# Bericht zur Prüfung 1996 über Schadenversicherungsmathematik (Spezialwissen)

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

Am 1. Juni 1996 fand in Köln eine Klausurveranstaltung statt, in der unter anderem das Spezialwissen in Schadenversicherungsmathematik geprüft wurde. Diese Prüfung war die erste im Rahmen des neuen Ausbildungssystems und des Zugangs zur Aktuarvereinigung DAV. Den erfolgreichen Teilnehmern wird die Prüfungsurkunde anläßlich der nächsten Mitgliederversammlung der DAV am 30. April 1997 in Würzburg überreicht. Die folgenden Aufgaben waren in 3 Stunden zu lösen. Für jede der Aufgaben waren maximal 30 Punkte erreichbar. Die Klausur wurde als bestanden bewertet, wenn mindestens 72 Punkte erreicht worden waren. Die Zusatzaufgabe wurde nur gewertet, wenn dafür eine der anderen Aufgaben nicht bearbeitet worden war. Als Hilfsmittel war zugelassen a) die schon früher benutzte Formelsammlung und b) ein nicht programmierbarer Taschenrechner.

*Aufgaben zu Grundlagen:*

Für die Modellierung des Gesamtschadens S eines Bestandes wird eine Poisson-verteilte Schadenzahl mit Parameter $\lambda = 0{,}1$ und eine Pareto-verteilte Schadenhöhe mit Parameter $a = 2{,}185$ verwendet.

a) Wie lautet die Normalapproximation für die Wahrscheinlichkeit $P(S > 3)$? Ist diese Approximation gut?
b) Wie lautet die Normal-Power-Approximation für dieselbe Wahrscheinlichkeit und obige Parameter? Was erhält man andererseits im Falle $\lambda = 1$ und $a = 3{,}3$?

*Lösung:*

Für die Momente von S gilt

$$E(S) = \lambda \frac{a}{a-1}, \quad a > 1,$$

$$\mathrm{Var}(S) = \lambda \frac{a}{a-2}, \quad a > 2,$$

$$E(S - E(S))^3 = \lambda \frac{a}{a-3}, \quad a > 3.$$

a) $P(S > 3)$ wird approximiert durch

$$1 - \Phi((3 - E(S))/\sqrt{\mathrm{Var}(S)}) = 1 - \Phi(2{,}5908) = 0{,}0048$$

wegen

$$E(S) = 0{,}18388, \quad \mathrm{Var}(S) = 1{,}1811 \quad \text{und} \quad (3 - E(S))/\sqrt{\mathrm{Var}(S)} = 2{,}5908.$$

Die Normalapproximation ist hier prinzipiell nicht gut wegen der großen absoluten Schiefe; die Paretoverteilung Q mit Parameter $a = 2{,}185$ hat unendlich große absolute Schiefe. Speziell gilt hier

$$P(S > 3) = \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} Q^{*n}(3, \infty) \geq (1 - e^{-\lambda})\, Q(3, \infty) = (1 - e^{-\lambda})\, 3^{-a} = 0{,}00863.$$

b) Für $a = 2{,}185$ kann man die Normal-Power-Approximation nicht berechnen, da $ES^3 = \infty$ gilt. Für die Parameter $\lambda = 1$ und $a = 3{,}3$ ergibt sich als Approximation für $P(S > 3)$ nach der Formelsammlung $1 - \Phi(z)$ mit

$$z = \left( \frac{9\lambda \mu_2^3(Q)}{\mu_3^2(Q)} + \frac{6\mu_2(Q)}{\mu_3(Q)} (3 - \lambda \mu_1(Q)) + 1 \right)^2 - 3\lambda^{1/2} \mu_2^{1/2}(Q).$$

