

Contents

1	Motivation und Aufgabenstellung	2
1.1	Motivation	2
1.2	Ziele - Aufgabenstellung	3
2	Daten	6
2.1	Die Daten der Sterbetafel 2004R	6
2.2	Öffentlich zugängliche Sterblichkeitsdaten	6
3	Schätzung von Basistafeln	8
3.1	Die Basistafel der Sterbetafel 2004R	8
3.2	Verschiedene Modelle für Basistafeln	8
4	Schätzen von Sterblichkeitstrends	9
4.1	Ergebnisse der Schätzung der Sterbetafel 2004R	9
5	Vergleich von Modellen	10
6	Weiter Fragestellung: Qualität der Schätzer	11
7	<i>optional</i> Internationalisierung	12

1 Motivation und Aufgabenstellung

Ziel dieses Kapitels ist es diese Arbeit zu motivieren und die Ziele dieser Arbeit aufzuzeigen.

1.1 Motivation

Die Idee für diese Arbeit ist entstanden, als ich im Winter 2018/2019 ein Praktikum bei einer Lebensversicherung gemacht habe. Dort habe ich eine Programmbibliothek geschrieben, die aus einer Sterbetafel Kommutationssymbole bestimmt.

Eine Sterbetafel ist eine Tafel, ist eine Tabelle, die für ein bestimmtes Kollektiv, die Ausscheidewahrscheinlichkeiten tabelliert. Um diese Aussage mit Inhalt zu füllen, definiere ich im folgenden die Begriffe. Bei einem Kollektiv handelt es sich um eine Menge an Personen, die im folgenden mit P_1, \dots, P_n bezeichnet werden. Jede dieser Personen hat einen Zeitpunkt T_i^g zu dem er geboren ist und einen zufälligen Zeitpunkt T_i^t zu dem er stirbt. Weiterhin ist das Alter X_i beim Ausscheiden aus dem Kollektiv interessant. Mit diesen Definitionen können wir nun 1-jährige Überlebenswahrscheinlichkeiten definieren:

$$q_x = \mathbb{P}(X \leq x + 1 | X > x), x \in \mathbb{N}_0 \quad (1)$$

Das heißt, q_x ist die Wahrscheinlichkeit, dass eine Person, die bereits x Jahre alt geworden ist, vor ihrem $x + 1$ -ten Geburtstag verstirbt. Diese Wahrscheinlichkeiten nennt man dann die Ausscheidewahrscheinlichkeiten durch Tod. Es gibt in der Personenversicherungsmathematik auch noch Invalidität als Ausscheidgrund, doch dieser Ausscheidgrund wird in dieser Arbeit nicht weiter untersucht.

Als nächstes werde ich darauf eingehen, warum Sterbetafeln für Lebensversicherer relevant sind und was Kommutationssymbole sind. In dieser Arbeit wird hauptsächlich die Sterbetafel 2004R verwendet. Dabei steht das R für Rente. Das heißt, diese Tafel wird für Versicherungsprodukte mit Erlebnisfallcharakter, wie die Rente, konzipiert. Erlebnisfallprodukte zeichnen sich dadurch aus, dass die versicherte Person nach eine Leistung erhält, wenn sie einen bestimmten Zeitraum n überlebt. In diesen Produkten, wird die Sterblichkeit als Rechnungsgrundlage neben dem Rechnungszins und den Kosten mit einbezogen, da nur bei Erleben geleistet wird. Konkret werden zwei Arten von Sterbetafeln generiert: Zum einen wird versucht die Sterblichkeiten möglichst genau zu schätzen, man spricht von einer Sterbetafel 2. Ordnung, zum anderen werden Risikoabschläge genommen, um sicher zu gehen, dass die Versicherung ihren Verpflichtungen auf alle Fälle nachkommen kann. Hier spricht man von einer Sterbetafel 1. Ordnung. Für die Konstruktion von Verträgen wird die Tafel 1. Ordnung verwendet. In dieser Arbeit interessiert uns der statistische Aspekt der Versicherungsmathematik, weshalb uns nur die 2. Ordnung interessiert.

