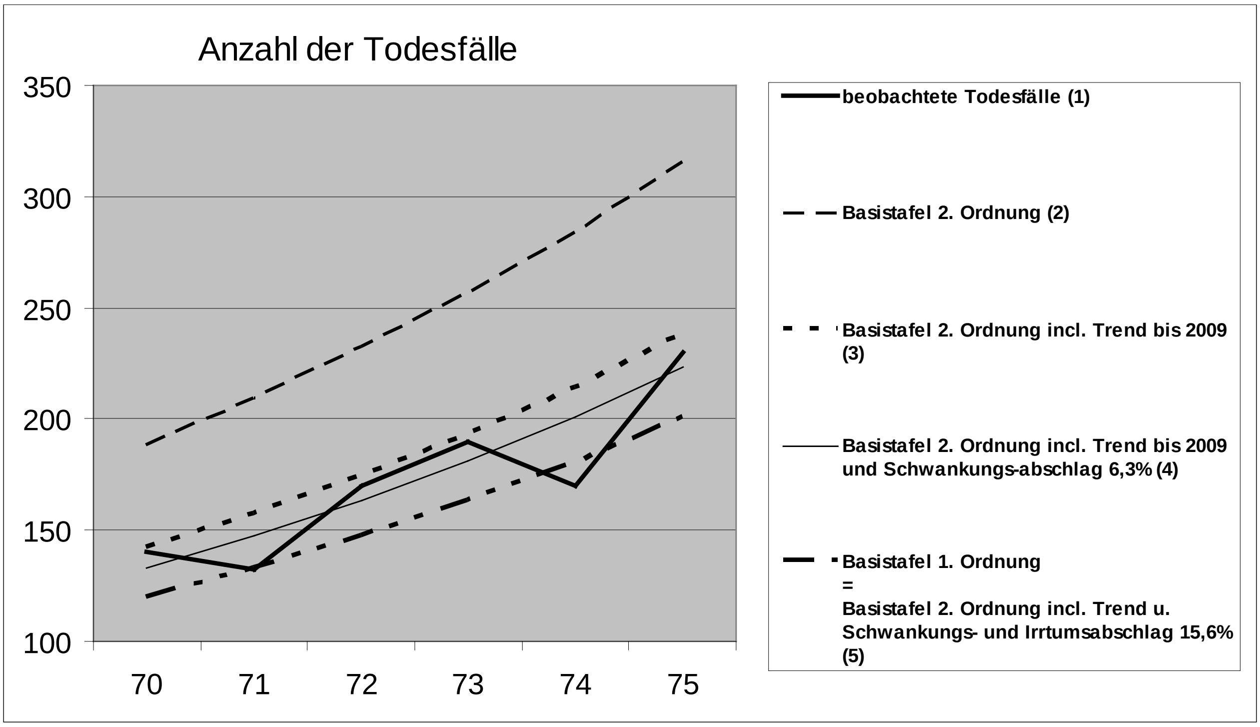
Bericht zur Prüfung im Mai 2010 über stochastische Risikomodellierung und statistische Methoden (Grundwissen)

Richard Herrmann, Dietmar Pfeifer, Dominik Schäfer, Viktor Sandor

Aufgabe 1 (32 Punkte): Sie sind Verantwortlicher Aktuar einer Lebensversicherung und müssen die Angemessenheit der biometrischen Rechnungsgrundlagen in einem Bestand von Rentenversicherungen in einem Tarif beurteilen. Für die Deckungsrückstellung wird die DAV-Sterbetafel 2004 R (Aggregattafel) verwendet. Die beobachteten Todesfälle sowie die erforderlichen Werte aus der DAV-Tafel sind:

			Sterbewahrscheinlichkeiten nach DAV 2004 R Aggregattafel				
Alter	Bestand	beobachtete Todesfälle	Basistafel 2. Ordnung	Trend-Faktor bis 2009	Basistafel 2. Ordnung incl. Trend bis 2009	Basistafel 2. Ordnung incl. Trend bis 2009 und Schwankungs-abschlag 6,3%	Basistafel 1. Ordnung = Basistafel 2. Ordnung incl. Trend u. Schwankungs- und Irrtumsabschlag 15,6%
		(1)	(2)		(3)	(4)	(5)
70	10000	140	0,0188	0,7538	0,0142	0,0133	0,0120
71	10000	132	0,0209	0,7519	0,0157	0,0148	0,0133
72	10000	170	0,0232	0,7511	0,0174	0,0163	0,0147
73	10000	190	0,0257	0,7515	0,0193	0,0181	0,0163
74	10000	170	0,0284	0,7532	0,0214	0,0201	0,0181
75	10000	230	0,0315	0,7562	0,0239	0,0223	0,0201

Tabelle 1



a) Überprüfen Sie, ob die Anwendung der biometrischen Rechnungsgrundlagen nicht mehr gerechtfertigt ist.

   i) Begründen Sie, mit welchen erwarteten Todesfällen die beobachteten Todesfälle verglichen werden.

   ii) Führen Sie die Überprüfung mit Hilfe dreier unterschiedlicher Testverfahren durch und interpretieren Sie Ihr Ergebnis. (Signifikanzniveau $\alpha = 5$ %)

b) Die biometrischen Rechnungsgrundlagen enthalten einen Sicherheitsabschlag zur Berücksichtigung des Schwankungsrisikos.

   i) Stellen Sie die Herleitung des Schwankungsabschlags der DAV 2004 R dar.

   ii) Wie und unter welchen zusätzlichen Bedingungen können Sie den Schwankungszuschlag in die Überprüfung einbeziehen?

