Albert-Ludwigs-Universität Freiburg Sommersemester 2009

Statistische Methoden der Datenanalyse

Markus Schumacher

Übung V

Matthew Beckingham und Markus Warsinsky

27.5.2009

Anwesenheitsaufgaben

Aufgabe 26 *Anpassung einer parabolischen Parametrisierung*

Benutzen Sie die Methode der kleinsten Quadrate, um eine parabolische Parametrisierung

$$f_i(\theta_1,\theta_2,\theta_3;x_i) = \theta_1 + \theta_2 x_i + \theta_3 x_i^2 \qquad (1)$$

an die beobachteten Daten

i	x_i	$y_i \pm \sigma_i$
1	-0.6	5 ± 2
2	-0.2	3 ± 1
3	$+0.2$	5 ± 1
4	$+0.6$	8 ± 2

anzupassen.

Geben Sie auch Schätzwerte auf die Fehler $\sigma(\hat{\theta}_1)$ und $\sigma(\hat{\theta}_2)$ sowie einen Schätzwert für den Korrelationskoeffizienten ρ_{13} an, da die Parameter θ_1 und θ_3 korreliert sind.

Benutzen Sie für die Berechnung die Matrixschreibweise. Die Theorievorhersage wird dann geschrieben als

$$\vec{f} = A\vec{\theta}$$

mit

$$\vec{f} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix}, \qquad \vec{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$$

und einer von Ihnen zu berechnenden Matrix A.

Die unabhängigen Messungen werden dargestellt durch einen Vektor $\vec{y}$ und eine diagonale Kovarianzmatrix $V(\vec{y})$:

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}, \qquad V(\vec{y}) = \begin{pmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{pmatrix}.$$

Die Schätzwerte $\hat{\vec{\theta}}$ für die Parameter und die zugehörige Kovarianzmatrix $V(\hat{\vec{\theta}})$ ergeben sich dann zu

$$\hat{\vec{\theta}} = (A^T V^{-1} A)^{-1} A^T V^{-1} \vec{y} \quad \text{und} \quad V(\hat{\vec{\theta}}) = (A^T V^{-1} A)^{-1}$$

mit $V = V(\vec{y})$.

Hausaufgaben

Aufgabe 27 *„Maximum Likelihood" für die Normalverteilung* **4 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

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

benutzen.

Aufgabe 28 *„Maximum Likelihood" für die Poissonverteilung* **3 Punkte**

Betrachten Sie eine einmalige Messung einer Meßgröße n, die gemäß der Poisson-Wahrscheinlichkeitsdichtefunktion $f(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}$ verteilt sein soll.

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

(ii) Ist der Schätzer erwartungstreu?

(iii) Benutzen Sie die MVB-Methode, um die Varianz des Schätzers $V(\hat{\nu})$ auszurechnen.

Aufgabe 29 *Beweis der Rao-Cremér-Frechet-Ungleichung* **7 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

$$V[\hat{\theta}] \geq \frac{(1 + \frac{\partial b}{\partial \theta})^2}{-E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right]} \tag{3}$$

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

$$\left(\int u^2 \mathrm{d}x_1 \ldots \mathrm{d}x_n\right)\left(\int v^2 \mathrm{d}x_1 \ldots \mathrm{d}x_n\right) \geq \left(\int uv \mathrm{d}x_1 \ldots \mathrm{d}x_n\right)^2, \tag{4}$$

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 30 *Messung der Polarisationsasymmetrie* **6 Punkte**

Eine wichtige Messung der Paritätsverletzung in der elektroschwachen Wechselwirkung des Standardmodells wurde durchgeführt, indem die inelastische Streuung von polarisierten Elektronen an einem Deuteriumtarget vermessen wurde. Die Paritätsverletzung führt dabei zu einer Differenz im Streuwirkungsquerschnitt zwischen links- und rechts-händig polarisierten Elektronen. Die Meßgröße war dabei die Polarisationsasymmetrie α, die gegeben ist durch

$$\alpha = \frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}, \tag{5}$$

wobei σ_L (σ_R) die Wirkungsquerschnitte für links- (rechts)-händige Elektronen bei der inelastischen Streuung an einem Deuteriumtarget sind.

Für eine gegebene Luminosität L ist die Anzahl der beobachteten Streuereignisse bei links- bzw. rechthändigen Elektronen eine poissonverteilte Zufallsvariable, n_L bzw. n_R, mit Mittelwerten ν_L bzw. ν_R. Die Mittelwerte stehen mit den Wirkungsquerschnitten über $\nu_L = \sigma_L L$ bzw. $\nu_R = \sigma_R L$ in Beziehung, wobei die Luminosität L für die beiden Datensätze identisch ist.

(i) Benutzen Sie die Tatsache, dass die beobachtete Anzahl von Streuereignissen poissonverteilt ist, um einen „Maximum Likelihood"-Schätzer $\hat{\alpha}$ für die Asymmetrie α aufzustellen.

(ii) Benutzen Sie Fehlerfortpflanzung (und nehmen Sie dabei an, dass die Messungen der Streuereignisse für links- bzw. rechtshändige Elektronen unabhängig voneinander sind), um die Standardabweichung $\sigma_{\hat{\alpha}}$ auf diesen Schätzer als Funktion von α und $n_{TOT} = n_R + n_L$ zu ermitteln.

(iii) Bevor das Experiment aufgebaut wurde, wurde erwartet, dass die Asymmetrie ungefähr 10^{-4} betragen würde. Welche Gesamtzahl von Streuereignissen musste man dafür beobachten, um die Asymmetrie mit einer relativen Genauigkeit von 10% zu messen?