Stochastische Eingangsprüfung, 17.05.2008

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

Aufgabe 1 (20 Punkte)
Sei $X : \Omega \longrightarrow [0, \infty)$ eine integrierbare Zufallsvariable mit $\int X d\mathbb{P} = 1$. Sei $Q : \mathcal{A} \longrightarrow \mathbb{R}$,

$$Q(A) := \int 1_A X d\mathbb{P}.$$

Sei $Y \geq 0$ eine Zufallsvariable. Beweisen Sie

$$\int Y \, dQ = \int XY d\mathbb{P}.$$

Hinweis: Beweisen Sie diese Gleichung erst für Treppenfunktionen.

Lösung
Sei $Y = \sum_{i=1}^{n} \alpha_i \cdot 1_{A_i}$, $\alpha_i \geq 0$, $A_i \in \mathcal{A}$, $i = 1, \ldots, n$ eine Treppenfunktion. Dann gilt laut Konstruktion des Integrals bezüglich Q

$$\int Y \, dQ = \sum_{i=1}^{n} \alpha_i Q(A_i) = \sum_{i=1}^{n} \alpha_i \int 1_{A_i} X d\mathbb{P} = \int \left(\sum_{i=1}^{n} \alpha_i \cdot 1_{A_i} X \right) d\mathbb{P} = \int YX \, d\mathbb{P}.$$

Sei nun $Y \geq 0$. Es gibt eine Folge $(Y_n)_{n \in \mathbb{N}}$ von Treppenfunktionen mit $Y_n \geq 0$ f.s. und $Y_n \nearrow Y$ f.s. Aus der Konstruktion des Integrals folgt

$$\int Y \, dQ = \lim_{n \to \infty} \int Y_n \, dQ.$$

Da Y_n eine Treppenfunktion ist, gilt wie oben gezeigt

$$\int Y_n \, dQ = \int Y_n X \, d\mathbb{P}.$$

Aus $Y_n \nearrow Y$ f.s. folgt auch $Y_n X \nearrow YX$ f.s. Mit dem Satz von der monotonen Konvergenz folgt

$$\lim_{n \to \infty} \int Y_n X \, d\mathbb{P} = \int YX \, d\mathbb{P}.$$

Insgesamt haben wir also gezeigt:

$$\int Y \, dQ = \lim_{n \to \infty} \int Y_n \, dQ = \lim_{n \to \infty} \int Y_n X \, d\mathbb{P} = \int YX \, d\mathbb{P}.$$

Aufgabe 2 (18 Punkte)
Eine Arbeitsunfähigkeitsversicherung übernimmt die aufgrund von Arbeitsunfähigkeit entstehende Finanzierungslücke in den Zahlungsverpflichtungen aus einer Kreditfinanzierung. Für eine solche Versicherung wird die Dauer der Arbeitsunfähigkeit T in Jahren als Zufallsvariable mit Dichte

$$f(x) = \begin{cases} 2\beta^2 x e^{-(\beta x)^2} & x > 0 \\ 0 & \text{sonst} \end{cases} \tag{1}$$

modelliert, wobei $\beta > 0$ ist.

(a) Beweisen Sie, dass f eine Dichte ist.

(b) Weisen Sie nach, dass

$$\mathbb{E}(T) = \frac{\sqrt{\pi}}{2\beta} \tag{2}$$

und bestimmen Sie die Varianz von T.

Hinweis: $\displaystyle\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$

(c) Es wird eine Karenzzeit von 0,25 Jahren eingeführt, d.h. die Versicherung leistet erst wenn die Dauer der Arbeitsunfähigkeit 0,25 Jahre überschreitet. Die Leistungdauer der Versicherung sei $\tilde{T}$.

(i) Bestimmen Sie die Verteilungsfunktion von $\tilde{T}$.

(ii) Besitzt die Zufallsvariable $\tilde{T}$ eine Dichte? Begründen Sie Ihre Antwort.

