Zerlegung von Stichproben und ihre Anwendung in der Tarifierung

Klaus Th. Hess

Institut für Mathematik
Universität Rostock

Festkolloquium
20 Jahre (neue) Versicherungsmathematik
an der Technischen Universität Dresden

21. Oktober 2011

Gliederung

Gliederung

Multiplikativer Tarif – Beispiel (1)

Beispiel Kraftfahrthaftpflicht mit zwei Tarifmerkmalen:

Grundprämie: 275,34			Jahresfahrleistung (in km)				
			unter 9.000	9.000 bis unter 15.000	15.000 bis unter 20.000	20.000 bis unter 28.000	mind. 28.000
		Tarif-faktor	0,85	0,93	1,00	1,10	1,20
Garage / Wohn-eigentum	ja / ja	0,85	198,93	217,66	234,04	257,44	280,85
	ja / nein	0,90	210,64	230,46	247,81	272,59	297,37
	nein / ja	0,90	210,64	230,46	247,81	272,59	297,37
	nein / nein	1,00	234,04	256,07	275,34	302,87	330,41

Abbildung: Multiplikativer Tarif (aus Zocher (2006))

Multiplikativer Tarif – Beispiel (2)

Vorteile eines multiplikativen Tarifs:

- Weniger Tariffaktoren als Tarifzellen.
- Trennung von Prämienniveau (Grundprämie) und Struktur des Bestandes.

Gliederung

Kollektives Modell – Definition

Das Paar

$$\left\langle N, \{X_j\}_{j\in\mathbb{N}} \right\rangle$$

mit $N : \Omega \to \mathbb{N}_0$ und $X_j : \Omega \to \mathbb{R}$ heißt kollektives Modell, falls gilt:

- Die Folge $\{X_j\}_{j\in\mathbb{N}}$ ist i.i.d.
- Die Folge $\{X_j\}_{j\in\mathbb{N}}$ ist unabhängig von N.

Interpretation:

- N ist die Schadenzahl des Bestandes (in einem Geschäftsjahr).
- X_j ist die Schadenhöhe des j–ten Schadens des Bestandes.

Kollektives Modell - Gesamtschaden

Wir betrachten den Gesamtschaden des kollektiven Modells:

$$S := \sum_{j=1}^{N} X_j = \sum_{n=0}^{\infty} \chi_{\{N=n\}} \sum_{j=1}^{n} X_j$$

Proposition.

- Es gilt $\varphi_S = m_N \circ \varphi_{X_1}$.
- Im Fall $P[X_1 \in \mathbb{N}_0] = 1$ gilt $m_S = m_N \circ m_{X_1}$.

Proposition (Waldsche Gleichungen).

$$\begin{aligned} E[S] &= E[N]E[X_1] \\ var[S] &= E[N]var[X_1] + var[N]E[X_1]^2 \end{aligned}$$

Gliederung

Abstraktes Kollektives Modell – Definition

Das Paar

$$\left\langle N, \{Y_j\}_{j\in\mathbb{N}} \right\rangle$$

mit $N : \Omega \to \mathbb{N}_0$ und $Y_j : \Omega \to \mathbb{R}^p$ heißt abstraktes kollektives Modell, falls gilt:

- Die Folge $\{Y_j\}_{j\in\mathbb{N}}$ ist i.i.d.
- Die Folge $\{Y_j\}_{j\in\mathbb{N}}$ ist unabhängig von N.

Interpretation:

- N ist die Schadenzahl des Bestandes.
- Y_j ist die Schadenvariable des j–ten Schadens des Bestandes.

Abstraktes Kollektives Modell – Transformation

Sei $g : \mathbb{R}^p \to \mathbb{R}$ eine meßbare Funktion, dann ist

$$\left\langle N, \{g \circ Y_j\}_{j \in \mathbb{N}} \right\rangle$$

ein kollektives Modell.

Abstraktes Kollektives Modell – Verdünnung (1)

Sei $C_k \in \mathcal{B}(\mathbb{R}^p)$ derart, dass für die Auswahlwahrscheinlichkeit $\eta_k := P[Y_1 \in C_k] > 0$ gilt. Die verdünnte Schadenzahl N_k ist durch

$$N_k := \sum_{j=1}^{N} \chi_{\{Y_j \in C_k\}}$$

und die verdünnten Schadenvariablen sind durch

$$Y_{k,j} := \sum_{h=1}^{\infty} \chi_{\{\nu_{k,j}=h\}} Y_h$$

definiert, wobei der zufällige Index $\nu_{k,j}$ rekursiv durch

$$\begin{aligned} \nu_{k,0} &:= 0 \\ \nu_{k,j} &:= \inf\{h \in \mathbb{N} | \nu_{k,j-1} < h, Y_h \in C_k\} \end{aligned}$$

gegeben ist.

