Fachbereich Mathematik
Prof. Dr. Michael Kohler
Dipl.-Math. Andreas Fromkorth
Dipl.-Inf. Jens Mehnert

TECHNISCHE UNIVERSITÄT DARMSTADT

SS 09
13.7.2009

# 11. Übungsblatt zur „Einführung in die Stochastik"

**Aufgabe 42** (3 Punkte)

Student S. vermutet, dass die zufällige Zeit (in Minuten), die Dozent K. bei seiner Statistik Vorlesung immer zu früh kommt, durch eine stetig verteilte Zufallsvariable $X$ mit Dichte

$$f_\alpha(x) = \begin{cases} \alpha & \text{für} \quad 0 \leq x \leq 10, \\ \frac{1}{10} - \alpha & \text{für} \quad 10 < x \leq 20, \\ 0 & \text{für} \quad x < 0 \text{ oder } x > 20 \end{cases}$$

beschrieben werden kann. Um den Parameter $\alpha \in \left[0, \frac{1}{10}\right]$ der Dichte von $X$ zu schätzen, notiert sich Student S., dass Dozent K. bei den letzten $n = 5$ Vorlesungen

$$x_1 = 5 \text{ bzw. } x_2 = 12 \text{ bzw. } x_3 = 3 \text{ bzw. } x_4 = 7 \text{ bzw. } x_5 = 19$$

Minuten zu früh kam.

a) Bestimmen Sie die Likelihood-Funktion

$$L(\alpha) = \prod_{i=1}^{5} f_\alpha(x_i), \quad (\alpha \in [0, \frac{1}{10}]).$$

b) Bestimmen Sie – ausgehend von den angegebenen Werten von $x_1, \ldots, x_5$ – die zugehörige Maximum-Likelihood-Schätzung von $\alpha$.

c) Seien $X_1, \ldots, X_n$ unabhängige und identisch verteilte Zufallsvariablen mit Dichte $f_\alpha$. Zeigen Sie: Der Schätzer

$$T_n(X_1, \ldots, X_n) = \frac{1}{10 \cdot n} \sum_{i=1}^{n} 1_{[0,10]}(X_i)$$

mit

$$1_{[0,10]}(x) = \begin{cases} 1 & \text{für} \quad 0 \leq x \leq 10, \\ 0 & \text{für} \quad x < 0 \text{ oder } x > 10, \end{cases}$$

ist ein erwartungstreuer Schätzer für $\alpha$.

d) Ist der Schätzer in c) auch stark konsistent ? Begründen Sie ihre Antwort.

**Lösung:** a) Es gilt

$$L(\alpha; x_1, \ldots, x_5) = \prod_{i=1}^{5} f_\alpha(x_i)$$
$$= \alpha^{\#\{x_i : 0 \leq x_i \leq 10, i \in \{1,\ldots,10\}\}} \cdot \left(\frac{1}{10} - \alpha\right)^{\#\{x_i : 10 < x_i \leq 20, i \in \{1,\ldots,10\}\}}$$

b) Es gilt

$$\begin{aligned} L(\alpha; 5, 12, 3, 7, 19) =& f_\alpha(5) \cdot f_\alpha(12) \cdot f_\alpha(3) \cdot f_\alpha(7) \cdot f_\alpha(19) \\ =& \alpha^3 \cdot \left(\frac{1}{10} - \alpha\right)^2 \\ =& \alpha^5 - \frac{2}{10}\alpha^4 + \frac{1}{100}\alpha^3 \end{aligned}$$

Es gilt nun

$$\begin{aligned} L^{'}(\alpha; 5, 12, 3, 7, 19) =& 5\alpha^4 - \frac{4}{5}\alpha^3 + \frac{3}{100}\alpha^2 \\ =& \alpha^2 \cdot \left(5\alpha^2 - \frac{4}{5}\alpha + \frac{3}{100}\right) \end{aligned}$$

und

$$L^{''}(\alpha; 5, 12, 3, 7, 19) = 20\alpha^3 - \frac{12}{5}\alpha^2 + \frac{6}{100}\alpha.$$

