Prof. Dr. E. Cramer
Dipl.-Math. K. Herle

RWTH Aachen, SS 2010
15. Juli 2010

Einführung in die angewandte Stochastik

Lösung zur 10. Übung

Lösung Aufgabe 33

(a) Seien $x_1, \ldots, x_n$ Realisationen von $X_1, \ldots, X_n$ mit $x_i \geq 1$ für alle $i \in \{1, \ldots, n\}$ (sonst gilt $f_\alpha(x_i) = 0$). Dann ist die zugehörige Likelihood-Funktion aufgrund der stochastischen Unabhängigkeit der Zufallsvariablen $X_1, \ldots, X_n$ gegeben durch

$$L(\alpha|x_1, \ldots, x_n) = \prod_{i=1}^{n} f_\alpha(x_i) = \prod_{i=1}^{n} \frac{\alpha}{x_i^{\alpha+1}} = \alpha^n \prod_{i=1}^{n} x_i^{-(\alpha+1)}, \quad \alpha > 0.$$

Betrachte Maximierung der Log-Likelihood-Funktion:

$$\begin{aligned} l(\alpha) &= \ln(L(\alpha|x_1, \ldots, x_n)) \\ &= \ln(\alpha^n) + \ln(\prod_{i=1}^{n} x_i^{-(\alpha+1)}) \\ &= n\ln(\alpha) + \sum_{i=1}^{n}(-(\alpha+1))\ln(x_i) = n\ln(\alpha) - (\alpha+1)\sum_{i=1}^{n}\ln(x_i), \quad \alpha > 0. \end{aligned}$$

Die Funktion l ist differenzierbar auf $(0, \infty)$ mit

$$l'(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^{n} \ln(x_i), \quad \alpha > 0.$$

Also gilt

$$l'(\alpha) \begin{cases} > \\ = \\ < \end{cases} 0 \quad \Longleftrightarrow \alpha \begin{cases} < \\ = \\ > \end{cases} \frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

für $\alpha > 0$. Für

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(x_i)}$$

gilt daher: l ist streng monoton wachsend auf $(0, \hat{\alpha}]$ und streng monoton fallend auf $[\hat{\alpha}, \infty)$. Damit ist $\hat{\alpha}$ die (eindeutige) ML-Schätzung für α (bei geg. $x_1, \ldots, x_n$).

(b) Einsetzen der $n = 12$ gegebenen Datenwerte liefert den folgenden Schätzwert für α:

$$\hat{\alpha} = \frac{12}{\sum_{i=1}^{12} \ln(x_i)} \approx 0,903.$$

(c) Ersetzen von α durch den Schätzwert $\hat{\alpha} \approx 0.903$ liefert:

$$P_{\hat{\alpha}}(X_1 > 2) = 1 - F_{\hat{\alpha}}(2) \approx 1 - (1 - \frac{1}{2^{0,903}}) = \frac{1}{2^{0,903}} \approx 0,535.$$

Lösung Aufgabe 34

Zur Beschreibung der Situation betrachte stochastisch unabhängige Zufallsvariablen $X_1, \ldots, X_n$, $n = 1000$, mit der Interpretation

$$X_i = \begin{cases} 1, & i\text{-ter Passagier tritt Flug nicht an,} \\ 0, & i\text{-ter Passagier tritt Flug an.} \end{cases}$$

Dann gilt $X_i \sim bin(1, p), i = 1, \ldots, n$ mit $p \in (0, 1)$ unbekannt.

Gemäß Vorlesung ist in dieser Situation ein approximatives zweiseitiges $(1 - \alpha)$-Konfidenzintervall für p gegeben durch

$$K = \left[\hat{p} - \frac{u_{1-\alpha/2}}{\sqrt{n}}\sqrt{\hat{p}(1-\hat{p})}, \hat{p} + \frac{u_{1-\alpha/2}}{\sqrt{n}}\sqrt{\hat{p}(1-\hat{p})}\right].$$

Hier bezeichnet $u_{1-\alpha/2}$ das $(1 - \alpha/2)$-Quantil der Standardnormalverteilung $N(0, 1)$ und $\hat{p} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ die übliche erwartungstreue Punktschätzung für p.