   iii) Zu welchen Ergebnissen führen die Tests in diesem Fall? (Signifikanzniveau $\alpha = 5$ %)

<u>Lösungshinweise:</u>

- Verteilungsfunktionen F(k) der Binomialverteilung:

Bin(4;0,5)

k	F(k)
0	0,12960
1	0,47520
2	0,82080
3	0,97440
4	1,00000

Bin(5;0,5)

k	F(k)
0	0,03125
1	0,18750
2	0,50000
3	0,81250
4	0,96875
5	1,00000

Bin(6;0,5)

k	F(k)
0	0,01563
1	0,10938
2	0,34375
3	0,65625
4	0,89063
5	0,98438
6	1,00000

- Quantile der χ^2-Verteilung:

Freiheitsgrade	**90%**	**92,5%**	**95%**	**97,5%**	**99%**	**99,5%**
1	2,70554	3,17005	3,84146	5,02389	6,63490	7,87944
2	4,60517	5,18053	5,99146	7,37776	9,21034	10,59663
3	6,25139	6,90464	7,81473	9,34840	11,34487	12,83816
4	7,77944	8,49628	9,48773	11,14329	13,27670	14,86026
5	9,23636	10,00831	11,07050	12,83250	15,08627	16,74960
6	10,64464	11,46595	12,59159	14,44938	16,81189	18,54758
7	12,01704	12,88343	14,06714	16,01276	18,47531	20,27774
8	13,36157	14,26974	15,50731	17,53455	20,09024	21,95495
9	14,68366	15,63094	16,91898	19,02277	21,66599	23,58935
10	15,98718	16,97137	18,30704	20,48318	23,20925	25,18818

- Die Teststatistiken für den χ^2-Test betragen für die Spalten (2) bis (5) der Tabelle 1:

(2)	(3)	(4)	(5)
144,38	13,73	7,58	16,15

Lösung

a)

i)

Zur Überprüfung, ob die Anwendung der Rechnungsgrundlagen nicht mehr gerechtfertigt ist, stehen verschiedene statistische Testverfahren zur Verfügung. Hierbei werden die beobachteten Todesfälle mit den nach den verwendeten Rechnungsgrundlagen erwarteten Todesfällen im Beobachtungszeitrum miteinander verglichen.

Der Vergleich der beobachteten Todesfälle ist zunächst mit den erwarteten Todesfällen, die sich unter Verwendung der Basistafel 2. Ordnung inkl. Trend bis 2009 („Jahrestafel 2009“) ergeben, durchzuführen, da nur diese nach unserer Annahme die tatsächlichen Sterblichkeiten im Beobachtungszeitraum (d.h. ohne Sicherheiten bez. auf das statistische Schwankungs- und Irrtumsrisiko) darstellen

Die Nullhypothese H_0 lautet also:

Die tatsächlichen Sterbewahrscheinlichkeiten stimmen mit denen der Basistafel 2. Ordnung inkl. Trend bis 2009 überein.

ii)

Die Überprüfung wird nun mittels drei verschiedener Testverfahren durchgeführt:

1. Vorzeichentest
2. Iterationstest
3. χ^2-Test

Vorzeichentest:

Die Teststatistik T ergibt sich aus der Anzahl der positiven Differenzen zwischen den beobachteten und den erwarteten Todesfällen, d.h. T = 0.

Unter Gültigkeit der Nullhypothese ist *T* binomialverteilt mit Wahrscheinlichkeit ½ für jedes Vorzeichen, d.h. es gilt

$$T \sim B\left(6, \frac{1}{2}\right).$$

und damit

$$P(T \leq 0) = \left(\frac{1}{2}\right)^6 = 0{,}015625 < 5\%/2 = 0{,}025$$

Bei einem Signifikanzniveau von 5 % wird die Nullhypothese daher abgelehnt, die rechnungsmäßige Sterblichkeit stimmt mit der tatsächlichen nicht überein.

Iterationstest:

Die Teststatistik ergibt sich aus der Anzahl der Vorzeichenwechsel, die sich bei den Differenzen zwischen den beobachteten und den rechnungsmäßig erwarteten Todesfällen ergeben, d.h. T = 0.

Unter Gültigkeit der Nullhypothese ist T binomialverteilt mit Wahrscheinlichkeit ½ für jedes Vorzeichen, d.h. es gilt

$$T \sim B\left(5, \frac{1}{2}\right).$$

und damit

$$P(T \leq 0) = \left(\frac{1}{2}\right)^5 = 0,03125 < 5\% = 0,05$$

Bei einem Signifikanzniveau von 5 % wird die Nullhypothese daher abgelehnt, die rechnungsmäßige Sterblichkeit stimmt mit der tatsächlichen nicht überein.

χ^2-Test:

Die Teststatistik lautet

$$T = \sum_{x=70}^{75} \frac{\left(Z_x - E_x\right)^2}{E_x},$$

wobei Z_x die Anzahl der beobachteten Todesfälle und E_x die Anzahl der rechnungsmäßig erwarteten Todesfälle bezeichne.

Der Wert der Teststatistik ist in der Aufgabenstellung angegeben mit 13,73.

Die Teststatistik ist unter der Nullhypothese näherungsweise χ^2-verteilt mit 6 Freiheitsgraden.

Das 95%-Quantil der χ^2-Verteilung mit 6 Freiheitsgraden ist 12,59 (vgl. Tabelle).

Da der Wert der Teststatistik dieses überschreitet, wird die Nullhypothese auch nach diesem Test abgelehnt.

