Angewandte Stochastik

Prof. Dr. Evgeny Spodarev | Vorlesungskurs | 7. Thema

Heutiges Thema

- Punktschätzer

Parametrisches Modell

- Sei $(x_1, \ldots, x_n)$ eine *konkrete* Stichprobe, d.h., eine Realisierung einer Zufallsstichprobe $(X_1, \ldots, X_n)$.
- $X_1, \ldots, X_n$ seien unabhängige identisch verteilte Zufallsvariablen mit der unbekannten Verteilungsfunktion F.
- F gehört zu einer bekannten parametrischen Familie $\{F_\theta : \theta \in \Theta\}$.

Parametrisches Modell

- $\theta = (\theta_1, \ldots, \theta_m) \in \Theta$ ist hier der *m-dimensionale Parametervektor* der Verteilung F_θ und $\Theta \subset \mathbb{R}^m$ der sogenannte *Parameterraum* (eine Borel-Teilmenge von $\mathbb{R}^m$, die die Menge aller zugelassenen Parameterwerte darstellt).
- *Voraussetzung:* Die Parametrisierung $\theta \to F_\theta$ ist *identifizierbar*, indem $F_{\theta_1} \neq F_{\theta_2}$ für $\theta_1 \neq \theta_2$ gilt.

Parametrisches Modell

- *Wichtige Aufgabe der Statistik:* Schätzung des Parametervektors θ (oder eines Teils von θ) an Hand von der konkreten Stichprobe $(x_1, \ldots, x_n)$.
- In diesem Fall spricht man von einem *Punktschätzer* $\hat{\theta} : \mathbb{R}^n \to \mathbb{R}^m$, der eine gültige Stichprobenfunktion ist.
- Meistens wird angenommen, dass

$$P\left(\hat{\theta}(X_1, \ldots, X_n) \in \Theta\right) = 1\,,$$

  wobei es zu dieser Regel auch Ausnahmen gibt.
- Um zu betonen, dass P vom Parameter θ abhängt, werden wir Bezeichnungen

$$P_\theta, \mathrm{E}_\theta \text{ und } Var_\theta$$

  für das Maß P, den Erwartungswert und die Varianz bzgl. P verwenden.

Beispiel

1.
   - Sei X die Dauer des fehlerfreien Arbeitszyklus eines technischen Systems.
   - Oft wird $X \sim Exp(\lambda)$ angenommen.
   - Dann stellt $\{F_\theta : \theta \in \Theta\}$ mit $m = 1$, $\theta = \lambda$, $\Theta = \mathbb{R}_+$ und
$$F_\theta(x) = (1 - e^{-\theta x})I(x \geq 0)$$
ein parametrisches Modell dar.
   - Der Parameterraum ist eindimensional.
   - Später wird für λ der (Punkt-) Schätzer $\hat{\lambda}(x_1, \ldots, x_n) = \frac{1}{\bar{x}_n}$ vorgeschlagen.

Beispiel

2.

- In den Fragestellungen der statistischen Qualitätskontrolle werden n Erzeugnisse auf Mängel untersucht.
- Falls $p \in (0, 1)$ die unbekannte Wahrscheinlichkeit des Mangels ist, so wird mit $X \sim Bin(n, p)$ die Gesamtanzahl der mangelhaften Produkte beschrieben.

Beispiel

- Dabei wird folgendes parametrische Modell unterstellt:

$$\Theta = \{(n, p) : \ n \in \mathbb{N}, \, p \in (0, 1)\} \, , \theta = (n, p) \, , m = 2 \, ,$$

$$F_\theta(x) = P_\theta(X \leq x) = \begin{cases} 1, & x > n \\ \sum_{k=0}^{[x]} \binom{n}{k} p^k (1-p)^{n-k}, & x \in [0, n] \\ 0, & x < 0. \end{cases}$$

- Falls n bekannt ist, kann die Wahrscheinlichkeit p des Ausschusses durch den Punktschätzer $\hat{p}(x_1, \ldots, x_n) = \bar{x}_n \, , x_i \in \{0, 1\}$ näherungsweise berechnet werden.

Eigenschaften von Punktschätzern

- Sei $(X_1, \ldots, X_n)$ eine Zufallsstichprobe, definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P_\theta)$.
- Seien X_i, $i = 1, \ldots, n$ unabhängige identisch verteilte Zufallsvariablen mit Verteilungsfunktion $F \in \{F_\theta : \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^m$.
- **Ziel:** Finde einen Schätzer $\hat{\theta}(X_1, \ldots, X_n)$ für den Parameter θ mit den vorgegebenen Eigenschaften.

