

Parameterschätzung im hierarchischen Credibility-Modell nach B. Sundt

Autor(en): **Mangold, Klaus-Peter**

Objektyp: **Article**

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

Band (Jahr): - **(1987)**

Heft 1

PDF erstellt am: **27.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-967147>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

KLAUS-PETER MANGOLD, München

Parameterschätzung im hierarchischen Credibility-Modell nach B. Sundt¹

Im zweistufigen hierarchischen Modell geht es darum, die Grössen M , F , G , H , die im dreistufigen Modell von *Bühlmann/Jewell* [1] noch vom Parameter ψ_l abhängen und $M(\psi_l)$, $F(\psi_l)$, $G(\psi_l)$, $H(\psi_l)$ heissen, zu schätzen.

Für die Darstellung hier erscheint mir eine etwas ausführlichere Bezeichnungsweise nützlich.

Es liegen s Kohorten vor. Zu jeder Kohorte k gehören N_k Risiken und für jedes Risiko j der Kohorte k wurden n_{jk} Beobachtungen X_{ijk} gemacht, die jeweils mit einer Gewichtung p_{ijk} versehen sind. Summationen über einen bestimmten Index werden durch einen an entsprechender Stelle stehenden Punkt gekennzeichnet, also z.B.

$$p_{\cdot jk} := \sum_{i=1}^{n_{jk}} p_{ijk} \quad (= : V_{jk}^{(1)})$$

Es bezeichnet

$$Y_{jk} := \frac{1}{p_{\cdot jk}} \sum_{i=1}^{n_{jk}} p_{ijk} X_{ijk} \quad \text{das Risikomittel}$$

$$Y_k := \frac{1}{p_{\cdot \cdot k}} \sum_{j=1}^{N_k} p_{\cdot jk} Y_{jk} \quad \text{das Kohortenmittel}$$

$z_{jk}^{(1)}$: = den Credibility-Faktor von Risiko j in Kohorte k ,

$z_k^{(2)}$: = den Credibility-Faktor der Kohorte k ,

$$B(\varphi_k) := \frac{1}{z_{\cdot k}^{(1)}} \sum_{j=1}^{N_k} z_{jk}^{(1)} Y_{jk} \quad \text{das Credibility-gewichtete Kohortenmittel}$$

¹ B. Sundt, Some Credibility Regression Models for the Classification of Individual Passenger Car Models, Kap. 4.3, S. 39 ff., ASTIN-Kolloquium, Tel Aviv, 1986.

und

$$B(\varphi) := \frac{1}{z_{\cdot\cdot}^{(1)}} \sum_{k=1}^s z_{\cdot k}^{(1)} B(\varphi_k).$$

Man sollte nun zwei Fälle unterscheiden:

Neben den üblichen Voraussetzungen der bedingten Unabhängigkeit und (in oberen Niveaus) auch Gleichverteiltheit gelte

$$\text{I) } \text{Var}(X_{ijk} | \vartheta_{jk}) = \frac{\sigma^2(\vartheta_{jk})}{p_{ijk}}$$

oder

$$\text{II) } \text{Var}(X_{ijk} | \vartheta_{jk}) = \sigma^2(\vartheta_{jk})$$

Für $p_{ijk} = 1$ decken sich beide Fälle. Dann ist $n_{jk} = p_{\cdot jk}$. Sundt setzt Fall I voraus und gibt folgende Schätzer an:

$$F^* = \frac{1}{s} \sum_{k=1}^s \frac{1}{N_k} \sum_{j=1}^{N_k} \left(\frac{1}{n_{jk} - 1} \sum_{i=1}^{n_{jk}} p_{ijk} (X_{ijk} - Y_{jk})^2 \right)$$

$$G^* = \frac{1}{p_{\dots}} \sum_{k=1}^s \left(\frac{1}{1 - \sum_{j=1}^{N_k} p_{\cdot jk}^2 / p_{\cdot\cdot k}^2} \left(\sum_{j=1}^{N_k} p_{\cdot jk} (Y_{jk} - Y_k)^2 - (N_k - 1) F^* \right) \right)$$