Mit den Werten

$$\mu_1(Q) = 1{,}4348, \quad \mu_2(Q) = 2{,}5385 \quad \text{und} \quad \mu_3(Q) = 11$$

ergibt sich

$$z = 0{,}9907498$$

und damit

$$1 - \Phi(z) = 1 - 0{,}8389 = 0{,}1611\,.$$

*Aufgabe zur Solvabilität:*

Für die Beurteilung der Ruinwahrscheinlichkeit $\psi(s)$ bei unendlichem Planungshorizont wird der Anpassungskoeffizient R und die Schranke $\exp(-R\,s)$ benutzt, wobei s die für das Versicherungsgeschäft gestellte freie Reserve ist.

a) Wie ist der Anpassungskoeffizient definiert, wann existiert er, wann existiert er nicht?
b) Wie kann man aus der Ruintheorie eine ausreichende Bestandsprämie ableiten, wenn s und die Verteilung des Gesamtschadens S sowie eine tolerierte Ruinwahrscheinlichkeit $\varepsilon$ bekannt sind?
c) Wie beurteilt man die Ruinwahrscheinlichkeit im Fall eines großschadenanfälligen Bestandes?

*Lösung:*

a) Der Anpassungskoeffizient im klassischen Risikoreserveprozeß ist die positive Lösung von

$$\lambda + r\,c = \lambda \int \exp(r\,x)\,Q(dx)\,.$$

Hierbei ist $\lambda$ die Intensität des Schadenzahlprozesses und c die Prämie pro Zeiteinheit. Q ist die Schadenhöhenverteilung. R existiert nur dann, wenn

$$c > \lambda\,\mu_1(Q)$$

und wenn

$$\int \exp(r\,x)\,Q(dx) < \infty$$

für ein $r > 0$ gilt. R existiert nicht für Großschadenverteilungen Q wie z. B. Lognormal-, Loggamma- und gewisse Weibullverteilungen.

b) Ist $\psi(s)$ die Ruinwahrscheinlichkeit zum Startkapital s, so wird eine Ruinwahrscheinlichkeit $\varepsilon$ eingehalten, wenn der Anpassungskoeffizient R so eingestellt wird, daß

$$\exp(-R\,s) = \varepsilon$$

gilt, also wenn

$$R = -\frac{1}{s}\,\log(\varepsilon)\,.$$

Dies erreicht man durch Wahl der Prämie c nach dem Exponentialprinzip mit Risikoaversion R:

$$c = \frac{1}{R}\log E\exp(R\,S)\,.$$

Hierbei ist S der Gesamtschaden des Bestandes in einem Jahr. Für ein kleines R kann man die Prämie c approximieren durch die Prämie nach dem Varianzprinzip:

$$c_1 = ES + \frac{R}{2}\mathrm{Var}(S)\,.$$

Die Genauigkeit dieser Approximation ist an dem Ergebnis

$$|c - c_1| \le \frac{R^2}{6} ES^3 \exp(R\,S)$$

ablesbar.

c) Für Großschadenverteilungen Q kann das obige Argument nicht benutzt werden, da der Anpassungskoeffizient dort nicht existiert. Man kann stattdessen die Ruinwahrscheinlichkeit $\psi(s)$ mit der Summenformel berechnen und/oder eine Approximation benutzen. Zunächst die Summenformel:

$$\psi(s) = \sum_{k=1}^{\infty} p^k (1-p) H^{*k}(s, \infty),$$

mit $p = \lambda\mu/c$ und H mit Dichte $x \to Q(x, \infty)/\mu$. Die Approximation basiert auf den Approximationen für die Teilwahrscheinlichkeit von Summenverteilungen, deren Schadenhöhe subexponentiell ist: Für jede der üblicherweise benutzten Großschadenverteilungen Pareto, Lognormal, Loggamma oder Weibull (mit Formparameter kleiner als eins) für Q gilt mit $s \to \infty$:

$$\psi(s) \sim \frac{p}{1-p} H(s, \infty) = \frac{\lambda\mu}{c - \lambda\mu} H(s, \infty).$$

