Wir gehen stets von einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathbb{P})$ aus. $\mathcal{B}^1$ sei die σ-Algebra der Borelmengen von $\mathbb{R}$. λ^1 sei das Lebesgue-Borel-Maß auf $(\mathbb{R}, \mathcal{B}^1)$.

Aufgabe 1 (40 Punkte)

Sei $X : \Omega \longrightarrow \mathbb{R}$ eine Zufallsvariable mit λ^1-Dichte $f : \mathbb{R} \longrightarrow [0, \infty)$,

$$f(x) = \begin{cases} ae^{-5x} & x > 0 \\ 0 & x \leq 0. \end{cases}$$

(a) Man bestimme a so, dass f eine Dichte ist, und bestimmen Sie die Verteilungsfunktion F von X.

(b) Sei $b > 0$ und $Y := \min(X, b)$. Man bestimme die Verteilungsfunktion F_Y von Y und $\mathbb{E}(Y)$.

(c) Sei Y wie in (b) und μ_Y das von F_Y erzeugte Maß auf $(\mathbb{R}, \mathcal{B}^1)$. Man beweise

$$\mu_Y = \mathbb{P}(X \geq b)\delta_b + f \cdot 1_{[0,b)} \cdot \lambda^1. \tag{0.1}$$

(δ_b sei das Dirac-Maß auf $(\mathbb{R}, \mathcal{B}^1)$).

(d) Seien μ_1, μ_2 Maße auf $(\mathbb{R}, \mathcal{B}^1)$. Beweisen Sie:

(i) $\mu := \mu_1 + \mu_2$ ist ein Maß.

(ii) Ist $g : \mathbb{R} \longrightarrow [0, \infty)$ messbar (d.h. $\mathcal{B}^1 - \mathcal{B}^1$-messbar), dann gilt $\int g \, d\mu = \int g \, d\mu_1 + \int g \, d\mu_2$. (Hinweis: man gehe auf die Konstruktion des Integrals zurück).

(e) Für $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) = e^x$ bestimme man $\int g \, d\mu_Y$ mit μ_Y aus (0.1).

Aufgabe 2 (10 Punkte)

Gegeben seien zwei Zufallsvariablen Y und N. Dabei sei $Y \sim \text{Exp}(1)$-verteilt und bei gegebenem $Y = y$ sei $N \sim \text{Poi}(y)$-verteilt.

(a) Bestimmen Sie $\mathbb{E}(N)$

(b) Bestimmen Sie die Wahrscheinlichkeitsfunktion von N. (Hinweis: $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} \, dt$)

Aufgabe 3 (36 Punkte)

Ein Versicherungsbestand wurde über fünf Jahre beobachtet. Jeder aufgetretene Schaden wird je nach Schadenhöhe einem von vier Schadenintervallen A, B, C, D zugewiesen. Folgende Schadenzahlen wurden erhoben:

	Jahr				
Schadenhöhe	1	2	3	4	5
A	12	10	15	6	10
B	15	19	8	11	8
C	15	12	12	11	13
D	10	10	17	12	9

Wir gehen davon aus, dass sich der Bestand in diesen fünf Jahren nicht verändert hat.

(a) Zunächst soll die Hypothese geprüft werden, ob Schadeneintrittszeitpunkt und Schadenhöhe unabhängig voneinander sind. Für die Testgröße des χ^2-Unabhängigkeitstests ergibt sich der Wert $11,82$. Wird die Hypothese zum Niveau 10 % verworfen? Begründen Sie Ihre Antwort.

(b) Vorbereitung

(i) Seien $n_1, \ldots, n_k$ unabhängige Realisierungen einer Poisson-verteilten Zufallsvariablen. Bestimmen Sie den zugehörigen Maximum Likelihood Schätzer für den Parameter der Poisson-Verteilung.

(ii) Ist dieser Schätzer erwartungstreu? Begründen Sie Ihre Antwort.

(c) In der Folge sei N die Anzahl aller Schäden, die pro Jahr im Bestand insgesamt eintreten. Es wird angenommen, dass $N \sim \text{Poi}(\lambda)$ gilt. Man bestimme zu obigen Daten den Maximum Likelihood Schätzwert $\hat{\lambda}$ für λ.

(d) Es sei weiter wie in (c) N die Anzahl aller Schäden pro Jahr. In der Folge nehmen wir an, dass N annähernd normalverteilt ist mit Erwartungswert μ und Varianz σ^2.

(i) Testen Sie zum Niveau 5 % die Hypothese $\mu = 50$.

(ii) Bestimmen Sie ein ein Konfidenzintervall für μ zum Niveau 95 %.