Interpretation der Ergebnisse:
Alle drei Tests gegen die Basistafel 2. Ordnung inkl. Trend bis 2009 lehnen die Nullhypothese bei einem Signifikanzniveau von 5% ab. Dies heißt aber nicht zwingend, dass die verwendeten Rechnungsgrundlagen nicht der Realität entsprechen, da diese Entscheidung nur mit einer

Fehlerwahrscheinlichkeit von 5% gilt. Vielmehr kann dieses Ergebnis auch Folge einer zufallsbedingten Schwankung sein. Um die Angemessenheit der Rechnungsgrundlagen abschließend beurteilen zu können, sind weitere Auswertungen (beispielsweise der vorangegangen Jahre) erforderlich.

b)

i)

Die Berechnung des Schwankungsabschlags erfolgt hier anhand eines Modellbestandes von 200.000 Versicherten (je 100.000 Männer bzw. Frauen). Es wird zwischen Aufschub- und Rentenbezugszeit unterschieden und angenommen, dass sich 90 % der Verträge in der Aufschubzeit befinden.

Bei Rentenversicherungen wird im Todesfall in der Aufschubzeit – wenn überhaupt – eine Leistung ausgezahlt, die kleiner als das zu diesem Zeitpunkt vorhandene Deckungskapital ist.

Ziel ist es, einen Schwankungsabschlag s_x^α auf die Sterbewahrscheinlichkeit q_x so zu ermitteln, dass der unter Einhaltung des Sicherheitsniveaus 1-α maximal zulässige Schaden, der durch eine geringere Anzahl von Todesfällen als rechnungsmäßig erwartet entsteht, ausgeglichen werden kann.

Wird mit einer Sterbewahrscheinlichkeit q_x kalkuliert, so geht man davon aus, dass im Durchschnitt ein Betrag von $q_x \cdot L_x^M \cdot V_x$ frei wird. Der Schwankungsabschlag muss also so gewählt werden, dass mit Wahrscheinlichkeit 1-α die durch Tod im Modellbestand freiwerdende Deckungsrückstellung $\sum_x T_x^M V_x$ größer als die mit den Sterbewahrscheinlichkeiten $q_x - s_x^\alpha$ berechnete freiwerdende Deckungsrückstellung ist:
D.h. es muss gelten:

$$\mathrm{P}\left(\sum_x T_x^M V_x \geq \sum_x \left(q_x - s_x^\alpha\right) L_x^M V_x\right) \overset{!}{=} 1-\alpha \,,$$

Wir können nach dem Zentralen Grenzwertsatz annehmen, dass

$$Z := \sum_x T_x^M V_x$$

eine normalverteilte Zufallsvariable ist mit Erwartungswert

$$\mathrm{E}(Z) = \sum_x q_x L_x^M V_x$$

und Varianz

$$\mathrm{Var}(Z) = \sum_x q_x (1-q_x) L_x^M V_x^2 \,.$$

Mit diesen Bezeichnungen sowie der Vorgabe $s_x^\alpha = s^\alpha q_x$ mit $0 < s^\alpha < 1$ ergibt sich aus obiger Bedingung analog die äquivalente Darstellung

$$\mathrm{P}\left(Z \geq (1-s^{\alpha})\mathrm{E}(Z)\right) \stackrel{!}{=} 1-\alpha$$

$$\Leftrightarrow \qquad \mathrm{P}\left(\frac{Z-\mathrm{E}(Z)}{\sqrt{\mathrm{Var}(Z)}} \geq -s^{\alpha}\frac{\mathrm{E}(Z)}{\sqrt{\mathrm{Var}(Z)}}\right) \stackrel{!}{=} 1-\alpha$$

$$\Rightarrow \qquad s^{\alpha} = u_{1-\alpha}\frac{\sqrt{\mathrm{Var}(Z)}}{\mathrm{E}(Z)} = u_{1-\alpha}\frac{\sqrt{\sum_x q_x(1-q_x)L_x^M V_x^2}}{\sum_x q_x L_x^M V_x}.$$

ii)

Haben die Tests gegen die Basistafel 2. Ordnung inkl. Trend bis 2009 ergeben, dass die Nullhypothese zu dem vorgegebenen Niveau abgelehnt wird, kann noch überprüft werden, ob sich bei Tests gegen die Basistafel 2. Ordnung inkl. Trend bis 2009 und Schwankungsabschlag das gleiche Ergebnis einstellt. Ergibt sich auch hier eine Ablehnung der Nullhypothese, kann dies ein Hinweis darauf sein, dass die in den Rechnungsgrundlagen enthaltenen Sicherheiten „aufgebraucht“ wurden. In diesem Fall ist eventuell eine Anpassung der Rechnungsgrundlagen geboten.

iii)

Vorzeichentest:

Die Teststatistik T ergibt sich aus der Anzahl der positiven Differenzen zwischen den beobachteten und den erwarteten Todesfällen, d.h. T = 4.

Unter Gültigkeit der Nullhypothese ist T binomialverteilt mit Wahrscheinlichkeit ½ für jedes Vorzeichen, d.h. es gilt

$$T \sim B\left(6, \frac{1}{2}\right).$$

und damit

$P(T \leq 4) = 0{,}89063$, d.h. $P(T \leq 4) > 5\%/2 = 0{,}025$ und $P(T \leq 4) < 1-5\%/2 = 0{,}975$

Bei einem Signifikanzniveau von 5 % kann die Nullhypothese daher nicht abgelehnt werden.

