

Schriftliche Prüfung im Fach

**Angewandte Stochastik  
Klausur mit Lösungen**

gemäß Prüfungsordnung 6  
der Deutschen Aktuarvereinigung e.V.

am 8. Mai 2026

*Hinweise:*

- Als Hilfsmittel für die Klausur sind der Leitfaden und die im Leitfaden angegebene Literatur, Seminarunterlagen (Skript und Aufgabensammlung ohne handschriftliche Notizen; Markierungen mit Textmarker sowie unbeschriftete Post-Its sind erlaubt) sowie frühere Klausuren jeweils in Papierform sowie ein nicht programmierbarer Taschenrechner zugelassen. Hinweise zu den MC-Fragen finden Sie in der jeweiligen Aufgabenstellung. Alle Antworten sind in deutscher Sprache zu erstellen (fremdsprachige Fachbegriffe ausgenommen).
- Die Gesamtpunktzahl beträgt 180 Punkte. Die Klausur ist bestanden, wenn mindestens 90 Punkte erreicht werden.
- Bitte prüfen Sie die Ihnen vorliegende Prüfungsklausur auf Vollständigkeit. Die Klausur besteht aus 27 Seiten.
- Alle Antworten sind zu begründen und bei Rechenaufgaben muss der Lösungsweg ersichtlich sein.
- Bitte vermeiden Sie bei der Lösungserstellung die nicht zusammenhängende Streuung der Lösungen zu den einzelnen Aufgabenteilen.
- Aus Gründen der besseren Lesbarkeit wird auf die gleichzeitige Verwendung der Sprachformen männlich, weiblich und divers (m/w/d) verzichtet.

*Mitglieder der Prüfungskommission:*

Prof. Torsten Becker, Dr. Richard Herrmann,  
Prof. Christian Heumann, Dr. Stefan Pilz,  
Prof. Viktor Sandor, Dr. Dominik Schäfer

**Aufgabe 1. [Lebensdauermodelle] [30 Punkte]**

Ein Großunternehmen hat zwei Teilbetriebe A und B, von denen der Teilbetrieb B veräußert werden soll. Aus diesem Anlaß sollen sämtliche bei der Veräußerung bereits laufende Rentenleistungen der betrieblichen Altersversorgung auf eine Pensionskasse ausgelagert werden. Die Leistungen für beide Teilbetriebe werden künftig von der Pensionskasse erbracht. Gehen Sie im weiteren davon aus, dass die Risiken der Versicherten stochastisch unabhängig sind und dass der Pensionsplan nur eine jährlich vorschüssige Altersrente zwischen 1.200€ und 6.000€ an den ehemaligen Mitarbeiter vorsieht.

- (a) [9 Punkte] Stellen Sie mithilfe der Rekursionsformel für die Deckungsrückstellung formelmäßig für einen einzelnen Versicherten im Rentenbezug das riskierte Kapital für das Erlebens- und das Todesfallrisiko dar.
- (b) [8 Punkte] Definieren Sie eine Zufallsvariable  $Y$ , die für einen einzelnen Versicherten auf Grundlage der riskierten Kapitale die möglichen Schadenhöhen angibt.
  - (i) [2 Punkte] Geben Sie die Wahrscheinlichkeiten für die Realisationen von  $Y$  an.
  - (ii) [6 Punkte] Zeigen Sie, dass  $E(Y) = 0$  und  $\text{Var}(Y) = (vV_1)^2 q(1 - q)$  ist.
- (c) [6 Punkte]
  - (i) [2 Punkte] Geben Sie die Zufallsvariable des Gesamtschadens für den Bestand eines Teilbetriebs unter Berücksichtigung von Geschlecht und Alter der Versicherten an.
  - (ii) [2 Punkte] Mit welchem Ansatz kann die Verteilungsfunktion des Gesamtschadens für einen Teilbetrieb hergeleitet werden?
  - (iii) [2 Punkte] Geben Sie Erwartungswert und Varianz der Verteilung des Gesamtschadens an.
- (d) [3 Punkte] Geben Sie eine hinreichende Bedingungen für die asymptotische Normalverteilung des Gesamtschadens an. Halten Sie im vorliegenden Fall die Bedingung erfüllt? Begründen Sie kurz Ihre Antwort ohne Beweis.
- (e) [4 Punkte] Zeigen Sie für den Fall normalverteilter Schadenssummen aus Teilaufgabe (c), dass aufgrund der Aufteilung in die Teilbestände A und B die erforderlichen Sicherheitskapitale  $S^{A;0,99}$  und  $S^{B;0,99}$  zu vorgegebener Sicherheitswahrscheinlichkeit  $(1 - \alpha) = 99\%$  für die Teilbestände A und B zusammen größer bzw. nicht kleiner sind als ohne die Aufteilung in die Teilbestände.

**Lösung**

(a) Nach der Rekursionsformel für die Deckungsrückstellung gilt

$$V_0 = -P + R + vV_1 \cdot (1 - q)$$

Hierbei bezeichnen

$V_0$  Deckungsrückstellung am Anfang des Versicherungsjahres

$V_1$  Deckungsrückstellung als Rentner am Ende des Versicherungsjahres

$P$  Prämie, jährlich vorschüssig

$R$  Versicherungsleistung, fällig am Beginn des Versicherungsjahres

$v = \frac{1}{1+z}$  Diskontfaktor ( $z = \text{Zins}$ )

$q$  Sterbewahrscheinlichkeit des Versicherten

Bei einer Rentenversicherung in der Rentenbezugszeit ist  $P = 0$  und  $R$  die vorschüssige Jahresrente.

Da bei Tod keine weitere Rente zu leisten ist, lässt sich  $V_0$  schreiben als

$$V_0 = R + vV_1(1 - q) + 0 \cdot q \Leftrightarrow V_0 - R = vV_1(1 - q) + 0 \cdot q \quad (1)$$

d. h. nach Abzug der sofort fälligen Rente muss die verbleibende Deckungsrückstellung für die erwartete und abgezinste Deckungsrückstellung am Ende des Versicherungsjahres ausreichen.

Das riskierte Kapital ist

$$\begin{aligned} K^E &= vV_1 - (V_0 - R) && \text{im Erlebensfall} \\ K^T &= 0 - (V_0 - R) && \text{im Todesfall} \end{aligned}$$

(b) Für jeden Versicherten sei

$$Y = \begin{cases} vV_1 - (V_0 - R) & \text{falls der Versicherte überlebt} \\ 0 - (V_0 - R) & \text{falls der Versicherte stirbt} \end{cases}$$

dann gibt  $Y$  den möglichen Schaden für den Versorgungsträger an.

(i) Die Verteilungsfunktion des Schadens ist gegeben durch

$$\begin{aligned}P(Y = vV_1 - (V_0 - R)) &= 1 - q \\P(Y = -V_0 + R) &= q\end{aligned}$$

(ii) Erwartungswert und Varianz lauten wegen (1):

$$\begin{aligned}E(Y) &= (1 - q)(vV_1 - (V_0 - R)) - q(V_0 - R) \\&= (vV_1 - (V_0 - R)) - q(vV_1 - (V_0 - R) + (V_0 - R)) \\&= vV_1 - (V_0 - R) - qvV_1 \\&= (1 - q)vV_1 - (V_0 - R) \\&\stackrel{(1)}{=} (1 - q)vV_1 - (1 - q)vV_1 \\&= 0\end{aligned}$$

$$\begin{aligned}\text{Var}(Y) &= (1 - q)(vV_1 - (V_0 - R))^2 + q(V_0 - R)^2 \\&= (1 - q)(vV_1 - vV_1(1 - q))^2 + q(vV_1(1 - q))^2 \\&= (1 - q)q^2(vV_1)^2 + q(1 - q)^2(vV_1)^2 \\&= (vV_1)^2 q(1 - q)\end{aligned}$$

(c) (i) Bezeichne  $Y_{g,x,i}$  den Schaden durch den  $i$ -ten Versicherten,  $i \in \{1, \dots, n\}$ , mit Geschlecht  $g \in \{1, 2\}$  und Alter  $x \in \{x_1, \dots, x_m\}$  für den Versorgungsträger. Der Schaden für einen (Teil-)Bestand ergibt sich dann durch

$$S_n = \sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n Y_{g,x,i}$$

(ii) Die Wahrscheinlichkeitsverteilung der Summe  $S_2 = X_1 + X_2$  von zwei stochastisch unabhängigen Zufallsvariablen  $X_1$  und  $X_2$  kann durch Faltung der Wahrscheinlichkeitsverteilungen der Summanden ermittelt werden. Entsprechendes gilt für  $S_n$ .

