Aufgabe 1 (30 Punkte)

Ein Lebensversicherungsunternehmen ermittelt die Rechnungsgrundlagen für eine gemischte Versicherung gegen Einmalbeitrag (zur Vereinfachung: Eintrittsalter: 31, Dauer: 3 Jahre, Unisex-Rechnungsgrundlagen). Das im Lebensversicherungsunternehmen vorliegende statistische Material zur Sterblichkeit ergibt folgende Ergebnisse aus einem geschlossenen Bestand.

Im Zeitraum 1.1.2010 bis 31.12.2011 werden für die Geburtsjahre 1977 bis 1979 folgende Bestände untersucht und Todesfälle festgestellt:

Geburtsjahr	Bestand	Todesfälle
1977	1.000	17
1978	1.500	25
1979	1.800	33

a) Ermitteln Sie die empirischen Sterbehäufigkeiten für jedes der Alter 31 bis 33 nach der Sterbejahrmethode. Falls erforderlich gehen Sie von einer Gleichverteilung der Todeszeitpunkte aus. Erläutern Sie Ihre Vorgehensweise anhand einer Graphik.

b) Man erwartet in den ersten drei Jahren nach Einführung des neuen Tarifs einen Bestand von jeweils 1.000 Neuzugängen. Ermitteln Sie einen einheitlichen relativen Zuschlag bzw. Abschlag auf die Sterbewahrscheinlichkeiten, so dass auf Dauer mit Wahrscheinlichkeit 95 % die Anzahl der tatsächlichen Todesfälle geringer als die Anzahl der erwarteten Todesfälle ist.

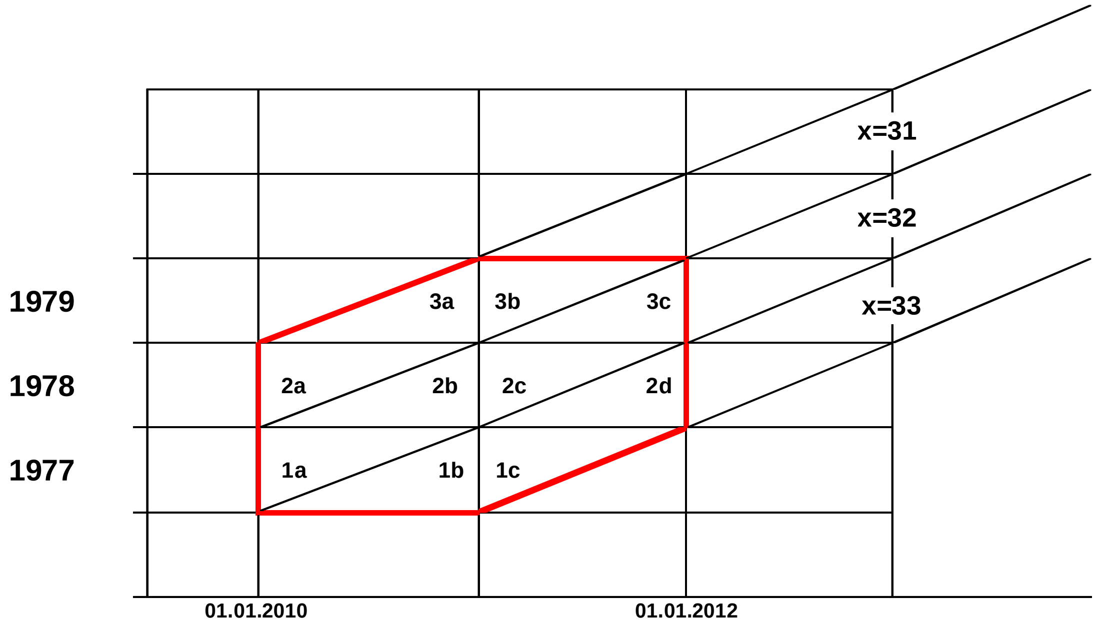
c) Sei K_x das riskierte Kapital für einen Versicherten im Alter x mit $K_{31} = 0{,}3$, $K_{32} = 0{,}2$ und $K_{33} = 0{,}1$. Ermitteln Sie den einjährigen Risikogewinn, der bei rechnungsmäßigem Risikoverlauf aufgrund der eingerechneten Sicherheiten entsteht.

d) Ermitteln Sie mit Hilfe der Gesamtschadenverteilung (Approximation durch Normalverteilung) das zusätzliche Kapital, das erforderlich ist, um sämtliche Leistungsfälle eines Jahres mit Wahrscheinlichkeit 95 % bezahlen zu können.

e) Vergleichen Sie die Ergebnisse von c) und d) und begründen Sie den Unterschied.

Lösung

zu a)



Geburtsjahr	Alter x	L_x am 1.1.2010	T_x [1.1.2010 , 1.1.2012)	Anteil	Bereich
1977	32	1000	17	25%	1a
	33	1000	17	50%	1b und 1c
1978	31	1500	25	25%	2a
	32	1500	25	50%	2b und 2c
	33	1500	25	25%	2d
1979	31	1800	33	50%	3a und 3b
	32	1800	33	25%	3c

Beobachtungszeitraum

$$B = [1.1.2010,\ 1.1.2012)$$

$$\begin{aligned}
q_{31} &= \frac{\left|T_{31}\left(B, 1978 \cup 1979\right)\right|}{\frac{1}{2}\left|L_{31}\left(B, 1978\right)\right| + \left|L_{31}\left(B, 1979\right)\right|} \\
&= \frac{\frac{1}{4} \cdot 25 + \frac{1}{2} \cdot 33}{\frac{1}{2} \cdot 1.500 + 1.800} \\
&= \frac{22,75}{2.550} \\
&= 0,00892
\end{aligned}$$

$$\begin{aligned}
q_{32} &= \frac{\left|T_{32}\left(B, 1977 \cup 1978 \cup 1979\right)\right|}{\frac{1}{2}\left|L_{32}\left(B, 1977\right)\right| + \left|L_{32}\left(B, 1978\right)\right| + \frac{1}{2}\left|L_{32}\left(B, 1979\right)\right|} \\
&= \frac{\frac{1}{4} \cdot 17 + \frac{1}{2} \cdot 25 + \frac{1}{4} \cdot 33}{\frac{1}{2} \cdot 1.000 + 1.500 + \frac{1}{2} \cdot 1.800} \\
&= \frac{25}{2.900} \\
&= 0,00862
\end{aligned}$$

$$\begin{aligned}
q_{33} &= \frac{\left|T_{33}\left(B, 1977 \cup 1978\right)\right|}{\left|L_{33}\left(B, 1977\right)\right| + \frac{1}{2}\left|L_{33}\left(B, 1978\right)\right|} \\
&= \frac{\frac{1}{2} \cdot 17 + \frac{1}{4} \cdot 25}{1.000 + \frac{1}{2} \cdot 1.500} \\
&= \frac{14,75}{1.750} \\
&= 0,00843
\end{aligned}$$

zu b) Für den Lebensversicherer besteht in den Altern 31 bis 33 ein Todesfallrisiko, da die Todesfallsumme in den ersten Versicherungsjahren höher als die bis dahin angesammelte Deckungsrückstellung ist. Um bzgl. des Schwankungsrisikos die geforderte Sicherheit von $1-\alpha = 0,95$ zu erreichen, muss daher für jedes Alter x der gleiche Zuschlag $s^{\alpha} \geq 0$ auf die Sterbewahrscheinlichkeit q_x ermittelt werden, so dass die Wahrscheinlichkeit, dass die Anzahl der Toten kleiner gleich als die erwartete Anzahl von Toten ist, 95% beträgt.

Für alle Alter x = 31,…,33 muss also gelten:

$$P\left(\sum_{x=31}^{33} T_x \leq \sum_{x=31}^{33} L_x \cdot q_x \cdot (1+s^{\alpha})\right) \stackrel{!}{=} 0,95 \qquad (*)$$

mit

L_x Anzahl der Lebenden im Alter x,

T_x Anzahl der Toten im Alter x,

s^{α} rel. Zuschlag

Die Anzahl der Toten T_x lässt sich auch darstellen als Summe von unabhängigen Bernoulli-verteilten Zufallsvariablen X_i, i=1,…, L_x, d.h.

$$T_x = \sum_{i=1}^{L_x} X_i \ \text{ mit } X_i \sim B(1;q_x).$$

Daher gilt $T_x \sim \text{Bin}(L_x;q_x)$ mit $ET_x = L_x q_x$ und $\text{Var}(T_x) = L_x q_x (1-q_x)$.

