Schadenversicherungsmathematik

Teil 1: Grundlagen

Dr. Ulrich Riegel

Mathematisches Institut
Ludwig-Maximilians-Universität München

Wintersemester 2023/24

Inhalt

Was ist Versicherung?

Versicherungsvertrag (Police):

- Der *Versicherungsnehmer (VN)* verpflichtet sich zur Zahlung eines im voraus fälligen Geldbetrags *(Prämie)*.
- Das *Versicherungsunternehmen (VU)* verpflichtet sich bei Eintritt von im Vertrag näher definierten ungewissen Ereignissen *(Schäden)* bestimmte Zahlungen an den VN zu leisten.
- Die Zahlungen hängen meist vom betreffenden Ereignis ab und sollen den aus dem Ereignis resultierenden Nachteil des VN reduzieren oder ausgleichen.

Was ist Versicherung?

Das VU übernimmt also ungewisse Zahlungen gegen feste Prämie.

- Der VN kann ungewissen Kosten gegen planbare Kosten tauschen.
- Ruinös hohe Schäden mit kleinen Eintrittswahrscheinlichkeiten werden durch Versicherung überhaupt erst tragbar.

Überblick über die wichtigsten Versicherungszweige

Beitragseinnahmen 2021 in Deutschland in Mrd. EUR:

Lebensversicherung:	103,2
Krankenversicherung:	45,4
Schadenversicherung:	77,3

Begriffe:

- SQ = Schadenquote = Schäden/Prämie
- HUK steht für Haftpflicht, Unfall & Kraftfahrt
- Kostenquote = Kosten/Prämie

Kostenquote 2021 in der Schadenversicherung insgesamt 18,8%
(Sach 23,6%, Kraftfahrt 10,8%, Haftpflicht 22,2%)

Quelle: GDV, Statistiken zur deutschen Versicherungswirtschaft, Sept. 2022

Versicherungszweige in der Schadenversicherung

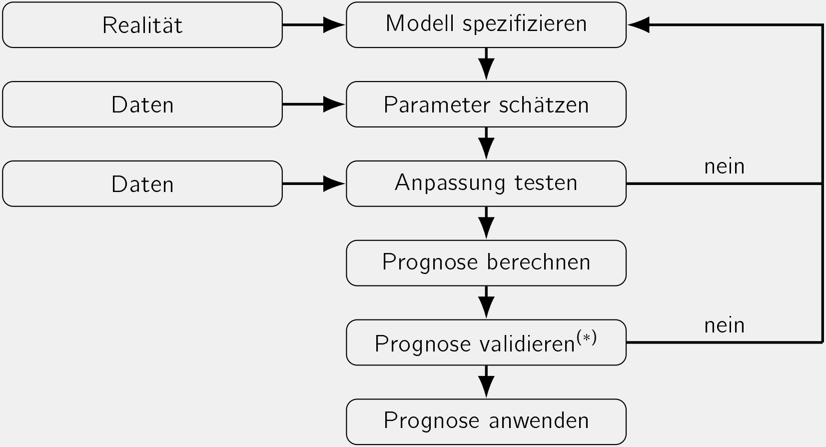
Beitrag 2021 (Mio. EUR)	Versicherungszweig		SQ 2019	SQ 2020	SQ 2021
17.047	Kfz-Haftpflicht (KH)	HUK	89,9%	79,4%	79,9%
10.237	Kfz-Vollkasko		88,8%	76,6%	92,4%
1.676	Kfz-Teilkasko		65,6%	56,0%	77,8%
8.410	Allg. Haftpflicht		63,9%	64,9%	64,4%
6.728	Unfall		59,1%	56,4%	56,6%
4.594	Rechtsschutz		70,0%	73,9%	69,7%
8.311	Industrielle und gewerbliche Sachversicherung	Sach	76,6%	90,6%	128,1%
2.557	Technische Versicherung		60,7%	58,1%	68,5%
9.337	Verbundene Wohngebäude		71,4%	66,6%	114,9%
3.296	Verbundene Hausrat		38,6%	35,1%	60,3%
536	sonstige Sachversicherung				
2.071	Transport und Luftfahrt		69,8%	65,6%	56,2%
1.953	Kredit		61,0%	56,2%	33,1%
244	Schutzbriefversicherung				
77.288	**Schadenversicherung gesamt**		**74,2%**	**70,3%**	**83,5%**

Einordnung der Schadenversicherungsmathematik

Unterschiede zwischen *Lebens*- und *Schadenversicherung*

Lebensversicherung	**Schadenversicherung**
max. 1 Schaden pro Risiko	mehrere Schäden pro Risiko/Jahr möglich
feste Schadenhöhe $\Rightarrow$ niedrige Variabilität	variable Schadenhöhe $\Rightarrow$ hohe Variabilität
lange Policenlaufzeit	kurze Policenlaufzeit (1 Jahr)
hohes Gewicht des Zinsertrags $\Rightarrow$ Finanzmathematik	niedrigeres Gewicht des Zinsertrags $\Rightarrow$ Stochastik

Modellbildung



(*) Plausibilitätskontrolle, Sensitivitätsprüfung, Ausreißereinfluss

Bezeichnungen und Begriffe

Alle Zufallsvariablen seien auf einem fest gewählten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mathrm{P})$ definiert. Für Zufallsvariable X und Y:

$\mathrm{E}(X)$	Erwartungswert von X bezüglich P
$\mathrm{Var}(X)$	Varianz von X bezüglich P
$\mathrm{Cov}(X, Y)$	Kovarianz von X und Y bezüglich P
$\mathrm{Sd}(X)$	Standardabweichung X, $\mathrm{Sd}(X) = \sqrt{\mathrm{Var}(X)}$
$\mathrm{Vco}(X)$	*Variationskoeffizient* von X, d.h. $\mathrm{Vco}(X) = \mathrm{Sd}(X)/\,\mathrm{E}(X)$
$\mathrm{Sch}(X)$	*Schiefe* von X, d.h. $\mathrm{Sch}(X) = \mathrm{E}((X - \mathrm{E}(X))^3)/\,\mathrm{Sd}(X)^3$
F_X	Verteilungsfunktion von X, d.h. $F_X(x) := \mathrm{P}(X \leq x)$
F_X^-	Quantilfunktion von F, d.h. $F_X^-(p) := \inf\{x \in \mathbb{R}_+ \mid F_X(x) \geq p\}$

Bezeichnungen und Begriffe

Risiko	Zufallsvariable $R \geq 0$ mit $0 < \mathsf{E}(R) < \infty$, $0 < \mathsf{Var}(R) < \infty$. Kleinste Einheit, die Gegenstand eines Versicherungsvertrags sein könnte. R gibt die Höhe des vom VU (unter diesem Vertrag) zu bezahlenden Jahresgesamtschadens an.
Police	Versicherungsvertrag für ein oder mehrere Risiken eines VN.
Risikogruppe (Kollektiv)	Menge von Risiken mit ähnlichen äußeren Merkmalen (z.B. Einfamilienhäuser in München in der Feuerversicherung).
Portfolio	Menge von beliebigen Risiken (z.B. alle Risiken eines VU).