(iii) Für den Erwartungswert gilt:

$$\begin{aligned}E(S_n) &= E\left(\sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n Y_{g,x,i}\right) \\&= \sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n E(Y_{g,x,i}) \\&= 0\end{aligned}$$

Für die Varianz gilt:

$$\begin{aligned}\text{Var}(S_n) &= \text{Var}\left(\sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n Y_{g,x,i}\right) \\&= \sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n \text{Var}(Y_{g,x,i})\end{aligned}$$

- (d) Nach dem Zentralen Grenzwertsatz ist die Erfüllung der Ljapunoff-Bedingung hinreichend für die Konvergenz der Verteilung von  $S_n = \sum_{g=1}^2 \sum_{x=x_1}^{x_m} \sum_{i=1}^n Y_{g,x,i}$  gegen die Normalverteilung.

Es ist also zu prüfen, ob die Ljapunoff-Bedingung erfüllt ist:

D. h.  $\exists \delta > 0$ , so dass

$$\lim_{n \rightarrow \infty} \frac{1}{\tau_n^{2+\delta}} \sum_{i=1}^n E(|Y_i - E(Y_i)|^{2+\delta}) = 0$$

mit  $\tau_n^2 = \text{Var}(S_n) = \sum_{i=1}^n \text{Var}(Y_i)$

Aufgrund der Aufgabenstellung gibt es eine Obergrenze  $R_{\max}$  für die Rentenleistungen. Dann ist die Folge der Einzelschäden gleichmäßig beschränkt und für die Folge der Standardabweichungen von  $S_n$  gilt:

$$\lim_{n \rightarrow \infty} \sqrt{\text{Var}(S_n)} = \infty$$

Dann konvergiert nach dem Zentralen Grenzwertsatz die Verteilung von  $S_n$  gegen die Normalverteilung.

- (e) Bezeichne  $S_w$  den Schaden im Bestand  $w \in \{A, B, G\}$  mit  $G = A \cup B$  für Gesamtbestand.

$S^A$  und  $S^B$  sind stochastisch unabhängig und es gilt  $S^G = S^A + S^B$ . Für die Verteilungen gilt  $S^w \sim \mathcal{N}(\mu_w, \sigma_w^2)$ .

Für den Value at Risk gilt dann mit  $\Phi$  als Verteilungsfunktion und  $u_{0,99}$  als 99%-Quantil der Standard-Normalverteilung

$$\begin{aligned} \text{VaR}_{0,01}(S^w) &= \mu_w + \sigma_w \cdot \Phi^{-1}(1 - 0,01) \\ &= \mu_w + \sigma_w \cdot u_{0,99} \end{aligned}$$

Das über den Erwartungswert hinausgehende Sicherheitskapital  $S^{w;0,99}$  ist dann

$$S^{w;0,99} = \text{VaR}_{0,01}(S^w) - \mu_w = \sigma_w \cdot u_{0,99}$$

Für die Summe  $S^G = S^A + S^B$  der unabhängigen normalverteilten Zufallsvariablen  $S^A$  und  $S^B$  gilt

$$\begin{aligned} S^G &= S^A + S^B \sim \mathcal{N}(\mu_A + \mu_B, \sigma_A^2 + \sigma_B^2) \\ \text{VaR}_{0,01}(S^G) &= \mu_A + \mu_B + \sqrt{\sigma_A^2 + \sigma_B^2} \cdot u_{0,99} \end{aligned}$$

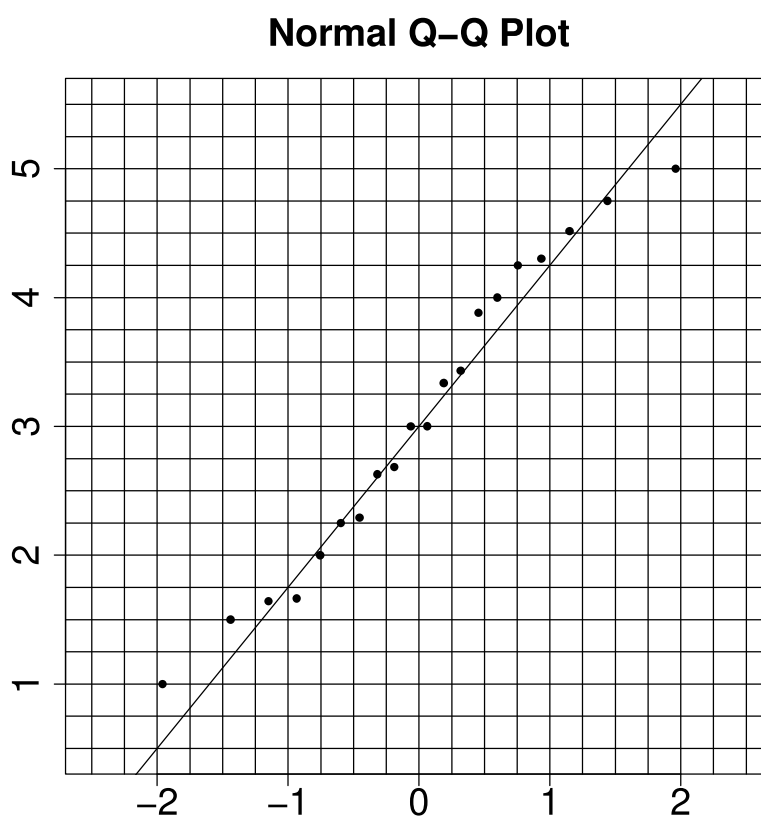
und

$$\begin{aligned} S^{G;0,99} &= \sqrt{\sigma_A^2 + \sigma_B^2} \cdot u_{0,99} \\ &\leq (\sigma_A + \sigma_B) \cdot u_{0,99} \\ &= S^{A;0,99} + S^{B;0,99} \end{aligned}$$

**Aufgabe 2.** [Deskriptive Statistik, Abhängigkeitsmaße] [30 Punkte]

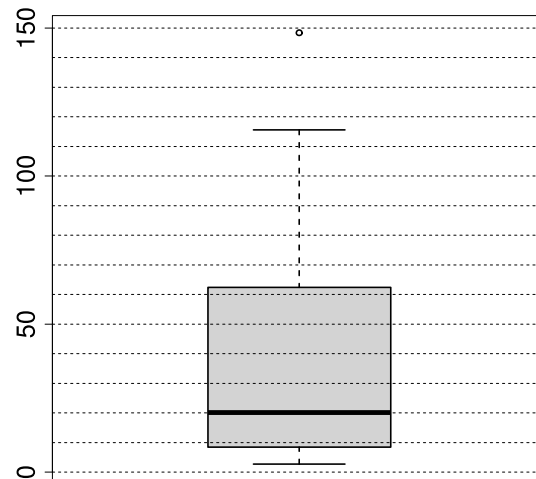
- (a) [3 Punkte] Im Aktuariat eines Sachversicherers wurden die Daten für den Gesamtschaden in Millionen Euro der letzten 20 Jahre  $x_1, \dots, x_{20}$  einer Sparte aufbereitet. Zu sehen ist ein Q-Q-Plot und die zugehörige Ausgleichsgerade der Werte  $y_k := \ln(x_k)$  gegen die Quantile der Standard-Normalverteilung.

Wie hoch waren der höchste und der niedrigste Gesamtschaden in den letzten 20 Jahren?