Einsetzen der vorgegebenen Zahlenwerte führt zum gewünschten Konfidenzintervall: Es gilt

$$n = 1000, \quad \hat{p} = \bar{x} = \frac{74}{1000} = 0.074, \quad 1 - \alpha = 0.9.$$

Insbesondere folgt $\alpha = 0.1$ und damit $1 - \alpha/2 = 0.95$. Einer Tabelle mit Quantilen der Standardnormalverteilung lässt sich entnehmen:

$$u_{1-\alpha/2} = u_{0.95} \approx 1.64.$$

Damit ergibt sich für das gesuchte Konfidenzintervall:

$$\begin{aligned} K &\approx [0.074 - \frac{1.64}{\sqrt{1000}}\sqrt{0.074 \cdot 0.926}, 0.074 + \frac{1.64}{\sqrt{1000}}\sqrt{0.074 \cdot 0.926}] \\ &\approx [0.060, 0.088]. \end{aligned}$$

Lösung Aufgabe 35

(a) (i) Gemäß D 5.6 ist ein zweiseitiges Konfidenzintervall für μ zum Konfidenz-Niveau $1 - \alpha$ bei bekannter Varianz σ^2 gegeben durch:

$$K_1 = \left[\bar{x} - u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right],$$

wobei $u_{1-\alpha/2}$ das $(1 - \alpha/2)$-Quantil der Standardnormalverteilung $N(0, 1)$ bezeichnet. Einsetzen der vorgegebenen Zahlenwerte liefert:

$$\begin{aligned} &n = 12, \\ &1 - \alpha = 0.9 \Leftrightarrow \alpha = 0.1 \Leftrightarrow 1 - \alpha/2 = 0.95 \\ &\Rightarrow u_{1-\alpha/2} = u_{0.95} \approx 1.64, \\ &\sigma^2 = 0.0324 \\ &\Rightarrow \sigma = \sqrt{\sigma^2} = 0.18, \\ &\bar{x} = \frac{1}{12}\sum_{i=1}^{12} x_i \approx 3.528. \end{aligned}$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_1 \approx [3.528 - 1.64\frac{0.18}{\sqrt{12}}, 3.528 + 1.64\frac{0.18}{\sqrt{12}}] \approx [3.443, 3.613].$$

(ii) Ein zweiseitiges Konfidenzintervall für μ zum Konfidenz-Niveau $1-\alpha$ bei unbekannter Varianz σ^2 ist gemäß D 5.8 gegeben durch

$$K_2 = \left[\bar{x} - t_{1-\alpha/2}(n-1)\frac{\hat{\sigma}}{\sqrt{n}},\ \bar{x} + t_{1-\alpha/2}(n-1)\frac{\hat{\sigma}}{\sqrt{n}}\right]$$

Hierbei bezeichne $t_{1-\alpha/2}(n-1)$ das $(1-\alpha/2)$-Quantil zu $t(n-1)$ (t-Verteilung mit $n-1$ Freiheitsgraden) und $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i-\bar{x})^2 = \frac{1}{n-1}(\sum_{i=1}^n x_i^2 - n\bar{x}^2)$ die erwartungstreue Punktschätzung für σ^2.

Einsetzen der vorgegebenen Zahlenwerte liefert:

$$\begin{aligned}
&n = 12,\\
&1-\alpha/2 = 0.95\\
&\Rightarrow t_{1-\alpha/2}(n-1) = t_{0.95}(11) \stackrel{Tab.}{\approx} 1.796,\\
&\bar{x} \approx 3.528,\\
&\hat{\sigma} = \sqrt{\frac{1}{11}(\sum_{i=1}^{12}(x_i-\bar{x})^2} \approx 0.185.
\end{aligned}$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_2 \approx [3.528 - 1.796\frac{0.185}{\sqrt{12}}, 3.528 + 1.796\frac{0.185}{\sqrt{12}}] \approx [3.432, 3.623].$$

(b) (i) Gemäß D 5.9 ist ein einseitiges unteres Konfidenzintervall für σ^2 bei unbekanntem μ zum Konfidenz-Niveau $1-\alpha$ gegeben durch

$$K_3 = \left[0, \frac{n-1}{\chi^2_\alpha(n-1)}\hat{\sigma}^2\right]$$

wobei $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i-\bar{x})^2 = \frac{1}{n-1}(\sum_{i=1}^n x_i^2 - n\bar{x}^2)$ die erwartungstreue Punktschätzung für σ^2 bezeichne.