Bezeichnungen und Begriffe

$\mathbb{R}_+$	Nichtnegative reelle Zahlen $\mathbb{R}_+ := [0, \infty)$
$\mathcal{L}^p(\mathbb{R}_+)$	Menge aller Zufallsgrößen X mit $X(\Omega) \subseteq \mathbb{R}_+$ und $\mathsf{E}(\lvert X\rvert^p) < \infty$ $(1 \leq p < \infty)$
$\mathcal{L}^R$	Menge aller Risiken; $\mathcal{L}^R = \{R \in \mathcal{L}^2(\mathbb{R}_+) \mid \mathsf{Var}(R) > 0\}$
$\mathcal{N}(\mu, \sigma^2)$	Normalverteilung mit Erwartungswert μ und Varianz σ^2
$\mathcal{N}_d(\mu, \Sigma)$	d-dimensionale Normalverteilung mit Erwartungswertvektor μ und Kovarianzmatrix Σ
Φ	Verteilungsfunktion der Standardnormalverteilung
P^X	Bildmaß von X
$P^{X\mid Y}$	reguläre bedingte Wahrscheinlichkeit von X gegeben Y
$X_n \xrightarrow{\mathsf{P}} Y$	Die Folge X_n konvergiert in P-Wahrscheinlichkeit gegen Y
$X_n \xrightarrow{\mathcal{D}} \mu$	Die Folge X_n konvergiert in Verteilung gegen das W-Maß μ

Inhalt

Erinnerung

Ungleichung von Tschebyscheff: Sei X eine Zufallsvariable mit endlicher Varianz, dann gilt für $\lambda > 0$

$$\mathrm{P}(|X - \mathrm{E}(X)| \geq \lambda) \leq \frac{\mathrm{Var}(X)}{\lambda^2}.$$

Schwaches Gesetz der großen Zahlen: Sei $X_1, X_2, \ldots$ eine Folge integrierbarer und paarweise unkorrelierter Zufallsvariabler mit $\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(X_i) = 0$. Dann gilt

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - \mathrm{E}(X_i)) = 0 \quad \text{in P-Wahrscheinlichkeit.}$$

Erinnerung

Starkes Gesetz der großen Zahlen: Für jede Folge $X_1, X_2, \ldots$ von paarweise unabhängigen, integrierbaren, identisch verteilten Zufallsvariablen gilt

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = \mathsf{E}(X_1) \quad \text{fast sicher.}$$

Definition: Eine Folge $X_1, X_2, \ldots$ von Zufallsvariablen heißt *unabhängig und identisch verteilt* (kurz *i.i.d. – independent and identically distributed*), wenn die Familie $(X_i)_{i\geq 1}$ unabhängig ist und die X_i alle die gleiche Verteilung haben.

Erinnerung

Zentraler Grenzwertsatz: Für i.i.d.-Folgen quadratintegrierbarer Zufallsvariablen $X_1, X_2, \ldots$ mit $\mu := \mathrm{E}(X_i)$ und $\sigma := \mathrm{Sd}(X_1)$ gilt

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1) \qquad (n \to \infty).$$

Cauchy-Schwarzsche Ungleichung: Für jede positiv semidefinite, symmetrische Bilinearform b auf einem Vektorraum V gilt

$$b(v,w)^2 \leq b(v,v)b(w,w) \quad \text{für alle } v, w \in V.$$

Spezialfall: $|\mathrm{Cov}(X,Y)| \leq \mathrm{Sd}(X)\,\mathrm{Sd}(Y)$ für alle quadratintegrierbaren Zufallsvariablen X und Y.

Ausgleich im Kollektiv

Frage: Wieso funktioniert Versicherung?

Satz (Ausgleich im Kollektiv): Sei $R_1, R_2, \ldots$ eine Folge von identisch verteilten und paarweise unkorrelierten Risiken. Dann gilt für den Gesamtschaden $S_I = R_1 + \cdots + R_I$ der ersten I Risiken

(1) $\lim\limits_{I \to \infty} \mathrm{Vco}(S_I) = 0$,
d.h. $\mathrm{Sd}(S_I)$ wächst bei wachsendem Kollektiv langsamer als $\mathrm{E}(S_I)$.

(2) $\lim\limits_{I \to \infty} \mathrm{P}\left(\left| \frac{S_I - \mathrm{E}(S_I)}{\mathrm{E}(S_I)} \right| > \varepsilon \right) = 0 \quad$ für alle $\varepsilon > 0$,
d.h. das Überschreiten einer prozentualen Maximalabweichung vom Erwartungswert wird bei wachsendem Kollektiv immer unwahrscheinlicher.

Ausgleich im Kollektiv

Bemerkung:

- Absolut wird die Standardabweichung mit wachsendem Portefeuille immer größer (vgl. auch folgendes Beispiel).
- Sie wächst aber langsamer als der Erwartungswert.
- Aus betriebswirtschaftilicher Sicht ist die *relative Abweichung* relevant!

Beispiel

Wir betrachten I Würfel. Es sei R_i die Augenzahl des Würfels i. Dann ist S_I die Summe der Augen aller I Würfel. Sei $\varepsilon = 0{,}1$.

I	$\mathrm{E}(S_I)$	$\mathrm{Sd}(S_I)$	$z_I = \frac{0{,}1 \cdot \mathrm{E}(S_I)}{\mathrm{Sd}(S_I)}$	$2\,\Phi(-z_I)$	$\mathrm{P}\left(\left\lvert\frac{S_I - \mathrm{E}(S_I)}{\mathrm{E}(S_I)}\right\rvert > \varepsilon\right)$
1	3.5	1.71	0.205	0.84	1
10	35	5.40	0.648	0.517	0.522
100	350	17.1	2.05	0.040	0.040
1000	3500	54.0	6.48	10^{-10}	10^{-10}

Hier haben wir den zentralen Grenzwertsatz für die Approximation

$$\mathrm{P}\left(\left\lvert\frac{S_I - \mathrm{E}(S_I)}{\mathrm{E}(S_I)}\right\rvert > \varepsilon\right) = \mathrm{P}\left(\left\lvert\frac{S_I - \mathrm{E}(S_I)}{\mathrm{Sd}(S_I)}\right\rvert > \frac{\varepsilon \cdot \mathrm{E}(S_I)}{\mathrm{Sd}(S_I)}\right) \approx 2 \cdot \Phi\left(-\frac{\varepsilon \cdot \mathrm{E}(S_I)}{\mathrm{Sd}(S_I)}\right).$$

verwendet.

Ausgleich im Kollektiv

Bemerkung:

- In der Realität sind die Risiken meist nicht identisch verteilt (z.B. verschieden große Häuser in der Feuerversicherung).
- Oft sind sie auch nicht unkorreliert (benachbarte Häuser in Feuer, Sturm, Erdbeben).

Definition: Eine Folge $R_1, R_2, \ldots$ von Risiken *genügt dem Ausgleich im Kollektiv*, wenn für $S_l = R_1 + \cdots + R_l$ gilt:

$$\lim_{l\to\infty} \mathrm{P}\left(\left|\frac{S_l - \mathrm{E}(S_l)}{\mathrm{E}(S_l)}\right| > \varepsilon\right) = 0 \quad \text{für alle } \varepsilon > 0\ .$$

Ausgleich im Kollektiv

Satz (Bühlmann): Seien $\mu_0, \sigma_0, n_0 > 0$. Jede Folge $R_1, R_2, \ldots$ von Risiken mit

$$\begin{aligned} \mathrm{E}(R_i) &\geq \mu_0 > 0, \\ \mathrm{Var}(R_i) &\leq \sigma_0^2 < \infty, \\ \mathrm{Cov}(R_i, R_j) &\leq 0 \quad \text{für } |i - j| > n_0 \end{aligned}$$

genügt dem Ausgleich im Kollektiv.

Fazit: Ein gewisser Ausgleich findet also fast immer statt. Natürlich bei positiver Korrelation langsamer als bei Unabhängigkeit. Ideal wären negativ korrelierte Risiken, was es aber praktisch nicht gibt.

Versicherungstechnisches Risiko

Definition:

Unter dem *versicherungstechnischen Risiko* versteht man die Möglichkeit, dass S größer ist als der geschätzte Erwartungswert $\widehat{\mathrm{E}}(S)$.

Das versicherungstechnische Risiko besteht aus drei Komponenten:

- *Zufallsrisiko*:
  Abweichnung der Schadenvariable S von $\mathrm{E}(S)$ unter der Annahme, dass die Verteilung bekannt ist.
- *Schätzrisiko* (auch *Diagnose*- oder *Irrtumsrisiko*):
  Abweichung der Schätzung $\widehat{\mathrm{E}}(S)$ vom wirklichen Erwartungswert $\mathrm{E}(S)$ unter der Annahme, dass die Vergangenheit aussagekräftig für die Zukunft ist.
- *Änderungsrisiko* (auch *Prognoserisiko*):
  Differenz zwischen $\mathrm{E}(S)$ in der Vergangenheit und $\mathrm{E}(S)$ in der Zukunft.