- (b) [4 Punkte] Auf welche Verteilung für die Zufallsvariable  $X = \text{Jahresgesamtschaden in Millionen Euro}$  überprüft das Aktuariat mit diesem Q-Q-Plot? Warum rechtfertigt der Q-Q-Plot diese Annahme?
- (c) [4 Punkte] Bestimmen Sie mit Hilfe der Grafik und auf Basis von Teil b) Schätzer für die beiden Verteilungsparameter der vermuteten Verteilung von  $X$ . Erläutern Sie Ihr Vorgehen. Bei den Angaben reicht eine Nachkommastelle.
- (d) [8 Punkte] Erstellen Sie den modifizierten Boxplot zu den Daten  $y_k$  im Q-Q Plot. Begründen Sie die gewählten bzw. ausgerechneten Werte.

- (e) [5 Punkte] In der folgenden Abbildung ist der modifizierte Boxplot der  $x_k$  gegeben. Erläutern Sie die Unterschiede zum Boxplot in (d).



- (f) [6 Punkte] Entscheiden Sie für jede Aussage, ob diese allgemein richtig oder falsch ist und begründen Sie Ihre Antworten. Schreiben Sie die Antworten auf die Lösungsblätter, die Sie abgeben.
- (i) Aus der empirischen Verteilungsfunktion für eine Stichprobe kann man einen Q-Q-Plot erstellen.
  - (ii) Aus einem Q-Q-Plot für eine Stichprobe kann man ein Histogramm erstellen.
  - (iii) Aus einem Box-Plot für eine Stichprobe kann man ein Histogramm erstellen.

## Lösung Aufgabe 2

- (a) Der kleinste Wert auf der  $y$ -Achse ist 1, der größte 5. Da es sich um logarithmierte Werte handelt, sind  $\exp(1) = 2,7$  Mio. EURO bzw.  $\exp(5) = 148,4$  Mio. EURO der kleinste bzw. größte Jahresgesamtschaden.
- (b) Da die logarithmierten Werte gegen eine Standard-Normalverteilung getestet werden, wird geprüft, ob  $\ln(X)$  eine Normalverteilung besitzt, d.h. ob  $X$  logarithmisch normalverteilt ist. Da die Punkte des Q-Q-Plots nahezu auf einer Geraden liegen, ist die Annahme plausibel.
- (c) Aus der Grafik liest man ab: Steigung der Anpassungsgeraden  $= 1,25 = \frac{5}{4}$ ; Achsenabschnitt der Anpassungsgeraden  $= 3$ . Da die Normalverteilung eine lineare Transformation von  $Z \sim N(0, 1)$  mit  $\mu + \sigma Z$  ist, folgt  $\ln(X) \sim N\left(3, \frac{25}{16}\right)$  mit den Schätzwerten  $\mu = 3$  und  $\sigma^2 = \left(\frac{5}{4}\right)^2$ . Daher ist  $X \sim LN\left(3, \frac{25}{16}\right)$ .
- (d) Die  $y$ -Werte im Q-Q Plot entsprechen der Ordnungsstatistik, man liest also ab  $y_{(5)} = 2, y_{(6)} = 2,25, y_{(15)} = 4, y_{(16)} = 4,25$ . Daraus ergeben sich die Quartile

$$y_{0,25} = \frac{1}{2}(y_{(5)} + y_{(6)}) \approx \frac{1}{2}(2 + 2,25) \approx 2,13$$

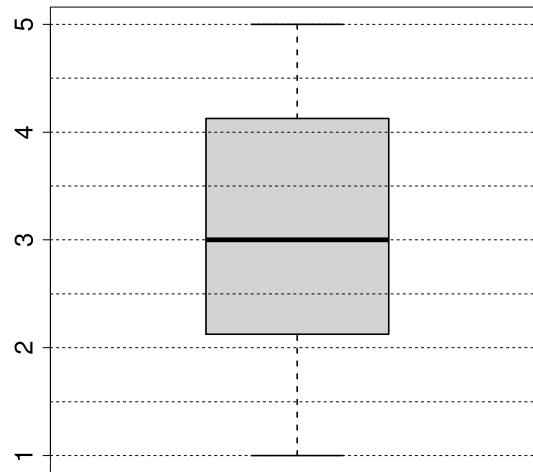
$$y_{0,75} = \frac{1}{2}(y_{(15)} + y_{(16)}) \approx \frac{1}{2}(4 + 4,25) = 4,13$$

Der Median ergibt sich aus

$$y_{0,5} = \frac{1}{2}(y_{(10)} + y_{(11)}) \approx \frac{1}{2}(3 + 3) = 3$$

Der Interquartilsabstand ist  $4,13 - 2,13 = 2$ . Die Box ist also im Bereich  $[2,1, 4,1]$ , die Whisker beim Minimum 1 und Maximum 5 da der Wert  $4,13 + 1,5 \cdot 2 = 7,13$  den Wert  $y_{(20)} = 5$  überschreitet.





- (e) Es handelt sich hier um die Werte  $x_k = \exp(y_k)$ ,  $k = 1, \dots, 20$ , also ändern sich die Werte auf der y-Achse exponentiell. Da  $\exp$  streng monoton wächst, gilt für die Ordnungsstatistik  $x_{(k)} = \exp(y_{(k)})$ .

**Variante 1:** ablesen des IQD aus dem Boxplot: der IQD beträgt etwa 52. Folglich liegt der obere Whisker maximal bei  $61 + 1,5 \cdot IQD \approx 139$ , kleiner als der größte Jahresschaden von 148,4, der damit ein Ausreißer ist. Der einzige Aureißer ist  $y_{(20)}$ , also liegt der obere Whisker beim größten Stichprobenwert, der kleiner ist als 133, also  $x_{(19)} = e^{y_{(19)}} = e^{4,75} \approx 115,6$ . Außerdem geht die Symmetrie verloren.

**Variante 2:** IQD berechnen mit den Werten aus (d). Es ergibt sich wegen  $x_{(k)} = \exp(y_{(k)})$

$$\begin{aligned}
 x_{0,25} &= \frac{1}{2} (\exp(y_{(5)}) + \exp(y_{(6)})) \approx \frac{1}{2} (e^2 + e^{2,25}) \approx 8,4 \\
 x_{0,75} &= \frac{1}{2} (\exp(y_{(15)}) + \exp(y_{(16)})) \approx \frac{1}{2} (e^4 + e^{4,25}) = 62,4 \\
 x_{0,5} &= \frac{1}{2} (\exp(y_{(10)}) + \exp(y_{(11)})) \approx \frac{1}{2} (e^3 + e^3) = 20,1
 \end{aligned}$$

und somit  $IQD = 62,4 - 8,4 = 54,0$ . Der obere Whisker würde bei maximal  $62,4 + 1,5 \cdot 54,0 = 144,3$  liegen. Ansonsten ist die Argumentation wie in Variante 1.

- (f) (i) falsch, da man für den Q-Q-Plot auch eine Verteilungsannahme benötigt.  
 (ii) wahr, da die Stichprobe auf der y-Achse gegeben ist.  
 (iii) falsch, da die Einzelwerte fehlen.

### Aufgabe 3. [36 Punkte] Induktive Statistik

*Hinweis: bei allen Ergebnissen genügen 2 Nachkommastellen*

Ein KFZ-Versicherungsbestand mit  $n$  Versicherungsnehmern wird daraufhin untersucht, inwiefern die Anzahl der Schäden in einem betrachteten Jahr von verschiedenen Einflussgrößen abhängt: Leistung des Motors in kW, Alter des Versicherungsnehmers in Kategorien und ob das Fahrzeug Teil einer Flotte ist oder nicht. Das Alter (in Jahren) ist dabei in folgende fünf Kategorien eingeteilt: 18 – 24, 25 – 39, 40 – 54, 55 – 64, 65 – 98.

