Am 5. Mai 2007 wurde zum dritten Mal eine stochastische Eingangsprüfung durchgeführt. Insgesamt konnten maximal 56 Punkte erreicht werden, zum Bestehen der Prüfung waren mindestens 26 Punkte notwendig. Von 26 Teilnehmern haben 19 die Prüfung bestanden. Zugelassen waren ein Taschenrechner, sowie eine ausgeteilte Formelsammlung und Wahrscheinlichkeitstabellen.

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 (12 Punkte)
Sei $\mathcal{A}' \subset \mathcal{A}$ eine σ-Algebra.

(a) Man zeige: Ist $X : \Omega \longrightarrow \mathbb{R}$ $\mathcal{A}'$-messbar, dann ist X $\mathcal{A}$-messbar.

(b) Gilt die Umkehrung von (a)?

(c) $\mathcal{A}'$ besitze folgende Eigenschaft:

$$\forall A' \in \mathcal{A}' : \mathbb{P}(A') \in \{0, 1\}.$$

Man zeige für $X : \Omega \longrightarrow \mathbb{R}$ $\mathcal{A}'$-messbar:

(i) Die Verteilungsfunktion F von X nimmt nur die Werte 0 und 1 an.

(ii) Man schließe daraus, dass X fast sicher konstant ist.

Lösung:
(a) Sei $B \in \mathcal{B}^1$. Da X messbar ist, gilt $X^{-1}(B) \in \mathcal{A}'$, also wegen $\mathcal{A}' \subset \mathcal{A}$ auch $X^{-1}(B) \in \mathcal{A}$.
(b) Die Umkehrung gilt nicht: Ω bestehe aus mindestens zwei Elementen. Sei $\mathcal{A} := \mathcal{P}(\Omega)$, $\mathcal{A}' := \{\emptyset, \Omega\}$ und für $\omega \in \Omega$ sei $X := 1_{\{\omega\}}$, also die Indikatorfunktion von $\{\omega\} \in \mathcal{A}$. X ist $\mathcal{A}$- aber nicht $\mathcal{A}'$-messbar.
(c) (i) Sei $x \in \mathbb{R}$. Wegen $\{X \leq x\} \in \mathcal{A}'$ folgt laut Annahme $\mathbb{P}(X \leq x) \in \{0, 1\}$. Somit gilt

$$F(x) = \mathbb{P}(X \leq x) \in \{0, 1\}.$$

(ii) F ist von rechts stetig und monoton wachsend, $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$. Wegen (i) sind die Mengen $F^{-1}(0)$ und $F^{-1}(1)$ nichtleer. Sei

$$\begin{aligned} x_0 &:= \sup\{x \in \mathbb{R} : F(x) = 0\} \\ x_1 &:= \inf\{x \in \mathbb{R} : F(x) = 1\}. \end{aligned}$$

Es gilt $x_0 = x_1$: Aufgrund der Monotonie von F gilt $x_0 \leq x_1$. Angenommen es gilt $x_0 < x_1$. Dann gibt es ein $x_2 \in]x_0, x_1[$ mit $F(x_2) \in]0, 1[$ im Widerspruch zu (i). Daher gilt $x_0 = x_1$.
Nun gilt wegen der Stetigkeit von rechts

$$\lim_{x \searrow x_1} F(x) = F(x_1)$$

und wegen der Definition von x_0, x_1

$$\mathbb{P}(X = x_1) = F(x_1) - \lim_{x \nearrow x_1} F(x) = 1 - 0 = 1.$$

Aufgabe 2 (16 Punkte)
Aus einem Versichertenkollektiv scheiden die Versicherungsnehmer durch Tod oder Storno aus. Es wurde festgestellt, dass die Anzahl N der Ausgeschiedenen pro Jahr Poisson-verteilt ist mit Parameter $\lambda > 0$. Die Wahrscheinlichkeit, dass die Ausscheideursache Storno ist, betrage $p \in (0,1)$. Die Anzahl der Toten bzw. Storni wird mit T bzw. S bezeichnet.

(a) Bestimmen Sie $\mathbb{P}(S = k|N = m + k)$ für $m, k \in \mathbb{N}_0$.

(b) Bestimmen Sie die Wahrscheinlichkeitsfunktion von S. Welche bekannte Verteilung ergibt sich?