Iterationstest:

Die Teststatistik ergibt sich aus der Anzahl der Vorzeichenwechsel, die sich bei den Differenzen zwischen den beobachteten und den rechnungsmäßig erwarteten Todesfällen ergeben, d.h. T = 4.

Unter Gültigkeit der Nullhypothese ist T binomialverteilt mit Wahrscheinlichkeit ½ für jedes Vorzeichen, d.h. es gilt

$$T \sim B\left(5, \frac{1}{2}\right).$$

und damit

$P(T \leq 4) = 0{,}96875$, d.h. $P(T \leq 4) > 5\%/2 = 0{,}025$ und $P(T \leq 4) < 1 - 5\%/2 = 0{,}975$

Bei einem Signifikanzniveau von 5 % kann die Nullhypothese daher nicht abgelehnt werden.

χ^2**-Test:**

Die Teststatistik lautet

$$T = \sum_{x=70}^{75} \frac{\left(Z_x - E_x\right)^2}{E_x},$$

wobei Z_x die Anzahl der beobachteten Todesfälle und E_x die Anzahl der rechnungsmäßig erwarteten Todesfälle bezeichne.

Der Wert der Teststatistik ist in der Aufgabenstellung angegeben mit 7,58.

Die Teststatistik ist unter der Nullhypothese näherungsweise χ^2-verteilt mit 6 Freiheitsgraden.

Das 95%-Quantil der χ^2-Verteilung mit 6 Freiheitsgraden ist 12,59 (vgl. Tabelle).

Da der Wert der Teststatistik dieses nicht überschreitet, wird die Nullhypothese auch nach diesem Test nicht abgelehnt.

Aufgabe 2 (22 Punkte): Die folgende Tabelle enthält die Jahres-Schäden einer Sachversicherung der letzten 7 Jahre in Tausend €:

Jahr	2003	2004	2005	2006	2007	2008	2009
Schaden	322	87	155	993	1366	814	324

a) Erstellen Sie für diesen Datensatz ein Histogramm mit drei Klassen, die folgendermaßen abgegrenzt sind:

i	0	1	2	3
v_i	0	300	1000	1500

b) Überprüfen Sie mit Hilfe eines P-P-Plots, ob die Schäden (ausgedrückt in Mio. €) als Realisierungen einer $\mathrm{Pa}a(2)$-Pareto-Verteilung angesehen werden können.

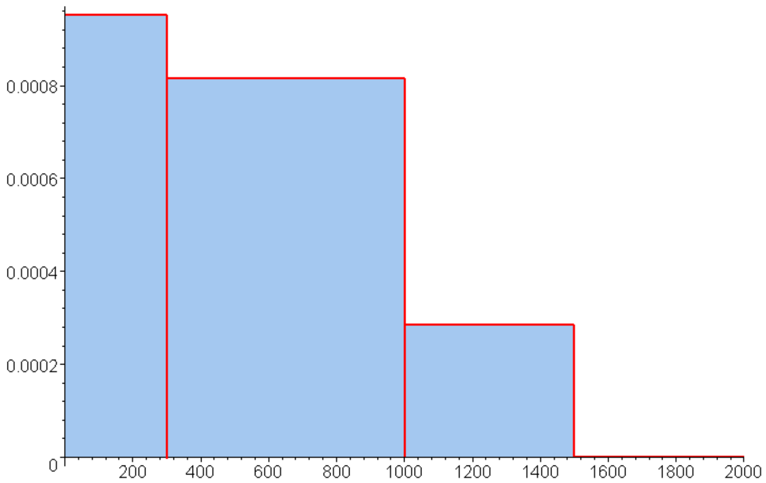
c) Berechnen Sie unter dieser Hypothese den Value at Risk sowie den Expected Shortfall zum Risikoniveau $\alpha = 0{,}005$.

d) Mit welcher Transformation kann man $\mathrm{Pa}a(2)$-verteilte Risiken in $\mathrm{E}(1)$-exponentialverteilte Risiken umrechnen, so dass der zugehörige P-P-Plot gleich bleibt?

Lösung:

a) Berechnung der Kenngrößen für das Histogramm:

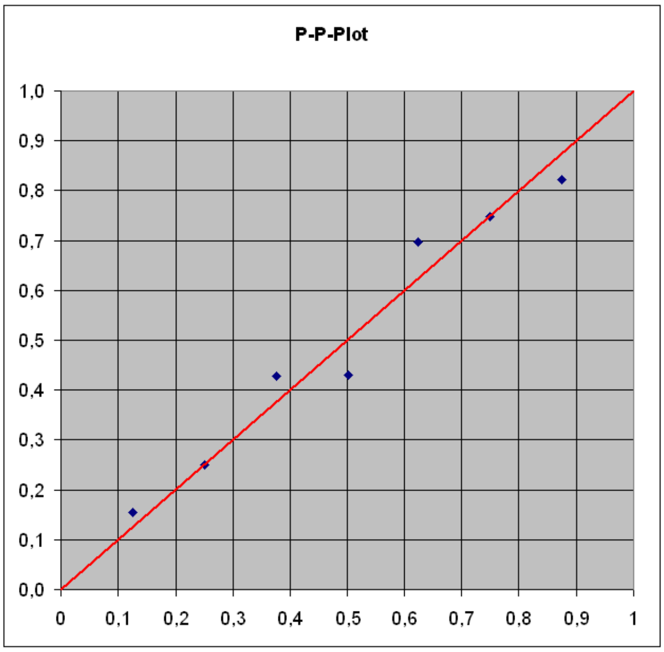
i	0	1	2	3
Klassengrenzen v_i	0	300	1000	1500
Intervallbreiten b_i		300	700	500
Rel. Klassenhäufigkeiten $f(K_i)$		$\frac{2}{7}$	$\frac{4}{7}$	$\frac{1}{7}$
Höhen $H_i = \frac{f(K_i)}{b_i}$		$\frac{2}{2100} = 0,000952$	$\frac{4}{4900} = 0,000816$	$\frac{1}{3500} = 0,000286$