Nulltellen der ersten Ableitung sind, 0 (zweifach), $\frac{3}{50}$ und $\frac{1}{10}$. Das gesuchte Maximum kann also nur $\alpha = \frac{3}{50}$ (sonst wäre $L(\alpha; 5, 12, 3, 7, 19) = 0$). Da die Randwerte ebenfalls eine 0 liefern und $L^{''} < 0$ für $\alpha = \frac{3}{50}$ ist gilt dies damit auch.
c) Es gilt

$$\begin{aligned} E\left[T_n\left(X_1, \ldots, X_n\right)\right] =& \frac{1}{10 \cdot n} \sum_{i=1}^{n} E\left[1_{[0,10]}\left(X_1\right)\right] \\ =& \frac{1}{10 \cdot n} \sum_{i=1}^{n} 1 \cdot P\left[1_{[0,10]}\left(X_1\right) = 1\right] \\ =& \frac{1}{10 \cdot n} \sum_{i=1}^{n} 10 \cdot \alpha \\ =& \alpha \end{aligned}$$

d) Es gilt nach dem starken Gesetz der großen Zahlen

$$T_n\left(X_1, \ldots, X_n\right) = \frac{1}{10 \cdot n} \sum_{i=1}^{n} 1_{[0,10]}\left(X_n\right) \rightarrow E\left[1_{[0,10]}\left(X_1\right)\right] = \alpha$$

fast sicher, womit der Schätzer auch stark konsistent ist.

**Aufgabe 43** (3 Punkte)

Bei der Shell Jugendstudie 2006 wurden 1231 Mädchen befragt. Dabei gaben 55 Prozent der befragten Mädchen an, das Abitur anzustreben. Leiten Sie aus dem zentralen Grenzwertsatz ein approximatives Konfidenzintervall $[u, 1]$ zum Konfidenzniveau 0.95 für den Anteil der Mädchen her, die das Abitur anstreben.

**Lösung:** Nach der Vorlesung ist

$$\frac{\sqrt{n}}{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i\left(1-\frac{1}{n}\sum_{i=1}^n X_i\right)}}\left(\frac{1}{n}\sum_{i=1}^n X_i - EX_1\right)$$

annähernd $N(0,1)$-verteilt. Sei $u_\alpha \in \mathbb{R}$ so, dass für $N(0,1)$-verteilte ZV $Z$ gilt

$$P[Z \le u_\alpha] = 1-\alpha.$$

Dann gilt wegen $EX_1 = p$:

$$\lim_{n\to\infty} P\left[\frac{\sqrt{n}}{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i\left(1-\frac{1}{n}\sum_{i=1}^n X_i\right)}}\left(\frac{1}{n}\sum_{i=1}^n X_i - EX_1\right) \le u_\alpha\right] = 1-\alpha.$$

Die Bedingung innerhalb der Wk. ist äquivalent zu

$$\begin{aligned} p &\ge \frac{1}{n}\sum_{i=1}^n X_i - u\alpha \cdot \frac{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i\left(1-\frac{1}{n}\sum_{i=1}^n X_i\right)}}{\sqrt{n}} \\ &= 0.55 - 1.65 \cdot \frac{\sqrt{0.55\cdot 0.45}}{\sqrt{1231}} \\ &= 0.527 \end{aligned}$$

also erhalten wir als gesuchtes Konfidenzintervall

$$\left[\frac{1}{n}\sum_{i=1}^n X_i - u\alpha \cdot \frac{\sqrt{\frac{1}{n}\sum_{i=1}^n X_i\left(1-\frac{1}{n}\sum_{i=1}^n X_i\right)}}{\sqrt{n}}, 1\right] \approx [0.527, 1].$$

**Aufgabe 44** (3 Punkte)

Ein Autohersteller behauptet, dass der Benzinverbrauch für einen neuentwickelten Typ im Mittel 6 Liter pro 100km beträgt. Dabei kann er davon ausgehen, dass der Verbrauch normalverteilt ist mit der Standardabweichung $\sigma = 0.3$.

a) Eine Verbraucherzentrale vermutet, dass der Hersteller einen zu niedrigen Mittelwert $\mu$ angegeben hat. Sie überprüft deshalb 20 Autos des neuen Typs auf ihren Verbrauch und berechnet aus diesen Werten das arithmetische Mittel $\bar{x} = 6.1$. Kann man hiermit die Behauptung des Herstellers widerlegen ?

b) Eine Autozeitschrift bekommt von 152 Käufern des neuen Typs Beschwerden über den zu hohen Verbrauch zugesandt. Sie errechnet aus diesen Werten das arithmetische Mittel $\bar{x} = 7.3$. Kann man hiermit die Behauptung des Herstellers widerlegen ?