Lösungen

Aufgabe 1
(a)

$$1 = \int_{\mathbb{R}} f(x)\,dx = a\int_0^\infty e^{-5x}\,dx = \frac{a}{5} \Longrightarrow a = 5$$

Für $x \in \mathbb{R}$ gilt

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

Somit gilt $F(x) = 0$ für $x \le 0$ und für $x > 0$

$$F(x) = 5\int_0^x e^{-5t}\,dt = 1 - e^{-5x}.$$

(b) Es gilt

$$\begin{aligned} y < b &\quad \Longrightarrow \{\omega \in \Omega | Y(\omega) \le y\} = \{\omega \in \Omega | X(\omega) \le y\} \\ y \ge b &\quad \Longrightarrow \{\omega \in \Omega | Y(\omega) \le y\} = \Omega \end{aligned}$$

und somit

$$F_Y(y) = \begin{cases} F(y) & y < b \\ 1 & y \ge b \end{cases} = \begin{cases} 0 & y \le 0 \\ 1 - e^{-5y} & 0 < y < b \\ 1 & y \ge b \end{cases}.$$

Somit folgt wegen $Y \ge 0$

$$\mathbb{E}(Y) = \int_0^\infty (1 - F_Y(y))\,dy = \int_0^b e^{-5y}\,dy = \frac{1}{5}\left(1 - e^{-5b}\right).$$

(c) Sei μ die rechte Seite von (0.1). Zeige zunächst, dass $\mu_Y((-\infty, x]) = \mu((-\infty, x])$ für alle $x \in \mathbb{R}$ gilt. Nun ist $\mathbb{P}(X \ge b) = e^{-5b}$,

$$\delta_b((-\infty, x]) = \begin{cases} 1 & b \le x \\ 0 & b > x, \end{cases}$$

$$\left(f \cdot 1_{[0,b)}\lambda^1\right)((-\infty, x]) = \int_0^b 1_{(-\infty,x]} f(t)\,dt = \begin{cases} 0 & x \le 0 \\ 1 - e^{-5x} & 0 \le x < b \\ 1 - e^{-5b} & x \ge b. \end{cases}$$

Insgesamt also

$$\begin{aligned} x \le 0 &\quad \Longrightarrow \mu((-\infty, x]) = e^{-5b} \cdot 0 + 0 = 0 = F_Y(x), \\ 0 < x < b &\quad \Longrightarrow \mu((-\infty, x]) = e^{-5b} \cdot 0 + 1 - e^{-5x} = 1 - e^{-5x} = F_Y(x), \\ x \ge b &\quad \Longrightarrow \mu((-\infty, x]) = e^{-5b} \cdot 1 + 1 - e^{-5b} = 1 = F_Y(x). \end{aligned}$$

d.h. $F_Y((-\infty,x]) = \mu((-\infty,x])$. Laut Definition von μ_Y gilt auch $F_Y((-\infty,x]) = \mu_Y((-\infty,x])$ also $\mu_Y((-\infty,x]) = \mu((-\infty,x])$ für alle $x \in \mathbb{R}$. Die Behauptung folgt nun wegen $\mu(\mathbb{R}) = \mu_Y(\mathbb{R}) = 1$ aus

$$\begin{aligned}
\mu_Y((a,b]) &= F_Y(b) - F_Y(a) = \mu_Y((-\infty,b]) - \mu_Y((-\infty,a]) = \mu((-\infty,b]) - \mu((-\infty,a]) \\
&= \mu((-\infty,b] \setminus (-\infty,a]) = \mu((a,b]).
\end{aligned}$$

Somit gilt (0.1) für alle halboffenen Intervalle $(a,b]$. Diese bilden einen $\cap$-stabilen Erzeuger von $\mathcal{B}^1$ und es folgt (0.1) auf $\mathcal{B}^1$.

(d)

(i) μ ist ein Maß: $\mu(\varnothing) = \mu_1(\varnothing) + \mu_2(\varnothing) = 0 + 0$. Seien $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkte Mengen $A_n \in \mathcal{B}^1$. Dann

$$\begin{aligned}
\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) &= \mu_1\left(\bigcup_{n\in\mathbb{N}} A_n\right) + \mu_2\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mu_1(A_n) + \sum_{n=1}^{\infty} \mu_2(A_n) \\
&= \sum_{n=1}^{\infty} (\mu_1(A_n) + \mu_2(A_n)) = \sum_{n=1}^{\infty} \mu(A_n).
\end{aligned}$$

Das vorletzte Gleichheitszeichen folgt daraus, dass man bei Reihen mit nicht negativen Gliedern den Grenzwert durch Umordnen nicht verändert.

(ii) 1. Schritt: g eine Treppenfunktion, also $g = \sum_{i=1}^{n} \alpha_i 1_{A_i}$ mit $\alpha_i \in [0,\infty)$, $A_i \in \mathcal{B}^1$, $i = 1,\dots,n$. Dann

$$\begin{aligned}
\int g\,d\mu &= \sum_{i=1}^{n} \alpha_i \int 1_{A_i} d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{i=1}^{n} \alpha_i \mu_1(A_i) + \sum_{i=1}^{n} \alpha_i \mu_2(A_i) \\
&= \int g\,d\mu_1 + \int g\,d\mu_2
\end{aligned}$$

2. Schritt: Sei nun $(g_n)_{n\in\mathbb{N}}$ eine Folge von Treppenfunktionen $0 \le g_n \uparrow g$. Dann gilt laut Konstruktion des Integrals

$$\lim_{n\to\infty} \int g_n d\mu = \int g\,d\mu \text{ und } \lim_{n\to\infty} \int g_n d\mu_i = \int g\,d\mu_i, i = 1,2.$$

Laut 1. Schritt gilt für die Treppenfunktionen

$$\forall n \in \mathbb{N}: \quad \int g_n d\mu = \int g_n d\mu_1 + \int g_n d\mu_2.$$

Geht man hier für $n \to \infty$ zur Grenze über, folgt die Behauptung.