Gesetz der großen Zahlen im Versicherungskontext

Seien $R_1, R_2, \ldots$ i.i.d. Risiken. Dann besagt das starke Gesetz der großen Zahlen

$$\mathrm{P}\left(\lim_{l \to \infty} \frac{R_1 + \cdots + R_l}{l} = \mathrm{E}(R_1)\right) = 1. \qquad (*)$$

Interpretation 1:

- $R_1, R_2, \ldots$ sind unabhängige Wiederholungen (= Versicherungsjahre) eines festen Risikos bei unveränderten äußeren Bedingungen.
- Dann besagt $(*)$, dass der durchschnittliche Schadenaufwand pro Jahr fast sicher gegen $\mathrm{E}(R_1)$ konvergiert.

Gesetz der großen Zahlen im Versicherungskontext

Interpretation 2:

- $R_1, R_2, \ldots$ sind i.i.d. Risiken im gleichen Jahr.
- Dann ist das arithmetische Mittel $\frac{1}{I}\sum_{i=1}^{I} R_i$ laut $(*)$ ein konsistenter Schätzer für $\mathsf{E}(R_1)$, d.h. man kann $\mathsf{E}(R_1)$ schätzen, indem man eine möglichst große Gruppe von wie R_1 verteilten unabhängigen Risiken $R_1, R_2, \ldots, R_I$ betrachtet und deren Durchschnittsschaden $\frac{1}{I}\sum_{i=1}^{I} R_i$ bildet.
- In der Praxis bilden die VU daher oft Gemeinschaftsstatistiken zur Schätzung von $\mathsf{E}(R_1)$.

Sicherheitskapital

Bemerkung:

- Wenn ein VU zur Bezahlung des Jahres-Gesamtschadens $S = \sum_{i=1}^{I} R_i$ seines Portefeuilles nur die theoretische Nettoprämie $E(S)$ zur Verfügung hätte, wäre VU insolvent, sobald $S > \mathrm{E}(S)$ (d.h. mit ca. 50% Wahrscheinlichkeit).
- Folglich braucht das VU *Sicherheitskapital (Eigenkapital)*.

Definition: Sei c das Sicherheitskapital des VU und G die Verteilungsfunktion des Gesamtschadens S. Dann heißt

$$\varepsilon = \mathrm{P}(S > \mathrm{E}(S) + c) = 1 - G(\mathrm{E}(S) + c)$$

Insolvenzwahrscheinlichkeit und $1 - \varepsilon = \mathrm{P}(S \leq \mathrm{E}(S) + c) = G(\mathrm{E}(S) + c)$ heißt *Sicherheitswahrscheinlichkeit*.

Sicherheitskapital

Bemerkung:

- Das Sicherheitskapital c muss so groß sein, dass die Insolvenzwahrscheinlichkeit ε „genügend klein“ ist (unter Solvency II: $\varepsilon \leq 0{,}5\%$).
- Die Ermittlung der Gesamtschaden-Verteilung G eines VU ist daher eine wichtige Aufgabe der Schadenversicherugsmathematik.

Kapitalkosten

Bemerkung:

- Das Sicherheitskapital c sollte „risikofrei" angelegt werden (risikofreier Zins, z.B. $r = 1\%$).
- Da c aber dennoch unter Verlustrisiko steht (durch S), verlangt der Kapitalmarkt/Geldgeber mehr Ertrag (Zins, Dividende) $r + s$, z.B. $r + s = 8\%$.
- Den Zusatzzins $s \cdot c$ (von z.B. $(8\% - 1\%)c = 7\% \cdot c$) nennt man *Kapitalkosten*.
- Diese müssen von den VN erbracht werden, da diese den Schutz des Sicherheitskapitals genießen:

  $\Rightarrow$ *Sicherheitszuschlag* (= *Schwankungszuschlag*)

Ungleichungen von Cantelli

Genauere Aussagen zur Größe des Sicherheitskapitals ermöglichen die Ungleichungen von Cantelli.

Satz (Ungleichungen von Cantelli):

Für jede quadratintegrierbare Zufallsvariable X gilt

$$\mathrm{P}\left(X \geq \mathrm{E}(X) + a\right) \leq \frac{\mathrm{Var}(X)}{\mathrm{Var}(X) + a^2} \quad \text{für } a > 0.$$

Ferner haben wir

$$\mathrm{P}\left(X > \mathrm{E}(X) - a\right) \geq \frac{a^2}{\mathrm{Var}(X) + a^2} \quad \text{für } a > 0.$$

Ungleichungen von Cantelli

Bemerkung: Um eine vorgegebene Sicherheitswahrscheinlichkeit $1-\varepsilon$ einzuhalten, können wir gemäß Cantelli

$$c^2 = \frac{1-\varepsilon}{\varepsilon} \operatorname{Var}(S_I)$$

wählen, d.h. zur Einhaltung einer fest vorgegebenen Insolvenzwahrscheinlichkeit sollte das Sicherheitskapital c gemäß Cantelli proportional zu $\operatorname{Sd}(S_I)$ wachsen.

Bemerkung: Im i.i.d.-Fall liefert der „Cantelli-Ansatz" $c^2 = \frac{1-\varepsilon}{\varepsilon} \operatorname{Var}(S_I)$

$$\left(I \cdot \frac{c}{I}\right)^2 = \frac{1-\varepsilon}{\varepsilon} \cdot I \cdot \operatorname{Var}(R_1), \quad \text{d.h.} \quad \frac{c}{I} = \sqrt{\frac{1-\varepsilon}{\varepsilon}} \cdot \frac{\operatorname{Sd}(R_1)}{\sqrt{I}},$$

d.h. bei wachsendem Portefeuille kann der Schwankungszuschlag pro Risiko reduziert werden (oder es ergibt sich eine höhere Sicherheitswahrscheinlichkeit).

Netto-, Risiko- und Bruttoprämie

Definition:

- Der erwartete Schaden $\mathsf{E}(S)$ bzw. $\mathsf{E}(R)$ wird auch als *Nettoprämie* bezeichnet.
- Die *Risikoprämie* setzt sich aus der Nettoprämie und dem Sicherheitszuschlag zusammen.
- Die *Bruttoprämie* besteht aus der Risikoprämie, einem Zuschlag für Betriebskosten (Durchlaufposten ca. 20%–40%) und einem Gewinnzuschlag.
- Auf die Bruttoprämie wird noch die Versicherungssteuer aufgeschlagen.

Inhalt

Problemstellung

Gemäß dem starken Gesetz der großen Zahlen möchte man zur Prämienkalkulation Gruppen von möglichst vielen i.i.d. Risiken bilden. In der Praxis wird

- „i.i.d." durch „ähnliche äußere Merkmale und ähnliche Gefährdung" ersetzt und
- „möglichst viele" durch Bildung von Marktstatistiken und Heranziehung mehrerer Beobachtungsjahre (nach Inflationskorrektur!) umgesetzt.

Um möglichst wenig Details über ihr Portefeuille preiszugeben, stellen die VU für Marktstatistiken häufig nur aggregierte Daten zur Verfügung, d.h. pro Risikogruppe die Anzahl Risiken, ggf. deren Gesamt-VS sowie Zahl und Gesamtbetrag der Schäden. Dann ist die Verteilung von einzelnen Risiken R_i nicht beobachtbar.

Problemstellung

VU kennen somit nur den Jahresgesamtschaden S für jede Risikogruppe, aber Volumen und Verteilung der Risikogruppe ändern sich von Jahr zu Jahr.

Daher müssen wir den Volumeneinfluss auf die Verteilungsparameter von S quantifizieren.