(c) Zeigen Sie $\mathbb{E}(S|N) = Np$, $\mathrm{Var}(S|N) = Np(1-p)$.

(d) Bestimmen Sie $\mathbb{E}(S)$ und $\mathrm{Var}(S)$.

Lösung:
(a) S gegeben $N = m + k$ ist $B(m + k, p)$-verteilt, also

$$\mathbb{P}(S = k|N = m + k) = \binom{m+k}{k} p^k (1-p)^m.$$

(b) Mit dem Satz von der totalen Wahrscheinlichkeit folgt

$$\begin{aligned}
\mathbb{P}(S = k) &= \sum_{m=0}^{\infty} \mathbb{P}(S = k|N = m + k)\mathbb{P}(N = m + k) \\
&= \sum_{m=0}^{\infty} \binom{m+k}{k} p^k (1-p)^m e^{-\lambda} \frac{\lambda^{m+k}}{(m+k)!} \\
&= p^k e^{-\lambda} \lambda^k \sum_{m=0}^{\infty} \frac{(m+k)!}{k!m!} (1-p)^m \frac{\lambda^m}{(m+k)!} \\
&= \frac{(p\lambda)^k}{k!} e^{-\lambda} \sum_{m=0}^{\infty} \frac{((1-p)\lambda)^m}{m!} \\
&= \frac{(p\lambda)^k}{k!} e^{-\lambda} e^{(1-p)\lambda} \\
&= \frac{(p\lambda)^k}{k!} e^{-p\lambda}
\end{aligned}$$

S ist Poisson-verteilt mit Parameter $p\lambda$.
(c) S gegeben $N = n$ ist $B(n, p)$-verteilt ist (s. (a)) und somit

$$\begin{aligned}
\mathbb{E}(S|N = n) &= np \\
\mathrm{Var}(S|N = n) &= np(1-p).
\end{aligned}$$

Folglich

$$\begin{aligned}
\mathbb{E}(S|N) &= Np \\
\mathrm{Var}(S|N) &= Np(1-p).
\end{aligned}$$

(d) Mit dem Satz über die iterierten Erwartungswerte bzw. der Varianzzerlegung folgt

$$\begin{aligned}\mathbb{E}(S) &= \mathbb{E}(\mathbb{E}(S|N)) = \mathbb{E}(Np) = p\mathbb{E}(N) = p\lambda \text{ bzw.}\\ \mathrm{Var}(S) &= \mathbb{E}(\mathrm{Var}(S|N)) + \mathrm{Var}(\mathbb{E}(S|N)) = \mathbb{E}(Np(1-p)) + \mathrm{Var}(Np)\\ &= p(1-p)\lambda + p^2\lambda = p\lambda.\end{aligned}$$

Aufgabe 3 (8 Punkte)
Ein Makler betreibt Telefonakquise. Er ruft solange bei potentiellen Kunden an, bis er einen Termin vereinbaren kann. Er trägt den Termin in seinen Kalender ein und gönnt sich eine Pause. Dann beginnt er mit neuen Anrufen. Die folgenden Zahlen geben an, wieviele Anrufe bei 5 solcher Anrufserien nötig waren:

$$5, 4, 10, 6, 3$$

Sei X die Anzahl der Anrufe, die nötig sind bis ein Termin zustande kommt. Es kann angenommen werden, dass

$$\mathbb{P}(X = k) = (1-\vartheta)^{k-1}\vartheta, k \in \{1, 2, \dots\}, \vartheta \in (0, 1).$$

gilt und dass die Anzahlen der Anrufe bei verschiedenen Kunden voneinander unabhängig sind. Bestimmen Sie mit den obigen Daten einen Maximum-Likelihood-Schätzer für ϑ.
Lösung:
Sei $L :]0, \infty[\longrightarrow]0, \infty[$ die Likelihood, $\ell :]0, \infty[\longrightarrow \mathbb{R}$ die Log Likelihood. Es gilt für die gegebene Stichprobe x_i, $i = 1, \dots, 5$:

$$L(\vartheta) = \prod_{i=1}^{5} (1-\vartheta)^{x_i - 1}\vartheta$$

$$\begin{aligned}\ell(\vartheta) &= \ln L(\vartheta) = \sum_{i=1}^{5} [(x_i - 1)\ln(1-\vartheta) + \ln\vartheta]\\ &= 5\ln\vartheta - 5\ln(1-\vartheta) + \ln(1-\vartheta)\sum_{i=1}^{5} x_i\end{aligned}$$

$$\ell'(\vartheta) = \frac{d}{d\vartheta}\ell(\vartheta) = \frac{5}{\vartheta} + \frac{5}{1-\vartheta} - \frac{\sum_{i=1}^{5} x_i}{1-\vartheta}.$$