*Aufgabe zur Tarifkalkulation I:*

Verallgemeinerte lineare Modelle spielen aktuell für die Tarifierung oder die Konstruktion neuer Tarifierungssysteme eine wichtige Rolle. Passen Sie an die folgenden Schadendaten ein solches Modell an: Die Daten sind beobachtete Schadenanzahlen X (i, j) mit Merkmalsausprägungen i, j der Merkmale A und B. Die Schadenanzahlen sind entstanden in Beständen mit N (i, j) Verträgen.

| X (i, j) | j = 1 | j = 2 | j = 3 |
|---|---|---|---|
| i = 1 | 168 | 395 | 498 |
| i = 2 | 7 | 44 | 56 |
| i = 3 | 300 | 439 | 469 |

| N (i, j) | j = 1 | j = 2 | j = 3 |
|---|---|---|---|
| i = 1 | 2000 | 1600 | 1500 |
| i = 2 | 100 | 80 | 40 |
| i = 3 | 1000 | 1500 | 2000 |

a) Schlagen Sie ein Modell für diese Daten vor. Beachten Sie die Anzahl der Parameter.
b) Welche Devianz muß minimiert werden, wenn Sie nur die Haupteffekte berücksichtigen wollen?
c) Wie sieht die natürliche Linkfunktion und die Varianzfunktion aus, wenn Sie Poisson-verteilte Größen X (i, j) unterstellen? Wie werden hierbei die Volumenmaße N (i, j) berücksichtigt?

*Lösung:*

Die Beobachtungen sind Schadenanzahlen X (i, j) in unterschiedlichen Beständen mit N (i, j) Verträgen. Die zugehörigen Schadenfrequenzen F (i, j) sind

| F (i, j) | j = 1 | j = 2 | j = 3 |
|---|---|---|---|
| i = 1 | 0,084 | 0,246875 | 0,332 |
| i = 2 | 0,07 | 0,55 | 1,4 |
| i = 3 | 0,3 | 0,29 | 0,23 |

a) 1. Ein *falscher* Ansatz: X (i, j) sind stochastisch unabhängig, $X(i,j) \sim \text{Bin}(N(i,j), p(i,j))$, $\log p(i,j) = \mu + \alpha_i + \beta_j$. In diesem Modell ist die Beobachtung X (2, 3) nicht möglich, das Modell also nicht passend.

2. X (i, j) sind stochastisch unabhängig,

$$X(i,j) \sim \text{NBin}(N(i,j), p(i,j)),$$
$$\log p(i,j) = \mu + \alpha_i + \beta_j.$$

Dies ist ein naheliegendes Modell, es hat 5 Parameter (wegen $\alpha_1 = \beta_1 = 0$), und dies ist bei der Anzahl der hier vorliegenden Beobachtungen kein Problem. Allerdings sind Zeilen- und Spalteneffekte hier nicht deutlich sichtbar; so ist z. B. $j \to F(1,j)$ steigend, $j \to F(3,j)$ jedoch fallend. Ein analoges Modell wäre:

3. X (i, j) sind stochastisch unabhängig, $X(i,j) \sim \text{Poisson}(N(i,j) * p(i,j))$, $\log p(i,j) = \mu + \alpha_i + \beta_j$.

4. X (i, j) sind stochastisch unabhängig,

$$X(i,j) \sim \text{NBin}(N(i,j), p(i,j)),$$
$$\log p(i,j) = \alpha_i + \beta_i * j.$$

In diesem Modell wird mögliche bzgl. i unterschiedliche Steigung der Frequenzen berücksichtigt.