Histogramm

b) Kenngrößen für den P-P-Plot, mit $F(x) = 1 - \frac{1}{(1+x)^2}$:

k	1	2	3	4	5	6	7
$\frac{k}{8}$	**0,125**	**0,25**	**0,375**	**0,5**	**0,625**	**0,75**	**0,875**
$x_{(k)}$	0,087	0,155	0,322	0,324	0,814	0,993	1,366
$F\left(x_{(k)}\right)$	**0,154**	**0,250**	**0,428**	**0,430**	**0,696**	**0,748**	**0,821**



Beurteilung: nach visuellem Eindruck akzeptabel

c) Es gilt: $F(x) = 1 - \frac{1}{(1+x)^2} = u \;\Leftrightarrow\; x = \frac{1}{\sqrt{1-u}} - 1$ für $0 < u < 1$, d.h. es gilt

$$\mathrm{VaR}_\alpha(X) = F^{-1}(1-\alpha) = \frac{1}{\sqrt{\alpha}} - 1$$

$$\mathrm{ES}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(X)\,du = \frac{1}{\alpha}\int_0^\alpha \left(\frac{1}{\sqrt{\alpha}} - 1\right) du = \frac{1}{\alpha}\left(2\sqrt{\alpha} - \alpha\right) = \frac{2}{\sqrt{\alpha}} - 1;$$

speziell: $\mathrm{VaR}_{0,005}(X) = 13{,}142$ und $\mathrm{ES}_{0,005}(X) = 27{,}284$.

d) Die streng monotone stetige Transformation $T(x) = -\ln(1-F(x)) = -\ln\left(\frac{1}{(1+x)^2}\right) = 2\cdot\ln(1+x)$ leistet das Verlangte, denn $F(X)$ ist $\mathrm{U}[0,1]$-verteilt.

Aufgabe 3 (36 Punkte): Die Krankheitsdauer X in Wochen wird als Zufallsvariable mit Dichte

$$f(x) = \begin{cases} 2\beta^2 x e^{-(\beta x)^2} & , x > 0 \\ 0 & , x \le 0 \end{cases}$$

modelliert, wobei $\beta > 0$ gilt. Gegeben seien Beobachtungen $x_1, \cdots, x_n$.

a) Gehört f zu einer Exponentialfamilie? Geben Sie gegebenenfalls den natürlichen Parameter an.

b) Bestimmen Sie $E\left(X^2\right)$ mit Hilfe von a).

c) Weisen Sie nach, dass $L : [0,\infty) \to \mathbb{R},\; L(\beta) = \beta^{2n} \exp\left(-\beta^2 \sum_{i=1}^n x_i^2\right)$ eine Likelihoodfunktion zu den gegebenen Daten ist.

d) Bestimmen Sie für den Parameter β einen Maximum Likelihood Schätzer zu den gegebenen Daten.

e) Bestimmen Sie die Informationsmatrix.

Für den Rest der Aufgabe gelte $n = 16$ und $\sum_{i=1}^{16} x_i = 20$, $\sum_{i=1}^{16} x_i^2 = 36$.

f) Bestimmen Sie den Maximum Likelihood Schätzwert für diese Daten.

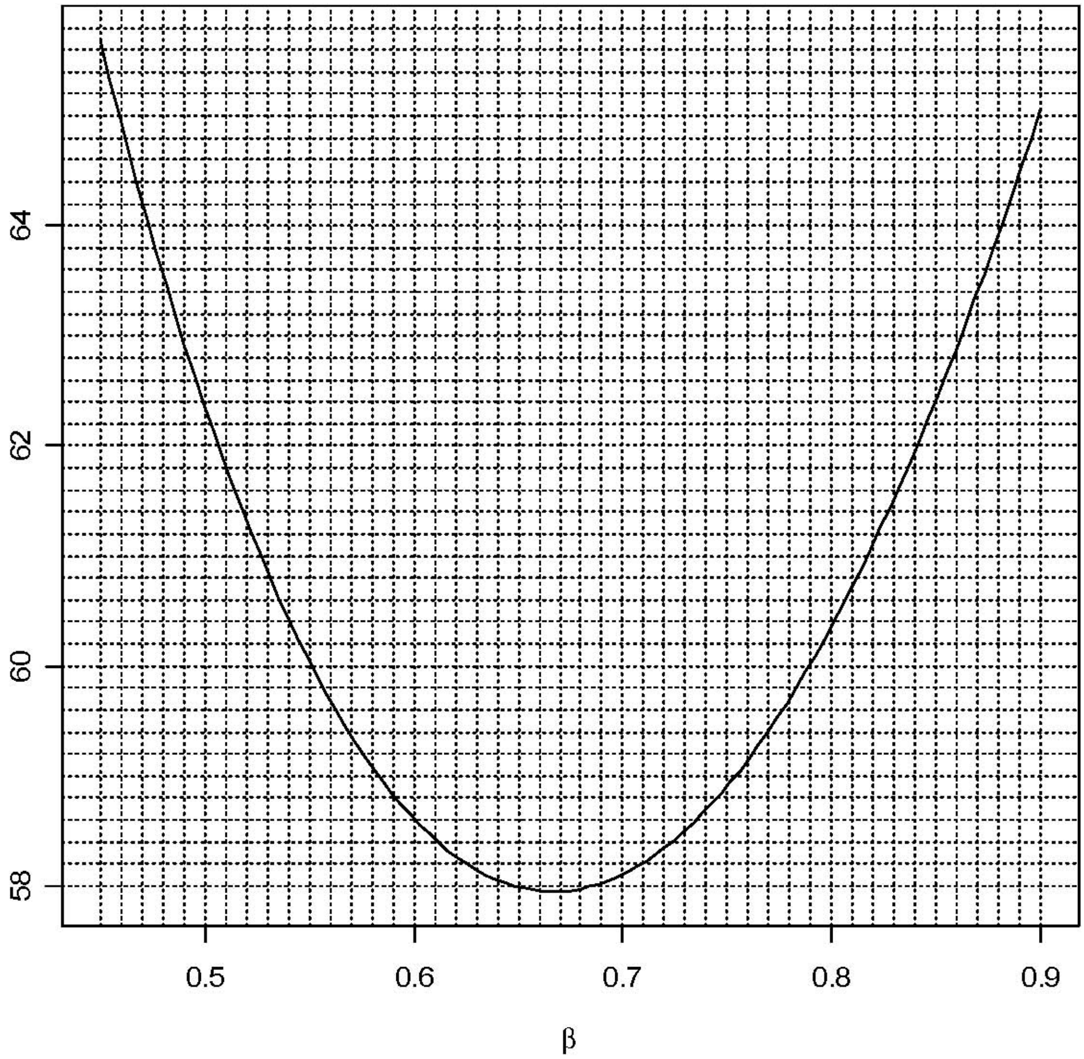
g) Testen Sie mit dem Likelihood Quotienten Test die Hypothese $H_0 : \beta = \beta_0 = 0,5$ zum Signifikanzniveau $\alpha = 10\%$.

