Albert-Ludwigs-Universität Freiburg Wintersemester 2009

Statistische Methoden der Datenanalyse

Markus Schumacher

Übung VI

Matthew Beckingham und Henrik Nilsen

26.11.2009

Anwesenheitsaufgaben

Aufgabe 23 *Erwartungstreue und Varianz des arithmetischen Mittelwerts*

Gegeben sei eine beliebige WDF mit Erwartungswert $E[x] = \mu$ und Varianz $V[x]$ und eine Stichprobe von Umfang n. Der arithmetiche Mittelwert der Stichprobe,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{1}$$

ist ein Schätzer für μ.

(i) Zeigen Sie, dass $\bar{x}$ erwartungstreu ist.

(ii) Zeigen Sie, dass die Varianz von $\bar{x}$, $V[\bar{x}]$, gleich $V[x]/n$ ist.

Aufgabe 24 *Erwartungstreuer Schätzer für σ^2*

Es werden n Messungen einer Zufallsvariablen x, die nach einer beliebigen Wahrscheinlichkeitsdichtefunktion mit Mittelwert μ und Varianz σ^2 verteilt ist, durchgeführt. Zeigen Sie, dass

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad \text{mit} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

ein erwartungstreuer Schätzer für σ^2 ist. Wie groß ist die Abweichung, wenn man $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ als Schätzer benutzen würde?
Tipp: $x_i - \bar{x} = x_i - \mu + \mu - \bar{x}$!

Aufgabe 25 *„Maximum Likelihood" für die Poissonverteilung*

Die Anzahl k der innerhalb einer Minute das Höllental passierenden LKW sei poissonverteilt, d.h. es gilt:

$$f(k;\nu) = \frac{\nu^k}{k!}e^{-\nu}, \quad k = 0,1,2,\cdots, \quad \nu > 0. \tag{3}$$

Eine Stichprobe mit Umfang $n = 100$ ($x_1, \cdots, x_{100}$) liefert folgende Ergebnis:

k	absolute Häufigkeit
0	10
1	20
2	30
3	20
4	10
5	5
6	5
7+	0

(i) Schätzen Sie den Wert des Parameters ν, indem Sie $L(\nu) = \prod_{i=1}^{100} f(k = x_i;\nu)$ maximieren für die gegebene Stichprobe ($x_1, \cdots, x_{100}$).

(ii) Berechnen Sie den ML-Schätzer $\hat{\nu}$ für eine allgemeine Stichprobe mit Umfang n, indem Sie $\log L$ maximieren. Die Messwerte sind dabei Poisson-verteilt.

(iii) Ist der Schätzer erwartungstreu?

(iv) Berechnen Sie die Varianz des Schätzers $V(\hat{\nu})$ für den Fall $n \to \infty$ unter der Annahme, dass der ML-Schätzer effizient ist. Verwenden Sie die MVB für erwartungstreue Schätzer

$$V[\hat{\nu}] = E\left[-\frac{\partial^2 \log L}{\partial \nu^2}\right]^{-1} \tag{4}$$

und dass für einen effizienten Schätzer gilt

$$E\left[\frac{\partial^2 \log L}{\partial \nu^2}\right] = \left(\frac{\partial^2 \log L}{\partial \nu^2}\right)_{\nu=\hat{\nu}} . \tag{5}$$

Hausaufgaben

Aufgabe 26 *„Maximum Likelihood" für die Normalverteilung* **6 Punkte**

Betrachten Sie n verschiedene Messungen x_i, wobei jede einen bekannten Fehler σ_i besitzt. Jede Variable ist also gemäß der Gaussverteilung $G(x_i;\mu,\sigma_i)$ verteilt. In dieser Übung werden wir den besten Schätzwert für den Mittelwert μ dieser Messungen ermitteln.

(i) Stellen Sie die Likelihoodfunktion für die n gaussischen Messungen auf.

(ii) Berechnen Sie den „Maximum Likelihood"-Schätzer $\hat{\mu}$ für den Mittelwert.

(iii) Berechnen Sie die Varianz des „Maximum Likelihood"-Schätzers $\hat{\mu}$, indem Sie die MVB-Formel

$$V(\hat{\mu}) = \frac{1}{\left(-\frac{\partial^2 \ln L}{\partial \mu^2}\right)_{\mu=\hat{\mu}}} \tag{6}$$

benutzen.