b) Ausgehend vom Modell a 3, in welchem gerade die Haupteffekte berücksichtigt werden, muß durch Wahl von $\mu, \alpha_i, \beta_j$ nach dem Maximum-Likelihood-Prinzip minimiert werden

$$-\sum_{i,j} (X(i,j) \log (N(i,j)\, p(i,j)) - N(i,j)\, p(i,j)\,.$$

Nach Weglassen des konstanten Terms

$$\sum_{i,j} X(i,j) \log N(i,j)$$

ist folgender Ausdruck zu minimieren:

$$L := -\sum_{i,j} N(i,j)\,(F(i,j)\,(\mu + \alpha_i + \beta_j) - \exp(\mu + \alpha_i + \beta_j))\,.$$

Sind $\hat{p}(i,j)$ die resultierenden Schätzer, so berechnet sich die Devianz nach der Formel

$$2(L - \hat{L}),$$

wobei L die oben definierte Größe mit den Werten $\hat{p}(i,j)$ ist, und $\hat{L}$ ist die L entsprechende Größe, bei der $\hat{p}(i,j)$ durch $F(i,j)$ ersetzt ist, also

$$-2\sum_{i,j} N(i,j)\,(F(i,j) \log (\hat{p}(i,j)/F(i,j)) - \hat{p}(i,j) + F(i,j))\,.$$

Die 2 ist der übliche Normierungsfaktor.

c) Die natürliche Linkfunktion für die Poissonverteilung ist $g(\mu) = \log(\mu)$, die Varianzfunktion ist $V(\mu) = \mu$. Die Volumenmaße gehen in die Modellierung als Faktoren der Parameter $\lambda(i,j)$ ein. Hierbei stellt man sich vor, daß die Verträge stochastisch unabhängig Schäden produzieren, daß also die Gesamtzahl aller Schäden im Teilbestand als Summe von $N(i,j)$ stochastisch unabhängigen identisch Poisson-verteilten Zufallsgrößen wieder Poisson-verteilt ist mit Parameter $N(i,j) * \lambda(i,j)$.

*Aufgaben zur Tarifkalkulation II:*

In einer Studie des VDS zum KH-Tarif heißt es sinngemäß, es seien mindestens 10000 Pkw-Jahreseinheiten erforderlich, wenn der beobachtete Schadenverlauf eines Jahres mit 95% statistischer Sicherheit um nicht mehr als $\pm 10\%$ von seinem wahren Wert abweichen soll.
Legen Sie dieser Aussage ein Kollektives Modell mit poissonverteilter Schadenzahl zugrunde und schließen Sie unter Annahme einer Schadenfrequenz von 0,1 pro Jahreseinheit auf den Variationskoeffizienten der Schadenhöhenverteilung. Kommentieren Sie das Ergebnis und die gemachten Annahmen bezüglich ihres Realitätsbezugs.

*Lösung der Aufgabe zur Tarifkalkulation II:*

Gemäß dem Kollektiven Modell gilt für den Gesamtschaden S bei poissonverteilter Schadenzahl N

$$E(S) = E(N)\,E(X)$$
$$\mathrm{Var}(S) = E(N)\,E(X^2)\,,$$

wenn X die Schadenhöhe ist. Daraus folgt

$$(\mathrm{Vko}(S))^2 = \mathrm{Var}(S)/(E(S))^2 = ((\mathrm{Vko}(X))^2 + 1)/E(N)\,.$$

Gegeben ist $E(N) = 0{,}1\,v$ mit v = Anzahl Jahreseinheiten und

$$P(|S - E(S)| \leq 0{,}10 \cdot E(S)) > 0{,}95$$

für $v \geq 11000$. Nimmt man näherungsweise an, daß S normalverteilt ist, so gilt

$$P(|S - E(S)| \leq 0{,}10 \cdot E(S)) = \Phi(a) - \Phi(-a)$$

mit der Standard-Normalverteilung $\Phi$ und

$$a = 0{,}10 \cdot E(S)/\mathrm{Sta}(S) = 0{,}10/\mathrm{Vko}(S)\,.$$

Nur für $a \geq 1{,}96$ ist $\Phi(a) - \Phi(-a) \geq 0{,}95$, also muß für $v \geq 11\,000$

$$0{,}10/\mathrm{Vko}(S) \geq 1{,}96$$

oder

$$\mathrm{Vko}(S) \leq 1/19{,}6$$

oder

$$(\mathrm{Vko}(X))^2 \leq 0{,}1\,v/19{,}6^2 - 1 \leq 1{,}86$$

gelten. Als muß $\mathrm{Vko}(X) \leq 1{,}365$ sein.