Abstraktes Kollektives Modell – Verdünnung (2)

Satz. Das verdünnte Modell

$$\left\langle N_k, \{Y_{k,j}\}_{j \in \mathbb{N}} \right\rangle$$

ist wiederum ein abstraktes kollektives Modell. Die Verteilung der verdünnten Schadenzahl ist durch ihre wahrscheinlichkeits-erzeugende Funktion

$$m_{N_k}(z) \;=\; m_N(1 - \eta_k + \eta_k z)$$

bestimmt und für die Verteilung der Schadenvariablen gilt

$$P\left[Y_{k,j} \leq y\right] \;=\; P\left[\,Y_1 \leq y \,|\, Y_1 \in C_k\right]$$

für alle $y \in \mathbb{R}$.

Insbesondere gilt $E[N_k] \;=\; \eta_k E[N]$.

Abstraktes Kollektives Modell – Verdünnung (3)

Beispiel. Sei N Poisson–verteilt mit Parameter λ. Dann gilt

$$m_N(z) = e^{-\lambda(1-z)}$$

und wir erhalten

$$\begin{aligned} m_{N_k}(z) &= m_N(1-\eta_k+\eta_k z) \\ &= e^{-\lambda(1-(1-\eta_k+\eta_k z))} \\ &= e^{-\lambda\eta_k(1-z)} \end{aligned}$$

Daher ist N_k Poisson–verteilt mit Parameter $\lambda\eta_k$.

Abstraktes Kollektives Modell – Zerlegung (1)

Sei $\{C_1, \ldots, C_m\}$ eine endliche Zerlegung des $\mathbb{R}^p$ mit $\eta_k := P[Y_1 \in C_k] > 0$. Verdünnen wir das abstrakte kollektive Modell

$$\left\langle N, \{Y_j\}_{j \in \mathbb{N}} \right\rangle$$

bezüglich jeder Zerlegungsmenge C_k, so erhalten wir die verdünnten abstrakten kollektiven Modelle

$$\left\langle N_1, \{Y_{1,j}\}_{j \in \mathbb{N}} \right\rangle, \quad \ldots\ , \quad \left\langle N_m, \{Y_{m,j}\}_{j \in \mathbb{N}} \right\rangle$$

Abstraktes Kollektives Modell – Zerlegung (2)

Satz.

- Alle verdünnten Schadenvariablen $Y_{k,j}$ sind unabhängig.
- Die verdünnten Schadenvariablen sind unabhängig von den verdünnten Schadenzahlen.
- Für die verdünnten Schadenzahlen gilt

$$\mathrm{P}_{N_1,\dots,N_m|N} = \mathbf{Mult}(N; \eta_1, \dots, \eta_m)$$

- Die verdünnten Schadenzahlen sind genau dann unabhängig, wenn die ursprüngliche Schadenzahl Poisson–verteilt ist.

Gliederung

Multiplikativer Tarif (1)

Wir betrachten das abstrakte kollektive Modell

$$\left\langle N, \left\{(X_j, U_j, T_j)\right\}_{j\in\mathbb{N}} \right\rangle$$

Interpretation:

- $X_j : \Omega \to \mathbb{R}$ ist die Schadenhöhe
- $U_j : \Omega \to \{1, \dots, I\}$ ist die Tarifklasse des 1. Tarifmerkmals
- $T_j : \Omega \to \{1, \dots, K\}$ ist die Tarifklasse des 2. Tarifmerkmals
- (U_j, T_j) ist die Tarifzelle

des j–ten Schadens.

Multiplikativer Tarif (2)

Weitere Annahmen:

- X_j und (U_j, T_j) sind unabhängig.
- Es gibt positive Parameter $\alpha_i, \beta_k, \pi_{i,k}$ mit

$$P\big[(U_j, T_j) = (i, k)\big] \;=\; \alpha_i \, \beta_k \, \pi_{i,k}$$

  für alle $i \in \{1, \ldots, I\}$ und $k \in \{1, \ldots, K\}$.

Dabei sind $\pi_{i,k}$ bekannte Volumenmaße und α_i, β_k unbekannte Parameter.