Gleichung (*) ist damit äquivalent zu

$$P\left(\frac{\sum_{x=31}^{33} T_x - E\left(\sum_{x=31}^{33} T_x\right)}{\sqrt{\text{Var}\left(\sum_{x=31}^{33} T_x\right)}} \leq \frac{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1+s^{\alpha}) - \sum_{x=31}^{33} L_x \cdot q_x}{\sqrt{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1-q_x)}}\right) \stackrel{!}{=} 0,95$$

Da die Zufallsvariablen $T_x = \sum_{i=31}^{33} X_i$ näherungsweise normal verteilt sind, ist die Zufallsvariable

$$\frac{\sum_{x=31}^{33} T_x - E\left(\sum_{x=31}^{33} T_x\right)}{\sqrt{\text{Var}\left(\sum_{x=31}^{33} T_x\right)}}$$

näherungsweise standardnormalverteilt und es gilt

$$\frac{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1+s^\alpha) - \sum_{x=31}^{33} L_x \cdot q_x}{\sqrt{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1-q_x)}} = \frac{\sum_{x=31}^{33} L_x \cdot q_x \cdot s^\alpha}{\sqrt{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1-q_x)}} = u_{0,95} \approx 1,64$$

$$\Leftrightarrow s^\alpha = 1,64 \cdot \frac{\sqrt{\sum_{x=31}^{33} L_x \cdot q_x \cdot (1-q_x)}}{\sum_{x=31}^{33} L_x \cdot q_x} = \frac{1,64}{\sqrt{1000}} \cdot \frac{\sqrt{\sum_{x=31}^{33} q_x \cdot (1-q_x)}}{\sum_{x=31}^{33} q_x} \approx 0,3204$$

zu c)

Der Risikogewinn G_x für das Alter x ergibt sich als

$$G_x = K_x (q_x (1+s^\alpha) L_x - q_x L_x)$$

$$= K_x q_x s^\alpha L_x$$

Und damit der Risikogewinn G insgesamt

$$G = \sum_{x=31}^{33} G_x = s^\alpha \sum_{x=31}^{33} K_x q_x L_x = 0,3204 \cdot (2,676 + 1,724 + 0,843) = 1,68$$

zu d)

Bezeichne S den Gesamtschaden für alle Alter insgesamt, dann gilt

$$S = \sum_{x=31}^{33} T_x \cdot K_x$$

und

$$E(S) = \sum_{x=31}^{33} L_x \cdot K_x \cdot q_x = 1.000 \cdot \sum_{x=31}^{33} K_x \cdot q_x = 5,24$$

$$Var(S) = \sum_{x=31}^{33} L_x \cdot K_x^2 \cdot q_x \cdot (1-q_x) = 1.000 \cdot \sum_{x=31}^{33} K_x^2 \cdot q_x \cdot (1-q_x) = 1,22$$

Gesucht ist das zusätzlich erforderliche Kapital S_0, so dass gilt:

$$P(S \le ES + S_0) = P\left(\sum_{x=31}^{33} T_x K_x \le \sum_{x=31}^{33} L_x \cdot q_x \cdot K_x + S_0\right) \overset{!}{=} 0,95$$

Dies ist äquivalent zu

$$P\left(\frac{S-ES}{\sqrt{Var(S)}} \leq \frac{S_0}{\sqrt{Var(S)}}\right) = 0{,}95$$

Da die Zufallsvariable $\frac{S-ES}{\sqrt{Var(S)}}$ näherungsweise standardnormalverteilt ist, folgt

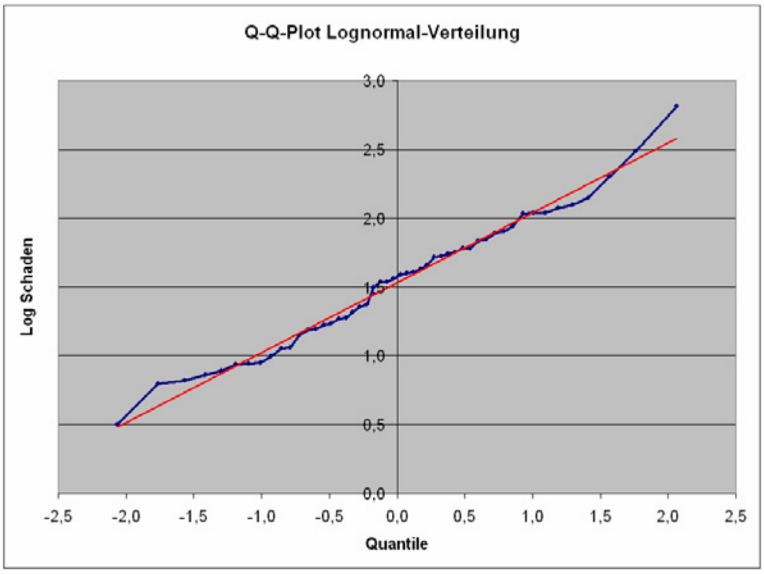
$$S_0 = \sqrt{Var(S)} \cdot u_{95\%} = \sqrt{1{,}22} \cdot 1{,}64 = 1{,}81$$

zu e)

Nach dem Ergebnis von c) führt der Ansatz einer Sicherheitswahrscheinlichkeit von 95% bei Modifikation der Sterbewahrscheinlichkeiten ohne Berücksichtigung des riskierten Kapitals bei rechnungsmäßigem Verlauf zu einem Risikogewinn von 1,68 und damit zu einem zusätzlichen Kapital in gleicher Höhe.

