Über das Testen von Sterblichkeitshypothesen über anomale Risiken

Autor(en): **Bühlmann, Hans**

Objektyp: **Article**

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

Band (Jahr): **61 (1961)**

PDF erstellt am: **16.05.2024**

Persistenter Link: <https://doi.org/10.5169/seals-966737>

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.

Über das Testen von Sterblichkeitshypothesen über anomale Risiken

Von Hans Bühlmann, Zürich

Herrn Prof. Jecklin zum 60. Geburtstag gewidmet

Zusammenfassung

Das Problem, ganze Sterblichkeitsverläufe [d. h. Sterblichkeitsfunktionen $q(x)$] gegeneinander zu testen, wird in dieser Arbeit theoretisch untersucht, und es werden verschiedene Tests angegeben. Es handelt sich dabei um Entscheidungskriterien, die in einem Funktionenraum (Raum der Sterblichkeitsfunktionen) operieren. Folgende Entscheidungsprobleme werden gelöst:

- Testen einer vorgewählten Funktion gegenüber ähnlichen Alternativ-Funktionen,
- Testen von verschiedenen «unähnlichen» Verläufen der Sterblichkeitsfunktion.

Alle angeführten Entscheidungsverfahren sind im Sinne von Neyman-Pearson optimal (most powerful).

1. Einleitung

Bei praktischen Untersuchungen stellt sich immer wieder die Frage, ob eine bestimmte Arbeitshypothese über die Anomalensterblichkeit zulässig sei oder nicht. Eine Möglichkeit, diese Frage zu beantworten, besteht darin, dass Beobachtungen aus der Anomalengesamtheit zu Sterblichkeitstabellen verarbeitet werden, d. h. dass für jede Anomalienklasse und jedes Versichertenalter «erfahrungsmässige» Sterblichkeitssätze bestimmt und diese mit den gemäss Arbeitshypothese entsprechenden «theoretischen» Sterblichkeitswahrscheinlichkeiten verglichen werden. Wie einwandfrei dieses Prozedere auch ist und wie aufschlussreich es auch sein mag, so krankt es doch an der praktischen Schwierigkeit, dass durch die Aufsplitterung des Beobachtungsmaterials in eine Unzahl von Gruppen diese meist nur kärglich dotiert werden.

Wenn es gerade diese Schwierigkeit ist, welche in der Praxis der Anomalienversicherung ohne gewisse Hypothesen über die Anomalensterblichkeit kaum lösbar ist, dann stellt sich notgedrungen die problematische Frage, ob und wann solche Annahmen zulässig seien.

Zur Illustration seien die beiden gebräuchlichsten Hypothesen über die Anomalensterblichkeit angeführt:

– die multiplikative Übersterblichkeitshypothese:

$$\text{Postulat } q'_x = (1 + \alpha) q_x \text{ für alle } x$$

– die additive Übersterblichkeitshypothese:

$$\text{Postulat } q'_x = q_x + \Delta \text{ für alle } x$$

welche wir in den folgenden Kapiteln als Beispiele benutzen werden.

Allgemein können die heute üblicherweise verwendeten Arbeitshypothesen wie folgt charakterisiert werden:

a) Sie postulieren einen funktionellen Zusammenhang zwischen normaler und anomaler Sterblichkeit,

b) Sie ermöglichen durch diesen funktionellen Zusammenhang den Vergleich von Erfahrungszahlen aus verschiedenen Anomalienklassen.

In praxi verzweigt sich deshalb die Frage nach der angepassten Anomalensterblichkeit in zwei Teilfragen:

I. Welcher funktionelle Zusammenhang zwischen anomaler und normaler Sterblichkeit ist am geeignetsten?

II. Wie werden richtigerweise die freien Parameter dieses postulierten Zusammenhanges festgelegt?

Dieses Frageschema zeigt deutlich, dass die Gegenstände unserer Entscheidung *Funktionen* sind im Gegensatz zu dem früher erwähnten «Sterblichkeitstafel-Verfahren», wo einzelne *Zahlwerte* (nämlich die einer bestimmten Klasse zugeordneten Sterbewahrscheinlichkeiten) untersucht werden. In knappster Form ausgedrückt ergibt die Antwort auf I. die *Form* der Sterblichkeitskurve, während diejenige auf II. das *Niveau* der Anomalensterblichkeit festlegt.