- [2 Punkte] Welche Verteilung wählen Sie für die Zielvariable  $Y$  (Anzahl der Schäden)? (kurze Begründung).
- [2 Punkte] Welches Regressionsmodell und welche Linkfunktion schlagen Sie vor?
- [3 Punkte] Stellen Sie die Likelihood als Funktion der Erwartungswerte  $\lambda_i$  der Zufallsvariablen  $Y_i$ ,  $i = 1, \dots, n$ , dar.
- [4 Punkte] Stellen Sie die Modellgleichung des Regressionsmodells mit kanonischer Linkfunktion auf, welche die drei Merkmale als Haupteffekte enthält. Beachten Sie dabei, dass das Alter in Kategorien vorliegt.
- [21 Punkte] Die Ausgabe des Regressionsmodells sieht folgendermaßen aus:

Coefficients:

|              | Estimate  | Std. Error | z value | Pr(> z ) |     |
|--------------|-----------|------------|---------|----------|-----|
| (Intercept)  | -1.617104 | 0.034088   | -47.44  | < 2e-16  | *** |
| leistung     | 0.002110  | 0.000361   | 5.84    | 5.2e-09  | *** |
| alter[25,40) | -0.434020 | 0.030729   | -14.12  | < 2e-16  | *** |
| alter[40,55) | -0.629992 | 0.031082   | -20.27  | < 2e-16  | *** |
| alter[55,65) | -0.800140 | 0.035032   | -22.84  | < 2e-16  | *** |
| alter[65,99) | -0.870273 | 0.035215   | -24.71  | < 2e-16  | *** |
| flotteja     | -0.089311 | ?          | -2.06   | 0.039    | *   |

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 90046 on 163211 degrees of freedom  
Residual deviance: 89153 on 163205 degrees of freedom  
AIC: 126859

- (i) [2 Punkte] Welche Kodierung wurde für die kategoriale Altersvariable gewählt? Was ist die Referenzkategorie?
- (ii) [1 Punkt] Geben Sie die Kodierung des Alters für eine Person im Alter von 22 Jahren an.
- (iii) [1 Punkt] Geben Sie die Kodierung des Alters für eine Person im Alter von 40 Jahren an.
- (iv) [1 Punkt] Welche Kodierung wurde für die Variable, die angibt, ob das Fahrzeug Teil einer Flotte ist oder nicht, gewählt?
- (v) [1 Punkt] Berechnen Sie den Standardfehler des geschätzten Koeffizienten für das Merkmal flotte.
- (vi) [1 Punkt] Interpretieren Sie den geschätzten Koeffizienten für das Merkmal Leistung hinsichtlich des Einflusses auf die Zielvariable.
- (vii) [3 Punkte] Interpretieren Sie die geschätzten Koeffizienten für das Merkmal alter hinsichtlich des Einflusses auf die Zielvariable.
- (viii) [1 Punkt] Interpretieren Sie den geschätzten Koeffizienten für das Merkmal flotte hinsichtlich des Einflusses auf die Zielvariable.
- (ix) [3 Punkte] Was ist die geschätzte erwartete Anzahl von Schäden eines 30-jährigen Versicherungsnehmers, der ein Flottenfahrzeug mit 80kW Leistung fährt?
- (x) [7 Punkte] Es soll statistisch überprüft werden, ob das Alter einen statistisch signifikanten Einfluss hat. Die folgende Ausgabe steht zur Verfügung:

|          | LR | Chisq | Df | Pr(>Chisq)  |
|----------|----|-------|----|-------------|
| leistung |    | 33    | 1  | 7.5e-09 *** |
| alter    |    | 856   | 4  | < 2e-16 *** |
| flotte   |    | 4     | 1  | 0.036 *     |

- i. [3 Punkte] Wie lauten die entsprechenden Hypothesen  $H_0$  und  $H_1$  (verbal und in Formeln)?
- ii. [2 Punkte] Welche Teststatistik verwenden Sie und welcher Verteilung folgt diese (asymptotisch), wenn  $H_0$  gilt?
- iii. [1 Punkt] Welchen konkreten Wert hat die Teststatistik in diesem Fall?
- iv. [1 Punkt] Wie lautet die Testentscheidung, wenn als Signifikanzniveau  $\alpha = 0.05$  angenommen wird?

- (f) [4 Punkte] Das Alter wird nun statt kategorisiert als numerischer Wert in ein Modell aufgenommen:

Coefficients:

|             | Estimate  | Std. Error | z value | Pr(> z )    |
|-------------|-----------|------------|---------|-------------|
| (Intercept) | -1.539107 | 0.030306   | -50.79  | < 2e-16 *** |
| leistung    | 0.001916  | 0.000361   | 5.32    | 1.1e-07 *** |
| alter       | -0.014411 | 0.000498   | -28.92  | < 2e-16 *** |
| flotteja    | -0.080731 | 0.043255   | -1.87   | 0.062 .     |

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 90046 on 163211 degrees of freedom  
Residual deviance: 89145 on 163208 degrees of freedom  
AIC: 126845

- (i) [2 Punkte] Das metrische Alter ist linear in den Prädiktor aufgenommen worden. Nennen Sie zwei weitere Möglichkeiten (außer Kategorien für das Alter zu bilden).
- (ii) [1 Punkt] Welchen Test und welche Teststatistik verwenden Sie bei diesem Modell, um zu testen ob Alter einen Einfluss hat?
- (iii) [1 Punkt] Welches Modell würde man bevorzugen: das Modell mit metrischem alter oder das Modell mit kategorisierter Altersvariable aus Teilaufgabe (e) (kurze Begründung)?

## Lösung

- (a) [2 Punkte] Eine geeignete Verteilung ist die Poisson-Verteilung, da es sich um die Anzahl von Ereignissen (hier: Schäden) innerhalb eines definierten Zeitintervalls (hier 1 Jahr) handelt.
- (b) [2 Punkte] Regressionsmodell: Poisson GLM mit (natürlichem) Logarithmus als natürlicher (kanonischer) Linkfunktion bzw. Exponentialfunktion als Response-Funktion.

$$E(Y_i|x_i) = \lambda_i = \exp(x_i^T \beta) \quad \log(\lambda_i) = x_i^T \beta .$$

- (c) [3 Punkte] Likelihood als Funktion der  $\lambda_i$ :

$$L(\lambda_1, \dots, \lambda_n) = \prod_{i=1}^n \frac{\lambda_i^{y_i}}{y_i!} \exp(-\lambda_i) .$$

- (d) [4 Punkte] Poisson GLM mit Leistung, Alter und Flotte als Haupteffekte. Alter kategorial mit 5 Kategorien, wähle Dummy-Kodierung mit Referenzkategorie, hier 18 – 24 (*jede andere Kategorie kann ebenfalls als Referenz gewählt werden*). Flotte binär, ja=1 (*nein=1 ebenfalls möglich*):

$$\log(\lambda_i) = \beta_0 + \beta_1 \text{Leistung} + \beta_2 \text{alter}_{(25-39)} + \dots + \beta_5 \text{alter}_{(65-98)} + \beta_6 \text{Flotte}_{(ja)}$$

- (e) [21 Punkte]

- (i) [2 Punkte] Kodierung: Dummy-Kodierung, 18 – 24 Referenzkategorie.
- (ii) [1 Punkt] Person 22 Jahre alt: fällt in Referenzkategorie, daher (0, 0, 0, 0)
- (iii) [1 Punkt] Person 40 Jahre alt: fällt in Kategorie 40 – 54, daher (0, 1, 0, 0)
- (iv) [1 Punkt] Flotte binär, ebenfalls Dummy-Kodierung: ja=1, 0=nein.
- (v) [1 Punkt]

$$z = \frac{\hat{\beta}_{\text{Flotte}}}{\hat{\sigma}_{\text{Flotte}}}$$

Daher

$$\hat{\sigma}_{\text{Flotte}} = \frac{\hat{\beta}_{\text{Flotte}}}{z} = \frac{-0.089311}{-2.06} = \mathbf{0.043(355)} .$$

- (vi) [1 Punkt] Interpretation: erhöht sich die Leistung um 1 kW, erhöht sich die erwartete Anzahl an Schäden multiplikativ um den Faktor  $\exp(0.002110) = 1.0021$ , also um 0.21%. Alternativ: erhöht sich die Leistung um 10 kW, erhöht sich die erwartete Anzahl an Schäden um 2.1 %

- (vii) [3 Punkte] Interpretation für Alter: jeweils in Bezug auf die Referenzkategorie. Im Vergleich zur Referenzkategorie 18 – 24 verringert sich die erwartete Anzahl an Schäden multiplikativ um die Faktoren/Prozente:

|                   |        |        |
|-------------------|--------|--------|
| $\exp(-0.434020)$ | 0.6479 | 35.21% |
| $\exp(-0.629992)$ | 0.5326 | 46.74% |
| $\exp(-0.800140)$ | 0.4493 | 55.07% |
| $\exp(-0.870273)$ | 0.4188 | 58.12% |

- (viii) [1 Punkt] Ist das Fahrzeug Teil einer Flotte, verringert sich die erwartete Anzahl an Schäden gegenüber Fahrzeugen, die nicht Teil einer Flotte sind, multiplikativ um den Faktor/Prozentwert 0.9146 oder 8.54%.