Einsetzen der vorgegebenen Zahlenwerte liefert:

$$\begin{aligned}
&n = 12,\\
&\alpha = 0.01 \Rightarrow \chi^2_\alpha(n-1) = \chi^2_{0.01}(11) \stackrel{Tab.}{\approx} 3.05,\\
&\hat{\sigma^2} \stackrel{\text{vgl. }(a)(ii)}{\approx} 0.0342.
\end{aligned}$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_3 = [0, \frac{n-1}{\chi^2_\alpha(n-1)}\hat{\sigma^2}] \approx [0, 0.123].$$

(ii) Gemäß D 5.11 erhält man ein einseitiges unteres Konfidenzintervall für σ zum Konfidenz-Niveau $1-\alpha = 99\%$ durch Ziehen der Quadratwurzel aus den Intervallgrenzen des Konfidenzintervalls aus (a):

$$K_4 = \left[\sqrt{0}, \sqrt{\frac{n-1}{\chi^2_\alpha(n-1)}\hat{\sigma}^2}\right] \stackrel{(i)}{\approx} [0, \sqrt{0.123}] \approx [0, 0.351].$$

(c) Versuchsplanung: Laut D 5.7 hat das zweiseitiges Konfidenzintervall für μ zum Konfidenz-Niveau $1-\alpha$ bei bekannter Varianz σ^2 höchstens die Länge $\mathfrak{L}_0$, falls

$$n \geq \frac{4\sigma^2 u_{1-\alpha/2}^2}{\mathfrak{L}_0^2}.$$

Einsetzen der vorgegebenen Zahlenwerte $\mathfrak{L}_0 = 0.1, \sigma^2 = 0.0324, u_{1-\alpha/2} = u_{0.95} \approx 1.64$ liefert:

$$n \geq \frac{4 \cdot 0.0324 \cdot 1.64^2}{0.01} \approx 34.86.$$

Also muss $n \geq 35$ gelten, damit die Länge des Konfidenzintervalls höchstens 0.1 beträgt.

Lösung Aufgabe 36

Gemäß Modell D 5.12 und Verfahren D 5.13 der Vorlesung ist ein zweiseitiges $(1-\alpha)$-Konfidenzintervall für die Differenz der Erwartungswerte $\Delta = \mu_1 - \mu_2$ bei unbekannter (gleicher) Varianz σ^2 gegeben durch:

$$K = \left[\hat{\Delta} - t_{1-\alpha/2}(n_1+n_2-2)\ \hat{\sigma}_{pool} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}, \hat{\Delta} + t_{1-\alpha/2}(n_1+n_2-2)\ \hat{\sigma}_{pool} \cdot \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right],$$

wobei $\hat{\Delta} = \bar{x} - \bar{y}$ eine Punktschätzung für die Differenz der Erwartungswerte Δ,

$$\hat{\sigma}_{pool}^2 = \frac{1}{n_1+n_2-2}\left(\sum_{i=1}^{n_1}(x_i-\bar{x})^2 + \sum_{j=1}^{n_2}(y_j-\bar{y})^2\right) = \frac{n_1-1}{n_1+n_2-2}\hat{\sigma}_1^2 + \frac{n_2-1}{n_1+n_2-2}\hat{\sigma}_2^2$$

eine kombinierte Varianzschätzung mit den Stichprobenvarianzen $\hat{\sigma}_1^2$ und $\hat{\sigma}_2^2$ der Stichproben $x_1, \ldots, x_{n_1}$ und $y_1, \ldots, y_{n_2}$ sowie $t_{1-\alpha/2}(n_1+n_2-2)$ das $(1-\alpha/2)$-Quantil der $t(n_1+n_2-2)$-Verteilung sind.

