# Klausur zur Mathematik der Schadenversicherung, Spezialwissen November 2008

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

Wieder sind sieben Aufgabe gestellt, und jede ist mit einer maximalen Punktzahl gekennzeichnet. Insgesamt sind 180 Punkte zu erreichen, und mit 72 Punkten wird die Klausur als bestanden bewertet. Die Zusatzaufgabe wird nur gewertet, wenn eine andere Aufgabe nicht bearbeitet oder als nicht bearbeitet gekennzeichnet wurde.

## 1 Aufgabe zur Modellierung (30 Punkte)

1. Die Gumbel-Copula ist die archimedische Copula mit erzeugender Funktion
$$\phi(u) = (-\log(u))^\theta, \theta > 1.$$
Sie hat asymptotische Tailabhängigkeit $\lambda = 2 - 2^{1/\theta}$. Bauen Sie die Verteilung eines Vektors $(X, Y)$ auf aus der Randverteilung $Pareto(2)$ mit Dichte
$$x \to 2x^{-3}, x > 1$$
für $X$ und für $Y$, und aus der Gumbel-Copula mit
$$\theta = 1/(\log_2(3) - 1)$$
und zeigen Sie, dass in diesem Modell
$$\mathbb{P}\{X > x \text{ und } Y > x\} \sim \frac{1}{2}\mathbb{P}\{X > x\}, x \to \infty.$$
**Hinweis:** $\sim$ bezeichnet hier die asymptotische Äquivalenz: $a(x) \sim b(x)$ genau dann, wenn $a(x)/b(x) \to 1$. (17 Punkte)

2. Welche Verteilung hat die folgende Darstellung als (stetige) Phasentypverteilung:
$$\pi = (1, 0),\ B = \begin{pmatrix} -1 & 1 \\ p & -1 \end{pmatrix}?$$
(13 Punkte)

**Lösung:**

1. Nach Aufgabenstellung hat die Gumbel-Copula eine Verteilungsfunktion $C(u,v), 0 \le u,v \le 1$, welche

$$\lim_{u\to 1} \frac{1-2u+C(u,u)}{1-u} = 2 - 2^{1/\theta} = 2 - 2^{(\log_2(3)-1)} = 2 - 3/2 = 1/2$$

erfüllt. Die gemeinsame Verteilung von $(X,Y)$ hat Verteilungsfunktion

$$\mathbb{P}\{X \le x \text{ und } Y \le y\} = C(F(x),F(y))$$

mit stetiger Verteilungsfunktion $F(x) = \mathbb{P}\{X \le x\}, x \ge 0$. Die Form der Verteilung ist für die Aussage in der Aufgabe und für den Beweis irrelevant. Wegen

$$\mathbb{P}\{X > x \text{ und } Y > x\} = 1 - 2\mathbb{P}\{X \le x\} + \mathbb{P}\{X \le x \text{ und } Y \le x\} \quad (1)$$

$$= 1 - 2F(x) + C(F(x),F(x)) \sim \frac{1}{2}(1-F(x)), x \to \infty,$$

erhält man die Behauptung

$$\mathbb{P}\{X > x \text{ und } Y > x\} \sim \frac{1}{2}\mathbb{P}\{X > x\}, x \to \infty.$$

Bei (1) wurde die *Siebformel*

$$\mathbb{P}(A^c \cap B^c) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B)$$

benutzt.

2. Der durch $\pi$ und $B$ beschriebene Markov-Prozess startet in Zustand 1, in dem er eine Zeit verweilt, welche die Exponentialverteilung Exp(1) hat. Von dort springt der Prozess zum Zustand 2, wo er wieder eine solche Wartezeit sitzen bleibt. Vom Zustand 2 geht es in den absorbierenden Zustand mit Wahrscheinlichkeit $p$, und mit Wahrscheinlichkeit $1-p$ springt er zum Zustand 1 (Neustart). Die dargestellte Verteilung $P$ hat demnach die Darstellung

$$P = (1-p)Exp(1)^{*2} + pExp(1)^{*2} * P.$$

Dies legt die Vermutung nahe, dass $P$ die Verteilung

$$P_0 = \sum_{n=1}^{\infty} p^{n-1}(1-p)Q^{*n}$$

mit $Q = Exp(1)^{*2}$ ist. Dies kann man nachrechnen mit der Formel für die Darstellung einer Summenverteilung als Phasentyp-Verteilung: die Darstellung für $Q$ ist