**Lösung:** a) Hier führen wir einen Gaußtest zum Niveau $\alpha = 0.05$ durch. Es soll überprüft werden

$$H_0 : \mu \le 6 \text{ vs. } H_1 : \mu > 6.$$

Es gilt

$$\begin{aligned} T(X_1, \ldots, X_n) &= \frac{\sqrt{n}}{\sigma_0}(\overline{X}_n - \mu_0) \\ &= \frac{\sqrt{20}}{0.3}(6.1 - 6) \\ &\approx 1.49 \end{aligned}$$

Es gilt $\Phi(c) = 0.95$ für $c = 1.65$, womit mit $1.49 < 1.65$ $H_0$ nicht abgelehnt werden kann. Damit ist die Angabe des Herstellers bzgl. des gegebenen Niveaus $\alpha = 0.05$ nicht wiederlegt.
b) Wir führen wir einen Gaußtest zum Niveau $\alpha = 0.05$ durch. Wir prüfen

$$H_0 : \mu \leq 6 \text{ vs. } H_1 : \mu > 6.$$

Damit gilt

$$\begin{aligned} T(X_1, \ldots, X_n) &= \frac{\sqrt{n}}{\sigma_0}(\overline{X}_n - \mu_0) \\ &= \frac{\sqrt{152}}{0.3}(7.3 - 6) \\ &= 53.42 > 1.65. \end{aligned}$$

Hier kann $H_0$ abgelehnt werden. Man könnte das Testergebnis trotzdem anzweifeln wegen einem sampling bias bei der Fahrzeugauswahl. Es haben sich eventuell nur die Leute gemeldet, die einen erhöhten Verbrauchswert festgestellt haben. Dies kann aus Übertreibung, Defekten des Fahrzeuges, Fahrweise des Fahrzeughalters und ähnlichem resultieren.

**Aufgabe 45** (3 Punkte)

In einer Studie wurde die Pulsfrequenz von 53 8-9-jährigen Jungen gemessen. Es ergab sich eine mittlere Pulsfrequenz von $86,7$ Schlägen / Minute. Langjährige Erfahrungen haben gezeigt, dass die Pulsfrequenz normalverteilt ist mit Mittelwert $\mu$ und Standardabweichung $\sigma = 10,3$ Schläge / Minute. Zu testen ist die Hypothese $H_0 : \mu \geq 90$ gegen $H_1 : \mu < 90$.
In der Praxis gibt man zur Beurteilung der Signifikanz der Daten häufig den sogenannten *p-Wert* der Daten an, d.h., das kleinste Fehlerlevel $\alpha$, mit dem man die Nullhypothese noch ablehnen kann. Berechnen Sie zu den obigen Daten den p-Wert, und interpretieren Sie das Ergebnis.

**Lösung:** Wür führen einen Gaußtest zum niveau $\alpha$ durch, wie in der Aufgabenstellung beschrieben mit

$$\begin{aligned} T(X_1, \ldots, X_n) =& \frac{\sqrt{n}}{\sigma_0}(\overline{X}_n - \mu_0) \\ =& \frac{\sqrt{53}}{10,3}(86,7 - 90) \\ \approx& -2,333 < c. \end{aligned}$$

Gesucht ist $\alpha$ mit $\Phi(c) = \alpha$ so, dass $H_0$ gerade noch abgelehnt werden kann. Es gilt $\Phi(-2,33) = 1 - \Phi(2,33) = 0.01$. Bzgl. der gegebenen Tabelle ist $\alpha = 0.01$ der gesuchte $p$-wert, denn $-2.333 < -2.33$.

**Einige Werte der Verteilungsfunktion ($\Phi$) der Standardnormalverteilung:**

| $x$ | 0 | 0.13 | 0.26 | 0.42 | 0.53 | 0.85 | 1.29 | 1.65 | 1.96 | 2.33 | 2.58 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Phi(x)$ | 0.5 | 0.55 | 0.6 | 0.66 | 0.7 | 0.8 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |