Bericht zur Prüfung im Oktober 2004 über Schadenversicherungsmathematik (Spezialwissen)

Christian Hipp (Karlsruhe) und *Thomas Mack* (München)

Wieder waren 6 Aufgaben gestellt worden. Angegeben ist die maximale Punktzahl für die jeweilige Aufgabe bzw. den Aufgabenteil. Insgesamt waren 180 Punkte möglich. Als Hilfsmittel zugelassen waren die klassische Formelsammlung, die verteilt wurde, sowie ein Taschenrechner. Von den 36 Teilnehmern haben 20 die Klausur bestanden.

1. Aufgabe (30 Punkte) Grundwissen

a) Seien X, Y stochastisch unabhängig mit jeweils einer Phasentypverteilung. Zeigen Sie, dass dann auch

$$\min(X, Y) \text{ und } \max(X, Y)$$

jeweils eine Phasentypverteilung besitzen. (10 Punkte)

b) Betrachten Sie die zweidimensionale Marshall-Olkin Kopula

$$C(u, v) = \min(u^{1-\alpha}v, uv^{1-\beta}), 0 \leq u, v \leq 1, 0 < \alpha, \beta < 1.$$

Um diese Kopula an Daten $(X_i, Y_i), i = 1, ..., n$, anzupassen, kann man Spearman's Rho für die Daten berechnen und dann α und β so wählen, dass das Spearman's Rho der Kopula mit dem Spearman's Rho der Daten übereinstimmt. Rechnen Sie nach, dass das Spearman's Rho der obigen Kopula

$$\frac{3\alpha\beta}{2\alpha + 2\beta - \alpha\beta}$$

ist. Passen Sie $\alpha = \beta$ so an, dass Spearman's Rho 0,5 ist. (20 Punkte)

Anleitung: Benutzen Sie die Beziehung

$$E[UV] = \int_0^1 \int_0^1 C(u, v)dudv,$$

welche für (U, V) mit Verteilungsfunktion $C(u, v)$ gilt.

Lösung:

Zu a)

Die Modelle für die Phasentypverteilung von X bzw. Y seien $\{\Delta_X, 1, ..., I_X\}$ mit Parametern π_X, B_X sowie $\{\Delta_Y, 1, ..., I_Y\}$ mit Parametern π_Y, B_Y. Seien $M_X(t)$ und $M_Y(t)$ die zugehörigen (unabhängigen) Markov-Prozesse so dass

$$X = \inf\{t \geq 0 : M_X(t) = \Delta_X\}$$

$$Y = \inf\{t \geq 0 : M_Y(t) = \Delta_Y\}.$$

Viele Klausurteilnehmer versuchten, mit

$$P\{\min(X, Y) > t\} = P\{X > t\}P\{Y > t\} = \pi_X^T \exp(tB_X)\mathbf{1}\pi_Y^T \exp(tB_Y)\mathbf{1}$$

die gewünschte Darstellung zu erhalten. Dies geht auch:

$$P\{\min(X,Y) > t\} = \pi^T \exp(tB)\mathbf{1} \tag{1}$$

mit $\pi = (\pi(i,k))_{i=1,...,I_X,k=1,...,I_Y}, \pi(i,k) = \pi_X(i)\pi_Y(k)$ sowie

$$B = (b(i,j,k,l))_{i,k=1,...,I_X,j,l=1,...,I_Y}$$

mit

$$b(i,j,k,l) = \left\{ \begin{array}{cc} B_Y(j,l) & \text{wenn } i=k, j \neq l \\ B_X(i,k) & \text{wenn } j=l, i \neq k \\ B_X(i,k) + B_Y(j,l) & \text{wenn } i=k, j=l \\ 0 & \text{sonst .} \end{array} \right.$$

Es ist nicht so einfach, in einer Klausur eine solche Darstellung zu *erraten*, und zudem ist es mühselig, die Gleichheit (1) herzuleiten. Ein anderer, kürzerer Lösungsweg geht über die folgende Definition der Darstellung von Phasentypverteilungen.
Der Prozess $M(t) = (M_X(t), M_Y(t))$ ist ein Markov-Prozess auf dem kartesischen Produkt

$$\{\Delta_X, 1, ..., I_X\} \times \{\Delta_Y, 1, ..., I_Y\}.$$

Dann ist durch

$$\max(X,Y) = \inf\{t : M(t) = (\Delta_X, \Delta_Y)\}$$

eine Darstellung der Verteilung von $\max(X,Y)$ als Phasentypverteilung gefunden. Beim Minimum betrachten wir die Menge

$$D = \{(i, \Delta_Y), i = 1, ..., I_X, (\Delta_X, j), j = 1, ..., I_Y\}.$$

Da D von $M(t)$ - einmal betreten - nicht wieder verlassen wird, können wir D als absorbierenden Zustand auffassen, und damit wird

$$\min(X,Y) = \inf\{t : M(t) \in D\}$$

eine Darstellung der Verteilung von $\min(X,Y)$ als Phasentypverteilung. Die Parameter dieser Phasentypverteilungen sind die für das Minimum oben angegebenen π und B. Für das Maximum sind die Zustände (i, Δ_Y) und (Δ_X, j) nicht absorbierend für $M(t)$, hier werden die Intensitäten

$$\begin{array}{ll} b(i,k,i,\Delta_Y) & = -\sum_{l=1}^{I_Y} B_Y(k,l), \\ b(i,\Delta_Y,j,\Delta_Y) & = B_X(i,j), \\ b(\Delta_X,k,\Delta_X,l) & = B_Y(k,l) \text{ und} \\ b(i,k,\Delta_X,k) & = -\sum_{j=1}^{I_X} B_X(i,j) \end{array}$$

zur Matrix B hinzugefügt.
Für den Fall, dass $X \sim Exp(\alpha)$ und $Y \sim Exp(\beta)$ hat $\max(X,Y)$ die Darstellung als Phasentypverteilung mit $\pi = (1,0,0)$ und

$$B = \begin{pmatrix} -\alpha-\beta & \alpha & \beta \\ 0 & -\beta & 0 \\ 0 & 0 & -\alpha \end{pmatrix}.$$

