{0.975}$ wird die Nullhypothese **nicht abgelehnt**.