$$H^* = \frac{1}{1 - \sum_{k=1}^s (z_{\cdot k}^{(1)})^2 / (z_{\cdot\cdot}^{(1)})^2} \left(\sum_{k=1}^s \frac{z_{\cdot k}^{(1)}}{z_{\cdot\cdot}^{(1)}} (B(\varphi_k) - B(\varphi))^2 - (s-1) \frac{G^*}{z_{\cdot\cdot}^{(1)}} \right)$$

$$M^* = \frac{1}{z_{\cdot\cdot}^{(2)}} \sum_{k=1}^s z_k^{(2)} B(\varphi_k)$$

Auf den Fall mit $n_{ijk} := p_{ijk} = 1$ aus [1] umgeschrieben:

$$\begin{aligned}
 F^* &= \frac{1}{s} \cdot \sum_{k=1}^s \frac{1}{N_k} \sum_{j=1}^{N_k} \left(\frac{1}{1 - n_{jk}/n_{jk}^2} \left(\sum_{i=1}^{n_{jk}} \frac{1}{n_{jk}} (X_{ijk} - Y_{jk})^2 \right) \right) \\
 &= \frac{1}{s} \cdot \sum_{k=1}^s \frac{1}{N_k} \sum_{j=1}^{N_k} \left(\frac{1}{1 - \sum_{i=1}^{n_{jk}} n_{ijk}^2 / n_{jk}^2} \left(\sum_{i=1}^{n_{jk}} \frac{n_{ijk}}{n_{jk}} (X_{ijk} - Y_{jk})^2 \right) \right) \\
 G^* &= \sum_{k=1}^s \frac{n_{\cdot k}}{n_{\cdot\cdot}} \left(\frac{1}{1 - \sum_{j=1}^{N_k} n_{jk}^2 / n_{\cdot k}^2} \right) \left(\sum_{j=1}^{N_k} \frac{n_{jk}}{n_{\cdot k}} (Y_{jk} - Y_k)^2 - (N_k - 1) \frac{F^*}{n_{\cdot k}} \right) \\
 H^* &= \frac{1}{1 - \sum_{k=1}^s (z_{\cdot k}^{(1)})^2 / (z_{\cdot\cdot}^{(1)})^2} \left(\sum_{k=1}^s \frac{z_{\cdot k}^{(1)}}{z_{\cdot\cdot}^{(1)}} (B(\varphi_k) - B(\varphi))^2 - (s - 1) \frac{G^*}{z_{\cdot\cdot}^{(1)}} \right)
 \end{aligned}$$

und M^* wie oben.

Dass bei *Sundt* der Schätzer F^* im Aufbau etwas "aus der Reihe tanzt", liegt an der speziellen Voraussetzung I). Liegt Fall II) vor, so kann man leicht einen gleichartigen Aufbau der einzelnen Schätzer erkennen. Sofern ich die von *Sundt* nicht explizit motivierte Herleitung richtig nachvollzogen habe, beruhen die Schätzer auf folgendem "Bildungsgesetz":

Man geht jeweils von den beobachteten Varianzen auf den einzelnen Niveaus aus. Also

$$\text{für } F: \quad \sum_{i=1}^{n_{jk}} \frac{p_{ijk}}{p_{\cdot jk}} (X_{ijk} - Y_{jk})^2 \quad \text{für } \begin{array}{l} k = 1, \dots, s \\ j = 1, \dots, N_k \end{array}$$

$$\text{für } G: \quad \sum_{j=1}^{N_k} \frac{p_{\cdot jk}}{p_{\cdot\cdot k}} (Y_{jk} - Y_k)^2 \quad \text{für } k = 1, \dots, s$$

$$\text{für } H: \quad \sum_{k=1}^s \frac{z_{\cdot k}^{(1)}}{z_{\cdot\cdot}^{(1)}} (B(\varphi_k) - B(\varphi))^2$$

Von diesen Grössen errechnet man den Erwartungswert und formt dann so um, dass man erwartungstreue Statistiken erhält. Die so erhaltenen Werte werden über alle darüber liegenden Niveaus gemittelt.