$$\pi^{(0)} = (1,0),\ B^{(0)} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix},$$

und damit hat $P_0$ die obige Darstellung wegen

$$b_{ij} = b_{ij}^{(0)} + pb_{i0}^{(0)}\pi_j^{(0)}.$$

Man kann auch die zugehörige momenterzeugende Funktion $f(x)$ ausrechnen: $f(x)$ erfüllt die Gleichung

$$f(x) = (1-p)(1-x)^{-2} + pf(x)(1-x)^{-2},$$

und dies ist äquivalent zu

$$f(x) = G((1-x)^{-2})$$

mit $G(z) = z(1-p)/(1-pz)$. Und $G(z)$ ist die erzeugende Funktion der verschobenen geometrischen Verteilung mit Punktwahrscheinlichkeiten $p_n = p^{n-1}(1-p), n = 1, 2, ...$ Dies führt mit der Fundamentalformel zur gleichen Summenverteilung $P_0$.

## 2 Aufgabe zur Solvabilität (30 Punkte)

a) Betrachten Sie die Ruinwahrscheinlichkeit $\psi(s)$ im klassischen Lundberg-Modell mit Startkapital $s$, der Prämienrate $c$, der Schadenintensität $\lambda = 2$, und der Schadenhöhenverteilung $Q = Pareto(a)$ mit Dichte

$$x \to ax^{-a-1}, x > 1.$$

1. Welche Ruinwahrscheinlichkeit ergibt sich für $a = 3/4$?
2. Berechnen Sie die Asymptotik für die Ruinwahrscheinlichkeit, wenn $a = 2$ und $c = 6$ ist.

Gemeint ist hier die Asymptotik von $\psi(s)$ für $s \to \infty$. (12 Punkte)

b) Der Gesamtschaden $S = X_1 + ... + X_N$ eines Versicherungsportefeuilles sei modelliert mit einem kollektiven Modell mit geometrischer Schadenzahlverteilung, also

$$\mathbb{P}\{N = k\} = (1-p)p^k, k = 0, 1, 2, ...$$

mit $0 < p < 1$ und einer Schadenhöhenverteilung mit Dichte

$$x \to \frac{1}{2}x\exp(-x) + \frac{1}{2}\exp(-x), x > 0.$$

Zeigen Sie:

$$\mathbb{P}\{S > t\} \leq \exp(-Rt), t > 0,$$

wobei $R$ die positive Lösung der Gleichung

$$p + 2r = p(1-r)^{-2}$$

ist. (18 Punkte)

**Hinweis zu b):** Die angegebene Verteilung der $X_i$ ist die Leiterhöhenverteilung in einem Lundbergmodell mit der Schadenhöhenverteilung, welche die Dichte

$$x \to x\exp(-x), x > 0$$

hat. In diesem Modell kann man dann die Lundberg-Ungleichung verwenden.

**Lösung:**

a) (a) Für $a = 3/4 < 1$ hat die Schadenhöhenverteilung $Q$ Erwartungswert Unendlich, und somit gilt $\psi(s) = 1$ für alle $s$.

(b) Der Erwartungswert von $Q$ ist für $a > 1$ gegeben durch $\mu = a/(a-1)$. Die Leiterhöhenverteilung $H$ zur Schadenhöhenverteilung $Q$ hat die Dichte

$$\begin{aligned} h(x) &= 1/\mu \text{ für } 0 < x < 1, \text{ und} \\ h(x) &= \frac{1}{\mu}x^{-a} \text{ für } x \geq 1. \end{aligned}$$

Damit unterscheidet sich die Tailwahrscheinlichkeit $H(x,\infty)$ für $x \geq 1$ nur durch einen konstanten Faktor von der Tailwahrscheinlichkeit der Pareto($a-1$)-Verteilung, $H$ ist also subexponentiell. Schließlich ist für $c = 6, \mu = 2, \lambda = 2$

$$p = \psi(0) = \lambda\mu/c = 2/3 < 1,$$

und daher gilt

$$\psi(s) \sim \frac{p}{1-p} H(s,\infty) = 2 \int_s^\infty \frac{1}{2} x^{-2} dx = s^{-1}.$$