Bei Berücksichtigung des riskierten Kapitals werden die Sterbewahrscheinlichkeiten nicht mehr gleich gewichtet, sondern mit dem jeweils riskierten Kapital. Da das höhere riskierte Kapital mit höheren Sterbewahrscheinlichkeiten einhergeht, wird bei diesem Ansatz ein höheres Risiko festgestellt und damit ein höheres Kapital erforderlich.

<u>Aufgabe 2 (25 Punkte):</u> Die nachfolgende Graphik zeigt einen Q-Q-Plot für logarithmisch transformierte Schäden aus einer trendbereinigten VGV-Historie von 50 Jahren (monetäre Einheit: 1 Mio. EUR).



a) Lesen Sie aus der Graphik (näherungsweise) plausible Schätzer für die Parameter μ und σ der angenommenen Lognormal-Verteilung ab. Begründen Sie Ihre Vorgehensweise.

b) Als Maximum-Likelihood-Schätzer erhält man aus dem Datensatz die Größen $\hat{\mu} = 1{,}53$ und $\hat{\sigma} = 0{,}48$. Bestimmen Sie auf der Grundlage beider Schätzmethoden jeweils einen Schätzer für den Value-at-Risk der zu Grunde liegenden Schadenverteilung zum Sicherheitsniveau 99,5% (Solvency II-Standard).

c) Der Schadenaktuar verwendet die Schätzer in einem Internen Modell, wo er die VGV-Schäden zusammen mit Haftpflicht-Schäden modellieren will, für die er eine Verteilung mit der Dichte

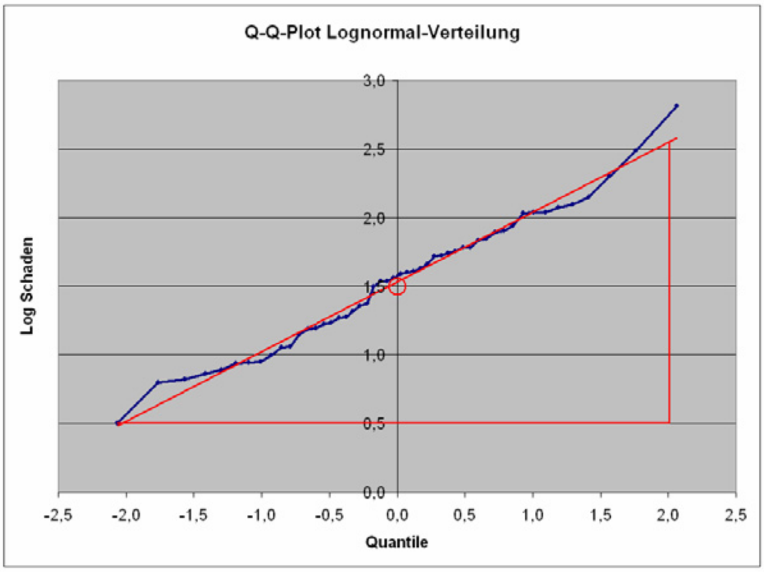
$$f(x) = \frac{2x}{(1+x^2)^2}, \quad x > 0 \text{ (Burr-Verteilung)}$$

unterstellt. Beschreiben Sie eine möglichst einfache Methode, mit der man Haftpflicht-Schäden mit dieser Verteilung simulieren kann. Wenden Sie diese Methode auf der Basis der folgenden Standard-Zufallszahlen U_k [stochastisch unabhängig und jeweils $\mathcal{U}[0,1]$-verteilt] konkret an:

k	1	2	3
U_k	0,4216	0,3041	0,7855

Lösung:

a) Bestimmung von Achsenabschnitt ($\rightarrow \mu$) und Steigung ($\rightarrow \sigma$):



$$\text{Näherungsweise: } \hat{\mu} = 1{,}5; \; \hat{\sigma} = \frac{2{,}5 - 0{,}5}{4} = 0{,}5$$

b) Die Formel für den Value-at-Risk zum Sicherheitsniveau $1-\alpha$ lautet bei Lognormal-Verteilung:

$$\mathrm{VaR}_\alpha(X) = \exp\left(\mu + u_{1-\alpha}\sigma\right)$$

mit dem $1-\alpha$-Quantil $u_{1-\alpha}$ der Standard-Normalverteilung, was zu den jeweiligen Schätzungen

$$\widehat{\mathrm{VaR}_\alpha(X)} = \exp\left(\tilde{\mu} + u_{1-\alpha}\tilde{\sigma}\right)$$

führt. Hier ist $u_{1-\alpha} = u_{0{,}995} = 2{,}58$ (siehe Tabellenanhang zu den Aufgaben); tabellarisch ergibt sich:

