

Zeitschrift: Mitteilungen / Vereinigung Schweizerischer Versicherungsmathematiker
= Bulletin / Association des Actuaires Suisses = Bulletin / Association of
Swiss Actuaries

Band: - (1981)

Heft: 1

Artikel: Ein Beitrag zur Intervallschätzung der Glaubwürdigkeitsparameter

Autor: Zagorac, Drasko

DOI: <https://doi.org/10.5169/seals-550882>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 13.10.2024

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

DRASKO ZAGORAC, Belgrad

Ein Beitrag zur Intervallschätzung der Glaubwürdigkeitsparameter¹

Vorbemerkungen

In den meisten Credibilitymodellen wird eine lineare Minimum-Quadrat-Schätzung des a posteriori Erwartungswerts gesucht. Ein anderes Vorgehen besteht in der sogenannten Minimax-Schätzung, bei der die Prämienberechnung nach Methoden der Spieltheorie erfolgt. Auf Anregung von Prof. Bühlmann habe ich in meiner Diplomarbeit nach Intervallschätzungen für die in der klassischen Credibilityformel auftretenden Parameter gesucht. Die entsprechenden Resultate sind im folgenden beschrieben anhand des einfachsten Modells mit konstanten Volumina sowie normal verteiltem Schadensatz und Strukturparameter.

1 Modell

Es werden N Risiken im Laufe von n Jahren betrachtet, und für das Risiko j im Jahr i folgende Bezeichnungen eingeführt:

- S_{ij} – Schadentotal
- P_{ij} – Prämienvolumen
- X_{ij} – Schadensatz.

Die Schadensätze können wie folgt als Matrix dargestellt werden:

| | | | |
|-----|----------|--------------|--------------|
| | 1 | ... j | ... N |
| 1 | X_{11} | ... X_{1j} | ... X_{1N} |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| i | X_{i1} | ... X_{ij} | ... X_{iN} |
| ⋮ | ⋮ | ⋮ | ⋮ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| n | X_{n1} | ... X_{nj} | ... X_{nN} |

¹ Auszug aus meiner Diplomarbeit 1979 unter der Leitung von Prof. H. Bühlmann

2 Voraussetzungen

- i) Das Prämienvolumen ist konstant, d.h. wir betrachten natürliche Risikoeinheiten, also, $P_{ij}=1$ ($i=1, \dots, n, j=1, \dots, N$).
- ii) Die Schadensätze X_{ij} sind unabhängige und normal verteilte Zufallsvariablen mit Parametern (m_{ij}, σ_{ij}) .
- iii) Die Fluktuationen sind homogen in der Zeit sowie im Bestand, d.h. $\sigma_{ij}=\sigma$.
- iv) Der Risikoparameter ist homogen in der Zeit, d.h. $m_{ij}=m_j$, was bedeutet, dass die Schadensätze für das betrachtete Risiko j in der Zeit gleich verteilt sind.
- v) Die m_j sind unabhängige Zufallsgrößen, normal und gleichverteilt mit Parametern (m_0, σ_0) , also mit der Strukturdichte

$$u(m_j) = \frac{1}{\sigma_0 \sqrt{2\pi}} e^{-\frac{(m_j - m_0)^2}{2\sigma_0^2}}$$

Gemäss den Voraussetzungen lässt sich für die bedingte Variable X_{ij}/m_j sowie für m_j schreiben:

$$X_{ij}/m_j \sim \mathcal{N}(m_j, \sigma)$$

$$m_j \sim \mathcal{N}(m_0, \sigma_0)$$

woraus hervorgeht, dass

$$X_{ij} \sim \mathcal{N}(m_0, \sqrt{\sigma_0^2 + \sigma^2}).$$

3 Parameter und ihre Schätzungen

Die Glaubwürdigkeitsparameter sind:

$$m = EE(X_{ij}/m_j) = E(m_j) = m_0$$

$$v = E \text{Var} (X_{ij}/m_j) = E(\sigma^2) = \sigma^2$$

$$w = \text{Var} E(X_{ij}/m_j) = \text{Var} (m_j) = \sigma_0^2$$

Wir führen folgende statistische Masszahlen ein:

– Den beobachteten Mittelwert des Bestandes:

$$M = \frac{1}{nN} \sum_{i=1}^n \sum_{j=1}^N X_{ij}$$

Wegen

$$E(M) = \frac{1}{nN} E\left(\sum_{i,j} X_{ij}\right) = \frac{1}{nN} \sum_{i,j} E(X_{ij}) = m_0,$$

ist M eine erwartungstreue Schätzung des Parameter m_0 .