Quantile der χ_m^2-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

h) Geben Sie mit Hilfe des Likelihood Quotienten Tests ein Schätzintervall für β zu einem Niveau von 95 % an. Verwenden Sie dazu den folgenden Funktionsgraphen und begründen Sie Ihre Antwort.

Graph der Funktion $\beta \mapsto -64\ln(\beta) + 72\beta^2$



Lösung

Zu a)

f gehört zu einer Exponentialfamilie, denn es gilt für $x > 0$:

$$f(x) = \exp\left(-\beta^2 x^2 + \ln \beta^2 + \ln(2x)\right) = \exp\left(c(\beta)t(x) + a(x) - b(\beta)\right)$$

mit $c(\beta) = -\beta^2, t(x) = x^2, a(x) = \ln(2x), b(\beta) = -\ln \beta^2$. Damit ist der natürliche Parameter $\vartheta = -\beta^2$. Mit dem natürlichen Parameter ergibt sich dann die Exponentialfamilie in kanonischer Form $f(x) = \exp\left(\vartheta x^2 + \ln(2x) - \ln(-\vartheta)\right)$.

Zu b)

Es gilt $b'(\vartheta) = \mathbb{E}\left(t(X)\right) = \mathbb{E}\left(X^2\right)$ also $b'(\vartheta) = -\frac{1}{\vartheta} = \frac{1}{\beta^2}$ und somit $\mathbb{E}\left(X^2\right) = \frac{1}{\beta^2}$.

Zu c)

Eine Maximum Likelihood ist bis auf eine Konstante, die nicht von β abhängt eindeutig bestimmt. Die gemeinsame Dichte der $X_1, ..., X_n$ ergibt sich zu

$$f\left(x_1,\dots,x_n\right)=2^n\beta^{2n}\left(\prod_{i=1}^{n}x_i\right)\cdot\exp\left(-\beta^2\sum_{i=1}^{n}x_i^{\ 2}\right)=2^n\left(\prod_{i=1}^{n}x_i\right)\cdot L\left(\beta\right).$$

Somit folgt die Behauptung, da $2^n\prod_{i=1}^{n}x_i$ eine Konstante ist, die nicht von β abhängt.

Zu d)

Es gilt:

$$\ell\left(\beta\right):=\ln L\left(\beta\right)=2n\ln\beta-\beta^2\sum_{i=1}^{n}x_i^{\ 2}$$

$$\ell'\left(\beta\right)=\frac{2n}{\beta}-2\beta\sum_{i=1}^{n}x_i^{\ 2}$$

$$\ell''\left(\beta\right)=-\frac{2n}{\beta^2}-2\sum_{i=1}^{n}x_i^{\ 2}$$

Auflösen der Gleichung $\ell'\left(\beta\right)=0$ ergibt wegen $\ell''<0$ den ML-Schätzer

$$\hat{\beta}=\sqrt{\frac{n}{\sum_{i=1}^{n}X_i^{\ 2}}}\,.$$

Zu e)

Die Informationsmatrix ergibt sich mit b) und d)

$$I\left(\beta\right)=\mathbb{E}\left(-\ell''\left(\beta\right)\right)=\frac{2n}{\beta^2}+2\cdot\frac{n}{\beta^2}=\frac{4n}{\beta^2}.$$

Zu f)

Einsetzen in d) führt zum Schätzwert $\hat{\beta}=\sqrt{\frac{16}{36}}=\frac{4}{6}=\frac{2}{3}$

Zu g)

Zu bestimmen ist die Testgröße $W=-2\left(\ell\left(\beta_0\right)-\ell\left(\hat{\beta}\right)\right)=-2\left(\ell\left(0{,}5\right)-\ell\left(\frac{2}{3}\right)\right)$. Bei der gegebenen Stichprobe ist $\ell\left(\beta\right)=32\ln\beta-36\beta^2$ und somit

$$W=-2\left(\ell\left(0{,}5\right)-\ell\left(\frac{2}{3}\right)\right)=-2\left(32\ln\left(\frac{0{,}5}{2/3}\right)-36\left(\frac{1}{4}-\frac{4}{9}\right)\right)=4{,}11$$

H_0 wird verworfen, wenn $W\geq\chi^2_{1;1-0{,}1}=2{,}71$ gilt. Da $W=4{,}11>2{,}71$, wird die Annahme $\beta=0{,}5$ verworfen.