$\tilde{\mu}$	$\tilde{\sigma}$	$\widehat{\mathrm{VaR}_{0{,}005}(X)}$
1,5	0,5	16,28
1,53	0,48	15,93

c) Die zugehörige Verteilungsfunktion ist

$$F(x) = \int_0^x f(u)\,du = \left.\frac{1}{1+u^2}\right|_0^x = 1 - \frac{1}{1+x^2}, \quad x > 0.$$

Einfachste Methode: Inversionsmethode mit $F^{-1}(u) = \sqrt{\frac{1}{1-u} - 1} = \sqrt{\frac{u}{1-u}}, \; 0 < u < 1.$

Ergebnis der Simulation:

k	1	2	3
U_k	0,4216	0,3041	0,7855
$F^{-1}(U_k)$	0,8538	0,6611	1,9136

Aufgabe 3 (35 Punkte)

Sei X eine Zufallsvariable mit der Verteilungsfunktion

$$F(x) = \begin{cases} 1-(1+x)^{-\alpha} & \text{falls } x > 0 \\ 0 & \text{sonst.} \end{cases}, \text{ mit } \alpha > 0.$$

a) Beweisen Sie, dass die Verteilung von X zu einer Exponentialfamilie gehört. Bestimmen Sie den natürlichen Parameter η und $b(\eta)$.

b) Bestimmen Sie die Fischerinformation $I(\alpha)$ von X.

c) Man kann zeigen, dass $\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \ln(X_i + 1)}$ ein Maximum Likelihoodschätzer für α ist, wobei $X_1, \ldots, X_n \sim X$ unabhängig seien. Geben Sie die asymptotische Verteilung von $\hat{\alpha}$ an.

d) Bestimmen Sie den Value at Risk $VaR_{0,995}(\alpha)$ von X zum Sicherheitsniveau von 99,5 %.

e) Begründen Sie, warum $VaR_{0,995}(\hat{\alpha})$ ein Maximumlikelihood Schätzer von $VaR_{0,995}(\alpha)$ ist. Geben Sie die asymptotische Verteilung von $VaR_{0,995}(\hat{\alpha})$ an.

Aus einem Datensatz mit 100 Schäden ergibt sich mit der obigen Verteilungsannahme ein Maximumlikelihood Schätzwert von $\hat{\alpha} = 3$.

f) Bestimmen Sie ein asymptotisches 95 %-Schätzintervall für α.

g) Geben Sie mit Hilfe der Maximum-Likelihood-Theorie ein asymptotisches 95 %-Schätzintervall für $VaR_{0,995}(\alpha)$ an.

Quantile der χ^2_m-Verteilung						
m \ ***p***	**0,025**	**0,05**	**0,1**	**0,9**	**0,95**	**0,975**
1	0,00	0,00	0,02	2,71	3,84	5,02
2	0,05	0,10	0,21	4,61	5,99	7,38
14	5,63	6,57	7,79	21,06	23,69	26,12
15	6,26	7,26	8,55	22,31	25,00	27,49
16	6,91	7,96	9,31	23,54	26,30	28,85
99	73,36	77,05	81,45	117,41	123,23	128,42
100	74,22	77,93	82,36	118,50	124,34	129,56

Quantile der Standardnormalverteilung $\mathcal{N}(0,1)$						
p	**0,900**	**0,950**	**0,975**	**0,990**	**0,995**	**0,999**
u_p	1,28	1,64	1,96	2,33	2,58	3,09

Lösung

Zu a) Wegen

$$\alpha(x+1)^{-\alpha-1} = \exp\left(\ln\alpha - (\alpha+1)\ln(x+1)\right) = \exp\left((\alpha+1)\left(-\ln(x+1)\right) - (-\ln\alpha)\right)$$

ergibt sich der natürliche Parameter $\eta = \alpha + 1$ und $b(\eta) = -\ln(\eta - 1)$.

Zu b)

$$\begin{aligned} f(x) &= \alpha(x+1)^{-(\alpha+1)}, x > 0 \\ \ell(\alpha, x) &= \ln\alpha - (\alpha+1)\ln(x+1) \\ \partial_\alpha \ell(\alpha, x) &= \frac{1}{\alpha} - \ln(x+1) \\ \partial^2{}_{\alpha\alpha}\ell(\alpha, x) &= -\frac{1}{\alpha^2} \\ I(\alpha) &= E\left(-\partial^2{}_{\alpha\alpha}\ell(\alpha, X)\right) = \frac{1}{\alpha^2} \end{aligned}$$

Zu c)

Es gilt für den Maximum Likelihood Schätzer asymptotisch $\hat{\alpha} \sim \mathcal{N}\left(\alpha, \frac{1}{nI(\alpha)}\right) = \mathcal{N}\left(\alpha, \frac{\alpha^2}{n}\right)$.