b) Wir betrachten, wie in der Anleitung vorgeschlagen, ein Lundberg-Modell mit der Schadenhöhenverteilung $Q$, welche die Dichte $x \to x\exp(-x), x > 0$, besitzt. Die Verteilung $Q$ hat momenterzeugende Funktion $MEF_Q(x) = (1-x)^{-2}$. Die zugehörige Leiterhöhenverteilung $H$ ist die in der Aufgabe angegebene Schadenhöhenverteilung der Modellierung des Gesamtschadens $S$. Die Lundberg-Gleichung für den Anpassungskoeffizienten lautet

$$\lambda + rc = \lambda MEF_Q(r),$$

und die Gleichung der Aufgabenstellung entspricht dieser Gleichung mit $\lambda = p$ und $c = 2$ :

$$MEF_Q(r) = (1-r)^{-2}.$$

Die Cramér-Lundberg Ungleichung für die zugehörige Ruinwahrscheinlichkeit $\psi(s)$ lautet

$$\psi(s) \leq \exp(-Rs), s \geq 0.$$

Die Summenformel für die Ruinwahrscheinlichkeit ist

$$\psi(s) = \sum_{n=0}^{\infty} q^n (1-q) H^{*n}(s,\infty),$$

wobei $q = \lambda\mu/c = p$ gilt. Das wiederum ist gerade dasselbe wie die Tailwahrscheinlichkeit $\mathbb{P}\{S > s\}$.

## 3 Aufgabe zur Tarifierung (30 Punkte)

Aus stochastisch unabhängigen Schadendaten, bei denen eine Pareto-Verteilung mit Dichte

$$f(x) = ax^{-a-1}, x > 1, a > 0$$

unterstellt wird, soll mit der Maximum-Likelihood-Methode der Mittelwert bestimmt werden.

1. Welchen Maximum-Likelihood-Schätzer für den Mittelwert $\mu = a/(a-1)$ erhält man?
2. Welche asymptotische Varianz hat dieser Schätzer?
3. Vergleichen Sie die asymptotische Varianz aus 2. mit der asymptotischen Varianz des arithmetischen Mittels der Beobachtungen.

**Hinweis :** Da der Maximum-Likelihood-Schätzer so gut sein soll, müsste seine asymptotische Varianz kleiner sein als die des arithmetischen Mittels.

**Lösung:**

1. Der Maximum-Likelihood Schätzer für $a$ ist bekannt, er lautet

   $$\hat{a} = \left( \frac{1}{n} \sum_{i=1}^{n} \log(X_i) \right)^{-1}.$$

   Der Maximum-Likelihood Schätzer für $\mu$ ist somit $\hat{\mu} = \hat{a}/(\hat{a}-1)$. Zu demselben Ergebnis gelangt man, wenn man die gegebene Familie von Dichten mit dem Parameter $\mu$ versieht, also die Dichten

   $$p(x, \mu) = \frac{\mu}{\mu - 1} x^{-\mu/(\mu-1)-1}, x > 1, \tag{2}$$

   betrachtet.

2. Der Schätzer $\hat{a}$ hat asymptotische Varianz $\sigma^2(a) = a^2$ wegen

   $$(\sigma^2(a))^{-1} = -E\left[\frac{\partial^2}{\partial a^2} \log(p(X, a)\right] = a^{-2}.$$

   Mit $\mu = a/(a-1) = G(a)$ hat $\hat{\mu}$ asymptotische Varianz

   $$G'(a)^2\sigma^2(a) = \frac{a^2}{(a-1)^4}.$$

   Dasselbe Ergebnis erhält man, wenn man von den Dichten (2) ausgeht.

3. Die asymptotische Varianz des arithmetischen Mittels ist (für $a > 2$)

$$\begin{aligned} Var(X) &= \frac{a}{a-2} - \left(\frac{a}{a-1}\right)^2 \\ &= \frac{a(a-1)^2 - a^2(a-2)}{(a-2)(a-1)^2} = \frac{a}{(a-2)(a-1)^2}. \end{aligned}$$

Dies unterscheidet sich von $\sigma^2(\mu)$ durch den Faktor

$$C = a(a-2)/(a-1)^2 < 1,$$

und damit ist $\sigma^2(\mu) = CVar(X) < Var(X)$, der Maximum-Likelihood Schätzer ist also tatsächlich besser als das arithmetische Mittel.