Zweck dieser Arbeit ist es, mathematische Kriterien (Tests) anzugeben, die im Lichte dieser beiden Teilfragen eine objektive Entscheidung für oder gegen eine gegebene Arbeitshypothese ermöglichen.

2. Vorgehen

a) Aus einer Gruppe medizinisch gleichartig und gleichgradig anomaler Risiken stehe folgendes Beobachtungsmaterial zur Verfügung:

L_x = Anzahl Lebende vom Alter x ,

T_x = Anzahl Tote vom Alter x .

Die L_x gelten als feste Parameter, die T_x als Zufallsvariablen. Im folgenden wird allerdings das Symbol T_x sowohl für eine bestimmte Realisation der Zufallsvariablen T_x als zur Bezeichnung dieser Zufallsvariablen selbst verwendet. Diese Doppelbedeutung des Symbols ist absichtlich nicht vermieden worden, da sie zu keinen Missverständnissen Anlass gibt, sondern im Gegenteil die Zusammenhänge besser erkennen lässt.

b) Dieses Beobachtungsmaterial wird dann durch die hier vorgeschlagenen Tests auf die Frage hin untersucht

I'. Ist ein bestimmter postulierter Zusammenhang zwischen normaler und anomaler Sterblichkeit zulässig?

II'. Ist bei gegebenem funktionellem Zusammenhang eine bestimmte Hypothese über das Niveau der Anomalensterblichkeit annehmbar?

c) Zur Abklärung der Güte dieser Tests wird von der Voraussetzung ausgegangen, dass die T_x poissonverteilt mit unbekanntem aber festem Parameter μ_x sind (um den Zusammenhang mit der üblichen aktuariellen Bezeichnungsweise zum Ausdruck zu bringen, wird statt μ_x das Symbol $q'_x L_x$ gebraucht). Selbstverständlich kann diese Annahme, dass die T_x poissonverteilt seien, der Realität darum nicht voll gerecht werden, da die Poissonverteilung, welche beliebig grosse Beobachtungswerte – wenn auch mit verschwindend kleinen Wahrscheinlichkeiten – zulässt, der naturgegebenen Beziehung $T_x \leq L_x$ nicht gerecht wird. Demgegenüber sei festgehalten, dass die Praxis der Poissonverteilung nicht nur die Einfachheit einer praktischen Arbeitsformel zugute hält, sondern in ihr auch deshalb eine sehr zweckmässige Annahme erblickt, weil durch den Prozess der Gewichtung des Parameters aus der Klasse der Poissonverteilungen eine so grosse Gesamtheit von Verteilungsfunktionen abgeleitet werden kann, dass eine Anpassung an die praktischen Bedürfnisse mehr als gewährleistet ist. Im Sinne einer weiteren Abklärung würde der Verfasser es deshalb sehr begrüßen, wenn die vorliegende Arbeit, deren Untersuchungen auf der Hypothese der einfachen Poissonverteilung aufbauen, im Hinblick auf gewichtete Poissonverteilungen ausgedehnt werden könnte.

d) Der Vollständigkeit halber seien noch folgende hier verwendete Bezeichnungen erörtert

q_x = Einjährige Sterbewahrscheinlichkeit des *normalen* x -jährigen,
 q'_x = Einjährige Sterbewahrscheinlichkeit des *anormalen* x -jährigen.