Zu d)

Zu bestimmen ist die Umkehrfunktion von F. Sei $u \in (0,1)$. Dann gilt

$$F(x) = u \Leftrightarrow x = (1-u)^{-1/\alpha} - 1.$$

Damit ergibt sich mit $u = 0{,}995$

$$VaR_{0{,}995}(\alpha) = 200^{1/\alpha} - 1.$$

Zu e)

Die Funktion $h:(0,\infty) \to \mathbb{R}, h(\alpha) = 200^{1/\alpha} - 1$ ist wegen $h'(\alpha) = -\frac{200^{1/\alpha}\ln 200}{\alpha^2} < 0$ streng monoton fallend, also injektiv. Da $\hat{\alpha}$ ein Maximum Likelihood Schätzer ist, folgt dass $h(\hat{\alpha})$ ein Maximum Likelihood Schätzer von $h(\alpha)$ ist mit $h(\hat{\alpha}) \sim \mathcal{N}\left(h(\alpha), \frac{h'(\alpha)^2}{nI(\alpha)}\right)$ asymptotisch.

Zu f)

Aus c) ergibt sich als 95 % Schätzintervall $\left(\hat{\alpha}-\frac{u_{0,975}}{\sqrt{nI(\hat{\alpha})}},\hat{\alpha}+\frac{u_{0,975}}{\sqrt{nI(\hat{\alpha})}}\right)$. Hier erhält man mit $n=100$ und $\hat{\alpha}=3$ $\left(3-\frac{1,96\cdot 3}{10},3+\frac{1,96\cdot 3}{10}\right)=(2,412;3,588)$

Zu g)

Mit der Funktion h aus e) gilt $h(\hat{\alpha})=VaR_{0,995}(\hat{\alpha})$. Es ergibt sich daraus das 95 % Schätzintervall

$$\left(h(\hat{\alpha})-\frac{u_{0,975}\left|h'(\hat{\alpha})\right|}{\sqrt{nI(\hat{\alpha})}},h(\hat{\alpha})+\frac{u_{0,975}\left|h'(\hat{\alpha})\right|}{\sqrt{nI(\hat{\alpha})}}\right)$$

mit $n=100$, $h(3)=4,848$, $h'(3)=-3,443$ also $(2,82;6,87)$.

Aufgabe 4 (35 Punkte):

In einem Kollektiv von 30.000 Kfz-Versicherungsnehmern verteilen sich die Policen wie folgt auf die Merkmale Region (i = 1, 2) und Typklasse (j = 1, 2):

$n_{i,j}$	Typklasse 1	Typklasse 2
Region 1	5.000	10.000
Region 2	5.000	10.000

Beobachtet wurden folgende Schadenanzahlen $N_{i,j}$

$N_{i,j}$	Typklasse 1	Typklasse 2	Gesamt
Region 1	20	130	150
Region 2	10	290	300
Gesamt	30	420	

Daraus ergeben sich folgende Schadenhäufigkeiten $y_{i,j} = N_{i,j} / n_{i,j}$

$y_{i,j}$	Typklasse 1	Typklasse 2
Region 1	0,4%	1,3%
Region 2	0,2%	2,9%

Ein Aktuar möchte mit Hilfe eines Poisson-GLMs (abhängige Variable: Schadenhäufigkeit) entscheiden, ob sich die Regionen 1 und 2 bezüglich der erwarteten Schadenhäufigkeit signifikant unterscheiden. Der Aktuar entscheidet sich dazu für eine Parametrisierung ohne Interaktionen mit dem

Beobachtungsvektor $\begin{pmatrix} 0{,}4\% \\ 1{,}3\% \\ 0{,}2\% \\ 2{,}9\% \end{pmatrix}$, der Designmatrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$ und dem Parametervektor $\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$.

a) Der Aktuar entscheidet sich, die kanonische Linkfunktion zu verwenden.

- Geben Sie die kanonische Link-Funktion $g(\mu)$ an.

- Mit der kanonischen Linkfunktion ergeben sich in den Tarifzellen folgende erwartete Schadenhäufigkeiten:

$\mu_{i,j} = E(y_{i,j})$	Typklasse 1	Typklasse 2
Region 1	$\mu_{1,1} = \exp(\beta_1)$	$\mu_{1,2}$ = ?
Region 2	$\mu_{2,1} = \exp(\beta_1 + \beta_3)$	$\mu_{2,2}$ = ?