# 7 Zusatzaufgabe (30 Punkte)

1. Ist die Lognormalverteilung eine Phasentyp-Verteilung? Begründen Sie Ihre Antwort. (7 Punkte)
2. Gibt es eine Dichte $f(x)$ einer Schadenhöhenverteilung mit der Eigenschaft, dass $f(x) > 0$ für alle $x > 0$ und
$$\int_0^\infty f(x)\log(f(x))dx > 0?$$
(7 Punkte)
3. Geben Sie ein Beispiel an für eine Verteilungsfunktion $F(x)$ mit
$$\int_0^\infty (1 - F(x))dx = \infty.$$
(7 Punkte)
4. Betrachten Sie das Lundberg-Modell mit den Parametern $\lambda$ als Schadenfrequenz, $c$ als Prämienrate und $Q = Exp(a)$ als Schadenhöhenverteilung. Unter welchen Bedingungen an $\lambda, c$ und $a$ existiert in diesem Modell der Anpassungskoeffizient? (9 Punkte)

**Lösung:**

1. Eine Phasentyp-Verteilung $P$ ist eine Kleinschadenverteilung, d.h. es gibt ein $r > 0$ mit
$$\int exp(rx)P(dx) < \infty.$$
Da die Lognormalverteilung diese Eigenschaft nicht hat, kann sie auch keine Phasentyp-Verteilung sein.

   Man kann auch so argumentieren: die momenterzeugende Funktion einer Phasentyp-Verteilung ist rational. Bei der Lognormalverteilung gibt es keine analytisch geschlossene Form der momenterzeugenden Funktion.
2. Da die Dichte $f(x)$ das Integral 1 ergibt, müssen die Werte von $f(x)$ überwiegend klein sein, der Logarithmus dieser Werte ist demnach überwiegend negativ. Dennoch sind für das Integral beliebige positive (wie auch negative) Werte möglich. Für eine Exp($a$)-Verteilung mit $f(x) = a\exp(-ax)$ erhält man
$$\int_0^\infty f(x)\log(f(x))dx = \log a - 1,$$
ein Wert, der für $a > 0$ beliebig zwischen $-\infty$ und $+\infty$ variiert.

3. Die Formel $\int(1-F(x))dx$ stellt den Erwartungswert dar. Ein Beispiel ist damit gegeben durch eine Pareto-Verteilung mit Parameter $a<1$, also beispielsweise durch

$$F(x)=1-x^{-1/2}, x>1.$$

4. Eine notwendige Voraussetzung für $\psi(s)<1$ für ein (und dann für alle) $s\geq 0$ ist

$$p=\lambda\mu/c<1, \tag{3}$$

wobei in unserem Beispiel $\mu=1/a$ ist. Die Gleichung für den Anpassungskoeffizienten lautet

$$\lambda+rc=\lambda MEF_X(r),$$

und in unserem Beispiel ist $MEF_X(r)=a/(a-r), r<a$. Einsetzen liefert die Gleichung

$$\lambda+rc=\lambda a/(a-r)$$

mit den Lösungen $r_1=0$ und $r_2=(ca-\lambda)/c$. Der Wert $r_2$ ist positiv (und $<a$) genau dann, wenn (3) gilt. Also ist Bedingung (3) notwendig und hinreichend für die Existenz des Anpassungskoeffizienten.

## 5 **Aufgabe zur Schadenreservierung** (Spezialwissen 2008) (30 Punkte)

Sei $C_{ik}$, $1 \leq i, k \leq n$, der Schadenstand von Anfalljahr i nach k Abwicklungsjahren und $S_{ik} = C_{ik} - C_{i,k-1}$ der Zuwachs. Benutzen Sie diese Bezeichnungen bei der Beantwortung folgender Fragen:

1a. Nennen Sie die drei Modellannahmen des Zuwachsquoten-Modells (ZQ). (3 Punkte)

1b. Nennen Sie die drei Modellannahmen des Chain-Ladder-Modells (CL). (3 Punkte)

1c. Ändern Sie nun die Modellierungen des Erwartungswerts in 1a und 1b in jeweils äquivalenter Weise so ab, dass zwei völlig identische Formulierungen entstehen bis auf den <u>einzigen</u> Unterschied, dass bei CL der Vorjahresstand genau dieselbe Rolle spielt wie das Volumenmaß bei ZQ. (6 Punkte)

