Eingangsprüfung Stochastik, 17.10.2008

Wir gehen stets von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ aus.

Aufgabe 1 (12 Punkte)
Sei $X : \Omega \longrightarrow [0, \infty)$ eine Zufallsvariable. Zeigen Sie

$$\int X \, d\mathbb{P} \geq 0.$$

Hinweis: Beweisen Sie diese Ungleichung zuerst für Treppenfunktionen.
Lösung
Sei $X = \sum_{i=1}^{n} \alpha_i \, 1_{A_i}$, $\alpha_i \geq 0$, $A_i \in \mathcal{A}$, $i = 1, \ldots, n$ eine Treppenfunktion. Dann gilt

$$\int X \, d\mathbb{P} = \sum_{i=1}^{n} \alpha_i \, \mathbb{P}(A_i)$$

und damit $\int X \, d\mathbb{P} \geq 0$.
Sei nun $X \geq 0$ beliebig. Dann gibt es eine monoton wachsende Folge $(X_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit $X_n \geq 0$ und $\lim_{n \to \infty} X_n = X$. Es gilt mit dem Satz von der monotonen Konvergenz

$$\int X \, d\mathbb{P} = \lim_{n \to \infty} \int X_n \, d\mathbb{P},$$

und aus dem vorher Gezeigten folgt nun $\int X \, d\mathbb{P} \geq 0$.
Aufgabe 2 (22 Punkte)

(a) Seien $X_1, \ldots, X_n$ mit $n \geq 2$ unabhängig und identisch verteilte Zufallsvariablen mit $X_i \sim \mathrm{Exp}(\lambda)$ und sei $G := \sum_{i=1}^{n} X_i$. Zeigen Sie, dass

$$\mathbb{E}\left(\frac{1}{G}\right) = \frac{\lambda}{n-1}$$

gilt.
Hinweis: Es gilt $G \sim \Gamma(n, \lambda)$. Für die Gamma-Funktion $\Gamma : (0, \infty) \to (0, \infty)$ mit $\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}$ gilt die Identität $\Gamma(x+1) = x\Gamma(x)$.

(b) Gegeben seien unabhängige Realisierungen $x_1, \ldots, x_n$ einer $\mathrm{Exp}(\lambda)$-verteilten Zufallsvariablen X, mit unbekanntem $\lambda > 0$. Zeigen Sie, dass $\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$ ein Maximum Likelihood Schätzwert zu $x_1, \ldots, x_n$ für λ ist. Welcher Maximum Likelihood Schätzer für λ ergibt sich?

(c) Prüfen Sie, ob der Schätzer aus (b) erwartungstreu ist, und geben Sie gegebenenfalls eine erwartungstreue Modifikation an.

Lösung
Zu (a)
Es gilt $G \sim \Gamma(n, \lambda)$ und damit

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{G}\right) &= \int_0^\infty \frac{1}{x} \cdot \frac{\lambda^n}{\Gamma(n)}\, x^{n-1} e^{-\lambda x}\, dx \\
&= \frac{\lambda^n}{\Gamma(n)} \int_0^\infty x^{(n-1)-1} e^{-\lambda x}\, dx \\
&= \frac{\lambda\, \Gamma(n-1)}{\Gamma(n)} \int_0^\infty \frac{\lambda^{n-1}}{\Gamma(n-1)} x^{(n-1)-1} e^{-\lambda x}\, dx \\
&= \frac{\lambda}{n-1}.
\end{aligned}$$

Zu (b)
Sei $x_1, \ldots, x_n$ eine Stichprobe. Aufgrund der Unabhängigkeit der X_i, $i = 1, \ldots, n$ gilt für die Likelihood

$$L(\lambda; x_1, \ldots, x_n) = \lambda e^{-\lambda x_1} \cdot \ldots \cdot \lambda e^{-\lambda x_n} = \lambda^n \;\exp\left(-\lambda \sum_{i=1}^n x_i\right).$$

Für die Loglikelihood-Funktion ℓ mit $\ell(\lambda) := \ln(L(\lambda))$ ergibt sich

$$\begin{aligned}
\ell(\lambda) &:= \ln L(\lambda) = \ln(\lambda^n) - \lambda \sum_{i=1}^n x_i = n \ln(\lambda) - \lambda \sum_{i=1}^n x_i \\
\ell'(\lambda) &= \frac{n}{\lambda} - \sum_{i=1}^n x_i \\
\ell''(\lambda) &= -\frac{n}{\lambda^2} < 0
\end{aligned}$$

und somit liegt in $\frac{n}{\sum_{i=1}^n x_i}$ ein Maximum der Likelihoodfunktion vor. Damit ist die Zufallsvariable