  Ermitteln Sie die noch fehlenden Werte $\mu_{1,2}$ und $\mu_{2,2}$ in Abhängigkeit von β_1, β_2 und β_3.

b) Für das betrachtete Poisson-GLM gilt allgemein die Form

$$f(y_{i,j}) = \exp\left\{\frac{n_{i,j}}{\psi}\left(y_{i,j}\vartheta_{i,j} - b(\vartheta_{i,j})\right) + c(y_{i,j}, \psi / n_{i,j})\right\}.$$

- Geben Sie die Funktion $b(\vartheta_{i,j})$ an.
- Welcher Zusammenhang besteht zwischen $\mu_{i,j}$ und $b(\vartheta_{i,j})$? Zeigen Sie, dass $\vartheta_{i,j} = \ln(\mu_{i,j})$.
- Ermitteln Sie die log-likelihood $\ell_{i,j} = \ln f(y_{i,j})$ in Abhängigkeit von $\mu_{i,j}$, $y_{i,j}$, $n_{i,j}$ und Ψ.

Hinweis: Ohne Einschränkung der Allgemeinheit setzen Sie hier und im Folgenden $c(y_{i,j}, \psi / n_{i,j}) = 0$.

c) Da man die kanonische Linkfunktion verwendet, können die Parameter β_1, β_2 und β_3 ohne Auswertung der log-likelihood durch sogenannte Marginalsummenbedingungen geschätzt werden.

- Mit der Tabelle in a) berechnet sich die erwartete Schadenanzahl für die gesamte Typklasse 1 als $E(N_{\text{Typklasse 1}}) = 5.000 \exp(\beta_1) + 5.000 \exp(\beta_1 + \beta_3) = 5.000 \cdot \exp(\beta_1) \cdot (1 + \exp(\beta_3))$. Stellen Sie in entsprechender Weise die erwarteten Schadenanzahlen in Typklasse 2 bzw. Region 1 dar.

- Stellen Sie nun die drei entsprechenden Marginalsummenbedingungen auf und schätzen Sie daraus die Größen $\exp(\beta_1)$, $\exp(\beta_2)$ und $\exp(\beta_3)$.

  Hinweis: Ohne Beweis können Sie verwenden, dass $\exp(\beta_2) = 7 \exp(\beta_1)$. *Setzen Sie dies in* $E(N_{Region\ 1})$ *ein, um* $\exp(\beta_1)$ *zu ermitteln. Damit ist auch* $\exp(\beta_2)$ *bekannt und Sie erhalten* $\exp(\beta_3)$ *aus* $E(N_{Typklasse\ 1})$.

d) Der Aktuar hat in c) die Schätzungen $\exp(\beta_1)$ =0,002; $\exp(\beta_2)$ = 0,014 und $\exp(\beta_3)$ = 2,0 erhalten. Aufgrund einer tiefer gehenden Analyse kann er zudem im Folgenden von $\Psi = 15$ ausgehen.

Berechnen Sie mit diesen Angaben den Schätzwert für $\mu_{2,1}$ (vgl. Tabelle in a) und den Wert der log-likelihood $\ell_{2,1}$ aus b).

e) Durch Auswertung aller $\ell_{i,j}$ erhält der Aktuar eine Gesamt-log-likelihood $\ell_1 = \sum_{i,j} \ell_{i,j}$ = -148,09.

Unter der Nullhypothese $\beta_3 = 0$ ergibt sich dagegen eine Gesamt-log-likelihood von ℓ_0 = -149,79.

Führen Sie einen Likelihood-Quotiententest durch, der die Nullhypothese zum Niveau 5% testet. Formulieren Sie das Testergebnis im Hinblick auf die Fragestellung, ob sich die Schadenhäufigkeiten zwischen den Regionen signifikant unterscheiden.

Hinweis: Quantilswerte für verschiedene χ^2*-Verteilungen sind*

	1%-Quantil	5%-Quantil	95%-Quantil	99%-Quantil
1 Freiheitsgrad	0,000	0,004	3,841	6,635
2 Freiheitsgrade	0,020	0,103	5,991	9,210
3 Freiheitsgrade	0,115	0,352	7,815	11,345
4 Freiheitsgrade	0,297	0,711	9,488	13,277

Lösung:

a) Für das Poisson-GLM ist g(μ) = ln(μ) die kanonische Linkfunktion.

Als lineare Prädiktoren ergeben sich $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_1 + \beta_3 \\ \beta_2 + \beta_3 \end{pmatrix}$.