1d. Formulieren Sie nun auch die Modellierungen der Varianz in 1a und 1b in jeweils äquivalenter Weise so um, dass zwei völlig identische Formulierungen entstehen bis auf den <u>einzigen</u> Unterschied, dass bei CL der Vorjahresstand genau dieselbe Rolle spielt wie das Volumenmaß bei ZQ. (3 Punkte)

2a. Das ZQ-Modell gelte für zwei disjunkte Teilbestände eines Versicherungsbestandes (jeweils mit der Bedarfsprämie als Volumenmaß, aber mit unterschiedlichen Parametern). Zeigen Sie, dass es dann im allgemeinen <u>nicht</u> auch für die Summe der beiden Teilbestände gilt. (4 Punkte)

2b. Ein entsprechendes Resultat wie 2a gilt auch für das CL-Modell. Deswegen schätzt man grundsätzlich die Reserve für die Summe zweier Teilbestände mittels der Summe der für jeden Teil ermittelten Reserveschätzer. Um auch den zugehörigen Prognosefehler zu ermitteln, braucht man eine Annahme über den stochastischen Zusammenhang zwischen den Teilportefeuilles. Wie lautet diese Modellannahme für das CL-Modell? (5 Punkte)

2c. Wie lautet die entsprechende Annahme für das ZQ-Modell? (3 Punkte)

2d. Wie kann man (z.B. für das ZQ-Modell) prüfen, ob zwischen zwei Teileportefeuilles ein im Sinne des Modells relevanter Zusammenhang besteht? (3 Punkte)

**Lösung der Aufgabe zur Schadenreservierung**

1a. (ZQ1) Die Zuwächse $S_{ik}$, $1 \le i, k \le n$, sind unabhängig.

(ZQ2) $E(S_{ik}) = v_i m_k$ mit bekanntem Volumen $v_i$ und unbekanntem Parameter $m_k$.

(ZQ3) $Var(S_{ik}) = v_i s_k^2$ mit unbekanntem Parameter $s_k^2$.

1b. (CL1) Die Anfalljahre $\{C_{i1}, \ldots, C_{in}\}$, $1 \le i \le n$, sind unabhängig.

(CL2) $E(C_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} f_k$ mit unbekanntem Parameter $f_k$.

(CL3) $Var(C_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} \sigma_k^2$ mit unbekanntem Parameter $\sigma_k^2$.

1c. Wegen ZQ1 ist im ZQ-Modell 1a die Modellierung $E(S_{ik}) = v_i m_k$ äquivalent mit

$E(S_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = v_i m_k$ . Wegen der allgemein gültigen Beziehung

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

ist die Modellannahme $E(C_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} f_k$ des CL-Modells äquivalent mit

$E(S_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} m_k$ (wobei $m_k = f_k - 1$).

1d. Wegen ZQ1 ist im ZQ-Modell die Annahme $Var(S_{ik}) = v_i s_k^2$ äquivalent mit

$Var(S_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = v_i s_k^2$. Wegen des allgemein gültigen

$$Var(S_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = Var(C_{ik} - C_{i,k-1} \mid C_{i1}, \ldots, C_{i,k-1}) = Var(C_{ik} \mid C_{i1}, \ldots, C_{i,k-1})$$

kann CL3 auch in der Form $Var(S_{ik} \mid C_{i1}, \ldots, C_{i,k-1}) = C_{i,k-1} s_k^2$ geschrieben werden.

2a. Seien $D_{ik}$, $T_{ik}$, $w_i$ und $n_k$ die zu $C_{ik}$, $S_{ik}$, $v_i$, $m_k$ analogen Bezeichnungen für das ZQ-Modell des zweiten Teilbestandes. Dann gilt

$$E(S_{ik} + T_{ik}) = E(S_{ik}) + E(T_{ik}) = v_i m_k + w_i n_k \,.$$

Wenn das ZQ-Modell auch für die Summe der beiden Teilbestände gälte (mit der Bedarfsprämie als Volumenmaß), müsste es Parameter $r_k$ geben mit

$$E(S_{ik} + T_{ik}) = (v_i + w_i) r_k \,.$$

Zusammengenommen hätte man dann die für alle i, k zu erfüllende Bedingung

$$v_i(r_k - m_k) = w_i(n_k - r_k) \quad \text{oder} \quad \frac{v_i}{w_i} = \frac{n_k - r_k}{r_k - m_k} \,,$$

was theoretisch höchst selten und praktisch nie erfüllt ist.