- (ix) [3 Punkte] Prädiktor  $\eta_i = -1.617104 + 80 \cdot 0.002110 - 0.434020 - 0.089311 = -1.9716$ . Damit:

$$E(Y_i|x_i) = \exp(-1.9716) = 0.13923.$$

Die erwartete Anzahl wird auf etwa 0.14 geschätzt.

- (x) [7 Punkte]

- i. [3 Punkte] Hypothesen verbal:

Nullhypothese: Alle Parameter der Variable Alter sind 0 (oder: Alter hat keinen Einfluss auf die erwartete Anzahl der Schäden) versus mindestens ein Parameter ist ungleich 0 (Alter hat Einfluss auf die erwartete Anzahl der Schäden).

Hypothesen formal (mit Notation oben),  $\beta_2 = \beta_{\text{alter}(25-39)}$ , etc.

$$H_0 : \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$$

versus

$$H_1 : \text{mindestens ein } \beta_j, j = 2, 3, 4, 5 \text{ ist ungleich } 0.$$

- ii. [2 Punkte] Man verwendet die  $\chi^2$ -Statistik des Likelihood-Quotienten-Tests. Unter  $H_0$  ist die Teststatistik  $\chi^2$ -verteilt mit 4 Freiheitsgraden (da 4 Parameter).

- iii. [1 Punkt] Der Wert ist 856.

- iv. [1 Punkt]  $H_0$  wird abgelehnt, da  $p$ -Wert ( $< 2e-16$ ) kleiner als  $\alpha$  ist.

(f) [4 Punkte]

- (i) [2 Punkte] Alter zusätzlich als Polynom, z.B. quadratisch oder GAM (generalized additive model, splines)
- (ii) [1 Punkt] In diesem Modell hat man nur einen Koeffizienten für Alter. Der Test beruht auf der asymptotischen (Standard-)Normalverteilung der Z-Statistik unter  $H_0$ . Hier wird konkret der Wert  $z = -28.92$  als beobachteter Wert für die Z-Statistik verwendet.
- (iii) [1 Punkt] Modell mit linearem Alter, da AIC-Wert 126845 kleiner ist als AIC-Wert des Modells mit kategorialem Alter (126859).

#### Aufgabe 4. [24 Punkte] Zeitreihenanalyse

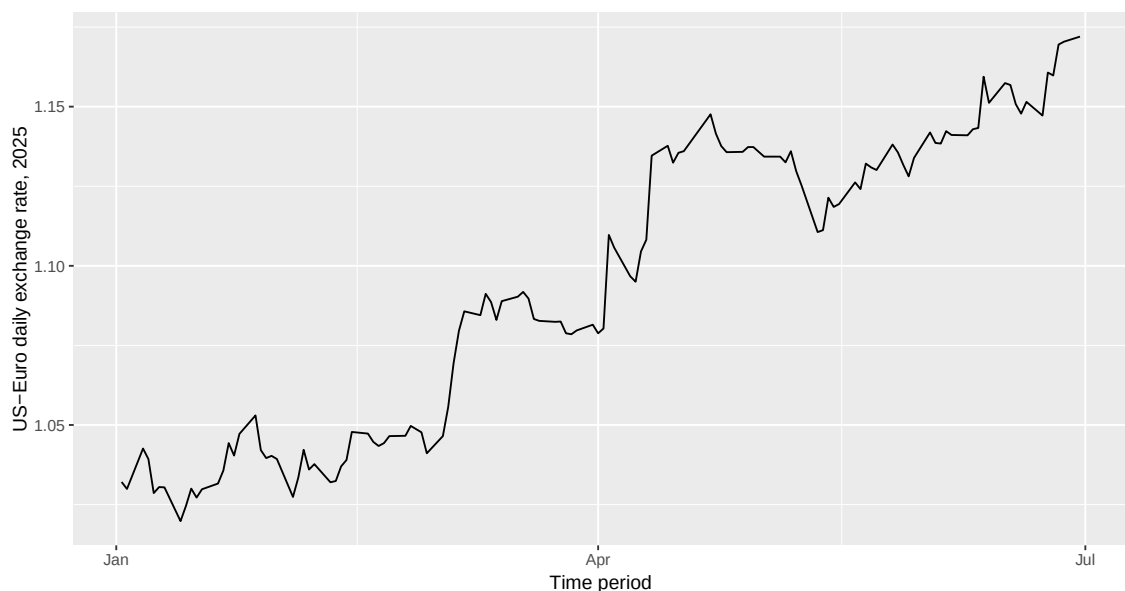
- (a) [3 Punkte] Nennen Sie drei Annahmen für schwach stationäre Prozesse.
- (b) [6 Punkte] Entscheiden Sie, ob die folgenden Aussagen richtig oder falsch sind. Geben Sie eine kurze Begründung für Ihre Entscheidung an.

Für jede richtige [falsche] Antwort gibt es zwei [Null] Punkte.

- (i) Bei  $MA(q)$  Modellen kann man zukünftige Werte schätzen und der Parameter  $q$  gibt die Anzahl der Vergangenheitswerte des MA-Modells an.
- (ii) Das  $AR(p)$  Modell ist für den Spezialfall  $p = 1$  immer stationär.
- (iii) Die partielle Autokorrelation ist ein Spezialfall der Autokorrelation.

(c) [15 Punkte]