In der hier vorliegenden Situation gilt:

$$n_1 = 12, \quad n_2 = 8,$$
$$1-\alpha = 0.95 \quad \Leftrightarrow \quad 1-\frac{\alpha}{2} = 0.975 \quad \Rightarrow \quad t_{1-\alpha/2}(n_1+n_2-2) = t_{0.975}(18) \approx 2.101,$$
$$\bar{x} = 50.4, \quad \bar{y} = 50.6 \quad \Rightarrow \quad \hat{\Delta} = -0.2,$$
$$\hat{\sigma}_1^2 \approx 4.433, \quad \hat{\sigma}_2^2 \approx 3.326 \quad \Rightarrow \quad \hat{\sigma}_{pool} \approx \sqrt{\frac{11 \cdot 4.433 + 7 \cdot 3.326}{18}} \approx 2.0006.$$

Nach Einsetzen dieser Werte erhalten wir als zweiseitiges Konfidenzintervall für $\Delta = \mu_1 - \mu_2$ zum Konfidenzniveau 0.95:

$$K = \left[-0.2 - 2.101 \cdot 2.0006 \cdot \sqrt{\frac{1}{12}+\frac{1}{8}}, -0.2 + 2.101 \cdot 2.0006 \cdot \sqrt{\frac{1}{12}+\frac{1}{8}}\right] \approx [-2.027, 1.627].$$

Lösung Aufgabe 37

(a) Gemäß D 7.4 der Vorlesung sind die mittels der Kleinsten Quadrate-Methode hergeleiteten Schätzfunktionen für die Steigung b und den Achsenabschnitt a gegeben durch

$$\hat{b} = \frac{s_{xY}}{s_x^2} \approx \frac{-0.0361}{0.0136} \approx -2.654.$$

und

$$\hat{a} = \bar{Y} - \hat{b}\bar{x} \approx 6.1 + 2.654 \cdot 2.65 \approx 13.133.$$

Hieraus ergibt sich die geschätzte Regressionsgerade

$$\widehat{f(x)} = \hat{a} + \hat{b}x \approx 13.133 - 2.654x, \quad x \in \mathbb{R}.$$

Speziell für $x = 2.7$ (Eur/Kg) erhält man die geschätzte Absatzmenge

$$\widehat{f(2.7)} = 13.133 - 2.654 \cdot 2.7 \approx 5.967.$$

(b) Gemäß D 7.13 (iii) der Vorlesung ist ein zweiseitiges $(1-\alpha)$-Konfidenz-Intervall für σ^2 gegeben durch

$$K = \left[\frac{(n-2)\hat{\sigma}^2}{\chi^2_{1-\alpha/2}(n-2)}, \frac{(n-2)\hat{\sigma}^2}{\chi^2_{\alpha/2}(n-2)} \right],$$

wobei $\chi^2_\beta(n-2)$ als β-Quantil zu $\chi^2(n-2)$ und (vgl. D 7.7 der Vorlesung)

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{i=1}^{n}(Y_i - \hat{a} - \hat{b}x_i)^2 = \frac{n}{n-2}s_Y^2(1 - r_{xY}^2), \quad r_{xY}^2 = \frac{s_{xY}^2}{s_x^2 s_Y^2}.$$

Zahlenwerte:

$$n = 7, \quad \hat{\sigma}^2 = \frac{7}{5}s_Y^2\left(1 - \frac{s_{xY}^2}{s_x^2 s_Y^2}\right) \approx \frac{7}{5}0.1657\left(1 - \frac{(-0.0361)^2}{0.0136 \cdot 0.1657}\right) \approx 0.098,$$

$$1 - \alpha = 0.9 \quad \Rightarrow \quad \chi^2_{1-\alpha/2}(n-2) = \chi^2_{0.95}(5) \approx 11.07,\ \chi^2_{\alpha/2}(n-2) = \chi^2_{0.05}(5) \approx 1.15.$$

Damit:

$$K \approx \left[\frac{5 \cdot 0.098}{11.07}, \frac{5 \cdot 0.098}{1.15} \right] \approx [0.044, 0.426].$$