Zu b)
Wir berechnen zunächst

$$E[UV] = \int_0^1 \int_0^1 C(u,v)dudv = \int_0^1 \int_0^1 \min(u^{1-\alpha}v, uv^{1-\beta})dudv.$$

Das Minimum wird vom ersten Ausdruck $u^{1-\alpha}v$ angenommen genau dann, wenn $u^{1-\alpha}v \leq v^{1-\beta}u$ oder wenn $v \leq u^{\alpha/\beta}$. Damit wird

$$E[UV] = \int_0^1 \int_0^{u^{\alpha/\beta}} u^{1-\alpha}vdvdu + \int_0^1 \int_{u^{\alpha/\beta}}^1 v^{1-\beta}udvdu.$$

Der erste Summand ist

$$\int_0^1 u^{1-\alpha}\frac{1}{2}u^{2\alpha/\beta}du = \frac{1}{2}\frac{1}{2-\alpha+2\alpha/\beta} = \frac{1}{2}\frac{\beta}{2\beta+2\alpha-\alpha\beta}.$$

Der zweite Summand ist

$$\int_0^1 \int_0^{v^{\beta/\alpha}} v^{1-\beta}ududv,$$

und das ist dasselbe wie der erste Term, nur mit vertauschtem α und β, also gleich

$$\frac{1}{2}\frac{\alpha}{2\beta+2\alpha-\alpha\beta}.$$

Spearman's ρ wird daher zu

$$12E[UV] - 3 = \frac{6\alpha+6\beta}{2\alpha+2\beta-\alpha\beta} - 3 = \frac{3\alpha\beta}{2\alpha+2\beta-\alpha\beta}.$$

Dieser Ausdruck nimmt für $0 < \alpha = \beta < 1$ den Wert $1/2$ an genau dann, wenn

$$\frac{3\alpha^2}{4\alpha-\alpha^2} = 1/2$$

oder wenn

$$7\alpha^2 = 4\alpha$$

was nur für $\alpha = 4/7$ möglich ist.

2. Aufgabe (30 Punkte) Tarifierung

Bei Versicherungsverträgen, die pro Jahr mehr als einen Schaden aufweisen können, wird der Gesamtschaden pro Jahr oft durch eine Summenverteilung

$$P = \sum_{k=0}^{\infty} R\{k\}Q^{*k}$$

dargestellt. Aus den Gesamtschäden $X_i, i = 1, ..., n$, (die Einzelschäden liegen nicht vor) sollen die Verteilungen R und Q geschätzt werden. Um die Aufgabe zu vereinfachen wird angenommen, dass $R = Poisson(\lambda)$ und $Q = Exp(a)$ mit $\lambda, a > 0$.

Entwickeln Sie den EM-Algorithmus für die Schätzung von λ und a aus den Beobachtungen $X_1, ..., X_n$. Stellen Sie hierfür jedes X_i als Summe der nicht beobachtbaren Einzelschäden dar,

$$X_i = Z_1^{(i)} + ... + Z_{N_i}^{(i)}$$

und leiten Sie konkret die Gleichungen für λ_{n+1}, a_{n+1} her, welche abstrakt

$$\theta_{n+1} = \arg\max_{\theta} E_{\theta_n}[\log f(X, Y, \theta)|X] \tag{2}$$

lauten. Nehmen Sie hierbei an, dass die in den Gleichungen auftretenden bedingten Erwartungswerte

$$E_{\lambda_n, a_n}[N_i|X_i]$$

gegeben sind, sie müssen nicht berechnet werden.

Lösung:

Der EM-Algorithmus basiert auf einer gemeinsamen Dichte $f(X, Y, \theta)$ der beobachtbaren Größen X und der nicht beobachtbaren Größen Y, die von einem unbekannten Parameter θ abhängt. Der Maximum-Likelihood-Schätzer für θ wird berechnet mit der Iteration (2). Die Iteration wird mit einem beliebigen Startwert θ_0 initialisiert.
Im vorliegenden Beispiel ist $\theta = (\lambda, a), X$ ist der Vektor der Beobachtungen $X_1, ..., X_n$, und Y ist der Vektor der nicht beobachtbaren Schadenanzahlen $N_i, i = 1, ..., n$, die zu den Schadensummen X_i gehören. Die bedingte Verteilung von X_i, gegeben $N_i = k$, ist eine Gammaverteilung mit Parametern a und k, also wird

$$f(X, Y, \theta) = \prod_{i=1}^{n} \frac{\lambda^{N_i}}{N_i!} \exp(-\lambda) \frac{a^{N_i}}{\Gamma(N_i)} x^{N_i - 1} \exp(-aX_i). \tag{3}$$

Der Logarithmus hiervon ist

$$\sum_{i=1}^{n} N_i \log(\lambda) - \log(N_i!) - \lambda + N_i \log(a) - \log(\Gamma(N_i)) + (N_i - 1)\log(x) - aX_i. \tag{4}$$

Weil die Terme

$$\sum_{i=1}^{n} \log(N_i!) \text{ und } \sum_{i=1}^{n} \log(\Gamma(N_i) \text{ sowie } \sum_{i=1}^{n} (N_i - 1)\log(x)$$

durch Veränderung von λ, a nicht beeinflusst werden, kann man sie bei der Maximierung fortlassen. Die Iteration lautet dann

$$\begin{aligned}
\theta_{n+1} &= \arg\max_{a,\lambda} E_{\theta_n}[\sum_{i=1}^{n} N_i \log(\lambda) - \lambda + N_i \log(a) - aX_i|X] \\
&= \arg\max_{a,\lambda}[\sum_{i=1}^{n} E_{\theta_n}[N_i|X] \log(\lambda) - \lambda + E_{\theta_n}[N_i|X] \log(a) - aX_i
\end{aligned}$$

Die bedingten Erwartungswerte $\bar{N}_i = E_{\theta_n}[N_i|X] = E_{\theta_n}[N_i|X_i]$ sind als gegeben vorausgesetzt. Damit wird das Maximum angenommen an den Werten

$$\lambda_{n+1} = \frac{1}{n} \sum_{i=1}^{n} \bar{N}_i \text{ und } a_{n+1} = \frac{\sum_{i=1}^{n} \bar{N}_i}{\sum_{i=1}^{n} X_i}.$$

3. Aufgabe (30 Punkte) Solvabilität

Die Gesamtschadenverteilung S eines Versicherungsbestandes wird häufig durch eine poissonsche Summenverteilung

$$P = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \exp(-\lambda) Q^{*k}$$

modelliert, wobei durch Q das Risiko von Großschäden abgebildet werden kann. In diesem Modell wird eine kritische Schwelle s berechnet, mit der $\mathbb{P}\{S > s\}$ hinreichend klein wird. Sei $\bar{s}$ die kritische Schwelle, für welche $\mathbb{P}\{S > \bar{s}\} = \frac{1}{2}\mathbb{P}\{S > s\}$ gilt. Für subexponentielles Q (also z.B. für eine Paretoverteilung, eine Weibullverteilung mit $\beta < 1$ oder eine Lognormalverteilung) gilt dann

$$\mathbb{P}\{S > s\} \sim \lambda Q(s, \infty),$$

wobei $a(s) \sim b(s)$ ($a(s)$ verhält sich wie $b(s)$ für $s \to \infty$) heißen soll: $a(s)/b(s) \to 1, s \to \infty$. Zeigen Sie mit dieser Asymptotik:

a) Ist Q eine Paretoverteilung mit $Q(t, \infty) = t^{-a}, t > 1, a > 0$, dann verhält sich $\bar{s} - s$ wie Cs für $s \to \infty$, wobei $C > 0$ eine Konstante ist. (8 Punkte)

b) Ist Q eine Weibullverteilung mit $Q(t, \infty) = \exp(-t^\beta), t > 0, 0 < \beta < 1$, dann verhält sich $\bar{s} - s$ wie $Cs^{1-\beta}$ für $s \to \infty$, wobei $C > 0$ eine (andere) Konstante ist. (8 Punkte)

c) Ist Q eine Lognormalverteilung mit $Q(t, \infty) = 1 - \Phi((\log t - \mu)/\sigma), t \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0$, dann gilt für beliebiges $0 < \beta < 1, C > 0$ und $\bar{s} = s + Cs^\beta$

$$\mathbb{P}\{S > \bar{s}\}/\mathbb{P}\{S > s\} \to 1, s \to \infty.$$

Ist andererseits $\bar{s} = s + Cs$ mit $C > 0$, dann gilt

$$\mathbb{P}\{S > \bar{s}\}/\mathbb{P}\{S > s\} \to 0, s \to \infty.$$

(14 Punkte)

Anleitung: Benutzen Sie die Asymptotik

$$1 - \Phi(s) \sim \frac{1}{s}\phi(s).$$

Lösung:

Die Gleichung

$$\mathbb{P}\{S > \bar{s}\} = \frac{1}{2}\mathbb{P}\{S > s\}$$

impliziert

$$\lambda Q(\bar{s}, \infty) \sim \frac{\lambda}{2} Q(s, \infty). \tag{5}$$

Zu a)

Für $Q(s, \infty) = s^{-a}$ ergibt sich

$$\bar{s}^{-a} \sim \frac{1}{2}s^{-a} \text{ oder } \left(\frac{\bar{s}}{s}\right)^{-a} \to \frac{1}{2} \text{ oder } \frac{\bar{s}}{s} \to 2^{1/a}$$

und damit die Behauptung mit $C = 2^{1/a}$.

Zu b)

Für $Q(s,\infty) = \exp(-s^\beta)$ ergibt sich

$$\exp(-\bar{s}^\beta) \sim \frac{1}{2}\exp(-s^\beta) \text{ oder } \exp(-(\bar{s}^\beta - s^\beta)) \to \frac{1}{2}$$

und damit

$$\bar{s}^\beta - s^\beta \to \log(2).$$

Daraus erhält man zunächst $\bar{s}/s \to 1$. Nach Taylor erhält man

$$\bar{s}^\beta - s^\beta = \beta s^{\beta-1}(\bar{s} - s) + R \tag{6}$$

mit

$$R = \beta(\beta-1)x^{\beta-2}(\bar{s}-s)^2,$$

wobei $s \leq x \leq \bar{s}$ und damit

$$|R| \leq \beta(1-\beta)s^{\beta-2}(\bar{s}-s)^2$$

und weiter

$$|R/(\beta s^{\beta-1}(\bar{s}-s))| \leq (1-\beta)\frac{\bar{s}-s}{s} \to 0.$$

Mit (6) erhält man daraus

$$\beta s^{\beta-1}(\bar{s}-s) \to \log(2)$$

und damit die Behauptung mit $C = \log(2)/\beta$.

Zu c)

Nach Anleitung gilt

$$\frac{P(\bar{s},\infty)}{P(s,\infty)} \sim \frac{\phi((\log(\bar{s})-\mu)/\sigma)}{\phi((\log(s)-\mu)/\sigma)}\frac{\log(s)-\mu}{\log(\bar{s})-\mu}. \tag{7}$$

Der zweite Bruch konvergiert in beiden Fällen gegen Eins wegen

$$\frac{\log(s)-\mu}{\log(\bar{s})-\mu} \sim \frac{\log(s)}{\log(\bar{s})} = 1 + \log(\frac{s}{\bar{s}})/\log(s) \to 1.$$

Somit muss nur der erste Bruch auf der linken Seite von (7) untersucht werden, nennen wir ihn A.
Dann wird

$$\begin{aligned} -2\sigma^2\log(A) &= (\log(\bar{s})-\mu)^2 - (\log(s)-\mu)^2 \\ &= (\log(\bar{s}) + \log(s) - 2\mu)(\log(\bar{s}/s)) \end{aligned}$$

Sei zunächst $\bar{s} = s + Cs^\beta$. Dann ist $\bar{s}/s = 1 + Cs^{\beta-1}$ und wegen $\log(1+x) < x$ auch $\log(\bar{s}/s)) \leq Cs^{\beta-1}$, und damit erhalten wir wegen $\bar{s} \leq s(1+C)$ – und weil $s^{\beta-1}$ schneller gegen Null konvergiert als $\log(s)$ – die Konvergenz $\log(A) \to 0$ oder $A \to 1$.
Sei nun $\bar{s} = s(1+C)$. Dann gilt $\log(\bar{s}/s) = \log(1+C) > 0$, und damit erhalten wir die Konvergenz $\log(A) \to -\infty$ oder $A \to 0$.

4. Aufgabe (30 Punkte) Tarifkalkulation

Bei der Tarifkalkulation werden häufig Verallgemeinerte Lineare Modelle (GLMs) eingesetzt. Diese modellieren für die Schadenvariable Z_i mehrerer Risikogruppen $i = 1, 2, ...$ in einer festen Periode (Jahr) einen individuellen Erwartungswert $E(Z_i) = \mu_i$ und eine Varianz der Form $Var(Z_i) = \phi V(\mu_i)/v_i$ mit bekanntem Volumen v_i.

a) Geben Sie für die PKW-Haftpflicht-Versicherung drei verschiedene sinnvolle konkrete Wahlmöglichkeiten für Paare $(v_i,\ Z_i)$ an. (3 Punkte)

b) Die Varianzannahme der GLMs unterstellt für jede Risikogruppe eine umgekehrte Volumen-Proportionalität der Varianz. Wählen Sie ein Paar (v, Z) aus a), formulieren Sie dafür ein geeignetes möglichst einfaches Mikromodell für eine einzelne Risikogruppe und weisen Sie damit die unterstellte Volumenabhängigkeit der Varianz nach. (4 Punkte)

c) In der PKW-Haftpflicht-Versicherung kommen häufig unterjährige Policenlaufzeiten vor. Geben Sie ein Mikromodell für den Gesamtschaden R_k von Policen $k = 1, 2, ...$ mit unterschiedlich langen unterjährigen Laufzeiten t_k, $0 < t_k \leq 1$, an und weisen Sie damit für ein geeignetes Paar (v, Z) die umgekehrte Volumenproportionalität der Varianz $Var(Z)$ innerhalb der Gruppe nach. (6 Punkte)

d) Jeder KH-Versicherer hat für die Risikogruppen seines Portefeuilles detailliertere Informationen als nur Volumen v und Schadenvariable Z pro Risikogruppe und Jahr, nämlich Höhe, Zeitpunkt und betroffenes Risiko jedes einzelnen Schadens. Geben Sie – basierend auf diesen detaillierteren Daten eines Jahres – für eine einzelne Risikogruppe einen erwartungstreuen Schätzer für den Parameter σ^2 der etwas allgemeineren Modellannahme $Var(Z) = \sigma^2/v$ an. (6 Punkte)

e) Wie kann man in einer einzelnen genügend großen Risikogruppe anhand der detaillierten Daten eines Jahres (vgl. d) die Gültigkeit der Modellannahme $Var(Z) = \sigma^2/v$ der umgekehrten Volumenproportionalität prüfen? (6 Punkte)

f) Nachdem die Volumenabhängigkeit pro Risikogruppe geklärt ist, kehren wir wieder zum eingangs angegebenen GLM-Modell für mehrere Risikogruppen zugleich zurück. Für die sogenannte Varianzfunktion V wird häufig der Ansatz $V(\mu) = \mu^\zeta$ gewählt mit $\zeta > 0$. Damit ergibt sich die Varianzannahme $Var(Z_i) = \phi\mu_i^\zeta/v_i$. Vor dem Einsatz des GLM soll die hierbei gemachte spezielle Annahme $\sigma_i^2 := \phi\mu_i^\zeta$ für die Form des Varianzparameters geprüft werden. Dazu liegen Ihnen (neben den Volumina v_i) noch Schätzer $\hat{\mu}_i$ und $\hat{\sigma}_i^2$ vor, die wie in d) innerhalb der jeweiligen Gruppe i anhand von detaillierteren Daten ermittelt wurden. Geben Sie an, wie man aus diesen Informationen einen Plot erstellen kann, mit dem zugleich die Geeignetheit der σ_i^2-Annahme geprüft wird und Schätzer für die unbekannten Parameter ϕ und ζ gewonnen werden. (5 Punkte)

Lösung:

Zu a)

Wegen der umgekehrten Volumenproportionalität der Varianz muss Z_i mittels Division durch v_i entstanden sein, z.B.:
v_{i1} = Anzahl Jahreseinheiten, Z_{i1} = Schadenbedarf = Gesamtschaden/v_{i1};
v_{i2} = Anzahl Jahreseinheiten, Z_{i2} = Schadenfrequenz = Anzahl Schäden/v_{i2};

v_{i3} = Anzahl Schäden, Z_{i3} = Durchschnittsschaden = Gesamtschaden/v_{i3}.

Zu b)

Die Risikogruppe bestehe aus n iid Risiken $R_1, R_2, ..., R_n$, wobei R_k den Gesamtschaden von Risiko k bezeichnet mit $E(R_k) = \mu$, $Var(R_k) = \sigma^2$. Für den Gesamtschaden $S = R_1 + ... + R_n$ der Gruppe gilt dann $E(S) = n\mu$, $Var(S) = n\sigma^2$, und die Anzahl Jahreseinheiten ist $v = n$. Für den Schadenbedarf $Z = S/v$ (gemäß dem 1. Beispiel von Teil a) gilt dann $Var(Z) = \sigma^2/v$.

Zu c)

Sei $E(R_k) = t_k\mu$ und $Var(R_k) = t_k\sigma^2$. Dann gilt für den Gesamtschaden $S = R_1 + ... + R_n$, $E(S) = v\mu$, $Var(S) = v\sigma^2$, wobei $v = t_1 + ... + t_n$ die Anzahl Jahreseinheiten ist. Für den Schadenbedarf $Z = S/v$ gilt dann $Var(Z) = \sigma^2/v$.

Zu d)

Im Mikromodell von c), dessen Daten dem Versicherer vorliegen (R_k = Gesamtschaden von Risiko k innerhalb seiner Policen-Laufzeit), ist jedes R_k/t_k ein erwartungstreuer Schätzer für μ mit $E\left(\frac{R_k}{t_k} - \mu\right)^2 = Var\left(\frac{R_k}{t_k}\right) = \frac{\sigma^2}{t_k}$. Daher ist der Erwartungswert von $t_k\left(\frac{R_k}{t_k} - \mu\right)^2$ ebenso wie der von $\frac{1}{n}\sum_{k=1}^{n} t_k\left(\frac{R_k}{t_k} - \mu\right)^2$ gleich σ^2. Ersetzt man μ durch $\hat{\mu} = \sum_{k=1}^{n} R_k \Big/ \sum_{k=1}^{n} t_k$ und korrigiert für den verlorenen Freiheitsgrad, so ergibt sich $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{k=1}^{n} t_k\left(\frac{R_k}{t_k} - \hat{\mu}\right)^2$ als erwartungstreuer Schätzer für σ^2. Der analoge Schätzer im Mikromodell von b) lautet $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{k=1}^{n}(R_k - \hat{\mu})^2 \; mit \; \hat{\mu} = \frac{1}{n}\sum_{k=1}^{n} R_k$ und ist ebenfalls eine gültige Lösung.

Zu e)

Man zerlegt die Risikogruppe nach einem von Volumen und Schadenanfall möglichst unabhängigen Kriterium, z.B. dem Geburtstag (1. bis 31.) des Versicherungsnehmers, in ca. 10 verschieden große Teilgruppen und schätzt den Parameter σ^2 von $Var(Z)$ für jede Teilgruppe wie in d). Wenn die so erhaltenen σ^2-Schätzer keine Abhängigkeit vom Teilgruppenvolumen aufweisen (und nicht signifikant unterschiedlich ausfallen), stützen die Daten die umgekehrte Volumenproportionalität der Varianz, denn dann ist $\hat{\sigma}^2_{(i)} \overset{(H_0)}{=} v_i \cdot \hat{V}ar(Z_i)$ bis auf Zufallseffekte konstant, d.h. $Var(Z_i) \approx \hat{\sigma}^2/v_i$.

Zu f)

Man plotte $y_i = \ln\left(\hat{\sigma}_i^2\right)$ gegen $x_i = \ln(\hat{\mu}_i)$. Wenn die Annahme $\sigma_i^2 = \phi\mu_i^\zeta$ zutrifft, sollten die Punkte zufällig um die Gerade $y = \ln(\phi) + \zeta x$ streuen. Dies würde die Annahme über σ_i^2 bestätigen. Schätzer für $\ln(\phi)$ und ζ ergeben sich dann direkt aus y-Abschnitt und Steigung der Regressionsgeraden.

5. Aufgabe (30 Punkte) Schadenreservierung

C_{ik} sei der Schadenstand (kumulativ) von Anfalljahr i nach k Entwicklungsjahren, $1 \leq i \leq n$, $1 \leq k \leq n$.

a) Wie lauten die 3 Grundannahmen des stochastischen Modells, mit dem das Chain-Ladder-Verfahren begründet werden kann? (3 Punkte)

b) Geben Sie die klassischen Schätzer $\hat{f}_k$ für den Abwicklungsfaktor f_k von k nach $k+1$ und $\hat{C}_{ik}$ für den noch nicht beobachteten Schadenstand C_{ik}, $i+k > n+1$, gemäß dem Chain-Ladder-Verfahren an. (2 Punkte)

c) Beweisen Sie, dass $\hat{f}_k = \sum\limits_{i=1}^{n} \hat{C}_{i,k+1} \Big/ \sum\limits_{i=1}^{n} \hat{C}_{ik}$ gilt (beachten Sie jeweils die obere Summationsgrenze), wobei zur Vereinfachung der Schreibweise $\hat{C}_{ik} := C_{ik}$ für $i+k \leq n+1$ gesetzt wurde. (5 Punkte)

d) Sei nun $\{D_{ik}\}$ ein zweites, ebenso großes Abwicklungsschema mit analoger Bezeichnungsweise D_{ik} wie C_{ik}, g_k wie f_k und analogen Schätzern. Beweisen Sie folgende Identität für $i+k > n+1$: (Hinweis: Formen Sie $\hat{C}_{ik}/\hat{D}_{ik}$ um.)