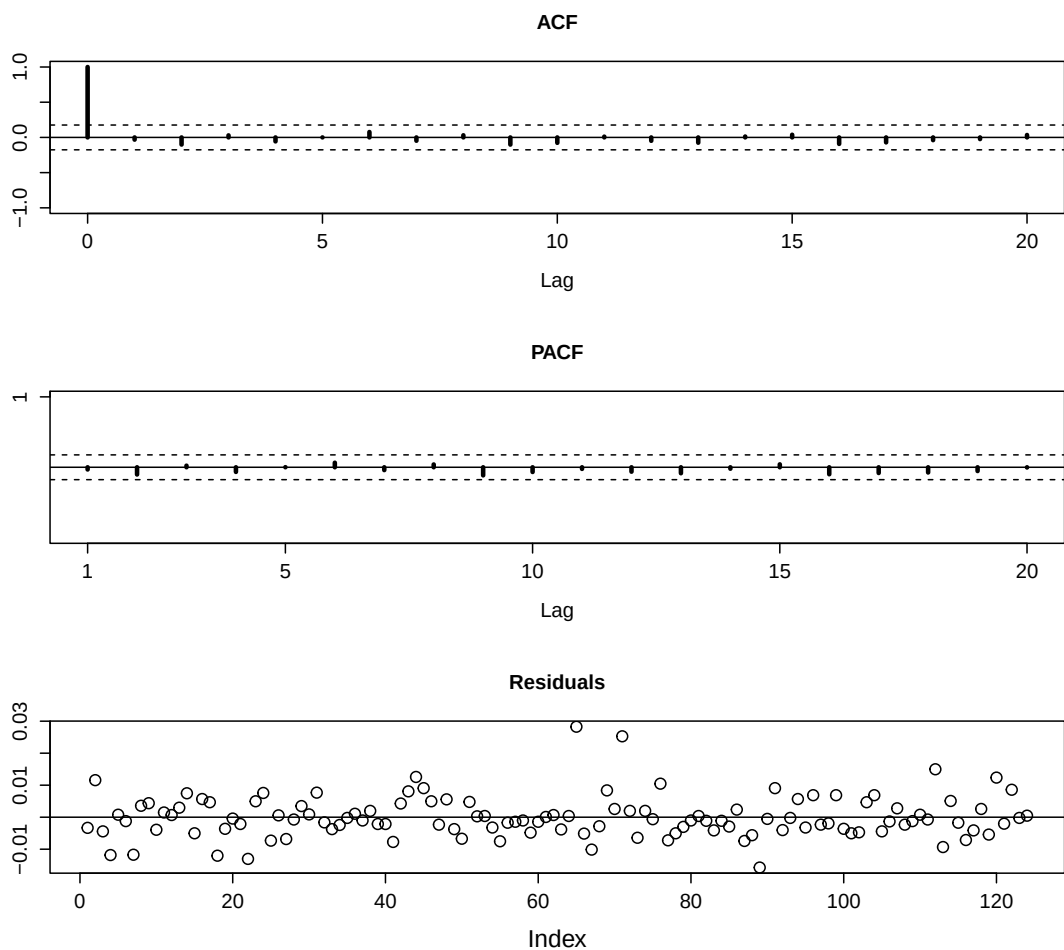
- (i) [3 Punkte] Die folgende Grafik enthält auf Tagesbasis den US Dollar - Euro Wechselkurs. Nennen Sie die **drei** Charakteristika, die geeignet sind den Verlauf der Zeitreihe zu beschreiben! (Abschätzung anhand der Grafik).



- (ii) [1 Punkt] Nennen Sie einen Grund warum die Zeitreihe nicht schwach stationär ist.
- (iii) [4 Punkte] Welche Transformationen bzw. Bereinigungen schlagen Sie vor um eine schwach stationäre Zeitreihe zu erreichen? (zwei Vorschläge mit kurzer Begründung)



- (iv) [3 Punkte] Für die bereinigte Zeitreihe wurde die Autokorrelationsfunktion bzw. die partielle Autokorrelationsfunktion bzw. die Residuen berechnet und visualisiert. Interpretieren Sie die Grafiken (jeweils 1 Punkt). Bitte beachten Sie, dass die PACF erst ab  $lag=1$  definiert ist.



- (v) [2 Punkte] Sie sollen für die bereinigte Zeitreihe ein Modell schätzen. Machen Sie einen Vorschlag und begründen ihn kurz!
- (vi) [2 Punkte] Wie beurteilen Sie den Einsatz von diesem Zeitreihenmodell (d.h. auf Basis der bereinigten Zeitreihe) für die kurzfristige Prognose (mit kurzer Begründung)?

## Lösung

(a) [3 Punkte]

- (i) Konstanter Erwartungswert.
- (ii) Konstante Varianz.
- (iii) Die Autokorrelationsfunktion hängt nur von dem Lag ab, nicht aber von den jeweiligen Zeitpunkten.

(b) [6 Punkte]

- (i) richtig. Im MA-Prozess besteht die Beobachtung  $y_t$  aus  $q$  Komponenten (gewichtet, weißes Rauschen) für die Zeitpunkte  $t - q, \dots, t - 1$  (und dem weißen Rauschen zum aktuellen Zeitpunkt  $t$ ). Eine Prognose ist mittels der Residuen möglich.
- (ii) falsch. Ein Gegenbeispiel ist der random walk.
- (iii) falsch. Es handelt sich um verwandte Konzepte der Abhängigkeit: bedingte versus marginale Betrachtung.

(c) [15 Punkte]

(i) [3 Punkte]

- i. Trend(s) vorhanden.
- ii. Volatilität vermutlich konstant.
- iii. Saisonalität nicht sichtbar bzw. nicht vorhanden.
- iv. auch akzeptabel: gleichabständige Beobachtungen bzw. metrisch skaliertes Merkmal. Die Bedingungen sind erfüllt.

(ii) [1 Punkt] Die Zeitreihe ist nicht stationär; Gründe: Trend (ein Grund genügt).

(iii) [4 Punkte] Transformationen

- i. Trendbereinigung durch Differenzen erster Ordnung.
- ii. Gleitende Durchschnitte.

(iv) [3 Punkte] Interpretation

- 
- i. Autokorrelationsfunktion (ACF). Die ACF misst (unter Annahme der schwachen Stationarität der Zeitreihe) die paarweisen Korrelationen zwischen zwei Beobachtungen zu den Lags (Zeitabständen) 0, 1, 2, 3, usw. Bei Lag 0 hat die ACF immer den Wert Eins. In der Grafik ist dies gut sichtbar. Die gestrichelten waagrechten Linien um die Null-Linie sind die Annahmehbereiche für einen entsprechenden Test mit  $H_0$ : Autokorrelation zu Lag  $l$  ist 0. Für alle Lags ist die Autokorrelation nicht signifikant von 0 verschieden, d.h., die Autokorrelationen befinden sich im Annahmehbereich des Test.
  - ii. Partielle Autokorrelationsfunktion: Die Werte sind für alle Lags nahe dem Wert Null (Lag 0 wird hier in der Grafik nicht abgetragen, die Funktion beginnt bei Lag 1), d.h. sie liegen wieder im Annahmehbereich des entsprechenden Tests mit  $H_0 : PACF(l) = 0$ . Die partielle ACF gibt die um Zwischenwerte bereinigte Korrelation an.
  - iii. Residuen-Plot: Die Residuen streuen um den Wert Null.
- (v) [2 Punkte] Die Werte der (partiellen) Autokorrelationsfunktion suggerieren ein Modell ohne (partielle) Korrelationen, dazu passt eine Zeitreihe mit unabhängigen Beobachtungen.
- (vi) [2 Punkte] Aufgrund der Struktur in der ACF, PACF und Residuen wird die Verwendung eines Zeitreihenmodelles nicht hilfreich sein für die Modellierung der kurzfristige Prognose.

**Aufgabe 5. [Credibility] [30 Punkte]**

Ein Versicherungstarif soll nach Regionen differenziert werden. Für die Schadenhöhen  $X$  in den Regionen  $i = 1, \dots, 6$  liegen jeweils Beobachtungen aus  $n = 5$  Jahren mit den in der folgenden Tabelle gegebenen empirischen Mittelwerten  $\bar{X}_i$  und Varianzen  $\hat{\sigma}_i^2$  vor. Als zusätzliche Information wurden die Quotienten  $\hat{\sigma}_i/\bar{X}_i$  ermittelt:

| $i$                        | 1   | 2   | 3   | 4   | 5   | 6   |
|----------------------------|-----|-----|-----|-----|-----|-----|
| $\bar{X}_i$                | 14  | 20  | 24  | 25  | 10  | 30  |
| $\hat{\sigma}_i^2$         | 49  | 36  | 144 | 100 | 16  | 324 |
| $\hat{\sigma}_i/\bar{X}_i$ | 0,5 | 0,3 | 0,5 | 0,4 | 0,4 | 0,6 |

- (a) [9 Punkte] Berechnen Sie auf dieser Basis den Wert der Credibility-Prämie nach Bühlmann für die Region 1.

(Hinweis: Ohne Nachzurechnen können Sie verwenden, dass die empirische Varianz der Messwerte 14, 20, 24, 25, 10, 30 von  $\bar{X}_i$  den Wert 55,1 hat.)

- (b) [19 Punkte] In Region 1 sei die Schadenhöhe  $X$  lognormalverteilt mit den Parametern  $\vartheta$  und  $s^2$ , so dass  $E(X) = \exp(\vartheta + 0,5 s^2)$  und  $Var(X) = (\exp(s^2) - 1) \cdot \exp(2\vartheta + s^2)$  gilt.