- Die beobachtete mittlere quadratische Abweichung vom Mittelwert $X_{.j} = \sum_i X_{ij}/n$ pro Vertrag:

$$V = \frac{1}{N} \sum_{j=1}^N \frac{1}{n-1} \sum_{i=1}^n (X_{ij} - X_{.j})^2$$

Wegen

$$(n-1)NE(V) = \sum_{i,j} E(X_{ij}^2) - n \sum_j E(X_{.j}^2) = (n-1)N\sigma^2$$

ist V eine erwartungstreue Schätzung des Parameters σ^2 .

- Die beobachtete mittlere quadratische Abweichung vom Mittelwert des Bestands, d. h. die beobachtete Portefeuillevarianz:

$$W = \frac{1}{nN-1} \sum_{i=1}^n \sum_{j=1}^N (X_{ij} - M)^2.$$

Wegen

$$E(W) = \sigma^2 + \frac{n(N-1)}{nN-1} \sigma_0^2 = \sigma^2 + \pi \sigma_0^2$$

ist W eine erwartungstreue Schätzung des Parameters $\sigma^2 + \pi \sigma_0^2$ mit $\pi = \frac{n(N-1)}{nN-1}$

4 Das Fisher-Lemma

Seien x_1, \dots, x_n unabhängige Zufallsgrößen, von denen jede normalverteilt ist mit Parametern $(0, \sigma^2)$. Für jede orthogonale Transformation

$$y_i = \sum_{j=1}^n c_{ij} x_j \quad (i=1, \dots, n),$$

ist die quadratische Form

$$Q(x_1, \dots, x_n) = \sum_{i=1}^n x_i^2 - y_1^2 - \dots - y_p^2 \quad (p < n)$$

als Zufallsvariable unabhängig von den y_1, \dots, y_p verteilt mit Dichte

$$\frac{1}{\sigma^2} k_{n-p}\left(\frac{x}{\sigma^2}\right) = \frac{x^{\frac{n-p}{2}-1} e^{-\frac{x}{2\sigma^2}}}{2^{\frac{n-p}{2}} \cdot \sigma^{n-p} \Gamma\left(\frac{n-p}{2}\right)},$$

wobei $k_m(x)$ die Dichte der χ^2 -Verteilung mit m Freiheitsgraden ist.

[Beweis: Siehe H. Cramer: "MATHEMATICAL METHODS OF STATISTICS" 1974, p. 379]

Sei $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ das Stichprobenmittel der normalen Stichprobe (x_1, \dots, x_n)

und $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$ die Stichprobenvarianz, dann gilt:

\bar{x} und s^2 sind unabhängig, und $\frac{ns^2}{\sigma^2}$ ist χ^2 -verteilt mit $(n-1)$ Freiheitsgraden, wobei σ^2 die Varianz der Grundgesamtheit ist.

Bemerkung: \bar{x} ist normal verteilt mit Parametern $(m, \sigma/\sqrt{n})$, wo m den Erwartungswert der Grundgesamtheit darstellt.

5 Die Bonfferoni-Ungleichung

Seien $A_i (i \in I)$ Ereignisse aus der σ -Algebra \mathcal{A} , und A_i^c ihre Komplemente, dann gilt

$$\text{Prob} \left(\bigcap_i A_i \right) \geq 1 - \sum_i \text{Prob} (A_i^c).$$

Beweis:

Für $B_i \in \mathcal{A}$

$$\text{Prob} \left(\bigcup_i B_i \right) = \text{Prob} \left(\bigcap_i B_i^c \right)^c = 1 - \text{Prob} \left(\bigcap_i B_i^c \right) \leq \sum_i \text{Prob} (B_i).$$

Also gilt

$$1 - \sum_i \text{Prob} (B_i) \leq \text{Prob} \left(\bigcap_i B_i^c \right).$$

Wenn wir noch $A_i^c = B_i$ setzen, erhalten wir die gesuchte Ungleichung.