An dieser Stelle sei noch kurz erwähnt, dass genau dieser Unterschied in den Sicherheitsabschlägen den Unterschied zwischen Tafeln mit Sterbe- und mit Rentencharakter ausmachen. Für Rentenversicherungen ist es sicherer eine etwas zu geringe Sterblichkeit anzunehmen, während es für Risikolebensversicherungen sicherer ist, eine etwas zu hohe Sterblichkeit anzunehmen.

Außerdem soll noch kurz auf Kommutationswerte eingegangen werden: Um Produkte mathematisch präzise beschreiben zu können, werden sogenannte Kommutationssymbole eingeführt. Dies wird am folgenden Beispiel anhand der Einmalprämie einer n -jährigen Erlebnisfallversicherung mit Versicherungssumme 1 Geldeinheit (GE) für eine Person mit Alter x erläutert. Sei diese Einmalprämie mit ${}_nE_x$ bezeichnet. Weiterhin bezeichne l_x die Lebenden im Alter x . Das heißt, es gelte $l_{x+1} = l_x \cdot (1 - p_x)$. Um an dieser Stelle Rundungsfehler zu vermindern, wird normalerweise eine recht große Startkollektivgröße angenommen. Zum Beispiel sei $l_0 = 1.000.000$. Dann ist

$${}_nE_x = \frac{l_{x+n} \cdot v^{x+n}}{l_x \cdot v^x}$$

wobei v den Diskontierungsfaktor bezeichnet. Er sei durch die folgende Gleichung definiert:

$$v^m = \left(\frac{1}{1+i} \right)^m$$

Bei dieser Gleichung bezeichne i den Rechnungszins. Insgesamt sieht man, dass die Einmalprämie also genau die diskontierte Anzahl an Personen am Ende der Aufschubzeit dividiert durch die diskontierte Anzahl an Personen am Beginn der Aufschubzeit ist. Deshalb spielt die Startgröße des Kollektivs keine Rolle und kann auf eine beliebige, numerisch angenehme Größe gesetzt werden.

Dieses Beispiel soll dieses Kapitel abschließen. Ich hoffe, dass ich eine gewisse Motivation geben konnte, warum Sterbetafeln von Interesse sind. Im nächsten Abschnitt wird auf die Ziele und die Aufgabenstellung dieser Arbeit eingegangen.

1.2 Ziele - Aufgabenstellung

Wir haben in dem obigen Abschnitt 1.1 gesehen, was Sterbetafeln sind und warum sie relevant sind. In diesem Abschnitt werden die Ziele dieser Arbeit vorgestellt.

Ein Problem bei der Verwendung von Sterbetafeln ist, dass sich die Sterblichkeit im Laufe der Zeit verändert. Genauer gesagt, hat sich die Lebenserwartung in den letzten Jahrzehnten erhöht. Diese Veränderung in der Lebenserwartung muss auch in Sterbetafeln berücksichtigt werden. Betrachten wir dazu zwei Säulendiagramme, die die Entwicklung der Sterblichkeit der Jahre 1990 und 2017 in der Bundesrepublik Deutschland zeigen:

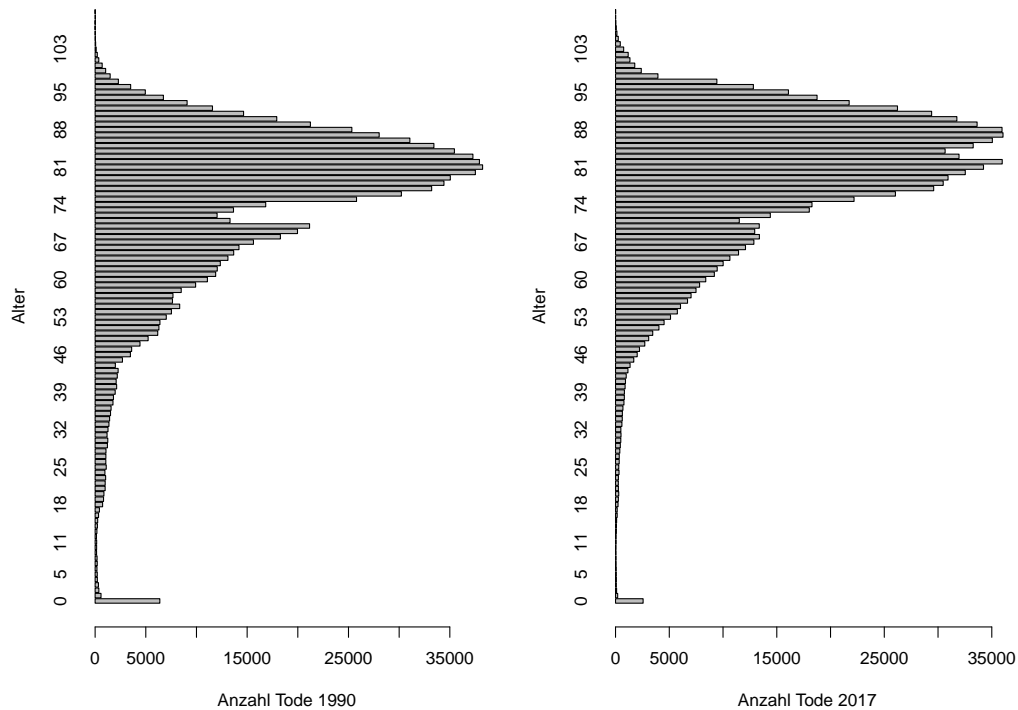


Figure 1: Anzahl an Toden in der Bundesrepublik Deutschland. Auf der linken Seite im Jahr 1990 und auf der rechten Seite im Jahr 2017

Wenn man die beiden Diagramme vergleicht, sieht man, dass im Jahr 2017 im Vergleich zum Jahr 1990 deutlich mehr Personen später versterben. Das Diagramm für das Jahr 2017 ist nach oben, zu den höheren Altern, verschoben. Diese Erhöhung der Lebenserwartung nennt man Sterbetrends. Um diese Trends zu schätzen, geht man in zwei Schritten vor. Im ersten Schritt wird eine Basistafel geschätzt und im zweiten Schritt werden dann Sterblichkeitstrends geschätzt.

Wie man diese Sterblichkeitstrends schätzt ist, wie bereits gesagt, auch in der Praxis sehr relevant. Entsprechend beschäftigt sich auch die Deutschen Aktuarvereinigung (DAV) mit diesem Problem und hat mit der Sterbetafel 2004R eine Richtlinie für deutsche Personenversicherer herausgegeben. Die Dokumentation für die Sterbetafel 2004R, sowie weiterer von der DAV veröffentlichte Sterbetafeln, können unter <https://aktuar.de/unsere-themen/lebensversicherung/Seiten/sterbetafeln.aspx> eingesehen werden.

Das Ziel dieser Arbeit ist es dann, sich Gedanken über die Eigenschaften dieser Schätzer zu machen und welche anderen Möglichkeiten man hat, um diese Trendfunktionen zu schätzen. Außerdem gibt es verschiedene Modelle und die Frage ist, wie man das beste Modell auswählt.

Diese Ziele geben der weitere Aufbau der Arbeit vor. Im nächsten Abschnitt, das heißt

in Abschnitt 2 gehen wir auf die Daten ein, die der Sterbetafel 2004R zugrunde liegen und auf Daten, die öffentlich verfügbar sind. In den nächsten drei Abschnitten 3, 4 und 5 gehen wir dann auf das Schätzen von Basistafeln, das Schätzen von Sterblichkeitstrends und auf die Auswahl von Modellen ein.

OPTIONAL In dem vorletzten Abschnitt wird die Qualität der Schätzer betrachtet und im letzten Abschnitt schauen wir uns die Situation im internationalen Vergleich an.

Abschließend möchte ich noch kurz auf die Daten und Codes, die zum erstellen der Diagramme (1) verwendet wurden, eingehen. Die Daten stammen von der Onlinedatenbank <https://www.mortality.org/cgi-bin/hmd/country.php?cntr=DEU&level=2>. Weitere Information zu diesen Daten können im Abschnitt 2 gefunden werden. Dieses Histogramm wurde in **R** erstellt. Der Code für dieses Diagramm, sowie alle anderen Codes, die in dieser Arbeit verwendet werden, können auf meiner Github-Seite unter <https://github.com/fake1884/moRtRe> eingesehen werden.