Kommentare:

1. Das Kollektive Modell ist für diese Aufgabenstellung eine durchaus realistische Modellannahme. Auch die Normalapproximation ist bei 1100 erwarteten KH-Schäden bzw. angesichts von $E(S) \geq 19{,}6 \cdot \mathrm{Sta}(S)$ ausreichend genau.
2. Ein Variationskoeffizient von 1,365 für die KH-Schadenhöhe ist unrealistisch nieder (realistische Werte liegen bei 5). Tatsächlich wurden der Aussage implizit nur Schadenhöhen, die bei etwa DM 50000 gestutzt waren, zugrunde gelegt.
3. Die Annahme einer Poissonverteilung für die Schadenzahl ist ebenfalls unrealistisch. Bei 1100 erwarteten Schäden beträgt bei der Poissonverteilung die Standardabweichung $\sqrt{1100} = 33$, so daß die tatsächliche Schadenzahl mit 99,7% Wahrscheinlichkeit zwischen 1000 und 1200 liegen müßte. In der Realität schwankt jedoch die Schadenfrequenz von Jahr zu Jahr ebenfalls, so daß sich höhere Abweichungen ergeben, die z. B. mit der Negativen Binomialverteilung realistischer modelliert werden können.

*Aufgabe zur Schadenreservierung:*

Sei $C_{ik}$ der Schadenstand von Anfalljahr i nach k Entwicklungsjahren, $S_{ik} = C_{ik} - C_{i,k-1}$ mit $C_{i0} = 0$ die jährliche Veränderung und $v_i$ ein Maß für das Volumen des Portefeuilles von Anfalljahr i. Das Modell der anfalljahrunabhängigen Schadenquotenzuwächse beruht auf der Annahme

$$E(S_{ik}) = v_i\,m_k, \tag{1}$$

das Chain-Ladder-Modell auf

$$E(C_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1}\,f_k, \tag{2}$$

jeweils mit einem unbekannten Parameter $m_k$ bzw. $f_k$.

1. Interpretieren Sie, für welche Art von in der Realität vorkommenden Änderungen bei einzelnen Schäden (die in $S_{ik}$ zusammengefaßt sind) eher das Modell (1) und für welche eher das Modell (2) in Betracht kommt.
2. Bilden Sie ein kombiniertes Modell, das beide genannten Modelle als Spezialfälle enthält – d. h. in dem ein Teil der in $S_{ik}$ steckenden Veränderung sich wie (1) verhält, der Rest wie (2) – und geben Sie an, wie Sie dann die Parameter schätzen würden (nur Angabe des Weges, nicht der Schätzer selbst).
3. Warum ist das kombinierte Modell, obwohl es beide Einzelmodelle (1) und (2) enthält, doch nicht in jedem Fall vorzuziehen, d. h. worin besteht ein Nachteil des kombinierten Modells gegenüber jedem der Einzelmodelle? Wie könnte dieser Nachteil beseitigt werden?