- (i) [2 Punkte] Weisen Sie nach, dass der zugehörige Variationskoeffizient  $\kappa$  (der Quotient aus Standardabweichung und Erwartungswert)

$$\kappa := \frac{\sqrt{Var(X)}}{E(X)} = \sqrt{\exp(s^2) - 1}$$

beträgt.

- (ii) [2 Punkte] Es wird angenommen, dass der Variationskoeffizient  $\kappa$  der Schadenhöhenverteilung in allen Regionen denselben Wert habe. Berechnen Sie den Schätzer

$$\hat{\kappa} = \frac{1}{6} \sum_{i=1}^6 \frac{\hat{\sigma}_i}{\bar{X}_i}$$

für diesen Wert.

- (iii) [15 Punkte] In einem Bayes'schen Credibility-Modell für die Region 1, bei dem die Schadenhöhen der oben angegebenen Lognormalverteilung folgen, ist  $\vartheta$  Realisierung eines zufälligen Strukturparameters  $\Theta$ . Als a-priori-Verteilung für den Strukturparameter wird die Gleichverteilung auf dem Intervall  $[2, 4]$  angesetzt.

Weisen Sie zunächst nach, dass sich mit dem Variationskoeffizienten  $\kappa$  aus Aufgabenteil (b)(i)

$$H(\vartheta) := E(X|\Theta = \vartheta) = \exp(\vartheta) \cdot \sqrt{1 + \kappa^2}$$

und

$$\text{Var}(X|\Theta = \vartheta) = \exp(2\vartheta) \cdot \kappa^2 \cdot (1 + \kappa^2)$$

ergibt.

Berechnen Sie damit den Wert der linearisierten Credibility-Prämie für die Region 1. Anstelle von  $\kappa$  können Sie dabei den Wert für  $\hat{\kappa}$  aus (b)(ii) ansetzen.

(Hinweis: Ohne Beweis können Sie verwenden, dass  $E(\exp(\Theta)) = 23,60$  und  $E(\exp(2\Theta)) = 731,59$  sowie  $\text{Var}(\exp(\Theta)) = 174,63$ . Falls Sie Aufgabenteil (b)(ii) nicht gelöst haben, verwenden Sie  $\kappa = 0,4$ .)

- (c) [2 Punkte] Welchen Vorteil bietet die Modellierung mit einem Bühlmann-Modell wie in Aufgabenteil (a) gegenüber dem Bayes'schen Credibility-Modell wie in Aufgabenteil (b)? Begründen Sie Ihre Antwort.

**Lösung:**

(a) Die Credibility-Prämie nach Bühlmann hat für Region 1 die Form

$$H_1^{**} = \hat{z} \cdot \bar{X}_1 + (1 - \hat{z}) \cdot \hat{E}(\mu(\Theta))$$

mit

$$\hat{E}(\mu(\Theta)) = \frac{1}{6} \sum_{i=1}^6 \bar{X}_i = 20,5.$$

Zur Ermittlung des Credibility-Faktors  $\hat{z}$  benötigt man die Schätzer

$$\hat{E}(\sigma^2(\Theta)) = \frac{1}{6} \sum_{i=1}^6 \hat{\sigma}_i^2 = 111,5$$

$$\widehat{Var}(\bar{\mathbf{X}}) = \frac{1}{5} \sum_{i=1}^6 (\bar{X}_i - \hat{E}(\mu(\Theta)))^2 = 55,1.$$

Mit  $n = 5$  erhält man

$$\hat{z} = 1 - \frac{\hat{E}(\sigma^2(\Theta))}{n \cdot \widehat{Var}(\bar{\mathbf{X}})} = 1 - \frac{111,5}{5 \cdot 55,1} = 0,5953.$$

Der Wert der Credibility-Prämie nach Bühlmann ist somit

$$H_1^{**} = 0,5953 \cdot 14 + 0,4047 \cdot 20,5 = 16,63.$$

(b) (i) Für den Variationskoeffizienten ergibt sich

$$\begin{aligned} \kappa &= \frac{\sqrt{Var(X)}}{E(X)} = \frac{\sqrt{(\exp(s^2) - 1) \cdot \exp(2\vartheta + s^2)}}{\exp(\vartheta + 0,5 s^2)} \\ &= \frac{\sqrt{(\exp(s^2) - 1) \cdot \exp(\vartheta + 0,5 s^2)}}{\exp(\vartheta + 0,5 s^2)} = \sqrt{\exp(s^2) - 1}. \end{aligned}$$

(ii) Auf Basis der Werte der Tabelle ergibt sich

$$\hat{\kappa} = \frac{1}{6} \sum_{i=1}^6 \frac{\hat{\sigma}_i}{\bar{X}_i} = 0,45.$$

- (iii) Zunächst berechnet man mit der angegebenen Formel für den Erwartungswert der Lognormalverteilung und der in Aufgabenteil (b)(i) gewonnenen Formel für  $\kappa$

$$\begin{aligned} H(\vartheta) &= E(X|\Theta = \vartheta) = \exp(\vartheta + 0,5 s^2) = \exp(\vartheta) \cdot \exp(0,5 s^2) \\ &= \exp(\vartheta) \cdot \sqrt{1 + \kappa^2}. \end{aligned}$$

Aus der Definition des Variationskoeffizienten und der obigen Formel für  $H(\vartheta)$  ergibt sich

$$Var(X|\Theta = \vartheta) = \kappa^2 \cdot (E(X|\Theta = \vartheta))^2 = \kappa^2 \cdot (H(\vartheta))^2 = \exp(2\vartheta) \cdot \kappa^2 \cdot (1 + \kappa^2).$$

Mit dem Hinweis erhält man

$$Var(H(\Theta)) = (1 + \kappa^2) \cdot Var(\exp(\Theta)) = (1 + 0,45^2) \cdot 174,63 = 209,99.$$

sowie

$$\begin{aligned} E(Var(X|\Theta)) &= \kappa^2 \cdot (1 + \kappa^2) \cdot E(\exp(2\Theta)) = 0,45^2 \cdot (1 + 0,45^2) \cdot 731,59 \\ &= 178,15. \end{aligned}$$

Der Credibility-Faktor beträgt somit

$$z = \frac{Var(H(\Theta))}{\frac{1}{n} E(Var(X|\Theta)) + Var(H(\Theta))} = \frac{209,99}{\frac{178,15}{5} + 209,99} = 0,8549.$$

Mit

$$E(X) = E(H(\Theta)) = \sqrt{1 + \kappa^2} \cdot E(\exp(\Theta)) = \sqrt{1 + 0,45^2} \cdot 23,60 = 25,88$$

erhält man schließlich den Wert der linearisierten Credibility-Prämie

$$H^{**} = z \cdot \bar{X}_1 + (1 - z) \cdot E(X) = 0,8549 \cdot 14 + 0,1451 \cdot 25,88 = 15,72.$$

(Mit  $\kappa = 0,4$  ergäbe sich  $H^{**} = 15,35$ .)

- (c) Der Vorteil des Bühlmann-Modells liegt darin, dass es ein verteilungsfreies Verfahren ist (also insbesondere keiner konkret spezifizierte Schadenhöhen- bzw. a-priori-Verteilung bedarf).