$$\frac{\hat{C}_{ik} \Big/ \hat{D}_{ik}}{\sum\limits_{j=1}^{n} \hat{C}_{jk} \Big/ \sum\limits_{j=1}^{n} \hat{D}_{jk}} = \frac{\hat{C}_{i,n+1-i} \Big/ \hat{D}_{i,n+1-i}}{\sum\limits_{j=1}^{n} \hat{C}_{j,n+1-i} \Big/ \sum\limits_{j=1}^{n} \hat{D}_{j,n+1-i}}.$$

(5 Punkte)

e) Legen Sie dar, welche unerfreuliche Folgerung sich aus d) ergibt, wenn $\{C_{ik}\}$ die Entwicklung der Zahlungen und $\{D_{ik}\}$ die Entwicklung der Aufwendungen (= Zahlungen + Einzelfallreserven) desselben Geschäftssegments sind. (5 Punkte)

f) Welcher qualitative statistische Zusammenhang zwischen den individuellen Abwicklungsfaktoren $C_{i,k+1}/C_{ik}$ im beobachteten Zahlungsdreieck und den beobachteten Zahlung/Aufwand-Ständen $\{C_{ik}/D_{ik}\}$ kann angesichts von d) in der Regel in der Praxis festgestellt werden? (5 Punkte)

g) Folgt aus dem Bisherigen, dass das unter a) dargestellte stochastische Modell in der Regel in der Praxis nicht erfüllt ist (mit Begründung)? (5 Punkte)

Lösung:

Zu a)

Es gibt Parameter $f_1, ..., f_{n-1}$ und $\sigma_1, ..., \sigma_{n-1}$ mit

$$E(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik} f_k,$$
$$Var(C_{i,k+1}|C_{i1}, ..., C_{ik}) = C_{ik} \sigma_k^2.$$

$\{C_{i1}, \ldots, C_{in}\}, 1 \leq i \leq n$, sind unabhängig.

Zu b)

$$\hat{f}_k = \sum_{i=1}^{n-k} C_{i,k+1} \Big/ \sum_{i=1}^{n-k} C_{ik}\,, \;\; \hat{C}_{ik} = C_{i,n+1-i} \hat{f}_{n+1-i} \cdot ... \cdot \hat{f}_{k-1}, \;\; \text{insbes. } \hat{C}_{i,k+1} = \hat{C}_{ik} \hat{f}_k.$$

Zu c)

$$\begin{aligned}\sum_{i=1}^{n}\hat{C}_{i,k+1} &= \sum_{i=1}^{n-k}C_{i,k+1}+\sum_{i=n-k+1}^{n}\hat{C}_{i,k+1}\\ &\overset{(b)}{=} \left(\sum_{i=1}^{n-k}C_{i,k}\right)\hat{f}_k+\sum_{i=n-k+1}^{n}\left(\hat{C}_{ik}\hat{f}_k\right)=\left(\sum_{i=1}^{n}\hat{C}_{ik}\right)\hat{f}_k\end{aligned}$$

Zu d)

$$\frac{\hat{C}_{ik}}{\hat{D}_{ik}} \overset{(b)}{=} \frac{C_{i,n+1-i}\hat{f}_{n+1-i}\cdot\ldots\cdot\hat{f}_{k-1}}{D_{i,n+1-i}\hat{g}_{n+1-i}\cdot\ldots\cdot\hat{g}_{k-1}} \overset{(c)}{=} \frac{C_{i,n+1-i}\cdot\left(\sum\limits_{j=1}^{n}\hat{C}_{jk}\right)\Big/\left(\sum\limits_{j=1}^{n}\hat{C}_{j,n+1-i}\right)}{D_{i,n+1-i}\cdot\left(\sum\limits_{j=1}^{n}\hat{D}_{jk}\right)\Big/\left(\sum\limits_{j=1}^{n}\hat{D}_{j,n+1-i}\right)}$$

Zu e)

Die Nenner $(\Sigma C)/(\Sigma D)$ in d) sind dann das durchschnittliche Verhältnis von Zahlungen zu Aufwendungen in den Entwicklungsjahren k bzw. $n+1-i$. Die Identität in d) besagt also, dass bei getrennter Chain-Ladder-Prognose von Zahlungen (Z) und Aufwendungen (A) das Z/A-Verhältnis $\hat{C}_{ik}/\hat{D}_{ik}$ der Prognosen für $k>n+1-i$ in Anfalljahr i in Relation zum durchschnittlichen Z/A-Verhältnis aller Anfalljahre so bleibt wie es im aktuellen Entwicklungsjahr $n+1-i$ ist, und zwar in allen Entwicklungsjahren $k>n+1-i$. Ein im Entwicklungsjahr $n+1-i$ überdurchschnittliches Z/A-Verhältnis $C_{i,n+1-i}/D_{i,n+1-i}$ bleibt also überdurchschnittlich bis zum Schluss, während in Wirklichkeit das Z/A-Verhältnis jedes Anfalljahrs am Ende der Abwicklung in der Regel $=1$ ist, ebenso wie das durchschnittliche Z/A-Verhältnis.

Zu f)

Im beobachteten Teil der Dreiecke nähern sich in der Regel überdurchschnittliche ebenso wie unterdurchschnittliche Z/A-Verhältnisse einzelner Anfalljahre im weiteren Verlauf dem durchschnittlichen Z/A-Verhältnis an. Damit bei einem überdurchschnittlichen Z/A-Verhältnis $C_{ik}/D_{ik}, k\geq n+1-i$, im nächsten Entwicklungsjahr das Z/A-Verhältnis $C_{i,k+1}/D_{i,k+1}$ näher an den Durchschnitt kommt, muss der individuelle C-Abwicklungsfaktor $C_{i,k+1}/C_{ik}$ unterdurchschnittlich (d.h. kleiner als $\hat{f}_k$) und/oder der individuelle D-Abwicklungsfaktor $D_{i,k+1}/D_{ik}$ überdurchschnittlich (d.h. größer als $\hat{g}_k$) ausfallen. Analog bei unterdurchschnittlichem Z/A-Verhältnis C_{ik}/D_{ik}. Dies lässt sich bei Daten aus der Praxis in der Regel tatsächlich nachweisen.