*Lösung der Aufgabe zur Schadenreservierung:*

1. In $S_{ik}$ stecken zwei Arten von Änderungen: zum einen die Endstandsschätzung von erstmals im Entwicklungsjahr k bekannt gewordenen Schäden (Gesamtbetrag $U_{ik}$), zum anderen die Änderung der Schadenstandsschätzung von bereits früher bekannten Schäden (Gesamtbetrag $T_{ik}$).
   Für $T_{ik}$ eignet sich eher das Modell (2), da es von Anzahl bzw. Gesamtbetrag der in $C_{i,k-1}$ steckenden offenen Schadenfälle abhängt, denn die in $T_{ik}$ zusammengefaßten Änderungen können im wesentlichen nur aus offenen Schäden resultieren.
   Für $U_{ik}$ kommt eher das Modell (1) in Betracht, da aufgrund der Unabhängigkeit der Einzelschäden und deren Meldezeitpunkt der in Entwicklungsjahr k gemeldete Schadenbetrag $U_{ik}$ nicht von

Zahl und Höhe der bereits früher gemeldeten Schäden abhängt, sondern lediglich von der Portefeuillegröße $v_i$.

2. Aus

$$E(U_{ik} | C_{i1}, \ldots, C_{i,k-1}) = E(U_{ik}) = v_i u_k \tag{1a}$$

und

$$E(T_{ik} | C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} t_k \tag{2a}$$

ergibt sich das kombinierte Modell

$$E(S_{ik} | C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} t_k + v_i u_k . \tag{3}$$

Die Parameter $t_k$ und $u_k$ kann man mittels gewöhnlicher linearer Regression der Zuwachsquoten $S_{ik}/v_i$ nach den Vorjahresständen $C_{i,k-1}/v_i$ schätzen, d.h. durch Minimieren von

$$\sum_{i=1}^{I+1-i} v_i \left\{ \frac{S_{ik}}{v_i} - \left( \frac{C_{i,k-1}}{v_i} \cdot t_k + u_k \right) \right\}^2 .$$

3. Der Nachteil des Modells (3) gegenüber (1) und (2) besteht darin, daß es bei denselben Daten doppelt soviele Parameter verwendet und daher weniger stabile Prognosen liefert. Dieser Nachteil kann beseitigt werden, wenn die Daten $T_{ik}$ und $U_{ik}$ bekannt sind, so daß die Modelle (1a) und (2a) getrennt geschätzt werden können.

*Aufgabe zur Risikoteilung:*

Für ein Portefeuille von 400000 KH-Risiken wird die Gültigkeit des Kollektiven Modells unterstellt. Die Anzahl Schäden pro Jahr wird als poissonverteilt mit mittlerer Frequenz 0,1 pro Risiko angenommen, die Schadenhöhe als lognormalverteilt mit Parametern $\mu = 7{,}1$ und $\sigma = 1{,}8$.

1. Die Schadenhöhen $X_1, X_2, \ldots$ werden durch die Schadengrenze $a = 100000$ DM in Normalschäden $X_n \leq a$ und Großschäden $X_n > a$ aufgeteilt. Berechnen Sie den Erwartungswert der Schadenhöhe $Y_2$ eines Großschadens.
2. Berechnen Sie Erwartungswert und Varianz der Jahressumme $S_1$ der Normalschäden und der Jahressumme $S_2$ der Großschäden. Benutzen Sie dabei zur Verkürzung der Rechenarbeit die folgenden Angaben zu Erwartungswert und Variationskoeffizient für die Normalschadenhöhe $Y_1$ bzw. die Großschadenhöhe $Y_2$:

$$\begin{aligned} E(Y_1) &= 4580, \\ \mathrm{Vko}(Y_1) &= 2{,}18, \\ E(Y_2) &= 222000, \\ \mathrm{Vko}(Y_2) &= 1{,}18. \end{aligned}$$

3. Sind $S_1$ und $S_2$ korreliert? Begründen Sie Ihre Antwort. Ändert sich das Korrelationsverhalten bei anderen in Betracht kommenden Verteilungsannahmen für Schadenzahl oder Schadenhöhe?