2 Daten

Das Konstruieren von Statistischen Modellen und Schätzern, ohne dass man Daten hat, auf deren Grundlage man TODO machen kann ist nicht sehr sinnvoll. Deshalb werden in diesem Abschnitt Datenbanken vorgestellt, die benutzt werden/man nutzen kann.

2.1 Die Daten der Sterbetafel 2004R

Die Sterbetafel 2004R basiert laut Kapitel 2 der Herleitung auf den folgenden Daten:

- Versichertendaten für die Jahre 1995 bis 2002, die von der Münchner Rück und der GenRe ausgewertet wurden,
- Abgekürzte Bevölkerungssterbetafeln des Statistischen Bundesamtes für das frühere Bundesgebiet ab 1971/73 bis 1998/200 (für 1986/88 stattdessen die Allgemeine Deutsche Sterbetafel 1986/88) und
- Daten der gesetzlichen Rentenversicherung (GRV) für das frühere Bundesgebiet von 1986 bis 2002 für die Alter von 66 bis 98 Jahren

Die Daten der Münchner Rück und der Gen Re sind Privateigentum der jeweiligen Unternehmen und deshalb nicht öffentlich zugänglich. Auch die Daten des Statistischen Bundesamtes bzw. der gesetzlichen Rentenversicherung sind auf ihren jeweiligen Webseiten nicht ohne weiteres öffentlich zugänglich. Allerdings gibt es auf der Internetseite Mortality.org Daten über die Sterblichkeit für viele Staaten. Auf diese Daten wird im nächsten Abschnitt genauer eingegangen.

Allerdings gibt es in dem 2. Kapitel der Herleitung der 2004R noch anderen interessante Punkte, die im folgenden Betrachtet werden sollen: TODO

Wiedervereinigung 1989 problematisch, da die Daten vorher separat erfasst wurden.

Ein weiteres Problem mit den Daten ist, dass es im Jahr 1969/1970 eine Grippewelle in Deutschland gab, die das Datenmaterial verzerrt.

2.2 Öffentlich zugängliche Sterblichkeitsdaten

In diesem Abschnitt werden die Daten von Mortality.org diskutiert. Auf der Website findet man vor allem die folgenden 4 Arten von Daten, die in dieser Arbeit benutzt werden:

- Geburtenzahlen pro Jahr aufgeteilt nach Geburten von Jungen und von Mädchen und ihre Summe.
- Gesamtbevölkerungsgröße pro Jahr und Alter aufgeteilt nach Männern, Frauen und Gesamt
- Anzahl an Toden pro Jahr und Alter aufgeteilt nach Männern, Frauen und Gesamt.

- Todesraten pro Jahr und Alter aufgeteilt nach Männern, Frauen und Gesamt.

Diese Daten sind von 1990 an für das ganze Bundesgebiet verfügbar und für Westbe beziehungsweise Ostdeutschland separat für die Jahre 1945 bis 1989.

Im folgenden eine kurze Erläuterung zu den Todesraten, die mortality.org entnommen sind:

TODO

Die folgende Tabelle soll einen kurzen Einblick in die Struktur der Daten geben:

TODO

Man sieht, dass die pro Jahr und Alter gegeben sind. In dieser Arbeit interessiert uns allerdings die Sterbewahrscheinlichkeiten einer Person, als Funktion ihres Geburtsjahres und ihres Alters. Wir wollen beispielsweise die Wahrscheinlichkeit schätzen, dass eine weibliche Person geboren im Jahr j , die bereits das Alter x erreicht hat, im Alter $x + 1$ stirbt. Also müssen die Daten in einem ersten Schritt umgewandelt werden. Auf diese Art erhält man die folgenden Daten:

TODO

Dies Daten nennt man ?Kohortendaten? im Gegensatz zu ...daten, die man mortality.org entnehmen kann. Jetzt gibt es noch das Problem, dass man dann nur Daten von 0-jährigen im Jahr 1990 bis zu 0-jährigen im Jahr 2017 benutzen kann. <- stimmt das? kann man nicht auch andere Daten benutzen?

Zensierte Daten, da nur Tode und Geburten bekannt <- TODO

Bei Sterbetrends handelt es sich um zensierte Daten. Das heißt, man hat noch keine Beobachtungen und versucht aus historischen Daten ein Modell zu entwickeln, mit dem man Trends in der Zukunft vorhersagen kann.