Aufgabe 27 *Beweis der Rao-Cremér-Frechet-Ungleichung* **10 Punkte**

Eine Zufallsvariable folge der Wahrscheinlichkeitsdichtefunktion $f(x;\theta)$, wobei θ ein unbekannter Parameter sei. Betrachten Sie eine Stichprobe $\vec{x} = (x_1,\ldots,x_n)$, die benutzt wird, um einen Schätzer $\hat{\theta}(\vec{x})$ für θ zu ermitteln. Gehen Sie von diesem Startpunkt aus, um die Rao-Cremér-Frechet-Ungleichung

$$V[\hat{\theta}] \geq \frac{(1+\frac{\partial b}{\partial \theta})^2}{-E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right]} \tag{7}$$

zu beweisen, wobei $b = E[\hat{\theta}] - \theta$ die Verzerrung (bzw. der Bias) des Schätzers ist. Gehen Sie dazu wie folgt vor:

(i) Benutzen Sie die Cauchy-Schwarz-Ungleichung für zwei Funktionen u und v

$$\left(\int u^2 \mathrm{d}x_1 \ldots \mathrm{d}x_n\right)\left(\int v^2 \mathrm{d}x_1 \ldots \mathrm{d}x_n\right) \geq \left(\int uv \mathrm{d}x_1 \ldots \mathrm{d}x_n\right)^2, \tag{8}$$

mit $u = (\hat{\theta} - E[\hat{\theta}])\sqrt{L}$ und $v = \frac{\partial \ln L}{\partial \theta}\sqrt{L}$, wobei $L = f_{\text{gem}}(\vec{x};\theta)$ die Likelihoodfunktion ist, welche auch gleichzeitig die gemeinsame Wahrscheinlichkeitsdichteverteilung für $\vec{x}$ ist.

Schreiben Sie die Ungleichung in Form von Erwartungswerten und Varianzen und ermitteln Sie daraus eine untere Grenze auf $V[\hat{\theta}]$.

(ii) Nehmen Sie an, dass $\vec{x} = (x_1,\ldots,x_n)$ nicht von θ abhängt, so dass θ und die Ableitung nach θ vor das Integral gezogen werden können. Zeigen Sie damit, dass

$$E\left[\frac{\partial \ln L}{\partial \theta}\right] = \int \frac{\partial \ln f_{\text{gem}}(\vec{x};\theta)}{\partial \theta} f_{\text{gem}}(\vec{x};\theta)\mathrm{d}x_1 \ldots \mathrm{d}x_n = 0.$$

(iii) Zum Beweis der Rao-Cremér-Frechet-Ungleichung zeigen Sie nun, dass

$$E\left[\hat{\theta}\frac{\partial \ln L}{\partial \theta}\right] = 1 + \frac{\partial b}{\partial \theta},$$

und in ähnlicher Weise:

$$E\left[\left(\frac{\partial \ln L}{\partial \theta}\right)^2\right] = -E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right].$$

Nehmen Sie dazu wieder an, dass die Ableitung nach θ und die Integration über $\vec{x}$ vertauscht werden können.

Aufgabe 28 *Erwartungswert für tau* **4 Punkte**

Die ML-Schätzer für den Parameter τ in der WDF $f(t_i;\tau) = \frac{1}{\tau}e^{-t_i/\tau}$ ist gegeben durch

$$\hat{\tau} = \frac{1}{n}\sum_{i=1}^{n} t_i. \tag{9}$$

Mit Hilfe der charakteristischen Funktion kann man die WDF des Schätzwerts bestimmen als

$$g(\hat{\tau};n,\tau) = \frac{n^n}{(n-1)!} \frac{\hat{\tau}^{n-1}}{\tau^n} e^{-n\hat{\tau}/\tau}. \tag{10}$$

In der Vorlesung wurde behauptet, dass $\hat{\tau}$ erwartungstreu ist, d.h. dass

$$\int_0^\infty \hat{\tau} g(\hat{\tau};n,\tau) d\hat{\tau} = \tau. \tag{11}$$

Zeigen Sie, dass die Behauptung wahr ist. Hinweis: Verwenden Sie $\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$ n mal, bis der Faktor $\hat{\tau}^n$ nicht mehr vorkommt.