Lösung
Zu (a)
f ist stetig und folglich messbar, und es gilt $f \geq 0$. Zu zeigen ist

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Für $a > 0$ ergibt sich

$$\int_0^a f(x)\,dx = \int_0^a 2\beta^2 x e^{-(\beta x)^2}\,dx = e^{-(\beta x)^2}\Big|_0^a = 1 - e^{-\beta^2 a^2}.$$

Mit $a \to \infty$ folgt

$$\int_{\mathbb{R}} f(x)\,d\lambda^1 = \int_0^\infty f(x)\,dx = 1.$$

Zu (b)
Die Verteilungsfunktion F ist wegen obiger Rechnung gegeben durch

$$F(x) = \begin{cases} 1 - e^{-\beta^2 x^2} & x > 0 \\ 0 & \text{sonst.} \end{cases}$$

Somit folgt

$$\mathbb{E}(T) \quad = \quad \int_0^\infty (1 - F(x))\,dx = \int_0^\infty e^{-\beta^2 x^2}\,dx = \frac{1}{\beta}\int_0^\infty e^{-y^2}\,dy = \frac{\sqrt{\pi}}{2\beta}.$$

Ferner erhält man mittels partieller Integration für $a > 0$

$$\begin{aligned}
\int_0^a 2\beta^2 x^3 e^{-(\beta x)^2}\,dx &= \int_0^a x^2 \left(2x\beta^2 e^{-(\beta x)^2}\right) dx \\
&= -x^2 e^{-(\beta x)^2}\Big|_0^a + \int_0^a 2xe^{-(\beta x)^2} dx \\
&= -a^2 e^{-(\beta a)^2} - \frac{1}{\beta^2} e^{-(\beta x)^2}\Big|_0^a \\
&= -a^2 e^{-(\beta a)^2} - \frac{1}{\beta^2} e^{-(\beta a)^2} + \frac{1}{\beta^2}
\end{aligned}$$

Der Grenzübergang $a \to \infty$ führt zu

$$\begin{aligned}\mathbb{E}(T^2) &= \int_0^\infty 2\beta^2 x^3 e^{-(\beta x)^2}\, dx = \frac{1}{\beta^2}\\ \mathrm{Var}(T) &= \frac{1}{\beta^2} - \frac{\pi}{4\beta^2} = \frac{4-\pi}{4\beta^2}.\end{aligned}$$

Zu (c)
(i) Sei F die Verteilungsfunktion von T. Es gilt

$$\tilde{T} = \max(0; T - 0,25).$$

Somit gilt $\mathbb{P}(\tilde{T} < 0) = 0$, $\mathbb{P}(\tilde{T} = 0) = \mathbb{P}(T \leq 0,25) = e^{-0,25^2\beta^2}$. Für $t > 0$ erhalten wir

$$\begin{aligned}\mathbb{P}(\tilde{T} \leq t) &= \mathbb{P}(\tilde{T} = 0) + \mathbb{P}(0 < \tilde{T} \leq t) = \mathbb{P}(\tilde{T} = 0) + \mathbb{P}(0 < T - 0,25 \leq t)\\ &= \mathbb{P}(\tilde{T} = 0) + \mathbb{P}(0,25 < T \leq t + 0,25)\\ &= \mathbb{P}(\tilde{T} = 0) + \mathbb{P}(T \leq t + 0,25) - \mathbb{P}(T \leq 0,25)\\ &= \mathbb{P}(T \leq t + 0,25) = 1 - e^{-\beta^2(t+0,25)^2}.\end{aligned}$$

Zusammengefasst ergibt sich

$$\mathbb{P}(\tilde{T} \leq t) = \begin{cases} 0 & t < 0 \\ 1 - \exp\left(-(\beta(t+0,25)^2\right) & t \geq 0 \end{cases}.$$

(ii) Wegen $\mathbb{P}(\tilde{T} = 0) > 0$ besitzt $\tilde{T}$ keine Dichte.

Aufgabe 3 (18 Punkte)
Wir betrachten weiter die Modellierung der Dauer der Arbeitsunfähigkeit T mittels (1) der vorigen Aufgabe.
Gegeben seien unabhängige Beobachtungen $t_1, \ldots, t_n$ der Dauer, wobei $n = 1000$, $\sum_{i=1}^{1000} t_i = 1773$, $\sum_{i=1}^{1000} t_i^2 = 3969$

(a) Bestimmen Sie einen Maximum-Likelihood Schätzer und den Schätzwert für β zu den gegebenen Beobachtungen.