3 Schätzung von Basistafeln

Hier wird das Schätzen von Basistafeln besprochen. Bei einer Basistafel handelt es sich um

TODO

3.1 Die Basistafel der Sterbetafel 2004R

Die Herleitung der Tafel 2004R erklärt in ihrem 3. Kapitel, wie die Basistafel geschätzt wird und welches Jahr als das Bezugsjahr angesetzt wird:

TODO

3.2 Verschiedene Modelle für Basistafeln

In dieser Arbeit werden dagegen zwei Modelle für das Schätzen von Basistafeln angegeben und mit einer Simulationsstudie verglichen: Zum einen wird ein intuitives Modell angegeben, zum anderen wird der Stein-Schätzer benutzt. Als Bezugsjahr wird das Jahr 1990 benutzt, da dort die meisten nicht-

Modell 1: intuitiv

Wir haben Personen P_1, \dots, P_{440296} mit korrespondierenden, zufälligen Todeszeitpunkten

$$X_1, \dots, X_{440296} \sim^{i.i.d} F \quad (2)$$

wobei F eine unbekannte Verteilung ist. Wir haben Realisierungen ...

Uns interessiert, wie in der Einleitung genannt

$$q_x = \mathbb{P}(X \leq x + 1 | X > x) \quad (3)$$

und schätzen

$$\hat{q}_x = \frac{\sum_{i \in I} x_i}{\#x_i}, I = \{i_1, \dots, i_n\} \text{ s.t. } X_{i_j} \leq x + 1 \text{ aber } X_{i_j} > x \quad (4)$$

Modell 2: Stein

Uns interessiert der Vektor

$$(q_0, \dots, q_n) \quad (5)$$

und wir beobachten

$$(Y_1, \dots, Y_n) = (q_0, \dots, q_n) + (\epsilon_1, \dots, \epsilon_n) \quad (6)$$

mit

Vergleich der Modelle mittels Simulation

4 Schätzen von Sterblichkeitstrends

Hier wird das Schätzen der Sterblichkeitstrends betrachtet. Im vorherigen Kapitel haben wir gesehen, wie man am besten eine Basistafel schätzt. Ausgehend von dieser Basistafel gibt es verschiedene Möglichkeiten die Sterbetrends zu schätzen. Im folgenden Betrachten wir zuerst das vorgehen bei der Sterbetafel 2004R und danach TODO

4.1 Ergebnisse der Schätzung der Sterbetafel 2004R

Dieser Abschnitt basiert auch dem 4. Kapitel der Herleitung zur Basistafel 2004R.
traditionelles Modell:

$$\frac{q_{x,t+1}}{q_{x,t}} = \exp(-F(x)) \quad (7)$$

laut Anhang 11 wird F mittels einer Methode der kleinsten Quadrate durch lineare Regression aus den Werten für $\ln(q_{x,t})$ für $t_0 \leq t \leq t_1$ geschätzt.

Kohortenmodell:

$$\frac{q_{x,t+1}}{q_{x,t}} = \exp(-G(t+1-x)) \quad (8)$$

Synthesenmodell

$$\frac{q_{x,t+1}}{q_{x,t}} = \exp(-F(x) - G(t+1-x)) \quad (9)$$

Die Ergebnisse der Trendschätzung können in DAV 2004R Anhang 1 Teil B gefunden werden.

Alle diese Modelle sind nicht parametrisch. Vielleicht wäre ein linear Parametrisches Modell besser? also $F(x) = ax + b$?

Dann gibt es noch die Trenddämpfung TODO

5 Vergleich von Modellen

Nachdem man verschieden Modelle geschätzt hat, muss man ein Modell auswählen, dass man nimmt.

-> Anhang 11 : Likelihood-Ratio-Test zum Modellvergleich

6 Weiter Fragestellung: Qualität der Schätzer

Eine mögliche Fragestellung ist dann, welche Qualität diese Modelle besitzen:

- Sind sie konsistent?
- Sind sie unverzert und vielleicht UMVU?
- Wie sind die Fehler verteilt?

7 *optional* Internationalisierung

Man könnte sich eventuell auch internationale Sterbetrends anschauen.