Die entsprechenden Erwartungswerte ergeben sich aus g^{-1}(linearer Prädiktor) = exp(linearer Prädiktor), so dass

	Typklasse 1	Typklasse 2
Region 1	$\mu_{1,1} = \exp(\beta_1)$	$\mu_{1,2} = \exp(\beta_2)$
Region 2	$\mu_{2,1} = \exp(\beta_1 + \beta_3)$	$\mu_{2,2} = \exp(\beta_2 + \beta_3)$

b) Für das Poisson-GLM gilt $b(\vartheta_{i,j}) = \exp(\vartheta_{i,j})$.

Außerdem besteht der Zusammenhang $\mu_{i,j} = b'(\vartheta_{i,j}) = \exp(\vartheta_{i,j})$, so dass $\vartheta_{i,j} = \ln(\mu_{i,j})$.

Als log-likelihood ergibt sich somit $\ell_{i,j} = \frac{n_{i,j}}{\Psi}\left(y_{i,j} \ln(\mu_{i,j}) - \mu_{i,j}\right)$.

c) Allgemein gilt $E(N_{i,j}) = n_{i,j} \cdot \mu_{i,j}$, so dass

$$E(N_{\text{Typklasse 1}}) = n_{1,1} \exp(\beta_1) + n_{2,1} \exp(\beta_1 + \beta_3) = 5.000 \exp(\beta_1) + 5.000 \exp(\beta_1 + \beta_3) = 5.000 \exp(\beta_1)(1 + \exp(\beta_3))$$
$$E(N_{\text{Typklasse 2}}) = n_{1,2} \exp(\beta_2) + n_{2,2} \exp(\beta_2 + \beta_3) = 10.000 \exp(\beta_2) + 10.000 \exp(\beta_2 + \beta_3) = 10.000 \exp(\beta_2)(1 + \exp(\beta_3))$$
$$E(N_{\text{Region 1}}) = n_{1,1} \exp(\beta_1) + n_{1,2} \exp(\beta_2) = 5.000 \exp(\beta_1) + 10.000 \exp(\beta_2)$$

Die Marginalsummenbedingungen ergeben sich dadurch, dass die Erwartungswerte durch die tatsächlich beobachteten Gesamtschadenanzahlen ersetzt werden, d.h.

$$30 = 5.000 \exp(\beta_1)(1 + \exp(\beta_3))$$
$$420 = 10.000 \exp(\beta_2)(1 + \exp(\beta_3))$$
$$150 = 5.000 \exp(\beta_1) + 10.000 \exp(\beta_2)$$

Dividiert man die zweite durch die erste Gleichung, so erhält man den Hinweis

$$2 \exp(\beta_2) / \exp(\beta_1) = 420 / 30 = 14$$
$$\exp(\beta_2) = 7 \exp(\beta_1)$$

In die dritte Gleichung eingesetzt erhält man:

$$150 = 5.000 \exp(\beta_1) + 70.000 \exp(\beta_1)$$
$$\exp(\beta_1) = 150 / 75.000 = 0{,}002$$

und damit

$$\exp(\beta_2) = 7 \exp(\beta_1) = 7 \cdot 0{,}002 = 0{,}014.$$

Aus der ersten Gleichung ergibt sich schließlich

$\exp(\beta_3) = 30 /(5.000 \exp(\beta_1)) - 1 = 30 / 10 - 1 = 2{,}0$.

d) Es gilt $\mu_{2,1} = \exp(\beta_1 + \beta_3) = 0{,}002 \cdot 2{,}0 = 0{,}4\%$.

Damit ergibt sich $\ell_{2,1} = n_{2,1}\left(y_{2,1} \ln(\mu_{2,1}) - \mu_{2,1}\right)/15 = 5.000 \cdot (0{,}2\% \cdot \ln(0{,}4\%) - 0{,}4\%) / 15 = -5{,}01$.

e) Die Teststatistik beträgt in diesem Fall $W = -2 \cdot (\ell_0 - \ell_1) = -2 \cdot (-149{,}79 + 148{,}09) = 3{,}40$. Unter der Nullhypothese ist W asymptotisch χ^2-verteilt mit einem Freiheitsgrad. Der kritische Wert des Tests beträgt somit $\chi^2_{1,0,95} = 3{,}841$. Da 3,841 > 3,40 kann die Nullhypothese $\beta_3 = 0$ nicht abgelehnt werden.

Die Nullhypothese $\beta_3 = 0$ ist gleichbedeutend damit, dass zwischen den Schadenhäufigkeiten der Regionen 1 und 2 kein signifikanter Unterschied besteht. Der Likelihood-Quotiententest lässt somit nicht erkennen, dass sich die Regionen bezüglich der Schadenhäufigkeit signifikant unterscheiden.