6 Das Problem

Es sei eine Intervallschätzung der Glaubwürdigkeitsparameter m, v, w zu finden bei vorgegebener Fehlerwahrscheinlichkeit ε .

Im allgemeinen Fall ist das Problem nicht gelöst, im angegebenen Modell und unter obigen Voraussetzungen ist jedoch eine Lösung wie folgt möglich.

7 Die Lösung

7.1. Vertrauensintervall für den Parameter $m = m_0$

Die Masszahl M gehorcht einer Normalverteilung mit Parametern

$$\left(m_0, \sqrt{\frac{\sigma_0^2 + \sigma^2}{nN}}\right),$$

so dass die Variable

$$(M - m_0) / \sqrt{\frac{\sigma^2 + \sigma_0^2}{nN}}$$

normalverteilt ist mit Parametern $(0, 1)$.

Die Masszahl W kann in folgender Form geschrieben werden:

$$(nN - 1)W = \sum_i \sum_j (X_{ij} - m_0)^2 - nN(M - m_0)^2,$$

woraus nach Division mit $\sigma^2 + \sigma_0^2$ folgt:

$$\frac{(nN - 1)W}{\sigma^2 + \sigma_0^2} = \sum_i \sum_j \frac{(X_{ij} - m_0)^2}{\sigma^2 + \sigma_0^2} - nN \frac{(M - m_0)^2}{\sigma^2 + \sigma_0^2}.$$

Die Variablen $(X_{ij} - m_0) / \sqrt{\sigma^2 + \sigma_0^2}$ sind normalverteilt mit Parametern $(0, 1)$, und

$$y = \sqrt{nN} \frac{M - m_0}{\sqrt{\sigma^2 + \sigma_0^2}} = \frac{1}{\sqrt{nN}} \sum_i \sum_j \frac{X_{ij} - m_0}{\sqrt{\sigma^2 + \sigma_0^2}}$$

ist eine orthogonale Transformation dieser Variablen.

Laut dem Fisher-Lemma für $p=1$ gilt, dass die Zufallsgrösse $(nN - 1)W / (\sigma^2 + \sigma_0^2)$ χ^2 -verteilt ist mit $(nN - 1)$ Freiheitsgraden.

Führen wir noch die Variable T ein:

$$T = \frac{M - m_0}{\sqrt{\sigma^2 + \sigma_0^2}} \sqrt{nN} / \sqrt{\frac{1}{nN-1} \frac{(nN-1)W}{\sigma^2 + \sigma_0^2}} = \frac{M - m_0}{\sqrt{W}} \sqrt{nN}$$

T hat eine Student-Verteilung mit $(nN - 1)$ Freiheitsgraden. Sei ein ε gegeben, $0 < \varepsilon < 1$. Wir bestimmen t , so dass

$$\text{Prob} (|T| \leq t) = 1 - \varepsilon,$$

d. h.

$$\begin{aligned} \text{Prob} (-t \leq T \leq t) &= \text{Prob} \left(-t \leq \frac{M - m_0}{\sqrt{W}} \sqrt{nN} \leq t \right) \\ &= \text{Prob} \left(t \sqrt{\frac{W}{nN}} + M \geq m_0 \geq -t \sqrt{\frac{W}{nN}} + M \right) = 1 - \varepsilon. \end{aligned}$$

Auf diese Weise kann man das Vertrauensintervall des Parameters m_0 bestimmen:

$$I_1 = \left[M - t \sqrt{\frac{W}{nN}}, \quad M + t \sqrt{\frac{W}{nN}} \right]$$

mit

$$\text{Prob} (m_0 \in I_1) = 1 - \varepsilon.$$

7.2 Vertrauensintervall für den Parameter $w = \sigma_0^2$

Es wurde gezeigt, dass die Masszahl $(nN - 1)W / (\sigma^2 + \sigma_0^2)$ eine χ^2 -Verteilung mit $(nN - 1)$ Freiheitsgraden hat. Für das gegebene ε suchen wir k , so dass

$$\text{Prob} \left(\frac{(nN - 1)W}{\sigma^2 + \sigma_0^2} \geq k \right) = 1 - \varepsilon.$$