Für die statistische Analyse benötigen wir einfaches parametrisches Modell für den Gesamtschaden $S = \sum_{i=1}^{I} R_i$ einer Risikogruppe.

Individuelles Modell

Definition:

Die Modellannahmen des *Individuellen Modells* einer Gruppe $R_1, \ldots, R_I$ von Risiken mit bekannten Volumina (VS, JE) $u_1, \ldots, u_I$ lauten:

(1) $R_1, \ldots, R_I$ sind unabhängig,

(2) $\mathrm{E}(R_i) = mu_i$

(3) $\mathrm{Var}(R_i) = s^2 u_i$

für $1 \leq i \leq I$. Hierbei sind m und s^2 unbekannte Parameter, die bei Inflationsbereinigung der Schäden über mehrere Jahre konstant angenommen werden, aber von Risikogruppe zu Risikogruppe verschieden sind.

Das Volumen $v = u_1 + \cdots + u_I$ heißt (je nach Kontext) auch *Jahres-Gesamt-VS* bzw. *Anzahl Jahreseinheiten*.

Spezialfälle

Homogene Risikogruppe:

Sind $R_1, \ldots, R_I$ i.i.d. mit $m := \mathrm{E}(R_i)$ und $s^2 := \mathrm{Var}(R_i)$, dann gilt das individuelle Modell mit $u_i = 1$. In diesem Fall ist das Volumen gleich der Anzahl Risiken, d.h. $v = I$.

Unterjährige Versicherungsdauer (z.B. KH):

Das Risiko R_i habe nun eine Versicherungsdauer von $t_i \leq 365$ Tagen. Die versicherten Tage sämtlicher Risiken seien i.i.d. mit Erwartungswert $m/365$ und Varianz $s^2/365$. Mit $u_i := t_i/365$ gilt dann $\mathrm{E}(R_i) = t_i m/365 = m u_i$ und $\mathrm{Var}(R_i) = t_i s^2/365 = s^2 u_i$, d.h. die Axiome des individuellen Modells sind erfüllt. Das Volumen $v = u_1 + \cdots + u_I$ ist dann gleich der Anzahl Jahreseinheiten.

Spezialfälle

Unterschiedliche Versicherungssummen (z.B. Sach):

Hat R_i die Versicherungssumme u_i, so modelliert man (analog zur unterjährigen Versicherungsdauer)

$$\mathrm{E}(R_i) = \frac{u_i}{u_1} \cdot \mathrm{E}(R_1) \quad \text{und} \quad \mathrm{Var}(R_i) = \frac{u_i}{u_1} \mathrm{Var}(R_1),$$

d.h. R_i besteht aus u_i/u_1 unabhängigen „Stücken", die wie R_1 verteilt sind. Das Volumen $v = u_1 + \cdots + u_I$ ist in dieser Situation gleich der Jahres-Gesamt-VS.

Dieses Modell kann realistisch sein, wenn die Versicherungssumme bei gewerblichen Policen i.W. durch die Anzahl Gebäude auf dem Firmengelände bestimmt wird. Ist hingegen immer nur ein Objekt gedeckt, so wäre oft $R_i \sim \frac{u_i}{u_1} R_1$ ein besseres Modell.

Parameterschätzung im Individuellen Modell

Satz: *Im Individuellen Modell gilt für den des Gesamtschaden* $S = R_1 + \cdots + R_I$:

$$\mathrm{E}(S) = mv \qquad \textit{und} \qquad \mathrm{Var}(S) = s^2 v.$$

Für den Schadensatz $Z = S/v$ *gilt*

$$\mathrm{E}(Z) = m \qquad \textit{und} \qquad \mathrm{Var}(Z) = \frac{s^2}{v}.$$

Liegen unabhängige jahrweise Beobachtungen $Z_j = S_j/v_j$ *bei Volumen* v_j *vor* $(j = 1, \ldots, J)$, *so haben wir folgende erwartungstreue Parameterschätzer:*

$$\widehat{m} = \sum_{j=1}^{J} \frac{v_j}{\sum_{k=1}^{J} v_k} \cdot Z_j = \frac{\sum_{j=1}^{J} S_j}{\sum_{j=1}^{J} v_j},$$

$$\widehat{s}^2 = \frac{1}{J-1} \sum_{j=1}^{J} v_j \cdot (Z_j - \widehat{m})^2.$$

Parameterschätzung im Individuellen Modell

Bemerkung:

Hat v die Dimension EUR, so sind m und $\widehat{m}$ dimensionslos, s^2 und $\widehat{s}^2$ haben die Dimension EUR. Ist v dimensionslos, so haben m und $\widehat{m}$ die Dimension EUR und s^2 sowie $\widehat{s}^2$ die Dimension EUR^2.

Der Schätzer

$$\widehat{m} = \sum_{j=1}^{J} \frac{v_j}{\sum_{k=1}^{J} v_k} \cdot Z_j$$

ist ein volumengewichtetes Mittel der Z_j. Warum betrachten wir statt $\widehat{m}$ nicht einfach

$$\frac{1}{J} \sum_{j=1}^{J} Z_j ?$$

Beachte: Jedes Z_j ist ein erwartungstreuer Schätzer für m mit Varianz $\mathrm{Var}(Z_j) = \frac{s^2}{v_j}$, d.h. mit unterschiedlicher Genauigkeit.

Schätzer mit minimaler Varianz

Satz: *Seien* $T_1, \ldots, T_J$ *unkorrelierte und erwartungstreue Schätzer für* $t \in \mathbb{R}$, *d.h.* $\mathrm{E}(T_j) = t$ *für* $1 \leq j \leq J$. *Betrachte die Menge der Konvexkombinationen*

$$\mathcal{T} = \left\{ \sum_{j=1}^{J} w_j T_j \text{ mit } \sum_{j=1}^{J} w_j = 1 \right\}.$$

Beachte, dass $\mathrm{E}(T) = t$ *für jedes* $T \in \mathcal{T}$ *gilt. Dann hat* $T = \sum_{j=1}^{J} w_j T_j \in \mathcal{T}$ *genau dann minimale Varianz unter allen Elementen von* $\mathcal{T}$ *wenn* w_j *indirekt proportional zu* $\mathrm{Var}(T_j)$ *ist, d.h. wenn*

$$w_j = w_j^* \quad \text{mit} \quad w_j^* := \frac{1}{\sum_{k=1}^{J} \mathrm{Var}(T_k)^{-1}} \cdot \mathrm{Var}(T_j)^{-1}$$

gilt.

Modellierung des Gesamtschadens

Motivation:

- Oft ist es nicht ausreichend, den Erwartungswert und die Varianz des Gesamtschadens S zu schätzen.
- Man braucht eine Verteilung für S.
- Im individuellen Modell gilt

$$\mathrm{E}(S) = v \cdot m \quad \text{und} \quad \mathrm{Var}(S) = v \cdot s^2$$

wobei das Volumen v bekannt, aber die Parameter m und s^2 unbekannt sind.

Frage: Welche Verteilungsfamilie mit zwei Parametern ist am besten geeignet?

Eine Normalverteilung ist vermutlich nicht geeignet (negative Werte möglich, Symmetrie).

Anforderungen an die Verteilungsfamilie

A priori bekannt:

- $R_i \geq 0$ und $S \geq 0$
- Die Verteilung für ein individuelles Risiko R_i liefert eine hohe Wahrscheinlichkeit für $R_i = 0$, da die meisten Risiken pro Jahr schadenfrei sind
- Die Risiken $R_1, R_2, \ldots$ sind unabhängig

Um von der Verteilung der $R_1, \ldots, R_I$ zur Verteilung von $S = \sum_{i=1}^{I} R_i$ zu kommen, muss man Faltungsprodukte berechnen können und diese sollten wieder in der Verteilungsfamilie liegen.

Eines der bekanntesten Beispiele ist die *Gamma-Verteilung*.