2b. In Verallgemeinerung von CL3 nimmt man an, dass

(CL4) $Cov(C_{ik}, D_{ik} \mid C_{i1}, \ldots, C_{i,k-1}, D_{i1}, \ldots, D_{i,k-1}) = \rho_k \sqrt{C_{i,k-1} D_{i,k-1}}$

mit unbekanntem Parameter $\rho_k$ gilt oder äquivalent

(CL4') $$Cov\left(\frac{C_{ik}}{C_{i,k-1}}, \frac{D_{ik}}{D_{i,k-1}} \;\middle|\; C_{i1}, \ldots, C_{i,k-1}, D_{i1}, \ldots, D_{i,k-1}\right) = \frac{\rho_k}{\sqrt{C_{i,k-1} D_{i,k-1}}} \,.$$

2c. Gemäß der in Aufgabe 1 gezeigten Analogie lautet die entsprechende ZQ-Annahme

(ZQ4) $\mathrm{Cov}(S_{ik}, T_{ik}) = r_k \sqrt{v_i w_i}$ .

2d. Man plottet die standardisierten Residuen des einen Bestandes gegen die des anderen Bestandes und prüft die entstehende Punktwolke auf einen (von Null verschiedenen) linearen Trend.

## 4 **Aufgabe zur Tarifkalkulation II** (Spezialwissen 2008) (30 Punkte)

1a. Begründen Sie durch ein möglichst einfaches Modell für Erwartungswert und Varianz eines (Einzel-)Risikos, dass die Varianz des Schadenbedarfs einer Gruppe von Risiken üblicherweise als umgekehrt proportional zum Volumen der Gruppe angenommen werden kann. (4 Punkte)

1b. Begründen Sie durch ein möglichst einfaches Modell für Erwartungswert und Varianz eines (Einzel-)Risikos mit Versicherungssumme u, dass die Varianz des Schadensatzes einer Gruppe von Risiken mit unterschiedlichen Versicherungssummen als umgekehrt proportional zum Versicherungssummenvolumen der Gruppe angenommen werden kann. (6 Punkte)

1c. Geben Sie zwei Beispiele von einfachen, vollständig spezifizierten Verteilungen (nicht nur E und Var wie in 1a und 1b) des Jahresgesamtschadens R eines Risikos in Abhängigkeit von der Versicherungssumme u, und zwar ein Beispiel, das die Modellannahme 1b erfüllt und eines, das sie nicht erfüllt. (8 Punkte)

2a. Beschreiben Sie, wie ein Versicherungsunternehmen an Hand seiner KH-Daten prüfen kann, ob die Gamma-Verteilung zu den Schadenbedarfsdaten passt. (Welche Daten nehmen Sie? Was machen Sie damit? Wie kommen Sie zur Teststatistik und zur Entscheidung?) (6 Punkte)

2b. Sie wollen prüfen, ob in der am stärksten besetzten Zelle eines bestehenden Tarifs ein zusätzliches Risikomerkmal zu einer signifikanten Prämiendifferenzierung führt. Nehmen Sie an, der Schadenbedarf jeder Jahreseinheit sei Gamma-verteilt mit bekanntem Formparameter $\alpha$. Dann hat der Schadenbedarf Z von v Jahreseinheiten bekanntlich eine Gammaverteilung mit Formparameter $v\alpha$, d.h. Verteilungsdichte $f(z \mid \mu, v\alpha)$ bei Mittelwert $\mu$. Geben Sie mit diesen Bezeichnungen (an Stelle der expliziten Gamma-Dichte) an, wie der Likelihood-Quotienten-Test aussieht um festzustellen, ob ein Risikomerkmal mit I Ausprägungsklassen mit beobachteten (einjährigen) Schadenbedarfen $z_1, \ldots, z_I$ bei Volumina $v_1, \ldots, v_I$ ein signifikantes zusätzliches Tarifmerkmal in dieser Risikogruppe darstellt (nichts ausrechnen, nur Ansatz aufschreiben!). Erläutern Sie alle dabei benutzten zusätzlichen Bezeichnungen und sagen Sie, wie man sie ermittelt. (6 Punkte)

**Lösung zur Tarifkalkulation II**

1a. Die Risikogruppe habe I ganzjährig versicherte unabhängige Risiken mit Jahresgesamtschaden $R_i$, $1 \leq i \leq I$. Sei $E(R_i) = m$ und $Var(R_i) = s^2$. Dann ist $Z := (R_1 + \ldots + R_I)/I$ der Schadenbedarf der Risikogruppe und I das zugehörige Volumen und es gilt $Var(Z) = \frac{1}{I^2}\sum_{i=1}^{I} Var(R_i) = \frac{s^2}{I}$.