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^n X_i}$$

ein Maximum Likelihood Schätzer.
Zu (c)
Aus (a) und (b) folgt

$$\mathbb{E}\left(\frac{1}{\sum_{i=1}^n X_i}\right) = \frac{\lambda}{n-1} \Longrightarrow \mathbb{E}(\hat{\lambda}) = n \cdot \frac{\lambda}{n-1}.$$

Somit ist $\hat{\lambda}$ nicht erwartungstreu. Dagegen ist offenbar $\hat{\hat{\lambda}} = \frac{n-1}{\sum_{i=1}^n X_i}$ ein erwartungstreuer Schätzer.

Aufgabe 3 (18 Punkte)
Gegeben seien zwei Zufallsvariablen X und Λ. Dabei sei $\Lambda \sim U[1,2]$ also gleichverteilt auf dem Intervall [1,2] und bei gegebenem $\Lambda = \lambda$ sei $X \sim \text{Exp}(\lambda)$-verteilt.

(a) Bestimmen Sie $\mathbb{E}(X|\Lambda)$ und $\text{Var}(X|\Lambda)$.

(b) Bestimmen Sie $\mathbb{E}(X)$ und $\text{Var}(X)$.

(c) Bestimmen Sie $\mathbb{P}(X \geq 1)$.

Lösung
Zu (a)
Bei gegebenem $\lambda \in [1,2]$ gilt $\mathbb{E}(X|\Lambda = \lambda) = \dfrac{1}{\lambda}$ und $\text{Var}(X|\Lambda = \lambda) = \dfrac{1}{\lambda^2}$, also folgt $\mathbb{E}(X|\Lambda) = \dfrac{1}{\Lambda}$ und $\text{Var}(X) = \dfrac{1}{\Lambda^2}$.
Zu (b)
Mit dem Satz vom iterierten Erwartungswert gilt

$$\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X|\Lambda)) = \mathbb{E}\left(\frac{1}{\Lambda}\right) = \int_1^2 \frac{1}{\lambda} d\lambda = \ln 2 - \ln 1 = \ln 2.$$

Für die Varianz erhalten wir aus der Varianz-Zerlegung

$$\begin{aligned}
\text{Var}(X) &= \text{Var}(\mathbb{E}(X|\Lambda)) + \mathbb{E}(\text{Var}(X|\Lambda)) \\
&= \text{Var}\left(\frac{1}{\Lambda}\right) + \mathbb{E}\left(\frac{1}{\Lambda^2}\right) \\
&= \mathbb{E}\left(\frac{1}{\Lambda^2}\right) - \mathbb{E}\left(\frac{1}{\Lambda}\right)^2 + \mathbb{E}\left(\frac{1}{\Lambda^2}\right) \\
&= 2\mathbb{E}\left(\frac{1}{\Lambda^2}\right) - \mathbb{E}\left(\frac{1}{\Lambda}\right)^2.
\end{aligned}$$

Für den ersten Term ergibt sich

$$\mathbb{E}\left(\frac{1}{\Lambda^2}\right) = \int_1^2 \frac{1}{\lambda^2} d\lambda = -\frac{1}{\lambda}\bigg|_1^2 = -\frac{1}{2} + 1 = \frac{1}{2}.$$

Insgesamt folgt also

$$\text{Var}(X) = 2 \cdot \frac{1}{2} - (\ln 2)^2 = 1 - (\ln 2)^2.$$

Zu (c)
Es gilt

$$\mathbb{P}(X \geq 1) = \int_1^2 \mathbb{P}(X \geq 1|\Lambda = \lambda) f_\Lambda(\lambda) d\lambda = \int_1^2 e^{-\lambda}\, d\lambda = -e^{-\lambda}\bigg|_1^2 = e^{-1} - e^{-2}.$$

Aufgabe 4 (38 Punkte) Herr K. fährt täglich um 08:30 in die Arbeit, Arbeitsbeginn ist spätestens 09:15. Die Reisezeit T in Minuten sei normalverteilt.

(a) Es gelte $T \sim N(40, 16)$.