Gamma-Verteilung

Definition:

Die Dichte der *Gamma-Verteilung* $\Gamma(\mu, \alpha)$ mit $\mu > 0, \ \alpha > 0$ ist gegeben durch

$$g(x) = g_{\mu,\alpha}(x) = \frac{\left(\frac{\alpha}{\mu}\right)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\left(-\frac{\alpha x}{\mu}\right)$$

für $0 < x < \infty$. Hierbei ist

$$\Gamma\colon (0,\infty) \longrightarrow (0,\infty), \quad x \longmapsto \int_0^\infty t^{x-1} e^{-t}\, dt$$

die *Gamma-Funktion*.

Eigenschaften der Gamma-Funktion

Lemma: *Es gilt*

(1) $\Gamma(x+1) = x\Gamma(x)$ *für alle* $x > 0$

(2) $\Gamma(n+1) = n!$ *für alle* $n \in \mathbb{N}_0$

(3) $\dfrac{\Gamma(x+n)}{\Gamma(x)} = x(x+1)\cdot\ldots\cdot(x+n-1)$ *für* $x > 0,\ n \in \mathbb{N}$

(4) $\displaystyle\int_0^\infty t^{x-1}e^{-at}\,dt = \frac{\Gamma(x)}{a^x}$ *für* $x > 0,\ a > 0$

Eigenschaften der Gamma-Verteilung

Lemma: *Sei $X \sim \Gamma(\mu, \alpha)$. Dann gilt*

$$\mathrm{E}(X) = \mu \quad \textit{und} \quad \mathrm{Var}(X) = \mu^2/\alpha.$$

Insbesondere sehen wir

$$\alpha = \frac{\mathrm{E}(X)^2}{\mathrm{Var}(X)} = \frac{1}{\mathrm{Vco}(X)^2} \quad \text{bzw.} \quad \mathrm{Vco}(X) = \frac{1}{\sqrt{\alpha}}.$$

Eigenschaften der Gamma-Verteilung

Der Parameter μ ist ein *Skalenparameter*, denn für $\lambda > 0$ gilt

$$\frac{1}{\lambda} \cdot g_{\mu,\alpha}\left(\frac{x}{\lambda}\right) = g_{\lambda\mu,\alpha}(x),$$

d.h. für $X \sim \Gamma(\mu, \alpha)$ ist $\lambda X \sim \Gamma(\lambda\mu, \alpha)$. Währungs- oder Einheiten-Änderung betreffen somit nur den Skalenparameter μ.

Der zweite Parameter α ist der *Formparameter*.

Die Verteilung ist rechtsschief mit Schiefe

$$\mathrm{Sch}(X) = \mathrm{E}\left(\frac{(X-\mu)^3}{\mathrm{Sd}(X)^3}\right) = \frac{\mathrm{E}(X^3) - 3\,\mathrm{Var}(X)\,\mathrm{E}(X) - \mathrm{E}(X)^3}{\mathrm{Var}(X)^{3/2}} = \frac{2}{\sqrt{\alpha}}.$$

Eigenschaften der Gamma-Verteilung

Fallunterscheidung:

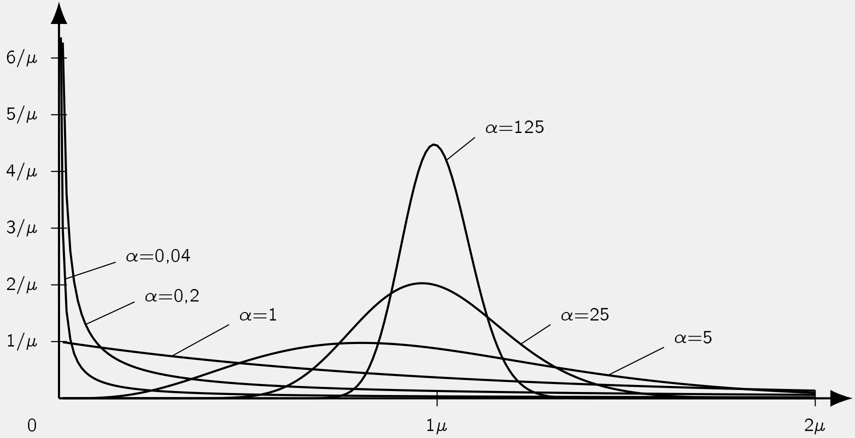
$\alpha = 1$	Exponentialverteilung, $g(x) = \frac{1}{\mu} \exp\left(-\frac{x}{\mu}\right)$, $\lim_{x \to 0} g(x) = 1/\mu$.
$\alpha < 1$	Die Dichte ist hyperbelartig, $\lim_{x \to 0} g(x) = \infty$.
$\alpha > 1$	Die Dichte ist unimodal mit Modus $\mu - \frac{\mu}{\alpha}$, $\lim_{x \to 0} g(x) = 0$.

Beispiel Kraftfahrt Hapftpflicht:

- Für $\alpha = 0{,}01$ gilt $\int_0^{\mu/200} g(x)\,dx \approx 0{,}91$.
- In KH liegt der Schadenbedarf bei etwa $E(R_i) \approx 200$ EUR, also $\frac{\mu}{200} \approx 1$ EUR.
- Sieht man dies als keinen Schaden an, so ist die Wahrscheinlichkeit für Schadenfreiheit also größer als 90%.

Gamma-Verteilung

Gamma-Dichten mit festem μ und verschiedenen Werten für α:



Faltung von Gamma-Verteilungen

Lemma: Es seien $X_1 \sim \Gamma(\mu_1, \alpha_1)$ und $X_2 \sim \Gamma(\mu_2, \alpha_2)$ unabhängig mit $\mu_1/\alpha_1 = \mu_2/\alpha_2$. Dann gilt

$$X_1 + X_2 \sim \Gamma(\mu_1 + \mu_2, \alpha_1 + \alpha_2).$$

Anpassung einer Gamma-Verteilung an die Momente von R_i gemäß Individuellem Modell ergibt $R_i \sim \text{Gamma}(\mu_i, \alpha_i)$ mit

$$\mu_i = \mathrm{E}(R_i) = m u_i \quad \text{und} \quad \alpha_i = \frac{(\mathrm{E}(R_i))^2}{\mathrm{Var}(R_i)} = \frac{m^2 u_i}{s^2}.$$

Zur späteren Validierung beachte, dass $\alpha_i \ll 1$ gelten sollte.

Gamma-Verteilung und Individuelles Modell

Korollar: *Seien* $R_1, \dots, R_I$ *unabhängig mit* $R_i \sim \Gamma(mu_i, m^2 u_i / s^2)$, *d.h.* R_i *Gamma-verteilt mit* $\mathrm{E}(R_i) = mu_i$ *und* $\mathrm{Var}(R_i) = s^2 u_i$. *Dann gilt für den Gesamtschaden* $S = R_1 + \cdots + R_I$

$$S \sim \Gamma\left(mv, \frac{m^2 v}{s^2}\right)$$

mit $v = u_1 + \cdots + u_I$. *Für den Schadensatz* $Z = S/v$ *erhalten wir*

$$Z \sim \Gamma\left(m, \frac{m^2 v}{s^2}\right)$$

mit dem gleichen Formarameter wie S.

Da Skalen- und Formparameter einer Gamma-Verteilung durch Erwartungswert und Varianz festgelegt sind, ist dieses Ergebnis nicht überraschend.

Gamma-Verteilung und Individuelles Modell

Bemerkung:

Im Folgenden benutzen wir die Parametrisierung $S/v = Z \sim \Gamma(\mu, v\alpha)$, d.h. $\mu := m$, $\alpha := m^2/s^2$, wobei das Volumen v bekannt ist. Wir nehmen an, dass in einer homogenen Risikogruppe die Parameter μ and α nach Inflationsbereinigung über mehrere Jahre konstant sind.