Definition (Erwartungstreue)

Ein Schätzer $\hat{\theta}(X_1, \ldots, X_n)$ für θ heißt *erwartungstreu* oder *unverzerrt*, falls

$$\mathbb{E}_\theta \hat{\theta}(X_1, \ldots, X_n) = \theta, \quad \theta \in \Theta\,.$$

Dabei wird vorausgesetzt, dass

$$\mathsf{E}_\theta |\hat{\theta}(X_1, \ldots, X_n)| < \infty, \quad \theta \in \Theta\,.$$

Der *Bias* (*Verzerrung*) eines Schätzers $\hat{\theta}(X_1, \ldots, X_n)$ ist gegeben durch

$$Bias(\hat{\theta}) = \mathsf{E}_\theta\, \hat{\theta}(X_1, \ldots, X_n) - \theta\,.$$

Falls $\hat{\theta}(X_1, \ldots, X_n)$ erwartungstreu ist, dann gilt $Bias(\hat{\theta}) = 0$ (kein systematischer Schätzfehler).

Definition (Asymptotische Erwartungstreue)

Der Schätzer $\hat{\theta}(X_1, \ldots, X_n)$ für θ heißt *asymptotisch erwartungstreu* (oder *asymptotisch unverzerrt*), falls (für große Datenmengen)

$$\mathsf{E}_\theta \, \hat{\theta}(X_1, \ldots, X_n) \underset{n\to\infty}{\longrightarrow} \theta, \quad \theta \in \Theta.$$

Definition (Konsistenz)

Falls

$$\hat{\theta}(X_1, \ldots, X_n) \underset{n\to\infty}{\longrightarrow} \theta, \quad \theta \in \Theta$$

in L^2, stochastisch bzw. fast sicher, dann heißt der Schätzer $\hat{\theta}(X_1, \ldots, X_n)$ ein *konsistenter Schätzer* für θ im *mittleren quadratischen, schwachen bzw. starken Sinne*.

- $\hat{\theta}$ *L^2-konsistent*: für $\mathbb{E}_\theta \, |\hat{\theta}(X_1, \ldots, X_n)|^2 < \infty$ gilt

$$\hat{\theta} \underset{n\to\infty}{\overset{L^2}{\longrightarrow}} \theta \iff \mathsf{E}_\theta |\hat{\theta}(X_1, \ldots, X_n) - \theta|^2 \underset{n\to\infty}{\longrightarrow} 0, \quad \theta \in \Theta.$$

Definition (Konsistenz)

- $\hat{\theta}$ *schwach konsistent* :

$$\hat{\theta} \xrightarrow[n\to\infty]{P} \theta \iff P_\theta(|\hat{\theta}(X_1,\dots,X_n)-\theta| > \varepsilon) \xrightarrow[n\to\infty]{} 0, \quad \varepsilon > 0, \theta \in \Theta.$$

- $\hat{\theta}$ *stark konsistent* :

$$\hat{\theta} \xrightarrow[n\to\infty]{\text{f.s.}} \theta \iff P_\theta\left(\lim_{n\to\infty} \hat{\theta}(X_1,\dots,X_n) = \theta\right) = 1, \quad \theta \in \Theta.$$

Daraus ergibt sich folgendes Diagramm:

$$\boxed{L^2 - \text{Konsistenz}} \Longrightarrow \boxed{\text{schwache Konsistenz}} \Longleftarrow \boxed{\text{starke Konsistenz}}$$

Definition (Mittlerer quadratischer Fehler (mean squared error))

Der mittlere quadratische Fehler eines Schätzers $\hat{\theta}(X_1, \dots, X_n)$ für θ ist definiert als

$$MSE(\hat{\theta}) = \mathsf{E}_\theta \left| \hat{\theta}(X_1, \dots, X_n) - \theta \right|^2 .$$

Lemma

Falls $m = 1$ und $\mathrm{E}_\theta\, \hat{\theta}^2(X_1, \ldots, X_n) < \infty\,, \quad \theta \in \Theta$, dann gilt

$$MSE(\hat{\theta}) = \mathrm{Var}_{\,\theta}\, \hat{\theta} + \left(Bias(\hat{\theta})\right)^2 .$$

Falls $\hat{\theta}$ *erwartungstreu* für θ ist, dann gilt $MSE(\hat{\theta}) = \mathrm{Var}_{\,\theta}\, \hat{\theta}$.