Allerdings ist mir nicht einleuchtend, wieso *Sundt* einmal (bei F^*) gewöhnlich mittelt, während er das andere Mal (bei G^*) gewichtet gemittelt hat. (Also einmal mit Gewicht $1/N_k$, das andere Mal mit $p_{..k}/p_{...}$.) Wie üblich begegnet man der Möglichkeit, dass $G^* < 0$ und/oder $H^* < 0$, indem man $\max(G^*, 0)$ bzw. $\max(H^*, 0)$ ansetzt.

Entsprechend dieser Vorgehensweise ergäbe sich demnach für das dreistufige Modell aus [1] mit c Gesellschaften l (der Index l bezeichnet die jeweilige Gesellschaft):

Oben aufgeführte Schätzer F^*_l , G^*_l , H^*_l werden über ein weiteres Niveau (gewichtet oder ungewichtet?) gemittelt. Zur Schätzung von I setzt man an:

$$T := \sum_{l=1}^c \frac{z_{.l}^{(2)}}{z_{..}^{(2)}} (B(\psi_l) - B(\psi))^2$$

$$B(\psi_l) := \frac{1}{z_{.l}^{(2)}} \sum_{k=1}^{s_l} z_{kl}^{(2)} B(\varphi_{kl})$$

$$B(\psi) := \frac{1}{z_{..}^{(2)}} \sum_{l=1}^c z_{.l}^{(2)} B(\psi_l)$$

Daraus leitet sich ab:

$$\begin{aligned} E(T) &= \frac{1}{z_{..}^{(2)}} \left(\sum_{l=1}^c z_{.l}^{(2)} \text{Var}(B(\psi_l)) - z_{..}^{(2)} \text{Var}(B(\psi)) \right) \\ &= \frac{1}{z_{..}^{(2)}} \cdot \left(\sum_{l=1}^c z_{.l}^{(2)} \left(\frac{H}{z_{.l}^{(2)}} + I \right) - z_{..}^{(2)} \left(\frac{H}{z_{..}^{(2)}} + \frac{\sum_{l=1}^c (z_{.l}^{(2)})^2}{(z_{..}^{(2)})^2} I \right) \right) \\ &= \frac{1}{z_{..}^{(2)}} (c-1) \cdot H + \left(1 - \frac{\sum_{l=1}^c (z_{.l}^{(2)})^2}{(z_{..}^{(2)})^2} \right) \cdot I \end{aligned}$$

Man erhält also

$$I^* = \frac{1}{1 - \sum_{l=1}^c (z_{\cdot l}^{(2)})^2 / (z_{\cdot\cdot}^{(2)})^2} \left(\sum_{l=1}^c \frac{z_{\cdot l}^{(2)}}{z_{\cdot\cdot}^{(2)}} (B(\psi_l) - B(\psi))^2 - (c-1) \frac{H^*}{z_{\cdot\cdot}^{(2)}} \right)$$

und wohl

$$M^* = \frac{1}{z_{\cdot\cdot}^{(3)}} \cdot \sum_{l=1}^c z_l^{(3)} B(\psi_l)$$

Der Berechnungsablauf bei einem dreistufigen hierarchischen Modell sieht demnach folgendermassen aus:

1. Errechne F^*
2. Errechne G^*
3. Errechne $z_{jkl}^{(1)}$ für $l = 1, \dots, c; k = 1, \dots, s_l; j = 1, \dots, N_{kl}$
4. Errechne H^*
5. Errechne $z_{kl}^{(2)}$ für $l = 1, \dots, c; k = 1, \dots, s_l$
6. Errechne I^*
7. Errechne $z_l^{(3)}$ für $l = 1, \dots, c$
8. Errechne M^*
9. Errechne $\widehat{M}(\psi_l)$ für $l = 1, \dots, c$
10. Errechne $\widehat{M}(\varphi_{kl})$ für $l = 1, \dots, c; k = 1, \dots, s_l$
11. Errechne $\widehat{\mu}(\vartheta_{jkl})$ für $l = 1, \dots, c; k = 1, \dots, s_l; j = 1, \dots, N_{kl}$

Klaus-Peter Mangold
 Münchener Rückversicherungs-Gesellschaft
 Königinstrasse 107
 Postfach 40 13 20
 D-8000 München 40

Literatur

- [1] Bühlmann, H./Jewell, W.S., Hierarchical Credibility Revisited, MVSVM, Heft 1, 1987.