Bemerkung:

Es seien $Z_j \sim \Gamma(\mu, v_j\alpha)$, $j = 1, \ldots, J$ unabhängige Beobachtungen mit bekannten Volumina v_j. Dann haben wir folgenden Momentenschätzer für μ und α:

$$\widehat{\mu} = \frac{\sum_{j=1}^J v_j Z_j}{\sum_{j=1}^J v_j}, \quad \widehat{\alpha} = \frac{\widehat{\mu}^2 (J-1)}{\sum_{j=1}^J v_j (Z_j - \widehat{\mu})^2}.$$

Der Schätzer $\widehat{\mu}$ hat minimale Varianz unter den Konvexkombinationen der Z_j. Für $\widehat{\alpha}$ beachte $\alpha = m^2/s^2 = \mu^2/s^2$.

Inhalt

Methoden zur Schätzung von Parametern

Wir verwenden vier Methoden zur Schätzung von Parametern:

- Momentenmethode
- Maximum-Likelihood-Methode (ML-Methode)
- Minimum-Chi-Quadrat-Methode (asymptotisch äquivalent zu ML-Methode)
- Kleinste-Quadrate-Methode
  (Minimierung des quadrierten Abstands; äquivalent zur ML-Methode bei Normal-Verteilung)

Maximum Likelihood-Schätzer

Gegeben seien unabhängige Beobachtungen $X_1, X_2, \ldots$, deren Dichten $g_i(x \mid \vartheta)$ bezüglich dominierender Maße ν_i von dem gleichen unbekannten Parametervektor $\vartheta := (\vartheta_1, \ldots, \vartheta_K)^t \in \Theta$ (mit $\Theta \subseteq \mathbb{R}^K$ offen) abhängen.

Definition: Ein Schätzer $\widehat{\vartheta}_n = \widehat{\vartheta}_n(X_1, \ldots, X_n)$ für $\vartheta = (\vartheta_1, \ldots, \vartheta_K)^t$ heißt *Maximum Likelihood-Schätzer*, wenn er die *Likelihoodfunktion*

$$L_n(\vartheta) := \prod_{i=1}^{n} g_i(X_i \mid \vartheta)$$

maximiert. Dies ist natürlich äquivalent dazu, dass $\widehat{\vartheta}_n$ die *log-Likelihoodfunktion*

$$\ell_n(\vartheta) := \ln L_n(\vartheta) = \sum_{i=1}^{n} \ln g_i(X_i \mid \vartheta)$$

maximiert.

Beispiel

Seien $X_1, X_2, \dots \sim$ Binomial$(1, p)$ mit $p \in \Theta := (0, 1)$. Dann ist

$$L_n(p) = \prod_{i=1}^{n} p^{X_i}(1-p)^{1-X_i}, \text{ d.h. } \ell_n(p) = \sum_{i=1}^{n} X_i \log(p) + \sum_{i=1}^{n}(1-X_i)\log(1-p).$$

Somit

$$0 = \frac{d\ell_n}{dp}(p) = \frac{\sum_{i=1}^n X_i}{p} - \frac{n - \sum_{i=1}^n X_i}{1-p} \quad \Leftrightarrow \quad p = \frac{\sum_{i=1}^n X_i}{n}.$$

Falls $\sum_{i=1}^n X_i \in \{0, n\}$ so besitzt die log-Likelihoodfunktion kein Maximum in $\Theta = (0, 1)$, d.h. es gibt keinen Maximum Likelihood-Schätzer. Für $n \to \infty$ geht jedoch die Wahrscheinlichkeit für $\sum_{i=1}^n X_i \in \{0, n\}$ gegen 0, d.h. für große n hat die log-Likelihood mit großer Wahrscheinlichkeit ein Maximum.

Für derartige Fälle brauchen wir den Begriff des *asymptotischen ML-Schätzers*.

Asymptotische Maximum Likelihood-Schätzer

Definition:

Eine Folge $\widehat{\vartheta}_n = \widehat{\vartheta}_n(X_1, \ldots, X_n)$, $n = 1, 2, \ldots$, heißt *asymptotischer Maximum Likelihood-Schätzer* für ϑ, wenn für jedes $\vartheta \in \Theta$

$$\mathrm{P}_\vartheta \left(L_n(\widehat{\vartheta}_n) = \max_{\vartheta^* \in \Theta} L_n(\vartheta^*) \right) \to 1 \quad (n \to \infty)$$

gilt.

Im vorhergehenden Beispiel erhält man einen asymptotischen ML-Schätzer, wenn man

$$\widehat{p}_n := \begin{cases} \sum_{i=1}^n X_i / n & \text{falls } \sum_{i=1}^n X_i \notin \{0, n\}, \\ p_0 & \text{sonst} \end{cases}$$

mit einem $p_0 \in (0, 1)$ setzt.

Vertauschbarkeitsbedingung

Notation: Wir schreiben

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial \vartheta_1} \\ \vdots \\ \frac{\partial}{\partial \vartheta_K} \end{pmatrix}.$$

Dann ist ∇f der Gradient und $\nabla \nabla^t f$ die Hessematrix einer Funktion $f : \Theta \to \mathbb{R}$.

Voraussetzung: Die g_i sind zweimal stetig nach ϑ differenzierbar und für $f_n(x, \vartheta) := \prod_{i=1}^n g_i(x_i \mid \vartheta)$, $\omega_n := \nu_1 \otimes \cdots \otimes \nu_n$ sind die Bedingungen

$$\nabla \int_{\mathbb{R}^n} f_n(x \mid \vartheta)\, \omega_n(dx) = \int_{\mathbb{R}^n} \nabla f_n(x \mid \vartheta)\, \omega_n(dx) \qquad \text{und}$$

$$\nabla \nabla^t \int_{\mathbb{R}^n} f_n(x \mid \vartheta)\, \omega_n(dx) = \int_{\mathbb{R}^n} \nabla \nabla^t f_n(x \mid \vartheta)\, \omega_n(dx)$$

erfüllt.

Score-Vektor und Fisher-Information

Definition: Der Zufallsvektor $\nabla \ell_n(\vartheta)$ heißt *Score-Vektor*. Die $K \times K$-Matrix

$$I_n(\vartheta) := \mathrm{E}_\vartheta \left(\nabla \ell_n(\vartheta) \cdot \nabla^t \ell_n(\vartheta) \right) = \mathrm{E}_\vartheta \left(\frac{\partial}{\partial \vartheta_i} \ell_n(\vartheta) \cdot \frac{\partial}{\partial \vartheta_j} \ell_n(\vartheta) \right)_{i,j}$$

heißt *Fisher-Informationsmatrix*.

Lemma: *Es gilt*

$$\mathrm{E}_\vartheta(\nabla \ell_n(\vartheta)) = 0$$

und

$$I_n(\vartheta) = \mathrm{Cov}_\vartheta(\nabla \ell_n(\vartheta)) = -\mathrm{E}_\vartheta(\nabla \nabla^t \ell_n(\vartheta)).$$

$\sqrt{n}$-Konsistenz und Exponentialfamilien

Definition: Ein asymptotischer ML-Schätzer heißt $\sqrt{n}$*-konsistent* für ϑ, wenn die Folge

$$\sqrt{n}(\widehat{\vartheta}_n - \vartheta)$$

P_ϑ-stochastisch beschränkt ist (d.h. zu jedem $\varepsilon > 0$ gibt es ein $C_\varepsilon > 0$, so dass für alle $n \in \mathbb{N}$ gilt $P_\vartheta(\|\sqrt{n}(\widehat{\vartheta}_n - \vartheta)\| \leq C_\varepsilon) > 1 - \varepsilon$).

Definition: Bei einer Verteilungsfamilie mit Dichten der Form

$$g(x \mid \vartheta) = \exp(\vartheta^t t(x) - b(\vartheta)) \cdot h(x),$$

spricht man von einer *Exponentialfamilie*.

Viele der wichtigen Verteilungsfamilien lassen sich als Exponentialfamilien schreiben, z.B. Normalverteilung, Poisson-Verteilung, Gamma-Verteiung und Inverse Gauß-Verteilung.