*Lösung der Aufgabe zur Risikoteilung:*

1. Der Erwartungswert der Höhe eines Großschadens beträgt

$$E(Y_2) = E(X_n | X_n > a) = \int_a^{\infty} x \, dF(x)/(1 - F(a)) = \exp(\mu + \sigma^2/2) \frac{1 - \Phi((\ln(a) - \mu - \sigma^2)/\sigma)}{1 - \Phi((\ln(a) - \mu)/\sigma)},$$

letzteres gemäß Formelsammlung mit der Lognormalverteilung F und der Standard-Normalverteilung $\Phi$. Ausrechnen ergibt $F(a) = 0{,}9929$ und $E(Y_2) = 6124 \cdot 0{,}2573/0{,}0071 = 222000$.

2. Für $i = 1$ und $i = 2$ gilt $E(S_i) = E(N_i)E(Y_i)$ mit $N_1$ bzw. $N_2$ = Anzahl der Normal- bzw. Großschäden. Mit $N = N_1 + N_2$ ist

$$\begin{aligned} E(N_1) &= E(N)F(a) = 0{,}1 \cdot 400000 \cdot 0{,}9929 = 39716, \\ E(N_2) &= E(N) - E(N_1) = 284, \\ E(S_1) &= 39716 \cdot 4580 = 181{,}9 \text{ Mio.}, \\ E(S_2) &= 284 \cdot 222000 = 63{,}05 \text{ Mio.} \end{aligned}$$

Wegen $E(Y_i^2) = \mathrm{Var}(Y_i) + (E(Y_i))^2 = ((\mathrm{Vko}(Y_i))^2 + 1) \cdot (E(Y_i))^2$ und da mit N auch $N_1$ und $N_2$ poissonverteilt sind, gilt

$$\begin{aligned}
\mathrm{Var}(S_i) &= E(N_i)\,E(Y_i^2),\\
\mathrm{Var}(S_1) &= 39716 \cdot (2{,}18^2 + 1) \cdot 4580^2 = 4{,}79 \cdot 10^{12},\\
\mathrm{Var}(S_2) &= 284 \cdot (1{,}18^2 + 1) \cdot 222000^2 = 33{,}49 \cdot 10^{12}.
\end{aligned}$$

3. $S_1$ und $S_2$ sind genau dann unkorreliert, wenn $\mathrm{Var}(S) = \mathrm{Var}(S_1) + \mathrm{Var}(S_2)$ für den Jahresgesamtschaden $S = S_1 + S_2$ gilt. Angesichts von

$$\mathrm{Var}(S) = E(N)\,E(X^2) = 40000 \cdot \exp(2\mu + \sigma^2) \cdot \exp(\sigma^2) = 38{,}31 \cdot 10^{12}$$

und

$$\mathrm{Var}(S_1) + \mathrm{Var}(S_2) = 38{,}28 \cdot 10^{12}$$

sind $S_1, S_2$ wohl unkorreliert. Dies läßt sich auch leicht allgemein zeigen: Wegen

$$\mathrm{Var}(S_i) = E(N_i)\,E(Y_i^2), \quad E(N_1) = E(N)\,F(a), \quad E(N_2) = E(N)(1 - F(a)),$$

$$E(Y_1^2) = E(X^2 \mid X \le a) = \int_0^a x^2\,dF(x)/F(a),$$

$$E(Y_2^2) = E(X^2 \mid X > a) = \int_a^\infty x^2\,dF(x)/(1 - F(a)),$$

wird

$$\begin{aligned}
\mathrm{Var}(S_1) + \mathrm{Var}(S_2) &= E(N)\int_0^a x^2\,dF(x) + E(N)\int_a^\infty x^2\,dF(x)\\
&= E(N)\int_0^\infty x^2\,dF(x)\\
&= E(N)\,E(X^2) = \mathrm{Var}(S).
\end{aligned}$$

Also sind $S_1$ und $S_2$ unkorreliert. Dies hängt offensichtlich nicht von der Schadenhöhenverteilung F ab, liegt aber entscheidend and der Poissonannahme für die Schadenzahl N. Würde die Schadenzahl N realistischer als gemischt poissonverteilt angenommen, so würden sich die Schwankungen im Poissonparameter sowohl auf $N_1$ als auch auf $N_2$ auswirken und so eine positive Korrelation von $N_1, N_2$ und damit auch von $S_1, S_2$ bewirken, was sich formelmäßig durch den in $\mathrm{Var}(S_i)$ entstehenden Zusatzterm $(\mathrm{Var}(N_i) - E(N_i))(E(Y_i))^2$ auswirkt.