Die Bezeichnung $q(x)$ resp. $q'(x)$ wird dann verwendet, wenn die funktionelle Betrachtungsweise im Vordergrund steht. Falls verschiedene Sterblichkeitssätze für den normalen oder anomalen x -jährigen auseinandergehalten werden müssen, wird ein oberer Index verwendet, z. B.

$$q_x^{(0)}, q_x^{(1)}, q_x^{(2)} \quad \text{oder} \quad q_x'^{(0)}, q_x'^{(1)}, q_x'^{(2)}.$$

3. Klassifizierung der Sterblichkeitsfunktionen

Wie schon erwähnt, folgt diese Arbeit nicht den üblichen Sterblichkeitsuntersuchungsmethoden, wo es darum geht, Sterbewahrscheinlichkeiten zu bestimmen. Die Objekte, welche hier der Untersuchung unterliegen, sind ganze Sterblichkeitsfunktionen (Sterblichkeitsverläufe). Die Tests, welche wir entwickeln, operieren deshalb in einem Funktionenraum $\mathfrak{Q}$.

$$\mathfrak{Q} = \text{Gesamtheit aller Sterblichkeitsfunktionen } q(x)$$

In diesem Funktionenraum definieren wir folgende Relationen:

- A. $q^{(1)}(x) \gg q^{(2)}(x)$ falls $q_x^{(1)} \geq q_x^{(2)}$ für alle x (Ordnungsrelation),
 B. $q^{(1)}(x) \approx q^{(2)}(x)$ falls $q_x^{(1)} = \lambda q_x^{(2)}$ für alle x (Ähnlichkeitsrelation).

Zwischen anomaler und normaler Sterblichkeit gilt dann sinngemäss Relation A.

$$q'(x) \gg q(x).$$

Die Relation B. erlaubt uns, Klassen ähnlicher Sterblichkeitsverläufe zu bilden (im folgenden kurz «Ähnlichkeitsklassen» genannt), womit sich die in den vorhergehenden Kapiteln erwähnte Problemstellung wie folgt neu formulieren lässt.

Ist bei gegebenem Beobachtungsmaterial

- I''. eine bestimmte Ähnlichkeitsklasse von Sterblichkeitsverläufen mit genügender Signifikanz zulässig?
 II''. eine bestimmte Funktion innerhalb derselben Ähnlichkeitsklasse auf Grund der Beobachtungen akzeptierbar?

4. Das Testen von Sterblichkeitsverläufen innerhalb der gleichen Ähnlichkeitsklasse

Wir beschäftigen uns hier mit dem Problem, das durch die Frage II'' aufgeworfen wird. Frage I'', deren Beantwortung etwas komplizierter ist, wird im nächsten Kapitel untersucht.

Es handelt sich also hier darum, ein Verfahren (Test) anzugeben, welches uns die Beurteilung ermöglicht, ob ein vorgewählter Sterblichkeitsverlauf gegenüber allen andern Sterblichkeitsfunktionen derselben Ähnlichkeitsklasse als richtig gewählt betrachtet werden kann. Nach dem Neyman-Pearson-Schema ¹⁾ verfahren wir wie folgt:

$$\begin{aligned} \text{Hypothese: } & q^{(0)}(x), \\ \text{Alternative: } & q'(x) = \lambda q^{(0)}(x). \end{aligned}$$

Auf Grund der Annahme, dass die T_x poissonverteilt seien, ergibt sich folgendes Likelihood-Verhältnis:

$$\frac{\prod_x \left[e^{-\lambda q^{(0)}(x) L_x} \frac{(\lambda q^{(0)}(x) L_x)^{T_x}}{T_x!} \right]}{\prod_x \left[e^{-q^{(0)}(x) L_x} \frac{(q^{(0)}(x) L_x)^{T_x}}{T_x!} \right]} = \prod_x [e^{-(\lambda-1)q^{(0)}(x)L_x} \lambda^{T_x}].$$

Der aussagekräftigste Test ergibt sich, falls die Verwerfungsgegend denjenigen Teil des Stichprobenraumes ausfüllt, wo das Likelihood-Verhältnis möglichst gross ist.