Asymptotik für i.i.d.-Folgen bei Exponentialfamilien

Satz: *Die $X_1, X_2, \ldots$ seien i.i.d. mit Dichten der Form*

$$g_i(x \mid \vartheta) = g(x \mid \vartheta) = \exp(\vartheta^t t(x) - b(\vartheta)) \cdot h(x).$$

Hierbei seien $t\colon R \to \mathbb{R}^K$ messbar und $b\colon \Theta \to \mathbb{R}$ zweimal stetig differenzierbar mit positiv definiter Hessematrix $\nabla\nabla^t b(\vartheta)$.

Dann gibt es einen $\sqrt{n}$-konsistenten asymptotischen ML-Schätzer für ϑ. Für jeden asymptotischen ML-Schätzer gilt die asymptotische Normalität

$$\sqrt{n}(\widehat{\vartheta}_n - \vartheta) \xrightarrow{\mathcal{D}_\vartheta} \mathcal{N}_K(0, I_1(\vartheta)^{-1})$$

(beachte: $I_n(\vartheta) = nI_1(\vartheta)$).

Wir benötigen ein etwas allgemeineres Resultat.

Hauptsatz über ML-Schätzer

Satz: *Für $n \to \infty$ gelte*

$$\frac{1}{n}\sum_{i=1}^{n} \mathrm{E}_\vartheta \left(\|\nabla \ln g_i(X_i,\vartheta)\|^2 \cdot \mathbb{1}_{\{\|\nabla \ln g_i(X_i,\vartheta)\| > \varepsilon\sqrt{n}\}} \right) \to 0 \quad \text{für alle } \varepsilon > 0 \qquad \text{(ML1)}$$

und

$$\frac{1}{n} I_n(\vartheta) \to \Sigma(\vartheta) \qquad \text{(ML2)}$$

mit einer positiv definiten Matrix $\Sigma(\vartheta)$. Weiter gelte für jede Folge $\vartheta_n^* = \vartheta_n^*(X_1,\ldots,X_n)$ mit der Eigenschaft

$$\sqrt{n}(\vartheta_n^* - \vartheta),\ n \geq 1 \quad \mathrm{P}_\vartheta\text{-stochastisch beschränkt},$$

dass

$$\frac{1}{n}\left(\nabla\nabla^t \ell_n(\vartheta_n^*)\right) \xrightarrow{\mathrm{P}_\vartheta} -\Sigma(\vartheta). \qquad \text{(ML3)}$$

Dann gibt es einen $\sqrt{n}$-konsistenten asymptotischen ML-Schätzer für ϑ. Für jeden $\sqrt{n}$-konsistenten asymptotischen ML-Schätzer gilt die asymptotische Normalität

$$\sqrt{n}(\widehat{\vartheta}_n - \vartheta) \xrightarrow{\mathcal{D}_\vartheta} \mathcal{N}_K(0, \Sigma(\vartheta)^{-1}).$$

Hauptsatz über ML-Schätzer

Bemerkung:

- Die Bedingungen (ML1) und (ML2) kommen aus dem multivariaten zentralen Grenzwertsatz und stellen sicher, dass $\nabla \ell_n(\vartheta)$ asymptotisch normalverteilt ist.
- (ML1) ist beispielsweise erfüllt, wenn die X_i nur endlich viele unterschiedliche Dichten $g_i(x \mid \vartheta)$ haben.
- Die Bedingung (ML3) besagt, dass die realisierte Hessematrix $\nabla \nabla^t \ell_n(\vartheta)$ für große n mit großer Wahrscheinlichkeit ähnlich zur erwarteten Hessematrix sein muss.
- Bei Exponentialfamilien ist (ML3) automatisch erfüllt, da die Hessematrix deterministisch ist.

Bedeutung der Fisher-Information

Folgerung aus dem Hauptsatz über ML-Schätzer:

Für große n ist

$$n \cdot I_n(\vartheta)^{-1} \approx \Sigma(\vartheta)^{-1}.$$

Folglich gilt näherungsweise

$$\widehat{\vartheta}_n - \vartheta \sim \mathcal{N}_K\left(0, I_n(\vartheta)^{-1}\right)$$

und somit

$$\mathrm{Cov}(\widehat{\vartheta}_n) \approx I_n(\vartheta)^{-1}.$$

Beispiel

Sei $X_i \sim \mathcal{N}(\mu, \sigma^2)$ und $\vartheta = (\mu, \sigma^2)$. Dann ist

$$\ell_n(\mu, \sigma^2) = \sum_{i=1}^{n} \left(-\frac{1}{2} \ln(2\pi\sigma^2) - \frac{(X_i - \mu)^2}{2\sigma^2} \right).$$

Durch Nullsetzen der partiellen Ableitungen

$$\frac{\partial}{\partial \mu} \ell_n(\mu, \sigma^2) = \sum_{i=1}^{n} \frac{(X_i - \mu)}{\sigma^2} \quad \text{und} \quad \frac{\partial}{\partial \sigma^2} \ell_n(\mu, \sigma^2) = \sum_{i=1}^{n} \left(-\frac{1}{2\sigma^2} + \frac{(X_i - \mu)^2}{2\sigma^4} \right)$$

erhalten wir die ML-Schätzer

$$\widehat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad \text{(erwartungstreu),}$$

$$\widehat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \widehat{\mu}_n)^2 \qquad \text{(asymptotisch erwartungstreu).}$$

Beispiel

Es gilt

$$\frac{\partial^2}{\partial\mu^2}\ell_n(\mu,\sigma^2) = -\sum_{i=1}^{n}\frac{1}{\sigma^2} = -\frac{n}{\sigma^2},$$

$$\left(\frac{\partial}{\partial\sigma^2}\right)^2\ell_n(\mu,\sigma^2) = \sum_{i=1}^{n}\left(\frac{1}{2\sigma^4} - \frac{(X_i-\mu)^2}{\sigma^6}\right),$$

$$\frac{\partial^2}{\partial\mu\partial\sigma^2}\ell_n(\mu,\sigma^2) = -\sum_{i=1}^{n}\frac{(X_i-\mu)}{\sigma^4}$$

und wir erhalten

$$I_n(\mu,\sigma^2) = \mathrm{E}(-\nabla\nabla^t\ell_n(\mu,\sigma^2)) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}.$$

Beispiel

Somit also

$$I_n(\mu, \sigma^2)^{-1} = \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}.$$

Zum Vergleich die exakte Kovarianzmatrix:

$$\mathrm{Cov}((\widehat{\mu}, \widehat{\sigma}^2)^t) = \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2(n-1)\sigma^4}{n^2} \end{pmatrix}.$$

Transformationssatz

Lemma (Transformationssatz):

Sei $T : \Theta \to \mathbb{R}^L$ *stetig differenzierbar und bei* ϑ *regulär (d.h.* $dT/d\vartheta$ *hat bei* ϑ *Rang* L*). Es seien die Voraussetzungen des Hauptsatzes über ML-Schätzer erfüllt und* $\widehat{\vartheta}_n$ *ein* $\sqrt{n}$*-konsistenter asymptotischer ML-Schätzer für* ϑ*.*

Dann gilt näherungsweise

$$T(\widehat{\vartheta}_n) - T(\vartheta) \sim \mathcal{N}_K\left(0, \left(\frac{dT}{d\vartheta}(\vartheta)\right) I_n(\vartheta)^{-1}\left(\frac{dT}{d\vartheta}(\vartheta)\right)^t\right).$$