*Zusatzaufgabe:*

Betrachten Sie ein allgemeines Bonus-Malus-System, welches von der Anzahl der Schäden pro Jahr abhängt.

a) Aus welchen Komponenten besteht ein solches Bonus-Malus-System?
b) Wie berechnet man, ausgehend von einer individuellen Schadenzahlverteilung, die individuelle Übergangsmatrix für einen Versicherungsvertrag?
c) Wie ist die zugehörige stationäre Verteilung definiert, und wie kann man sie berechnen?
d) Das Äquivalenzprinzip verlangt, daß über lange Vertragsdauer der erwartete Gesamtschaden übereinstimmt mit den erwarteten Prämienzahlungen. Approximativ kann man das Äquivalenzprinzip als Gleichung mit der stationären Verteilung formulieren. Wie lautet diese Gleichung? Kann man sie für alle Verträge im Bestand erfüllen? Wie geht man vor, um die Gleichung für möglichst viele Verträge zu erfüllen?

*Lösung:*

a) Klassen $i = 1, \ldots, I$. Umstufungsregeln $j(i, k)$; sie geben an, in welche Klasse j man aus Klasse i bei k Schäden kommt. Prämien $\pi(i)$ für Stufe $i = 1, \ldots, I$.

b) R sei die Schadenanzahl eines Jahres. Dann wird die Übergangswahrscheinlichkeit $p(i,j)$ von Zustand (Stufe) i nach Zustand j definiert durch

$$p(i,j) = \sum_{\substack{k:\\ j(i,k)=j}} P(R = k).$$

c) Die zur Markovschen Übergangsmatrix gehörende stationäre Verteilung $q(i)$, $i = 1, \ldots, I$, erfüllt

$$q(j) = \sum_{i=1}^{I} q(i)\, p(i,j)$$

und die Nebenbedingung

$$\sum_{j=1}^{I} q(j) = 1.$$

Sie ist die (in der Regel eindeutige) Startverteilung, mit der die Markov-Kette stationär wird. Die stationäre Verteilung ist der Grenzwert der n-Schritt-Übergangswahrscheinlichkeiten $p^{(n)}(i,j)$:

$$q(j) = \lim_{n \to \infty} p^{(n)}(i,j).$$

d) Die Gleichung lautet, wenn S der jährliche Gesamtschaden aus dem individuellen Vertrag ist,

$$E(S) = \sum_{i=1}^{I} \pi(i)\, q(i).$$

Diese Gleichung kann man nicht für alle Verträge gleichzeitig erfüllen, da $q(i) = q(i,\theta)$ und $S = S(\theta)$ über die Verteilung von R vom einzelnen Vertrag abhängt (dies wird durch den angefügten Risikoparameter $\theta$ ausgedrückt). Als Ausweg minimiert man den Ausdruck

$$\int \left( E(S(\theta)) - \sum_{i=1}^{I} \pi(i)\, q(i,\theta) \right)^2 \nu(\theta)\, d\theta.$$

Hierbei ist $\nu(\theta)$ die Dichte der Verteilung des Risikoparameters $\theta$ im Bestand. Hierdurch wird sichergestellt, daß das Äquivalenzprinzip für die meisten Verträge annähernd erfüllt wird.
Anmerkung: In der Praxis wird die stufenabhängige Bedarfsprämie nicht mit der stationären Verteilung, sondern durch Gruppenbildung ermittelt.