- a) Für den rechtsseitigen Test ($\lambda > 1$) lautet dann die Regel:
 «Verwirf die Hypothese falls $\sum T_x > K_1$ ». K_1 ist dabei so festzulegen, dass $P(\sum T_x > K_1) \leq \delta$ mit $\delta =$ Fehler erster Art,
- b) Für den linksseitigen Test ($\lambda < 1$) gilt analog:
 «Verwirf die Hypothese falls $\sum T_x < K_2$ mit $P(\sum T_x < K_2) \leq \delta$ ».

Die Verteilungsfunktion von $\sum T_x$ ergibt sich leicht als Poissonverteilung mit dem Parameter $\sum q^{(0)} L_x$. Hervorzuheben bleibt, dass – selbstverständlich immer unter der Annahme poissonverteilter T_x – die hier angegebenen Tests die stärkste Aussagekraft haben (uniformly most powerful tests), d.h. sofern wir nur innerhalb der gleichen Ähnlichkeitsklasse zu wählen haben, liefert die Funktion $\sum T_x$ das beste Kriterium für Entscheidungen gegenüber einseitigen Alternativen.

¹⁾ Die in dieser Arbeit benützte Technik zur Bestimmung des optimalen Tests geht auf die Neyman und Pearson'sche Testtheorie zurück. Für den Leser, der mit dieser Theorie nicht vertraut ist, sei auf folgende Literatur verwiesen: J. Neyman: First Course in Probability and Statistics, Henry Holt and Company, New York, 1950. Der Leser, der eine kurze und präzise Zusammenfassung der Neyman-Pearson-Theorie sucht, sei auf das Buch «Mathematische Statistik» von B. L. van der Waerden verwiesen (§ 59 Allgemeine Prinzipien. Möglichst mächtige Test).

Der in diesem Kapitel betrachtete Fall liegt in der Praxis immer dann vor, falls a priori die multiplikative Übersterblichkeitsmethode diskussionslos akzeptiert wird. Wie man leicht einsieht, gilt dann

$$q^{(0)}(x) = (1 + \alpha_0) q(x)$$

und

$$q'(x) = (1 + \alpha) q(x),$$

somit

$$\frac{q'(x)}{q^{(0)}(x)} = \frac{1 + \alpha}{1 + \alpha_0} = \lambda \text{ (unabhängig von } x\text{)}.$$

5. Das Testen gegenüber andern Ähnlichkeitsklassen

Die Frage nach dem aussagekräftigsten Entscheidungskriterium ist auf das engste mit den zugelassenen Alternativen verbunden. In diesem Kapitel beschäftigt uns der Fall, dass die Alternativ-Sterblichkeitsfunktionen nicht notwendigerweise der gleichen Ähnlichkeitsklasse wie die Nullfunktion ($q^{(0)}(x)$) angehören.

Das Testschema sieht dann folgendermassen aus:

Hypothese: $q^{(0)}(x)$,

Alternative: $q'(x)$,

Likelihood-Verhältnis:

$$\frac{\prod_x \left[e^{-q'(x) L_x} \frac{(q'(x) L_x)^{T_x}}{T_x!} \right]}{\prod_x \left[e^{-q^{(0)}(x) L_x} \frac{(q^{(0)}(x) L_x)^{T_x}}{T_x!} \right]} = \prod_x e^{-[q'(x) - q^{(0)}(x)] L_x} \left(\frac{q'(x)}{q^{(0)}(x)} \right)^{T_x}.$$

Die Testregeln lauten demnach:

$$\left(\text{Abkürzung in der Schreibweise: } \frac{q'(x)}{q^{(0)}(x)} = \lambda_x \right).$$

a) für den rechtsseitigen Test ($q'(x) \geq q^{(0)}(x)$):

$$\text{«Verwirf die Hypothese falls } \sum T_x \log \lambda_x > K_3\text{»},$$

b) für den linksseitigen Test ($q'(x) \leq q^{(0)}(x)$):

$$\text{«Verwirf die Hypothese falls } \sum T_x |\log \lambda_x| < K_4\text{»},$$