Definition (Vergleich von Schätzern)

Seien für θ zwei Schätzer durch $\hat{\theta}_1(X_1, \ldots, X_n)$ und $\hat{\theta}_2(X_1, \ldots, X_n)$ definiert. Man sagt, dass $\hat{\theta}_1$ *besser* ist als $\hat{\theta}_2$, falls

$$MSE(\hat{\theta}_1) < MSE(\hat{\theta}_2) \quad , \theta \in \Theta.$$

Falls $m = 1$ und die Schätzer $\hat{\theta}_1$, $\hat{\theta}_2$ erwartungstreu sind, so ist $\hat{\theta}_1$ *besser als* $\hat{\theta}_2$, falls $\hat{\theta}_1$ die kleinere Varianz besitzt. Dabei wird stets vorausgesetzt, dass $\mathrm{E}_\theta \, \hat{\theta}_i^2 < \infty \, , \quad \theta \in \Theta$.

Definition (Asymptotische Normalverteiltheit)

Sei $\hat{\theta}(X_1, \ldots, X_n)$ ein Schätzer für θ ($m = 1$). Falls
$0 < \operatorname{Var}_\theta \hat{\theta}(X_1, \ldots, X_n) < \infty\,, \quad \theta \in \Theta$ und

$$\frac{\hat{\theta}(X_1, \ldots, X_n) - \mathrm{E}_\theta\, \hat{\theta}(X_1, \ldots, X_n)}{\sqrt{\operatorname{Var}_\theta \hat{\theta}(X_1, \ldots, X_n)}} \xrightarrow[n \to \infty]{d} Y \sim N(0, 1)\,,$$

dann ist $\hat{\theta}(X_1, \ldots, X_n)$ *asymptotisch normalverteilt*.

Definition (Bester erwartungstreuer Schätzer)

Der Schätzer $\hat{\theta}(X_1, \ldots, X_n)$ für θ ist der *beste erwartungstreue Schätzer* , falls

$$\mathrm{E}_\theta\, \hat{\theta}^2(X_1, \ldots, X_n) < \infty\,, \quad \theta \in \Theta\,, \qquad \mathrm{E}_\theta\, \hat{\theta}(X_1, \ldots, X_n) = \theta\,, \quad \theta \in \Theta\,,$$

und $\hat{\theta}$ die minimale Varianz in der Klasse aller erwartungstreuen Schätzer für θ besitzt.
Das heißt, dass für einen beliebigen erwartungstreuen Schätzer $\tilde{\theta}(X_1, \ldots, X_n)$ mit

$$\mathrm{E}_\theta\, \tilde{\theta}^2(X_1, \ldots, X_n) < \infty \quad \text{gilt} \quad \mathrm{Var}_{\,\theta}\, \hat{\theta} \leq \mathrm{Var}_{\,\theta}\, \tilde{\theta}\,, \qquad \theta \in \Theta\,.$$

Empirische Momente

- Sei $X \stackrel{d}{=} X_i\,, \quad i = 1, \ldots, n$ ein statistisches Merkmal.
- Sei weiter $\mathrm{E}|X_i|^k < \infty$ für ein $k \in \mathbb{N}$, $m = 1$ und der zu schätzende Parameter $\theta = \mu_k = \mathrm{E}X_i^k$.
- Insbesondere gilt im Fall $k = 1$, dass $\theta = \mu_1 = \mu$ der Erwartungswert ist.

Definition

Das *k-te empirische Moment* von X wird als

$$\hat{\mu}_k = \frac{1}{n} \sum_{i=1}^{n} X_i^k$$

definiert. Unter dieser Definition gilt, dass $\hat{\mu}_1 = \bar{X}_n$, also das erste empirische Moment gleich dem Stichprobenmittel ist.

Satz (Eigenschaften der empirischen Momente)

Unter obigen Voraussetzungen gelten folgende Eigenschaften:

1. $\hat{\mu}_k$ ist erwartungstreu für μ_k (insbesondere $\bar{X}_n$).
2. $\hat{\mu}_k$ ist stark konsistent.
3. Falls $E_\theta |X|^{2k} < \infty\,, \quad \forall \theta \in \Theta$, dann ist $\hat{\mu}_k$ asymptotisch normalverteilt.
4. Es gilt Var $\bar{X}_n = \frac{\sigma^2}{n}$, wobei $\sigma^2 =$ Var $_\theta X$. Falls $X_i \sim N(\mu, \sigma^2)\,, \quad i = 1, \ldots, n$ (eine normalverteilte Stichprobe), dann gilt:

$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right) .$$