Es gilt

$$\sum_{i=1}^{I} \sum_{k=1}^{K} \alpha_i \, \beta_k \, \pi_{i,k} \;=\; 1$$

Multiplikativer Tarif (3)

Wir setzen $C_{i,k} := \mathbb{R} \times \{i\} \times \{k\}$ und zerlegen das abstrakte kollektive Modell

$$\left\langle N, \left\{(X_j, U_j, T_j)\right\}_{j\in\mathbb{N}} \right\rangle$$

nach den Tarifzellen. Dann gilt

$$\eta_{i,k} \;=\; \alpha_i\,\beta_k\,\pi_{i,k}$$

und wir erhalten die abstrakten kollektiven Modelle

$$\left\langle N_{i,k;j}, \left\{(X_{i,k;j}, U_{i,k;j}, T_{i,k;j})\right\}_{j\in\mathbb{N}} \right\rangle$$

Auf Grund der Konstruktion gilt

$$U_{i,k;j} = i \qquad \text{und} \qquad T_{i,k;j} = k$$

Multiplikativer Tarif (4)

Daher können wir uns auf die kollektiven Modelle

$$\left\langle N_{i,k}, \{X_{i,k;j}\}_{j\in\mathbb{N}} \right\rangle$$

beschränken.

Für die Verteilung der Schadenhöhen gilt:

$$\begin{aligned} P\left[X_{i,k;j} \leq x\right] &= P\left[X_{i,k;j} \leq x, U_{i,k;j} = i, T_{i,k;j} = k\right] \\ &= P\left[X_j \leq x, U_j = i, T_j = k \middle| U_j = i, T_j = k\right] \\ &= P\left[X_j \leq x\right] \end{aligned}$$

Damit sind alle verdünnten Schadenhöhen i.i.d. und ihre Verteilung kann mit den Methoden der klassischen Statistik geschätzt werden.

Multiplikativer Tarif (5)

Für die verdünnten Schadenzahlen gilt

$$E[N_{i,k}] = \alpha_i \, \beta_k \, \pi_{i,k} \, E[N]$$

Durch Summation erhalten wir

$$\begin{aligned} \sum_{i=1}^{I} E[N_{i,k}] &= \sum_{i=1}^{I} \alpha_i \, \beta_k \, \pi_{i,k} \, E[N] \qquad && k \in \{1, \ldots, K\} \\ \sum_{k=1}^{K} E[N_{i,k}] &= \sum_{k=1}^{K} \alpha_i \, \beta_k \, \pi_{i,k} \, E[N] \qquad && i \in \{1, \ldots, I\} \end{aligned}$$

Multiplikativer Tarif (6)

Daraus ergeben sich die Marginalsummengleichungen

$$
\begin{aligned}
\sum_{i=1}^{I} \widehat{\alpha}_i\, \widehat{\beta}_k\, \pi_{i,k} &= \sum_{i=1}^{I} \frac{N_{i,k}}{N} && k \in \{1, \ldots, K\} \\
\sum_{k=1}^{K} \widehat{\alpha}_i\, \widehat{\beta}_k\, \pi_{i,k} &= \sum_{k=1}^{K} \frac{N_{i,k}}{N} && i \in \{1, \ldots, I\} \\
\sum_{i=1}^{I} \sum_{k=1}^{K} \widehat{\alpha}_i\, \widehat{\beta}_k\, \pi_{i,k} &= 1
\end{aligned}
$$

Positive Lösungen der Marginalsummengleichungen heißen Marginalsummenschätzer.

Multiplikativer Tarif (7)

Die Marginalsummenschätzung auf der Basis der Schadenzahlen lässt sich durch das Maximum–Likelihood Prinzip begründen.

Da die verdünnten Schadenzahlen bedingt multinomialverteilt sind, erhalten wir mit $n := \sum_{i,k} n_{i,k}$

$$\begin{aligned} P\left[\bigcap_{i,k}\{N_{i,k} = n_{i,k}\}\right] &= P\left[\bigcap_{i,k}\{N_{i,k} = n_{i,k}\}\,\middle|\, N = n\right] P[N = n] \\ &= \frac{n!}{\prod_{i,k} n_{i,k}!} \prod_{i,k} \left(\alpha_i \beta_k \pi_{i,k}\right)^{n_{i,k}} \cdot\, p(n) \end{aligned}$$

Dabei sei $p : \mathbb{N}_0 \to [0,1]$ gegeben durch $p(n) := P[N = n]$.