1b. Mit den Bezeichnungen von 1a gelte jetzt $E(R_i) = u_i\mu$ und $Var(R_i) = u_i\sigma^2$ mit der Versicherungssumme $u_i$ von Risiko i. Dann ist $v := u_1 + \ldots + u_I$ das VS-Volumen der Gruppe, $Z := (R_1 + \ldots + R_I)/v$ ihr Schadensatz und es gilt

$$Var(Z) = \frac{1}{v^2}\sum_{i=1}^{I} Var(R_i) = \frac{1}{v^2}\sum_{i=1}^{I} u_i\sigma^2 = \frac{\sigma^2}{v}.$$

1c. Sei $R(u) = N(u)\cdot m$ mit $N(u) \sim \text{Poisson}(u\lambda)$, $\lambda > 0$, $m > 0$. Dann gilt $E(R(u)) = u\lambda m$ und $Var(R(u)) = Var(N(u))m^2 = u\lambda m^2$ , d. h. es ist ein Spezialfall von 1b.
Hingegen erfüllt die (Unfalltod-)Verteilung $R(u) = u\cdot B$ mit Bernoulli-verteiltem B (d. h. B = 1 mit Wahrscheinlichkeit p und B = 0 sonst) nicht die Annahmen von 1b, denn $Var(R(u)) = u^2p(1-p)$ ist nicht proportional zur Versicherungssumme u.

2a. Man nehme alle KH-Risiken, die das ganze letzte Jahr versichert waren (oder besser nur die aus der am stärksten besetzten Tarifgruppe, wobei auch unterjährige Risiken zu je einjährigen Risiken zusammengefasst werden können). Dann teile man sie nach der Höhe des Jahresschadens in mindestens I = 5 Intervalle so auf, dass jedes Intervall mindestens 10 Jahreseinheiten enthält (für die Asymptotik der Testgröße). Nun kann man eine Gammaverteilung anpassen, wobei deren Parameter mit dem zur Intervall-Aufteilung passenden ML-Schätzer berechnet werden sollten (oder alternativ mit dem Minimum-Chi-Quadrat-Schätzer). Daraus ergibt sich die Wahrscheinlichkeit $w_i$, dass der Jahresschaden eines Risikos in das Intervall i fällt. Damit berechnet man dann die Teststatistik

$$T = \sum_{i=1}^{I}\frac{(n_i - w_i n)^2}{w_i n}$$

des Chi-Quadrat-Anpassungstests, wobei $n_i$ die Anzahl Jahreseinheiten in Intervall i ist und n deren Gesamtzahl. T ist (asymptotisch) chi-quadrat-verteilt mit I-3 Freiheitsgraden. Wenn der Wert von T unter dem 95%-Perzentil liegt, kann die Gammaverteilung beibehalten werden.

2b. Der Likelihood-Quotienten-Test hat die Testgröße

$$Q = 2 \cdot \ln\left(\frac{L_1}{L_0}\right) \text{ mit } L_1 = \prod_{i=1}^{I} f(z_i \mid \hat{\mu}_i, v_i \alpha) \text{ und } L_0 = \prod_{i=1}^{I} f(z_i \mid \hat{\mu}, v_i \alpha)\,.$$

Dabei ist $L_1$ die gemeinsame Wahrscheinlichkeit der Beobachtungen $z_1, \ldots, z_I$ bei unterschiedlichen Erwartungswerten $\mu_i$ und $L_0$ die gemeinsame Wahrscheinlichkeit der Beobachtungen $z_1, \ldots, z_I$ bei einheitlichem Erwartungswert $\mu$. $\hat{\mu}_i$ bzw. $\hat{\mu}$ sind die zugehörigen ML-Schätzer, die man durch Maximieren von $L_1$ bzw. $L_0$ erhält. Q ist (asymptotisch) chi-quadrat-verteilt mit I-1 Freiheitsgraden, d. h. wenn der Wert von Q über dem 95%-Perzentil liegt, ergibt das Risikomerkmal eine signifikante Differenzierung.