Zu h)

Es wird die Funktion $f(\beta) = 2\left(\ell\left(\hat{\beta}\right) - \ell(\beta)\right) = 2\left(\ell\left(\frac{2}{3}\right) - \ell(\beta)\right) = 2\left(\ell\left(\frac{2}{3}\right) - 32\ln\beta + 36\beta^2\right) =$

$$= -64\ln\beta + 72\beta^2 - 57{,}95$$

betrachtet, das Schätzintervall ist dann durch $f^{-1}\left(\left(0;\chi^2_{1;0,95}\right]\right) = f^{-1}\left(\left(0;3{,}84\right]\right)$ bestimmt. Im gegebenen Funktionsgraphen ist f bis auf eine Verschiebung gegeben, also kann man das Schätzintervall $\left[0{,}51;0{,}85\right]$ ablesen.

Aufgabe 4 (30 Punkte): Die Zufallsvariable X bezeichne den Jahresgesamtschaden eines Risikos. Der Jahresgesamtschaden hänge dabei von einem Strukturparameter in Form einer Zufallsvariablen Θ ab. Bei bekanntem Θ ergibt sich damit als Netto-Jahresprämie

$$H(\Theta) = E[X \mid \Theta].$$

Der Strukturparameter Θ sei allerdings nicht direkt beobachtbar, so dass der Aktuar auf Approximation von $H(\Theta)$ angewiesen ist (so genannter Credibility-Ansatz).

Im folgenden wird davon ausgegangen, dass für das betrachtete Risiko n unabhängige Replikationen $X_1, \cdots, X_n$ von X vorliegen, mit arithmetischem Mittel $\overline{X}$.

Im Rahmen des „linearen" Credibility-Ansatzes wird H durch H_α^{**} mit

$$H_\alpha^{**} = \alpha\overline{X} + (1-\alpha)E(X)$$

approximiert, wobei

$$\alpha = \frac{Var\left(H(\Theta)\right)}{Var\left(\overline{X}\right)} = \frac{Var(\overline{X}) - n^{-1}\,E\left(Var[X \mid \Theta]\right)}{Var(\overline{X})}.$$

a) Gegeben seien Jahresschäden für 6 Risiken aus den Jahren 2006 bis 2009 (Werte in T€):

Risiko / Jahr	2006	2007	2008	2009	empirischer Mittelwert μ_i (pro Risiko i)	empirische Varianz s_i^2 (pro Risiko i)
1	1,16	1,76	1,85	1,62	1,60	0,09
2	1,72	0,50	1,05	0,87	1,04	0,26
3	0,49	1,59	0,48	1,51	1,02	0,38
4	0,20	1,53	0,35	1,04	0,78	0,38
5	1,62	0,10	0,18	0,82	0,68	0,50
6	1,87	0,97	0,54	0,30	0,92	0,48
				Mittelwert (pro Spalte)	1,01	0,35
				Varianz (pro Spalte)	0,10	0,02

Schätzen Sie in diesem Beispiel $E(X)$, $Var\left(\bar{X}\right)$ und $E\left(Var[X\mid\Theta]\right)$. Begründen Sie Ihre Vorgehensweise. Berechnen Sie α und die lineare Credibility-Prämie H_α^{**} für Risiko 1.

<u>Hinweis:</u> Dabei können Sie davon ausgehen, dass die Risiken 1 bis 6 sowie die zugehörigen Strukturparameter Θ_1 bis Θ_6 unabhängig und identisch verteilt sind. Die Jahresschäden jedes Risikos *i* seien zudem bedingt unabhängig bei gegebenem Θ_i.

<u>Kontrollergebnis:</u> $\alpha \approx 0{,}125$

b) Der Aktuar geht nun davon aus, dass der Strukturparameter Θ gleichverteilt auf [0, *a*] ist, mit einem $a > 0$. Des Weiteren sei *X* bei gegebenem $\Theta = \vartheta$ gleichverteilt auf $[0, \vartheta]$. Berechnen Sie die allgemeine Credibility-Prämie $H^* := E\left[H(\Theta)\mid X_1,\ldots,X_n\right]$.

<u>Kontrollergebnis:</u> Im Verlauf der Rechnung ermitteln Sie

$$f_{\Theta\mid X_1=x_1,\ldots,X_n=x_n}(\vartheta) = \frac{n-1}{m^{1-n}-a^{1-n}}\cdot\frac{1}{\vartheta^n}\mathbb{1}_{[m,a]}(\vartheta) \text{ mit } m := \max_{i=1,\ldots,n} x_i.$$

c) Unter den gemachten Annahmen gilt $\mathrm{E}(X) = \mathrm{E}\left(H(\Theta)\right) = \frac{a}{4}$. Welche Schätzung für *a* ergibt sich, wenn Sie dies mit der Schätzung für $E(X)$ aus a) vergleichen?

d) Berechnen Sie damit den Wert der allgemeinen Credibility-Prämie H^* für Risiko 1.