(i) Bestimmen Sie die Wahrscheinlichkeit, dass er zwischen 09:10 und 09:20 ankommt.

(ii) Wie groß ist die Wahrscheinlichkeit des Zuspätkommens?

(iii) An wievielen Arbeitstagen im Jahr (200 Tage pro Jahr) wird er im Mittel zu spät kommen?

(iv) Welchen Abfahrtszeitpunkt (minutengenau) schlagen Sie vor, damit K. bei unveränderter Reisezeit T im Mittel an höchstens 2 Tagen im Jahr zu spät kommt?

(b) Herr K. will weiterhin erst um 08:30 starten und probiert eine neue Route. Die Reisezeit sei weiter normalverteilt. An 30 Arbeitstagen kommt er im Mittel um 09:05 an, die empirische Varianz beträgt 12,25.

(i) Herr K. vermutet, dass seine Reisezeit sich verkürzt hat. Überprüfen Sie anhand einer geeigneten Nullhypothese seine Vermutung zum Niveau von 5 %.

(ii) Geben Sie für den Erwartungswert der Reisezeit ein Schätzintervall zum Niveau 5 % an.

(c) Diskutieren Sie die Annahme der Normalverteilung.

Lösung
Zu (a)
(i) Ankunftszeit liegt zwischen 09:10 und 09:20 genau dann wenn $T \in [40, 50]$ gilt.

$$\begin{aligned}\mathbb{P}(40 \leq T \leq 50) &= \mathbb{P}(T \leq 50) - \mathbb{P}(T \leq 40) = \Phi\left(\frac{50-40}{4}\right) - \Phi\left(\frac{40-40}{4}\right) \\ &= 0,4938.\end{aligned}$$

(ii) Herr K. verspätet sich genau dann wenn $T > 45$ gilt.

$$\mathbb{P}(45 > T) = 1 - \mathbb{P}(T \leq 45) = 1 - \Phi\left(\frac{5}{4}\right) = 0,1056.$$

(iii) Bezeichnet man mit N die Anzahl der Tage, an denen sich K. verspätet, und mit $p := 0,1056$, dann ist laut (ii) $N \sim B(200, p)$-verteilt und damit

$$\mathbb{E}(N) = 200p = 21,12.$$

Im Mittel kommt Herr K. an 21,12 Tagen im Jahr zu spät.
(iv) Die Wahrscheinlichkeit q, dass Herr K. sich verspätet, soll so bestimmt werden, dass $200q \leq 2$ gilt, also $q \leq 0,01$. Sei t_0 die Reisezeit, die mit einer Wahrscheinlichkeit von maximal $0,01$ überschritten wird, also

$$\mathbb{P}(T \geq t_0) = 0,01.$$

Wegen

$$\mathbb{P}(T \geq t_0) = 1 - \mathbb{P}(T \leq t_0) = 1 - \Phi\left(\frac{t_0 - 40}{4}\right)$$

ergibt sich zunächst

$$\Phi\left(\frac{t_0 - 40}{4}\right) = 0,99$$

und damit

$$t_0 = 40 + 4u_{0,99} = 40 + 4 \cdot 2,34 = 49,36 \approx 50.$$

Herr K. sollte also um 08:25 starten.
Zu (b)
(i) Zu testen ist die Nullhypothese $H_0 : \mu \leq \mu_0$ mit $\mu_0 := 40$. Es handelt sich um einen einseitigen t-Test, Testgröße ist

$$T = \frac{\overline{x} - \mu_0}{s}\sqrt{n}$$

und besitzt bei den vorliegenden Daten den Wert

$$\frac{-5}{3,5}\sqrt{30} = -7,82.$$

Die Hypothese wird im Fall $T > t_{29;0,95} = 1,699$ abgelehnt. Also wird H_0 nicht verworfen.
(ii) Das Schätzintervall lautet mit $t_{29;0,975} = 2,045$, $s = \sqrt{12,25} = 3,5$:

$$\left[\overline{x} - \frac{s \cdot t_{29;0,975}}{\sqrt{30}}, \overline{x} + \frac{s \cdot t_{29;0,975}}{\sqrt{30}}\right] = [33,69; 36,31]\,.$$

Zu (c):
Unter der Annahme der Normalverteilung gilt $\mathbb{P}(T < 0) > 0$. Die Annahme der Normalverteilung kann daher nur näherungsweise erfüllt sein.