Also ist

$$\text{Prob} \left(0 < \sigma^2 + \sigma_0^2 \leq \frac{(nN - 1)W}{k} \right) = 1 - \varepsilon.$$

Damit wird das Vertrauensintervall für den Parameter $\sigma^2 + \sigma_0^2$ bestimmt:

$$I_2 = \left(0, \frac{(nN - 1)W}{k} \right].$$

Aus $\text{Prob}(\sigma^2 + \sigma_0^2 \in I_2) = 1 - \varepsilon$ folgt

$$\text{Prob}(\sigma^2 \in I_2) \geq 1 - \varepsilon,$$

und

$$\text{Prob}(\sigma_0^2 \in I_2) \geq 1 - \varepsilon.$$

Für die Intervallschätzung des Parameters σ_0^2 können wir nichts weiter sagen, ausser dass

$$\text{Prob}(\sigma_0^2 \in I_2) \geq 1 - \varepsilon$$

ist.

7.3. Vertrauensintervall für den Parameter $v = \sigma^2$

Wir hatten $\text{Prob}(\sigma^2 \in I_2) \geq 1 - \varepsilon$. Ausserdem hat die Variable $N(n-1)V/\sigma^2$ eine χ^2 -Verteilung mit $N(n-1)$ Freiheitsgraden. Sei nämlich

$$\sigma^2 z_j = \sum_{i=1}^n (X_{ij} - X_{.j})^2 = \sum_{i=1}^n (X_{ij} - m_j + m_j - X_{.j})^2 = \sum_{i=1}^n (X_{ij} - m_j)^2 - n(X_{.j} - m_j)^2.$$

Die Variablen $(X_{ij} - m_j)/\sigma$ (für feste j) gehorchen einer Normalverteilung mit Parametern $(0, 1)$ und $y = \sqrt{n}(X_{.j} - m_j)/\sigma$ ist ihre orthogonale Transformation. Laut dem Fisher-Lemma haben die Z_j eine χ^2 -Verteilung mit $(n-1)$ Freiheitsgraden.

Sei $Z = \sum_{j=1}^N Z_j = \sum_{j=1}^N \sum_{i=1}^n \left(\frac{X_{ij} - X_{.j}}{\sigma} \right)^2$. Z ist χ^2 -verteilt mit $N(n-1) = \Sigma(n-1)$ Freiheitsgraden, weil die Z_j unabhängige χ^2 -Variablen mit $(n-1)$ Freiheitsgraden sind.

Für das gegebene ε werden q_1 und q_2 gesucht, so dass

$$\text{Prob} \left(\frac{N(n-1)V}{\sigma^2} \geq q_1 \right) = 1 - \varepsilon/2,$$

und

$$\text{Prob} \left(\frac{N(n-1)V}{\sigma^2} \leq q_2 \right) = 1 - \varepsilon/2,$$

also

$$\text{Prob} \left(\frac{N(n-1)V}{q_1} \leq \sigma^2 \leq \frac{N(n-1)V}{q_2} \right) = 1 - \varepsilon.$$

Als Vertrauensintervall für σ^2 ergibt sich daher:

$$I_3 = \left[\frac{N(n-1)V}{q_1}, \frac{N(n-1)V}{q_2} \right].$$

Somit gelten die Relationen:

$$\text{Prob}(\sigma^2 \in I_3) = 1 - \varepsilon \Rightarrow \text{Prob}(\sigma^2 \notin I_3) = \varepsilon,$$

und

$$\text{Prob}(\sigma^2 \in I_2) \geq 1 - \varepsilon \Rightarrow \text{Prob}(\sigma^2 \notin I_2) \leq \varepsilon.$$