<u>Hinweis:</u> Gemäß b) gilt $H^* = \frac{1}{2}\cdot\frac{n-1}{n-2}\cdot\frac{m^{2-n}-a^{2-n}}{m^{1-n}-a^{1-n}}$.

Lösung

a) Aufgrund der Unabhängigkeitsannahmen der Credibility-Theorie kann die vorletzte Spalte als unabhängige, identisch verteilte Realisierungen μ_i des empirischen Mittels $\bar{X}$ angesehen werden. Damit kann man

$E(X) = E(\bar{X})$ durch das empirische Mittel 1,01 und $Var\left(\bar{X}\right)$ durch die empirische Varianz 0,1 schätzen. Die letzte Spalte enthält unabhängige, identisch verteilte Realisierungen s_i^2 der empirischen Varianz s^2 von $X_1, \ldots, X_n$. Wegen der Erwartungstreue der empirischen Varianz gilt $E[s^2\mid\Theta] = Var[X\mid\Theta]$. Folglich ist $E\left(Var[X\mid\Theta]\right) = E\left(E[s^2\mid\Theta]\right) = E\left(s^2\right)$.

Daher kann man $E\left(Var[X\mid\Theta]\right)$ durch das empirische Mittel 0,35 schätzen. Mit n = 4 ergibt sich

$$\alpha = \frac{Var(\bar{X}) - n^{-1}\,E\left(Var[X\mid\Theta]\right)}{Var(\bar{X})} \approx \frac{0{,}10 - 0{,}35/4}{0{,}10} = 0{,}125$$

und

$$H_\alpha^{**} = \alpha\bar{X} + (1-\alpha)E(X) \approx 0{,}125\cdot 1{,}60 + 0{,}875\cdot 1{,}01 = 1{,}08\,.$$

b) Unter der gegebenen Verteilungsannahmen gilt nach der Formel von Bayes

$$
\begin{aligned}
& f_{\Theta|X_1=x_1,\dots,X_n=x_n}(\vartheta) \\
&= c\cdot f_{X_1,\dots,X_n|\Theta}(x_1,\dots,x_n)\cdot f_\Theta(\vartheta) \\
&= c\cdot\prod_{i=1}^{n}\left\{\frac{1}{\vartheta}1_{[0,\vartheta]}(x_i)\right\}\cdot\frac{1}{a}1_{[0,a]}(\vartheta) \\
&= c\cdot\frac{1}{\vartheta^n}1_{[\max_{i=1,\dots,n}x_i,\infty)}(\vartheta)\cdot\frac{1}{a}1_{[0,a]}(\vartheta) \\
&= c\cdot\frac{1}{a\vartheta^n}1_{[\max_{i=1,\dots,n}x_i,a]}(\vartheta)
\end{aligned}
$$

mit einer geeigneten Normierungskonstanten c. Dabei beachte man, dass $x_i \le \Theta \le a$ für alle i.Die Normierungskonstante c folgt aus

$$1 \overset{!}{=} \int_0^a f_{\Theta|X_1=x_1,\dots,X_n=x_n}(\vartheta)\,d\vartheta = \int_m^a c\cdot\frac{1}{a\vartheta^n}\,d\vartheta = \frac{c}{a(n-1)}\cdot\left(m^{1-n}-a^{1-n}\right),$$

wobei $m := \max\limits_{i=1,\dots,n} x_i$. Insgesamt ergibt sich

$$f_{\Theta|X_1=x_1,\dots,X_n=x_n}(\vartheta) = \frac{n-1}{m^{1-n}-a^{1-n}}\cdot\frac{1}{\vartheta^n}1_{[m,a]}(\vartheta).$$

Gleichzeitig ist $\mathrm{H}(\Theta)=\mathrm{E}[\mathrm{X}\,|\,\Theta]=\frac{\Theta}{2}$, und die allgemeine Credibility-Prämie ist

$$
\begin{aligned}
H^* &:= E[H(\Theta)\,|\,X_1,\dots,X_n] = \int_m^a \frac{\vartheta}{2}\cdot f_{\Theta|X_1=x_1,\dots,X_n=x_n}(\vartheta)\,d\vartheta \\
&= \frac{n-1}{2\left(m^{1-n}-a^{1-n}\right)}\cdot\int_m^a \vartheta^{1-n}\,d\vartheta = \frac{n-1}{2\left(m^{1-n}-a^{1-n}\right)}\cdot\frac{m^{2-n}-a^{2-n}}{n-2} \\
&= \frac{1}{2}\cdot\frac{n-1}{n-2}\cdot\frac{m^{2-n}-a^{2-n}}{m^{1-n}-a^{1-n}}.
\end{aligned}
$$

c) Es gilt in der Tat (dies war nicht nachzurechnen)

$$E(X) = E\left(H(\Theta)\right) = \int_0^a \frac{\vartheta}{2}\cdot f_\Theta(\vartheta)\,d\vartheta = \frac{1}{2a}\int_0^a \vartheta\,d\vartheta = \frac{1}{4a}\cdot a^2 = \frac{a}{4}.$$

Gleichsetzen mit der Schätzung $E(X)\approx 1{,}01$ ergibt $a \approx 4{,}04$.

d) Für Risiko 1 ist m = 1,85, und mit n = 4 und $a\approx 4{,}04$ folgt

$$H^* = \frac{1}{2}\cdot\frac{4-1}{4-2}\cdot\frac{1{,}85^{-2}-4{,}04^{-2}}{1{,}85^{-3}-4{,}04^{-3}} = 1{,}21.$$