## 6 Aufgabe zur Risikoteilung (Spezialwissen 2008) (30 Punkte)

Sei X die Höhe des Einzelschadens eines Portefeuilles, $F(x) = P(X \leq x)$ die zugehörige Verteilungsfunktion und $f(x) = dF(x)/dx$ die (als existent angenommene) Dichte. Sei weiter $0 < a < b < \infty$. Schäden mit $a \leq X < b$ werden als Normalschäden bezeichnet.

Geben Sie

a) die Höhe des Normalschadens (6 Punkte),

b) die Höhe des bei b kupierten Schadens (6 Punkte),

c) die Höhe des Überschadens über der Grenze b (6 Punkte),

d) die Höhe des (echten) Layer(einzel)schadens bei Priorität a und Layerhaftung b – a (6 Punkte)

jeweils die folgenden drei Ausdrücke an:

(i) den Erwartungswert als Funktion von X, a, b,

(ii) den Erwartungswert als Funktion von f, a, b,

(iii) die Verteilungsfunktion als Funktion von F, a, b.

Beispiel: Für X selbst lauten die drei gesuchten Ausdrücke E(X), $\int_0^\infty x\, f(x)\, dx$, F(x) (wobei hier die Schadengrenzen a und b nicht benötigt werden).

e) Rechnen Sie die Erwartungswertdarstellung $\int_0^\infty t\, d\tilde{F}(t)$ mit der in d)(iii) angegebenen Verteilungsfunktion $\tilde{F}$ aus und formen sie so um, dass sie sich letztlich als identisch mit der in d)(ii) angegebenen Darstellung des Erwartungswerts als Funktion der Dichte f(x) erweist. (6 Punkte).

**Lösung zur Risikoteilung**

a) $E(X \mid a \le X < b)$ ; $\dfrac{\int_a^b x\,f(x)\,dx}{\int_a^b f(x)\,dx}$ ;

$F_{normal}(t) = 0$ für $t < a$, $F_{normal}(t) = \dfrac{F(t) - F(a)}{F(b) - F(a)}$ für $t \in [a, b[$ , $F_{normal}(t) = 1$ für $t \ge b$.

b) $E(\min(X, b))$ ; $\int_0^b x\,f(x)\,dx + \int_b^\infty b\,f(x)\,dx$ ;

$F_{kup}(t) = F(t)$ für $t < b$ , $F_{kup}(t) = 1$ sonst.

c) $E(X-b \mid X > b)$ ; $\dfrac{\int_b^\infty (x-b)\,f(x)\,dx}{\int_b^\infty f(x)\,dx}$ ; $F_{über}(t) = \dfrac{F(b+t) - F(b)}{1 - F(b)}$.

d) $E(\min(X-a, b-a) \mid X > a) = E(\min(X, b) \mid X>a) - a$ ; $\dfrac{\int_a^b (x-a)\,f(x)\,dx + \int_b^\infty (b-a)\,f(x)\,dx}{\int_a^\infty f(x)\,dx}$ ;

$F_{layer}(t) = \dfrac{F(a+t) - F(a)}{1 - F(a)}$ für $t < b-a$, $F_{layer}(t) = 1$ für $t \ge b-a$.

e) Wegen $dF_{layer}(t) = \dfrac{dF(a+t)}{1 - F(a)}$ für $t < b-a$ und weil $F_{layer}$ bei $b-a$ einen Sprung der Höhe

$1 - \dfrac{F(b) - F(a)}{1 - F(a)} = \dfrac{1 - F(b)}{1 - F(a)}$ macht, ist

$$\int_0^\infty t\,dF_{layer}(t) = \int_0^{b-a} \frac{t\,dF(a+t)}{1 - F(a)} + (b-a)\frac{1 - F(b)}{1 - F(a)}$$

$$= \frac{1}{1 - F(a)}\left( \int_a^b (x-a)\,dF(x) + (b-a)\int_b^\infty dF(x) \right),$$

und Letzteres ist wegen $dF(x) = f(x)\,dx$ identisch mit der in Teilaufgabe d angegebenen Erwartungswertdarstellung als Funktion von f, a und b.