(b) Bestimmen Sie mittels (2) aus den Daten einen Schätzwert für β mit der Momentenmethode.

Lösung Zu (a) Die gemeinsame Dichte der unabhängigen Zufallsvariablen $T_i \sim T$, $i = 1, \ldots, n$ führt zu einer Likelihood L und ihrer Loglikelihood ℓ:

$$\begin{aligned}L(\beta) &= \prod_{i=1}^n 2\beta^2 t_i e^{-\beta^2 t_i^2} = 2^n \beta^{2n} e^{-\beta^2 \sum\limits_{i=1}^n t_i^2} \prod_{i=1}^n t_i\\ \ell(\beta) &= \ln(2^n) + 2n \ln\beta - \beta^2 \sum_{i=1}^n t_i^2 + \sum_{i=1}^n \ln(t_i)\\ \ell'(\beta) &= -2\beta \sum_{i=1}^n t_i^2 + \frac{2n}{\beta}\\ \ell''(\beta) &= -2\sum_{i=1}^n t_i^2 - \frac{2n}{\beta^2} < 0\end{aligned}$$

Es ergibt sich für $\hat{\beta} = \sqrt{\frac{n}{\sum_{i=1}^{n} t_i^2}}$ das Maximum von ℓ und somit der ML-Schätzer $\frac{\sqrt{n}}{\sqrt{\sum_{i=1}^{n} T_i^2}}$ für β und der Schätzwert

$$\hat{\beta} = \sqrt{\frac{1000}{3696}} = 0,5.$$

Zu (b)
Schätzer des Erwartungswerts ist der Mittelwert der Stichprobe

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i = \frac{1773}{1000} = 1,773.$$

Mit der Momentenmethode ergibt sich

$$\bar{t} = \frac{\sqrt{\pi}}{2\hat{\beta}} \Longrightarrow \hat{\beta} = \frac{\sqrt{\pi}}{2 \cdot 1,773} = 0,5.$$

Aufgabe 4 (11 Punkte)
Wir betrachten weiter die Modellierung der Dauer der Arbeitsunfähigkeit T mittels (1), wobei der Parameter β in (1) nicht mehr fest sondern zufallsabhängig sei.
Bei jedem Versicherungsfall wird angenommen, dass ein anderer Parameter β vorliegt (das wäre beispielsweise der Fall, wenn das Kollektiv nicht homogen ist). Der Parameter β wird nun als Ergebnis einer Zufallsvariablen B modelliert, die $\Gamma(2, \lambda)$-verteilt ist mit $\lambda > 0$ (die Dichte finden Sie in der Formelsammlung).

(a) Bestimmen Sie $\mathbb{E}(T|B)$.

(b) Bestimmen Sie $\mathbb{E}(T)$.

(c) Geben Sie die gemeinsame Dichte von (T, B) an.

Lösung
Zu (a)
Laut (2) gilt $\mathbb{E}(T|B) = \frac{\sqrt{\pi}}{2B}$.
Zu (b)
Die Dichte f_B von B ist gegeben durch

$$f_B(b) = \lambda^2 b e^{-\lambda b} 1_{(0,\infty)}(b).$$

$$\begin{aligned} \mathbb{E}(T) &= \mathbb{E}(\mathbb{E}(T|B)) = \mathbb{E}\left(\frac{\sqrt{\pi}}{2B}\right) = \frac{\sqrt{\pi}}{2}\int_0^\infty \frac{1}{b}\lambda^2 b e^{-\lambda b}\, db \\ &= \frac{\lambda^2\sqrt{\pi}}{2}\int_0^\infty e^{-\lambda b}\, db = \frac{\lambda\sqrt{\pi}}{2}. \end{aligned}$$

Zu (c)
Die gemeinsame Dichte g ist für $(t, b) \in (0, \infty)^2$ gegeben durch

$$g(t, b) = g(t|b) f_B(b) \stackrel{(1)}{=} 2b^2 t e^{-(bt)^2} \lambda^2 b e^{-\lambda b} = 2\lambda^2 b^3 t e^{-(bt)^2 - \lambda b},$$

und andernfalls ist sie gleich 0.