Zu g)

Nein. <u>Ohne</u> Kenntnis des jeweils anderen Dreiecks widerspricht das Modell a) nicht dem in d) – f) dargestellten Sachverhalt. Kleidet man nämlich die Feststellung f) in das Modell

$$E\left(\frac{C_{i,k+1}}{C_{ik}}\middle| C_{i1},...,C_{ik},D_{i1},...,D_{ik}\right)=f_k+a_k\left(\frac{D_{ik}}{C_{ik}}-E\left(\frac{D_{ik}}{C_{ik}}\middle| C_{i1},...,C_{ik}\right)\right)$$

(und analog für $D_{i,k+1}/D_{ik}$), dann muss man ohne Kenntnis des D-Dreiecks von durchschnittlichen Gegebenheiten ausgehen, d.h. den Erwartungswert über $D_{i1},...,D_{ik}$ bilden. Dies führt zu $E\left(\frac{C_{ik+1}}{C_{ik}}\middle| C_{i1},...,C_{ik}\right) = f_k$ wie in a).

6. Aufgabe (30 Punkte) Risikoteilung

F sei die Verteilungsfunktion der Schadenhöhe X.

1) (a) Geben Sie drei äquivalente Schreibweisen bzw. Definitionen für die zu F bzw. X gehörende Entlastungseffektfunktion r an. (3 Punkte)

   (b) Geben Sie eine direkte Formel für deren Ableitung r' an. (2 Punkte)

   (c) Welche Verteilungsfunktion F_0 liegt der Entlastungseffektfunktion $r_0(x) = 1 - (1-x)^4,\ 0 \le x \le 1$, zu Grunde? (4 Punkte)

2) Bei der rechtsseitigen Zensorierung von F bei $c > 0$ wird die Wahrscheinlichkeitsmasse rechts von c auf der Schadenhöhe c konzentriert, d.h. man erhält die Verteilungsfunktion $\tilde{F}$ mit

$$\tilde{F}(x) = F(x) \text{ für } x < c,$$
$$\tilde{F}(x) = 1 \text{ für } x \ge c.$$

   (a) Welche Funktion $\tilde{X} = T_1(X)$ von X entspricht dieser Zensorierung, d.h. hat $\tilde{F}$ als Verteilungsfunktion (ohne Beweis)? (2 Punkte)

   (b) Geben Sie die zu $\tilde{F}$ gehörende Entlastungseffektfunktion $\tilde{r}$ in Abhängigkeit von X oder F oder r an (eine der drei Möglichkeiten genügt). (3 Punkte)

3) Bei der rechtsseitigen Stutzung von F bei $c > 0$ wird die Wahrscheinlichkeitsmasse rechts von c auf die links von c verbleibende Wahrscheinlichkeitsmasse proportional umverteilt, d.h. alle $F(x)$, $x \le c$, werden mit demselben Faktor multipliziert.

   (a) Wie lautet die so entstehende Verteilungsfunktion F^* (ohne Beweis)? (2 Punkte)

   (b) Welche Funktion $X^* = T_2(X)$ von X entspricht dieser Stutzung, d.h. hat F^* als Verteilungsfunktion (ohne Beweis)? (2 Punkte)

   (c) Zeigen Sie mit Hilfe der allgemeinen Eigenschaften von Entlastungseffektfunktionen, dass für die zu F^* gehörende Entlastungseffektfunktion r^* die Relation $r^*(x) \ge x/c$ für alle $x < c$ gilt. (3 Punkte)

4) Benutzen Sie die bisherigen Resultate, um allgemein gültige Größer-Kleiner-Relationen zwischen r, $\tilde{r}$ und r^* herzuleiten. Hinweis: Überlegen Sie, welchen Einfluss anschaulich die Verschiebung der W-Masse rechts von c auf die Entlastungseffektfunktion hat, um eine Idee zu bekommen, welche Relationen vermutlich gelten, und versuchen Sie diese dann zu beweisen. (9 Punkte)

Lösung:

Zu 1a)

$$r(a) = \frac{E(\min(X,a))}{E(X)} = \frac{\int\limits_0^a x dF(x) + a \int\limits_a^\infty dF(x)}{\int\limits_0^\infty x dF(x)} = \frac{\int\limits_0^a (1-F(x))\,dx}{\int\limits_0^\infty (1-F(x))\,dx}.$$

Zu 1b)

$$r'(a) = (1-F(a))/E(X)$$

(am einfachsten aus der 3. Schreibweise von a).

Zu 1c)

Wegen 1b) gilt:

$$r_0'(x) = 4(1-x)^3 \Rightarrow (E(X_0))^{-1} = r_0'(0) = 4 \Rightarrow E(X_0) = 1/4.$$

$$1 - F_0(x) \;=\; E(X_0)\, r_0'(x) \;=\; (1-x)^3 \;\Rightarrow\; F_0(x) \;= 1-(1-x)^3\,, \quad 0 \le x \le 1$$

(Beta-Verteilung mit $p = 1, q = 3$).