(e) Für messbares $g : \mathbb{R} \longrightarrow \mathbb{R}$

$$\int g\,d\mu_Y \overset{(d)}{=} e^{-5b} \int g\,d\delta_b + \int_0^b 5e^{-5x} g(x)\,dx = e^{-5b} g(b) + \int_0^b 5e^{-5x} g(x)\,dx.$$

Setzt man $g(x) = e^x$ ein, dann folgt

$$\int e^x d\mu_Y(x) = e^{-4b} + \frac{5}{4}\left(1 - e^{-4b}\right).$$

Aufgabe 2
(a) Da $N \sim \text{Poi}(y)$ bei gegebenem y folgt $\mathbb{E}(N|Y=y) = y$ also $\mathbb{E}(N|Y) = Y$.

$$\mathbb{E}(N) = \mathbb{E}(\mathbb{E}(N|Y)) = \mathbb{E}(Y) = 1.$$

(b) Sei $k \in \mathbb{N}_0$. Laut Angabe besitzt Y die Dichte $f_Y(y) = 1_{(0,\infty)}e^{-1}(y)$.

$$\begin{aligned}\mathbb{P}(N=k) &= \int_{\mathbb{R}} \mathbb{P}(N=k|Y=y) f_Y(y)\,dy = \int_0^\infty e^{-y}\frac{y^k}{k!}e^{-y}\,dy = \frac{1}{k!}\int_0^\infty e^{-2y}y^k\,dy \\ &= \frac{1}{2^{k+1}k!}\int_0^\infty e^{-t}t^k\,dt = \frac{1}{2^{k+1}k!}\Gamma(k+1) = \frac{1}{2^{k+1}}.\end{aligned}$$

Aufgabe 3
(a) Relevant ist bei $k=4, l=5$ das Quantil $\chi^2_{3\cdot 4,0.9} = 18,549$. Die Annahme der Unabhängigkeit wird nicht verworfen.
(b)

(i) Sei $L : (0,\infty) \longrightarrow \mathbb{R}$ die Likelihoodfunktion

$$L(\lambda) = \prod_{j=1}^k e^{-\lambda}\frac{\lambda^{n_j}}{n_j!}$$

Wir setzen $l := \ln L$. Dann

$$\begin{aligned} l(\lambda) &= \sum_{j=1}^k (-\lambda + n_j \ln(\lambda) - \ln(n_j!)) \\ \frac{d}{d\lambda} l(\lambda) &= \sum_{j=1}^k \left(-1 + n_j \frac{1}{\lambda}\right) \\ \frac{d^2}{d\lambda^2} l(\lambda) &= -\sum_{j=1}^k n_j \frac{1}{\lambda^2} < 0. \end{aligned}$$

Setzt man die erste Gleichung 0 erhält man ein Maximum von L,

$$\hat{\lambda} = \frac{1}{k}\sum_{j=1}^k n_j.$$

(ii) Dieser Schätzer ist erwartungstreu: Sind $N_1, \dots, N_k \sim \text{Poi}(\lambda)$ unabhängig, $\lambda > 0$, dann

$$\mathbb{E}(\frac{1}{k}\sum_{j=1}^k N_j) = \frac{1}{k}\sum_{j=1}^k \underbrace{\mathbb{E}(N_j)}_{=\lambda} = \lambda.$$

(c) Es gilt $k=5$, $n_1 = 52$, $n_2 = 51$, $n_3 = 52$, $n_4 = 40$, $n_5 = 40$. Eingetzt ergibt das

$$\hat{\lambda} = 47.$$

(d)

(i) Es wird der t-Test angewendet. Testgröße ist

$$t = \frac{\overline{n} - 50}{\sqrt{s^2}} \sqrt{5}$$

wobei $\overline{n}$ der Mittelwert der n_j und

$$s^2 = \frac{1}{5-1} \sum_{i=1}^{5} (n_j - 47)^2 = 41$$

gilt. Eingesetzt ist $t = -1,047$. Wegen $t_{4,0,975} = 2,776$ wird die Hypothese $\lambda = 50$ nicht abgelehnt.

(ii) Das Konfidenzintervall wird gegeben durch $\left[\overline{n} - t_{4,0,975}\frac{\sqrt{s^2}}{\sqrt{5}}, \overline{n} - t_{4,0,975}\frac{\sqrt{s^2}}{\sqrt{5}}\right] =$ $[39,05; 54,95]$.