Beispiel

Seien jetzt $X_i \sim \mathcal{N}(\mu_i, \sigma_0^2)$ unabhängig mit bekanntem σ_0^2. Sei $\vartheta = (\mu_1, \ldots, \mu_n)$. Dann gilt

$$\frac{\partial}{\partial \mu_i} \ell_n(\mu_1, \ldots, \mu_n) = \frac{(X_i - \mu_i)}{\sigma_0^2},$$

sowie

$$\frac{\partial^2}{\partial \mu_i^2} \ell_n(\mu_1, \ldots, \mu_n) = -\frac{1}{\sigma_0^2} \quad \text{und} \quad \frac{\partial^2}{\partial \mu_i \partial \mu_j} \ell_n(\mu_1, \ldots, \mu_n) = 0$$

für $i \neq j$. Somit $\widehat{\vartheta}_n = (X_1, \ldots, X_n)$ und

$$I_n(\vartheta)^{-1} = \begin{pmatrix} \sigma_0^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_0^2 \end{pmatrix}.$$

Beispiel

Wir betrachten

$$T(\vartheta) = \frac{1}{n}(\mu_1 + \cdots + \mu_n),$$

und erhalten $T(\widehat{\vartheta}_n) = \frac{1}{n}(X_1 + \cdots + X_n)$. Wegen

$$\frac{\partial T}{\partial \mu_i} = \frac{1}{n}$$

erhalten wir aus dem Transformationssatz

$$\mathrm{Var}(T(\widehat{\vartheta}_n)) \approx \left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \begin{pmatrix} \sigma_0^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_0^2 \end{pmatrix} \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix} = \frac{\sigma_0^2}{n}$$

in Übereinstimmung mit dem exakten Wert $\mathrm{Var}(\widehat{\mu}) = \sigma_0^2/n$.

Likelihood-Quotiententest

Satz (Likelihood-Quotiententest): Es seien $M < K$, $\vartheta^0_{M+1}, \ldots, \vartheta^0_K$ fixiert,

$$\Delta := \{(\eta_1, \ldots, \eta_M) \in \mathbb{R}^M \mid (\eta_1, \ldots, \eta_M, \vartheta^0_{M+1}, \ldots, \vartheta^0_K) \in \Theta\}$$

und

$$\iota\colon \Delta \to \Theta, \quad (\eta_1, \ldots, \eta_M) \mapsto (\eta_1, \ldots, \eta_M, \vartheta^0_{M+1}, \ldots, \vartheta^0_K).$$

Sind dann $\widehat{\vartheta}_n$ und $\widehat{\eta}_n$ asymptotische ML-Schätzer für $\vartheta = \iota(\eta) \in \Theta$ bzw. $\eta \in \Delta$, so gilt unter den Voraussetzungen des Hauptsatzes über ML-Schätzer

$$T_n := 2 \cdot \ln\left(\frac{L_n(\widehat{\vartheta}_n)}{L_n(\iota(\widehat{\eta}_n))}\right) = 2 \cdot \left(\ell_n(\widehat{\vartheta}_n) - \ell_n(\iota(\widehat{\eta}_n))\right) \xrightarrow{\mathcal{D}_\vartheta} \chi^2_{K-M}.$$

Unter den Voraussetzungen des Satzes auf Folie 58 gilt die Aussage sogar für beliebige Immersionen $\iota\colon \Delta \to \Theta$.

Beispiel

Seien wieder $X_i \sim \mathcal{N}(\mu_i, \sigma_0^2)$ unabhängig mit bekanntem σ_0^2. Sei $\vartheta = (\mu_1, \dots, \mu_n)$. Wir betrachten die injektive Immersion

$$\iota \colon \mathbb{R} \to \mathbb{R}^n, \quad \mu \mapsto (\mu, \dots, \mu)^t.$$

Dann ist $\widehat{\vartheta}_n = (X_1, \dots, X_n)$ und $\widehat{\eta}_n = \widehat{\mu} := \frac{1}{n}(X_1 + \dots + X_n)$ und somit

$$\begin{aligned} 2 \cdot (\ell_n(\widehat{\vartheta}_n) - \ell_n(\iota(\widehat{\eta}_n))) &= 2 \cdot \sum_{i=1}^{n} \left(-\frac{1}{2} \ln(2\pi\sigma_0^2) - 0 + \frac{1}{2} \ln(2\pi\sigma_0^2) + \frac{(X_i - \widehat{\mu})^2}{2\sigma_0^2} \right) \\ &= \sum_{i=1}^{n} \frac{(X_i - \widehat{\mu})^2}{\sigma_0^2} \end{aligned}$$

annähernd χ^2_{n-1}-verteilt. Mit bekanntem σ_0^2 ergibt das einen Test auf Gleichheit der Erwartungswerte $\mu_1 = \dots = \mu_n$.

χ^2-Test

Es sei nun $\Theta := \{(p_1, \ldots, p_{K+1}) \in \mathbb{R}^{K+1} \mid p_i > 0, \sum_{i=1}^{K+1} p_i = 1\}$.

Beachte: Durch Weglassen der letzten Koordinate kann man Θ als offene Teilmenge von $\mathbb{R}^K$ auffassen. Mit $\mathcal{M}_{K+1}(n, p)$ bezeichnen wir die Multinomialverteilung mit den Parametern $n \in \mathbb{N}$ und $p \in \Theta$.

Seien $M < K$, $\Delta \subset \mathbb{R}^M$ offen und $p\colon \Delta \to \Theta$ eine zweimal stetig differenzierbare injektive Immersion. Sei $\eta \in \Delta$. Wir betrachten eine Folge $\mathcal{M}_{K+1}(n, p(\eta))$-verteilter Zufallsvektoren $X^{(n)} = (X_1^{(n)}, \ldots, X_{K+1}^{(n)}) \in \mathbb{N}_0^{K+1}$.

Satz (χ^2-Test): *Bezeichnet $\widehat{\eta}_n$ einen $\sqrt{n}$-konsistenten asymptotischen ML-Schätzer, so gilt für die Pearson-Fisher-Teststatistik:*

$$\sum_{i=1}^{K+1} \frac{(X_i^{(n)} - np_i(\widehat{\eta}_n))^2}{np_i(\widehat{\eta}_n)} \xrightarrow{\mathcal{D}_{p(\eta)}} \chi^2_{K-M}.$$

Minimum χ^2-Schätzer

Definition: In dieser Situation heißt $\widetilde{\eta}_n = \widetilde{\eta}_n(X^{(n)})$ *Minimum-χ^2-Schätzer*, wenn er die χ^2-Abstandsfunktion

$$Q_n(X^{(n)}, \eta) := \sum_{i=1}^{K+1} \frac{(X_i^{(n)} - np_i(\eta))^2}{np_i(\eta)}$$

minimiert, d.h. wenn

$$Q_n(X^{(n)}, \widetilde{\eta}_n) = \min_{\eta \in \Delta} Q_n(X^{(n)}, \eta)$$

gilt.

Satz: *Ein Minimum-χ^2-Schätzer $\widetilde{\eta}_n$ ist stets konsistent und mit jedem konsistenten asymptotischen ML-Schätzer $\widehat{\eta}_n$ asymptotisch äquivalent, d.h.*

$$\sqrt{n} \cdot (\widetilde{\eta}_n - \widehat{\eta}_n) \to 0 \quad \textit{nach } \mathrm{P}_\vartheta\textit{-Wkt.}$$

Kleinste-Quadrate-Schätzer

Lemma (KQ-Schätzer):

Seien jetzt $\mu : \Theta \to \mathbb{R}^n$, $\sigma_0^2 > 0$ *und* $X_1, \ldots, X_n \sim \mathcal{N}(\mu_i(\vartheta), \sigma_0^2)$.

Ein Schätzer $\widehat{\vartheta}$ *ist genau dann ein ML-Schätzer, wenn er ein Kleinste-Quadrate-Schätzer ist, d.h.*

$$\ell_n(\widehat{\vartheta}) = \max_{\vartheta \in \Theta} \ell_n(\vartheta) \quad \Longleftrightarrow \quad \sum_{i=1}^{n} (X_i - \mu_i(\widehat{\vartheta}))^2 = \min_{\vartheta \in \Theta} \sum_{i=1}^{n} (X_i - \mu_i(\vartheta))^2.$$