Einzige Nullstelle von ℓ' ist $\hat{\vartheta} = \dfrac{5}{\sum_{i=1}^{5} x_i}$. ℓ' wechselt in $\hat{\vartheta}$ das Vorzeichen und ist auf $]0, \hat{\vartheta}[$ bzw. $]\hat{\vartheta}, \infty[$ streng monoton steigend bzw. fallend. Somit ergibt sich die ML-Schätzung

$$\hat{\vartheta} = \frac{5}{28}.$$

Aufgabe 4 (20 Punkte)
Ein Versicherungsunternehmen legt seiner Prämienberechnung einen Schadendurchschnitt von 50 zugrunde. Die Schäden seien unabhängig und identisch verteilt. Folgende Schadendaten $x_1, \dots, x_{11}$ liegen vor:

3 12 16 24 34 38 46 87 97 134 260.

Es gilt $\sum_{i=1}^{11} x_i = 751$, $\sum_{i=1}^{11} x_i^2 = 108235$.

(a) Angenommen die Schäden sind $N(\mu, \sigma^2)$ verteilt.

(i) Geben Sie ein Konfidenzintervall für μ zum Konfidenzniveau 95 % an.

(ii) Wieviele Daten bräuchte man, um ein Konfidenzintervall zum Konfidenzniveau 95 % der Länge von höchstens 20 zu erhalten?

(b) Die Prämie soll erhöht werden. Ein Treuhänder besteht darauf, dass der Aktuar nachweist, dass der Durchschnittsschaden nun höher als 50 ist.

(i) Formulieren Sie Null- und Alternativhypothese?

(ii) Wie lautet die Entscheidung unter Normalverteilungsannahme und einem Signifikanzniveau von 5 %?

Lösung:
(a) (i) Es ergeben sich folgende Werte:

$$\begin{aligned} \overline{x} &= \frac{751}{11} = 68,27 \qquad \text{[Mittelwert]} \\ s^2 &= \frac{1}{10}\left(\sum_{i=1}^{11} x_i^2 - 11 \cdot (\overline{x})^2\right) = 5696,22 \qquad \text{[empirische Varianz]} \\ s &= 75,47. \end{aligned}$$

Das Konfidenzintervall lautet wegen $t_{10;0,975} = 2,228$ [97,5 % Quantil der t-Verteilung mit 10 Freiheitsgraden]

$$\left[\overline{x} - s \cdot t_{10;0,975} \frac{1}{\sqrt{11}}, \overline{x} + s \cdot t_{10;0,975} \frac{1}{\sqrt{11}}\right] = [17,57; 118,97].$$

(ii) Für den Stichprobenumfang n ist die Länge des Konfidenzuntervalls gleich

$$2 \cdot s \cdot t_{n-1,0,975} \cdot \frac{1}{\sqrt{n}}.$$

Wir nehmen zusätzlich an, dass die empirische Standardabweichung mit $s = 75,47$ unverändert bleibt. Da $(t_{n-1;0,975})_{n\in\mathbb{N}}$ monoton fällt, gilt $t_{n-1,0;975} \leq t_{10;0,975} = 2,228$ für $n \geq 11$. Somit reicht es

$$\sqrt{n} \geq \frac{2 \cdot s \cdot t_{n-1;0,975}}{20} = 16,81.$$

zu wählen. Für $n \geq 283$ erhält man eine Intervalllänge, die kleiner als 20 ist.
(b) (i) Sei $\mu := \mathbb{E}(X)$. Die Nullhypothese lautet

$$H_0 : \mu \leq 50$$

und die Alternative

$$H_1 : \mu > 50.$$

(ii) Es wird der einseitige t-Test verwendet.
Testgröße ist

$$t = \frac{\overline{x} - 50}{s}\sqrt{11} = 0,803.$$

H_0 wird abgelehnt, wenn $t > t_{10;0,95} = 1,812$. Wegen $t = 0,803$ wird H_0 nicht abgelehnt.