Aufgrund der Bonfferoni-Ungleichung ist:

$$\begin{aligned} \text{Prob}((\sigma^2 \in I_2) \cap (\sigma^2 \in I_3)) &\geq 1 - \text{Prob}(\sigma^2 \notin I_3) - \text{Prob}(\sigma^2 \notin I_2) \\ &= 1 - \varepsilon - \text{Prob}(\sigma^2 \notin I_2) \geq 1 - 2\varepsilon, \end{aligned}$$

also

$$\text{Prob}(\sigma^2 \in I_2 \cap I_3) \geq 1 - 2\varepsilon.$$

7.4 Vertrauensquader für $(m, v, m) = (m_0, \sigma^2, \sigma_0^2)$

Wir hatten:

$$\text{Prob}(m_0 \in I_1) = 1 - \varepsilon \quad \Rightarrow \quad \text{Prob}(m_0 \notin I_1) = \varepsilon,$$

$$\text{Prob}(\sigma^2 \in I_2 \cap I_3) \geq 1 - 2\varepsilon \quad \Rightarrow \quad \text{Prob}(\sigma^2 \notin I_2 \cap I_3) \leq 2\varepsilon,$$

$$\text{Prob}(\sigma_0^2 \in I_2) \geq 1 - \varepsilon \quad \Rightarrow \quad \text{Prob}(\sigma_0^2 \notin I_2) \leq \varepsilon.$$

Durch wiederholte Anwendung der Bonfferoni Ungleichung erhalten wir:

$$\begin{aligned} \text{Prob}(m_0 \in I_1, \sigma^2 \in I_2 \cap I_3, \sigma_0^2 \in I_2) &\geq 1 - \text{Prob}(m_0 \notin I_1) - \text{Prob}(\sigma^2 \notin I_2 \cap I_3) \\ &\quad - \text{Prob}(\sigma_0^2 \notin I_2) = 1 - \varepsilon - \text{Prob}(\sigma^2 \notin I_2 \cap I_3) - \text{Prob}(\sigma_0^2 \notin I_2) \\ &\geq 1 - \varepsilon - 2\varepsilon - \varepsilon = 1 - 4\varepsilon. \end{aligned}$$

Wenn wir die Bezeichnungen einführen:

$$[m_1, m_2] = I_1$$

$$[v_1, v_2] = I_2 \cap I_3$$

$$[w_1, w_2] = I_2$$

$$I = [m_1, m_2] \times [v_1, v_2] \times [w_1, w_2],$$

dann ist I ein Vertrauensquader für den Parameterpunkt $(m_0, \sigma^2, \sigma_0^2)$.
Die Wahrscheinlichkeit, dass dieser Punkt zur Menge I gehört, die der Stichprobe (X_{ij}) entspricht, ist grösser oder gleich $1-4\varepsilon$.

Drasko Zagorac
Vlajkovicva 30/3, stan 11
Y-11000 Beograd
Yugoslavia

Literatur

- H. Bühlmann: Minimax Credibility, Scandinavian Actuarial Journal, 1975.
H. Cramer: Mathematical Methods of Statistics, Princeton University Press, 1974.
A. Marazzi: Minimax Credibility, ETH Zürich, Diss. Nr. 5695/76.
D. Zagorac: Eine Übersicht über die Modelle der Glaubwürdigkeitstheorien, Diplomarbeit 1979, Belgrad

Zusammenfassung

Um die Methode der Intervallschätzung möglichst ohne Komplikationen illustrieren zu können, wird ein einfaches Credibility-Modell mit konstanten Volumina und normal verteilten Schadensätzen gewählt. Basierend auf einem Lemma von Fisher und der sogenannten Bonfferoni-Ungleichung werden dann Vertrauensintervalle für die drei Parameter m (= Mittelwert über das Kollektiv), v (= Variation in der Zeit) und w (= Variation im Kollektiv = Heterogenitätsmass) hergeleitet.

Résumé

En vue d'illustrer, si possible sans complications, la méthode de l'estimation des intervalles, l'auteur propose un modèle simple basé sur un volume constant et des taux de sinistres distribués selon une loi normale. Se fondant sur un lemme de Fisher et sur l'inéquation dite de Bonfferoni, il délimite des intervalles de confiance pour les 3 paramètres m (= moyenne sur le collectif), v (= variation dans le temps) et w (= variation dans le collectif, mesure d'hétérogénéité).

Summary

In order to illustrate the method of interval estimation of credibility parameters with as few complications as possible a simple model has been chosen with constant volumes and normally distributed loss ratios. Based on a lemma by Fisher and the so-called Bonfferoni inequality confidence intervals are then deduced for the parameters m (= collective mean), v (= variation in time) and w (= variation in the collective = measure of heterogeneity).