**Aufgabe 6.** [Stochastische Prozesse] [30 Punkte]

- (a) [21 Punkte] Die Zugehörigkeit eines versicherten Risikos zu einer der drei gegebenen Risikoklassen 1, 2 und 3 wird jedes Jahr neu ermittelt. Die mathematische Beschreibung geschieht mit Hilfe einer Markovkette  $\{X_t : t \in \mathbb{N}_0\}$  mit Zustandsraum  $S = \{1, 2, 3\}$ . Dabei bezeichnet  $X_t$  die Risikoklasse im Jahr  $t$ . Die Übergangsmatrix  $\Pi$  des Prozesses sei

$$\Pi = \begin{pmatrix} 0,6 & 0,3 & 0,1 \\ 0,2 & 0,5 & 0,3 \\ 0,1 & 0,2 & 0,7 \end{pmatrix}$$

- (i) [5 Punkte] Ein neues Risiko wurde im Jahr  $t = 0$  in Klasse 2 eingestuft. Mit welcher Wahrscheinlichkeit ist das Risiko im Jahr  $t = 2$  in Klasse 3?
- (ii) [4 Punkte] Mit welcher Wahrscheinlichkeit ist ein Risiko im Jahr  $t = 5$  in Klasse 3, wenn es zu Beginn zufällig (d.h. gleichverteilt) in eine der drei Klassen eingestuft wurde? Sie dürfen verwenden, dass (gerundet)

$$\Pi^5 = \begin{pmatrix} 0,2888 & 0,3328 & 0,3784 \\ 0,2650 & 0,3246 & 0,4104 \\ 0,2490 & 0,3168 & 0,4342 \end{pmatrix}.$$

- (iii) [2 Punkte] Eine Berechnung der Eigenwerte von  $\Pi$  in R ergab folgenden Output:

```
> eig <- eigen(P)
> eig$values
[1] 1.0000000 0.5414214 0.2585786
```

Untersuchen Sie, ob die Markovkette eine asymptotische Verteilung besitzt.

- (iv) [10 Punkte] Eine Berechnung der Eigenwerte und Eigenvektoren von  $\Pi^T$  (der transponierten Matrix von  $\Pi$ ) in R ergab folgenden Output:

```
> eigT <- eigen(t(P))
> eigT$values
[1] 1.0000000 0.5414214 0.2585786
> eigT$vectors
      [,1]      [,2]      [,3]
[1,] -0.4511292 -0.5615167  0.3365568
[2,] -0.5513802 -0.2325878 -0.8125199
[3,] -0.7017566  0.7941045  0.4759631
```



Bestimmen mit Hilfe dieser Angaben die asymptotische Verteilung  $\pi = (\pi_1, \pi_2, \pi_3)$  der Markovkette, wobei die  $\pi_i$  auf vier Nachkommastellen anzugeben sind. Erläutern Sie Ihr Vorgehen.

*Hinweis:* Sie dürfen verwenden, dass die Links-Eigenvektoren einer Matrix  $A$  den Rechts-Eigenvektoren von  $A^T$  zum jeweils selben Eigenwert entsprechen.

- (b) [9 Punkte] Entscheiden Sie, ob die folgenden Aussagen richtig oder falsch sind. Begründen Sie Ihre Entscheidungen und im Falle von falschen Aussagen geben Sie zudem die richtige Aussage an. Für jede richtige (falsche) Antwort gibt es drei (null) Punkte.

- (i) Das stochastische Integral

$$\int_0^t e^{t-s} dW_s$$

definiert eine Zufallsvariable, die normalverteilt ist mit Erwartungswert 0 und Varianz  $\frac{1}{2}(1 - e^{-2t})$ .

- (ii) Sei  $\{B_t : t \geq 0\}$  eine Brownsche Bewegung mit Drift  $\mu = 1$  und Volatilität  $\sigma = 3$ . Dann ist  $(B_2, B_3, B_5)^T$  eine 3-dimensionale Normalverteilung mit Kovarianzmatrix

$$\begin{pmatrix} 18 & 18 & 18 \\ 18 & 27 & 27 \\ 18 & 27 & 45 \end{pmatrix}$$

- (iii) Es sei  $\{N_t : t \geq 0\}$  ein Poisson-Prozess mit Intensitätsparameter  $\lambda = 1,25$ . Der Prozess springe zum Zeitpunkt  $t = 7,2$  in den Zustand  $N_{7,2} = 9$ . Dann springt er im Mittel zum Zeitpunkt  $t = 8,45$  in den Zustand 10.

## Lösung Aufgabe 6

(a) (i) Es gilt

$$\Pi^2 = \begin{pmatrix} 0,43 & 0,35 & 0,22 \\ 0,25 & 0,37 & 0,38 \\ 0,17 & 0,27 & 0,56 \end{pmatrix}.$$

Daraus ergibt sich die Verteilung von  $X_2$  unter der Startbedingung  $\mathbf{p}(0) = (0, 1, 0)$  zu

$$\mathbf{p}(2) = \mathbf{p}(0) \cdot \Pi^2 = (0,25, 0,37, 0,38).$$

Also

$$P(X_2 = 3 | X_0 = 2) = 0,38.$$

(ii) Die Startverteilung lautet nun  $\mathbf{p}(0) = (1/3, 1/3, 1/3)$ , daher (auf 4 Nachkommastellen gerundet)

$$\mathbf{p}(2) = \mathbf{p}(0) \cdot \Pi^5 = (0,2676, 0,3247, 0,4077).$$

Die gesuchte Wahrscheinlichkeit lautet 0,4077.

(iii) Da es genau einen Eigenwert vom Betrag 1 gibt und alle Eigenwerte verschieden sind (und somit alle die algebraische Vielfachheit 1 haben), ist der Konvergenzsatz für Markovketten anwendbar, so dass eine asymptotische Verteilung existiert.

(iv) Wenn eine asymptotische Verteilung existiert, kann man sie aus der Lösungsmenge des linearen Gleichungssystems  $\mathbf{p} = \mathbf{p} \cdot \Pi$  gewinnen, indem man den Lösungsvektor sucht, der nur nichtnegative Komponenten mit Summe 1 besitzt. Nun gilt

$$\mathbf{p} = \mathbf{p} \cdot \Pi \Leftrightarrow \mathbf{p}^T = (\mathbf{p} \cdot \Pi)^T = \Pi^T \cdot \mathbf{p}^T.$$

Es reicht somit, einen (Rechts)Eigenvektor zum Eigenwert 1 von  $\Pi^T$  zu kennen. Laut Aufgabenstellung ist dies der Vektor

$$\begin{pmatrix} -0,4511292 \\ -0,5513802 \\ -0,7017566 \end{pmatrix}$$

Da

$$(-0,4511292) + (-0,5513802) + (-0,7017566) = -1,704266$$

lautet die asymptotische Verteilung (auf vier Nachkommastellen gerundet)

$$\begin{aligned} \boldsymbol{\pi} &= \frac{1}{-1,704266} \cdot (-0,4511292, -0,5513802, -0,7017566) \\ &= (0,2647, 0,3235, 0,4118). \end{aligned}$$

- (b) (i) Falsch. Für eine deterministische Funktion  $h$  gilt

$$\int_0^t h(s) dW_s \sim N(0, \sigma^2)$$

mit

$$\sigma^2 = \int_0^t h(s)^2 ds.$$

Wegen

$$\int_0^t (e^{t-s})^2 ds = e^{2t} \left[ -\frac{1}{2} e^{-2s} \right]_0^t = e^{2t} \cdot \frac{1}{2} (1 - e^{-2t})$$

folgt

$$\int_0^t e^{t-s} dW_s \sim N(0, \frac{1}{2} (e^{2t} - 1)).$$

- (ii) Richtig. Für eine Brownsche Bewegung  $\{B_t : t \geq 0\}$  mit Drift  $\mu$  und Volatilität  $\sigma$  ist jeder Zufallsvektor der Form  $(B_{t_1}, \dots, B_{t_n})^\top$   $n$ -dimensional normalverteilt mit Kovarianzen

$$\text{Cov}(B_{t_i}, B_{t_k}) = \sigma^2 \min\{t_i, t_k\}.$$

Mit  $\sigma^2 = 9$ ,  $n = 3$  sowie  $t_1 = 2$ ,  $t_2 = 3$  und  $t_3 = 5$  folgt die angegebene Kovarianzmatrix.

- (iii) Falsch. Beim Poissonprozess ist die Verweildauer im Zustand  $k \in \mathbb{N}_0$  exponentialverteilt mit Parameter  $\lambda$ . Eine solche Verteilung hat den Erwartungswert  $\lambda^{-1}$ , hier also  $1/1,25 = 0,8$ . Der mittlere Zeitpunkt für den Sprung von 9 auf 10 ist daher  $t = 7,2 + 0,8 = 8$ .