c) für den beidseitigen Test:

$$\text{«Verwirf die Hypothese falls } \sum T_x \log \lambda_x > K_5\text{»}.$$

Die konstanten K_3 , K_4 und K_5 sind wiederum so festzusetzen, dass der Fehler erster Art kleiner als ein vorgegebenes δ wird. Diese Testverfahren sind für eine bestimmte einfache Alternative (simple alternative) optimal, können aber keinen Anspruch auf maximale Aussagekraft (maximum power) für alle rechtsseitigen und/oder linksseitigen Alternativen erheben. Nun sind wir aber in diesem Kapitel eigentlich nur an der Form der Sterblichkeitskurve interessiert, d. h. wir möchten wissen, ob unabhängig vom Sterblichkeitsniveau die Form einer Sterblichkeitsfunktion den Beobachtungen angepasst ist oder nicht. Zur Eliminierung der Abweichungen, die von Niveauschwankungen herrühren, geht man nach einem von Neyman für zusammengesetzte Hypothesen (composite hypotheses) angegebenen Verfahren so vor, dass man statt der üblichen Wahrscheinlichkeiten bedingte verwendet. Sinngemäß lautet hier die Bedingung: «bei gegebenem Niveau» oder nachdem wir im letzten Kapitel $\sum T_x$ als für das Niveau erschöpfend aufgezeigt haben: «bei gegebenem $\sum T_x$ ».

Dieser Übergang zu den bedingten Wahrscheinlichkeiten ändert an dem zu Beginn dieses Kapitels angegebenen Testschema nichts. Das Likelihood-Verhältnis muss zwar jetzt konsequenterweise als Quotient der bedingten Wahrscheinlichkeiten definiert werden, führt aber zu den gleichen Testregeln. Der wesentliche Unterschied zu den vorhergehenden Überlegungen dieses Kapitels liegt aber darin, dass die Verteilungsfunktion der Teststatistik $\sum T_x \log \lambda_x$ resp. $\sum T_x |\log \lambda_x|$ als bezüglich $\sum T_x$ bedingt zu gelten hat.

Folgende Berechnung zeigt, wie diese Verteilungsfunktion bestimmt werden kann:

$$\begin{aligned}
 P[Y_1 = T_1, Y_2 = T_2, \dots, Y_\omega = T_\omega / \sum Y_x = \sum T_x] &= \\
 &= \frac{\prod_x [e^{-q_x^{(0)} L_x} (q_x^{(0)} L_x)^{T_x} / T_x!]}{e^{-\sum_x q_x^{(0)} L_x} (\sum_x q_x^{(0)} L_x)^{\sum T_x} / (\sum_x T_x)!} = \frac{\prod_x (q_x^{(0)} L_x)^{T_x} / T_x!}{(\sum_x q_x^{(0)} L_x)^{\sum T_x} / (\sum_x T_x)!} .
 \end{aligned}$$

Es stellt sich für die bedingte Wahrscheinlichkeitsverteilung der T_x eine solche vom hypergeometrischen Typus heraus. Aus der hier angegebenen Verteilungsfunktion diejenige von $\sum T_x \log \lambda_x$ (resp. $\sum T_x |\log \lambda_x|$) zu berechnen ist zwar theoretisch möglich, der Praktiker wird aber auch mit Hilfe von Elektronenrechnern den enormen Rechenaufwand nicht bewältigen können, ohne dass er gewisse Approximationen in

Kauf nimmt. Es wird in dieser Arbeit aber bewusst darauf verzichtet, eine solche Approximation vorzuschlagen: soll doch hier in erster Linie das theoretische Fundament und nicht das technische Fachwerk der Entscheidungsverfahren beschrieben werden.