Zu 2a)

$$\tilde{X} = \min(X, c).$$

Zu 2b)

$$\tilde{r}(a) = \frac{E(\min(\tilde{X}, a))}{E(\tilde{X})} = \frac{E\left(\min(X, a)\right)}{E\left(\min(X, c)\right)} = \frac{r(a)}{r(c)} \text{ für } a \le c \quad \text{und} \quad \tilde{r}(a) = 1 \text{ für } a \ge c.$$

Anderer Lösungsweg: Einsetzen der Definition von $\tilde{F}$ in die letzte Schreibweise von 1a mit Ergebnis

$$\tilde{r}(a) = \frac{\int\limits_0^a (1 - F(x))\, dx}{\int\limits_0^c (1 - F(x))\, dx}.$$

Zu 3a)

$$F^*(x) = F(x)/F(c) \text{ für } x \le c \text{ und } F^*(x) = 1 \text{ für } x \ge c.$$

Zu 3b)

$$X^* = X|(X \le c)$$

(Dies ist nicht gleich $X \cdot 1_{\{X \le c\}}$).

Zu 3c)

$$r^* \text{ konkav} \Rightarrow r^*(x) = r^*\left(\frac{x}{c} \cdot c + \left(1 - \frac{x}{c}\right) \cdot 0\right) \ge \frac{x}{c} r^*(c) + \left(1 - \frac{x}{c}\right) r^*(0) = \frac{x}{c}.$$

Oder direkt verbal: Wegen der Konkavität verläuft r^* oberhalb der Diagonalen $y = x/c$. Die gleiche Ungleichung gilt natürlich auch für $\tilde{r}$, nicht aber für r.

Zu 4)

Zunächst sieht man sofort wegen 2b), dass $\tilde{r}(a) = \frac{r(a)}{r(c)} \ge r(a)$ gilt.
Für das Weitere braucht man eine Formel für r^* für $a < c$. Wieder gibt es drei Schreibweisen (nur mit der 2. oder 3. geht es weiter):

$$r^*(a) = \frac{E\left(\min(X^*, a)\right)}{E(X^*)} = \frac{E\left(\min(X, a)|X \le c\right)}{E(X|X \le c)},$$

$$r^*(a) = \frac{\int\limits_0^a x dF^*(x) + a \int\limits_a^\infty dF^*(x)}{\int\limits_0^\infty x dF^*(x)} = \frac{\int\limits_0^a x dF(x) + a \int\limits_a^c dF(x)}{\int\limits_0^c x dF(x)}.$$

$$r^*(a) = \frac{\int\limits_0^a (1 - F^*(x))\, dx}{\int\limits_0^\infty (1 - F^*(x))\, dx} = \frac{\int\limits_0^a (F(c) - F(x))\, dx}{\int\limits_0^c (F(c) - F(x))\, dx} = \frac{\int\limits_0^a (1 - F(x))\, dx - a\,(1 - F(c))}{\int\limits_0^c (1 - F(x))\, dx - c\,(1 - F(c))}.$$

Letzteres lässt sich leicht mit der 3. Schreibweise für $\tilde{r}$ vergleichen (vgl. 2b):

$$\tilde{r}(a) = \frac{\int\limits_0^a (1 - F(x))\, dx}{\int\limits_0^c (1 - F(x))\, dx} = \frac{Z^* + a\,(1 - F(c))}{N^* + c\,(1 - F(c))},$$

wobei $Z^* = \text{Zähler}(r^*)$ und $N^* = \text{Nenner}(r^*)$, je gemäß 3. Schreibweise.
Mit $r^*(a) = Z^*/N^*$ gilt also

$$\tilde{r}(a) \le r^*(a) \Leftrightarrow a\,(1 - F(c))\,N^* \le c\,(1 - F(c))\,Z^* \Leftrightarrow \frac{a}{c} \le \frac{Z^*}{N^*} = r^*(a).$$

Und da Letzteres wegen 3c) richtig ist, ist $\tilde{r} \le r^*$ bewiesen.

Anschaulich kann man die Relation $r \le \tilde{r} \le r^*$ auf mehrere Arten plausibel machen. Am einfachsten durch Zeichnung von F, $\tilde{F}$ und F^* in ein gemeinsames Diagramm. Dort ist der Entlastungseffekt durch das Verhältnis der Fläche über der Verteilungsfunktion links von a zur gesamten Fläche über der Verteilungsfunktion gegeben.

Eine zweite Möglichkeit benutzt die Tatsache $r(0) = \tilde{r}(0) = r^*(0) = 0$ und die Steigung der Tangente im Nullpunkt, die gemäß 1b) gerade gleich der Inversen des Erwartungswerts ist. Aus der trivialen Relation $E(X) \ge E(\tilde{X}) \ge E(X^*)$ folgt $r'(0) \le \tilde{r}'(0) \le r^{*\prime}(0)$. Dies legt $r \le \tilde{r} \le r^*$ nahe.

Als Drittes kann man auch den empir. Entlastungseffekt $\hat{r}(a) = \sum\limits_{i=1}^{n} \min(x_i, a) \Big/ \sum\limits_{i=1}^{n} x_i$ einer Stichprobe $x_1, ..., x_n$ aus F betrachten. $\hat{r}(a)$ ist offenbar ein gewichtetes Mittel der Einzel-Entlastungseffekte $\min(x_i, a)/x_i$ der einzelnen Schäden, wobei der Einzel-Entlastungseffekt umso kleiner ist, je größer der Schaden ist. Werden also die größten Schäden $X > c$ zu $\tilde{X} = \min(X, c) = c$ verkleinert, vergrößert sich offenbar der Entlastungseffekt. Und werden die Schäden $X > c$ umverteilt, so ist das für den Entlastungseffekt dasselbe, wie das Weglassen dieser Schäden; und das vergrößert den Entlastungseffekt natürlich noch mehr.

Eine solche intuitive Begründung war nicht gefordert, wurde aber bei Fehlen des exakten Beweises mit bis zu 6 Punkten honoriert.