Aufgabe 5 (23 Punkte)

Es wurden bei je drei Personen der Alter 30,35,40,45,50 die Dauer der Arbeitsunfähigkeit untersucht. Folgende Daten wurden erhoben:

Nr.	Alter	Krankheitsdauer y_i
1	30	5,04
2	30	4,66
3	30	3,48
4	35	3,55
5	35	1,86
6	35	4,48
7	40	4,25
8	40	3,93
9	40	5,03
10	45	4,82
11	45	4,79
12	45	4,84
13	50	5,70
14	50	5,39
15	50	4,80

(a) Bezeichnet man mit x_i das Alter und y_i die Krankheitsdauer, so wird das Modell

$$Y_i = a + bx_i + \varepsilon_i,\ i = 1, \dots, 15$$

untersucht. Die ε_i seien unabhängig und $N(0, \sigma^2)$-verteilt. Testen Sie zum Niveau $\alpha = 1$ % die Hypothese $b = 0$. Interpretieren Sie das Ergebnis. Sie können verwenden:

$$\begin{aligned}
&\overline{x} = 40,\ \sum_{i=1}^{15}(x_i - \overline{x})^2 = 750 \\
&\overline{y} = 4,44,\ \sum_{i=1}^{15}(y_i - \overline{y})^2 = 12,47,\ \sum_{i=1}^{15}(y_i - \overline{y})(x_i - \overline{x}) = 49,9.
\end{aligned} \tag{3}$$

(b) Es wird nun angenommen, dass für die Krankheitsdauer $Y \sim N(\mu, \sigma^2)$ gilt. Geben Sie Konfidenzintervalle zum Niveau $1 - \alpha = 99\%$ für μ und σ^2 an. Verwenden Sie dabei die Ergebnisse aus (3).

Lösung

Zu (a)

Es ergeben sich folgende Schätzwerte:

$$\begin{aligned}
\hat{b} &= \frac{49,9}{750} = 0,0665 \\
\hat{a} &= \overline{y} - \hat{b}\overline{x} = 4,44 - 0,0665 \cdot 40 = 1,78 \\
\hat{\sigma}^2 &= \frac{1}{13}\left(12,47 - \frac{49,9^2}{750}\right) = 0,7038 \\
\mathrm{se}(\hat{b}) &= \sqrt{\frac{0,7038}{750}} = 0,0306 \\
T &= \frac{\hat{b} - b_0}{\mathrm{se}(\hat{b})} = \frac{0,0665}{0,0306} = 2,173.
\end{aligned}$$

Die Hypothese $b = 0$ wird nicht abgelehnt, denn $|T| < t_{13;0,995} = 3,012$. Somit wird die Hypothese, dass die Dauer der Arbeitsunfähigkeit nicht vom Alter abhängt, angenommen.
Zu (b)
Es handelt sich um eine normalverteilte Zufallsvariable mit unbekannter Varianz.

$$\begin{aligned}
\hat{\mu} &= 4,44 \\
\hat{\sigma}^2 &= \frac{12,47}{14} = 0,8907.
\end{aligned}$$

Es ergeben sich mit den Quantilen

$$t_{14;0,995} = 2,977,\ \chi^2_{14;0,005} = 4,075,\ \chi^2_{14;0,995} = 31,319$$

für μ bzw. σ^2 die Konfidenzintervalle

$$\begin{aligned}
\left[\hat{\mu} - \hat{\sigma}\frac{t_{14;0,995}}{\sqrt{15}}, \hat{\mu} + \hat{\sigma}\frac{t_{14;0,995}}{\sqrt{15}}\right] &= \left[4,44 - 0,944\frac{2,977}{\sqrt{15}}, 4,44 + 0,944\frac{2,977}{\sqrt{15}}\right] \\
&= [3,71; 5,17] \text{ bzw.} \\
\left[\frac{14 \cdot \hat{\sigma}^2}{\chi^2_{14;0,995}}; \frac{14 \cdot \hat{\sigma}^2}{\chi^2_{14;0,005}}\right] &= [0,398; 3,060]\,.
\end{aligned}$$