Multiplikativer Tarif (8)

Likelihood–Funktion:

$$L\Big(\widehat{\alpha}_1,\ldots,\widehat{\alpha}_I,\widehat{\beta}_1,\ldots,\widehat{\beta}_K \;\Big|\; \{N_{i,k}\}_{i,k}\Big) = \frac{N!}{\prod_{i,k} N_{i,k}!}\prod_{i,k}(\widehat{\alpha}_i\widehat{\beta}_k\pi_{i,k})^{N_{i,k}}\cdot p(N)$$

Log–Likelihood–Funktion:

$$(\log\circ L)\Big(\widehat{\alpha}_1,\ldots,\widehat{\alpha}_I,\widehat{\beta}_1,\ldots,\widehat{\beta}_K \;\Big|\; \{N_{i,k}\}_{i,k}\Big) = \sum_{i,k} N_{i,k}\log(\widehat{\alpha}_i\widehat{\beta}_k\pi_{i,k}) + C$$

Multiplikativer Tarif (9)

Lagrange–Ansatz:

$$
\begin{aligned}
& h\Big(\widehat{\alpha}_1, \ldots, \widehat{\alpha}_I, \widehat{\beta}_1, \ldots, \widehat{\beta}_K, \lambda\Big) \\
&= \sum_{i,k} N_{i,k} \log(\widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k}) + \lambda \left(1 - \sum_{i,k} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k}\right)
\end{aligned}
$$

Partielle Differentiation und Nullsetzen ergibt

$$
\begin{aligned}
\sum\nolimits_{k=1}^{K} \frac{N_{i,k}}{\widehat{\alpha}_i} - \lambda \sum\nolimits_{k=1}^{K} \widehat{\beta}_k \pi_{i,k} &= 0 \qquad i \in \{1, \ldots, I\} \\
\sum\nolimits_{i=1}^{I} \frac{N_{i,k}}{\widehat{\beta}_k} - \lambda \sum\nolimits_{i=1}^{I} \widehat{\alpha}_i \pi_{i,k} &= 0 \qquad k \in \{1, \ldots, K\} \\
1 - \sum\nolimits_{i,k} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k} &= 0
\end{aligned}
$$

Multiplikativer Tarif (10)

und damit

$$\begin{aligned}
\sum\nolimits_{k=1}^{K} N_{i,k} &= \lambda \sum\nolimits_{k=1}^{K} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k} \qquad i \in \{1, \dots, I\} \\
\sum\nolimits_{i=1}^{I} N_{i,k} &= \lambda \sum\nolimits_{i=1}^{I} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k} \qquad k \in \{1, \dots, K\} \\
1 &= \sum\nolimits_{i,k} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k}
\end{aligned}$$

Wegen

$$N = \sum\nolimits_{i,k} N_{i,k} = \lambda \sum\nolimits_{i,k} \widehat{\alpha}_i \widehat{\beta}_k \pi_{i,k} = \lambda$$

ergeben sich daraus die Marginalsummengleichungen

$$\begin{aligned}
\sum\nolimits_{k=1}^{K} \widehat{\alpha}_i \, \widehat{\beta}_k \, \pi_{i,k} &= \sum\nolimits_{k=1}^{K} \frac{N_{i,k}}{N} \qquad i \in \{1, \dots, I\} \\
\sum\nolimits_{i=1}^{I} \widehat{\alpha}_i \, \widehat{\beta}_k \, \pi_{i,k} &= \sum\nolimits_{i=1}^{I} \frac{N_{i,k}}{N} \qquad k \in \{1, \dots, K\}
\end{aligned}$$

Gliederung

Literatur (1)

- R. A. Bailey (1963): Insurance Rates with Minimum Bias. Proceedings CAS Vol L, 4–11.
- S. Dietze, T. Riedrich, K.D. Schmidt (2006): On the solution of marginal–sum equations. DSVM 1/2006.
- J. Jung (1968): On Automobile Insurance Ratemaking. ASTIN Bull. 5, 41–48.
- T. Mack (1991): A Simple Parametric Model for Rating Automobile Insurance of Estimating and IBNER Claims Reserves. ASTIN Bull. 21, 93–109.
- J. van Eeghen, E. K. Greup, J. A. Nijssen (1983): Rate Making. Surveys of actuarial studies no. 2, Nationale–Nederlanden, Rotterdam.
- M. Zocher (2006): Risikoadäquate Tarifierung in der Kraftfahrt-haftpflichtversicherung. Wiss. Z. TU Dresden 55, 131–135.

DSVM = Dresdner Schriften zur Versicherungsmathematik.

Literatur (2)

- T. Franke, W. Macht (1995): Decomposition of risk processes. DSVM 2/1995.
- K.T. Hess (2000): Random partitions of samples. DSVM 1/2000.
- K.T. Hess (2008): Marginal–sum and maximum–likelihood estimation in a multiplicative tariff. AStA Advances in Statistical Analysis 93, 221–233.
- K.T. Hess, W. Macht, K.D. Schmidt (1995): Thinning of risk processes. DSVM 1/1995.
- K.T. Hess, K.D. Schmidt (2002): A comparison of models for the chain–ladder method. Insurance Math. Econom. 31, 351–364.
- K.D. Schmidt (1996): Lectures on Risk Theory. Stuttgart: Teubner.

DSVM = Dresdner Schriften zur Versicherungsmathematik.