Zur Illustration der in diesem Kapitel beschriebenen Tests sei angegeben, wie aus zwei verschiedenen Sterblichkeitsklassen die geeignetere ausgewählt wird. Wir testen zu diesem Zweck die Hypothese der multiplikativen Übersterblichkeit gegenüber der Alternative der konstanten additiven Sterblichkeitserhöhung. Nach den vorhergehenden Überlegungen ergibt sich folgende Teststatistik:

$$\sum T_x \log \left(\frac{\Delta}{q_x} \right) \quad \text{mit der Regel}$$

«Verwirf die multiplikative Sterblichkeitshypothese falls

$$\sum T_x \log \left(\frac{\Delta}{q_x} \right) > K(\sum T_x, \delta)''\text{»}.$$

Das auf Grund dieser Regel definierte Verwerfungsgebiet (rejection region) bleibt invariant, falls irgendein anderer Verschiebungsparameter Δ' für Δ eingesetzt wird. Es gilt nämlich

$$\begin{aligned} \sum T_x \log \left(\frac{\Delta'}{q_x} \right) &= \sum T_x \log \left(\frac{\Delta}{q_x} \right) + \log \left(\frac{\Delta'}{\Delta} \right) \sum T_x, & \text{somit} \\ \sum T_x \log \left(\frac{\Delta}{q_x} \right) > K(\sum T_x, \delta) &\Leftrightarrow \sum T_x \log \left(\frac{\Delta'}{q_x} \right) > K(\sum T_x, \delta) + \log \left(\frac{\Delta'}{\Delta} \right) \sum T_x \\ &= K'(\sum T_x, \delta). \end{aligned}$$

Analog gilt für eine verschiedene Wahl der repräsentativen Funktion aus der Klasse der zu $q(x)$ ähnlichen Sterblichkeitsfunktionen

$$\begin{aligned} \sum T_x \log \left(\frac{\Delta}{q_x} \right) > K(\sum T_x, \delta) &\Leftrightarrow \sum T_x \log \left(\frac{\Delta}{\lambda q_x} \right) > K(\sum T_x, \delta) - \log \lambda \sum T_x \\ &= K''(\sum T_x, \delta). \end{aligned}$$

Durch diese beiden Äquivalenzen ist nachgewiesen, dass unser Testverfahren das gleichmässig aussagekräftigste Entscheidungskriterium (uniformly most powerful) für die Wahl zwischen additiver und multiplikativer Übersterblichkeitshypothese darstellt. Der Beweis, dass diese Optimaleigenschaft auch für zwei beliebige Ähnlichkeitsklassen gilt, verläuft analog und ist deshalb hier nicht explizit angegeben.

Résumé

Le présent travail a pour objet l'étude théorique des tests applicables aux courbes de mortalité (c'est-à-dire aux fonctions de mortalité $q(x)$;) différents tests y sont présentés. Il s'agit en fait de critères de décision valables dans un espace de fonctions (espace des fonctions de mortalité). Les problèmes résolus sont les suivants:

- Test d'une fonction donnée par rapport aux fonctions alternatives semblables,
- Test de différentes courbes non semblables de la fonction de mortalité.

Tous les procédés de décision présentés sont du type «most powerful», au sens de Neyman-Pearson.

Summary

The problem of testing the whole curve represented by the mortality rates against any other such curve is theoretically examined in this paper. The criteria described operate in a function space [i. e. in the space of all mortality functions $q(x)$]. The following problems are solved:

- Test of a specific function against all similar mortality functions,
- Test of differently shaped mortality functions against each other.

All tests described are most powerful in the sense of Neyman-Pearson.

Riassunto

L'autore si propone con questa pubblicazione di esaminare il problema teorico di tests applicabili a curve di mortalità [vale a dire a funzioni di mortalità $q(x)$] e di presentarne alcuni. Si tratta infatti in questo caso di criteri di decisione che operano in uno spazio di funzioni [cioè nello spazio di tutte le funzioni di mortalità $q(x)$]; egli risolve i problemi seguenti:

- Test di una funzione data rispetto a tutte le funzioni di mortalità aventi andamento simile,
- Test di funzioni di mortalità aventi andamento dissimile.

I criteri di decisione descritti sono del tipo «most powerful» nel senso di Neyman-Pearson.