5 Punktschätzung

Zoltán Zomotor

Versionsstand: 5. Februar 2015, 13:50

Die nummerierten Felder bitte während der Vorlesung ausfüllen.

Bitte hier notieren, was beim Bearbeiten unklar geblieben ist:

Inhaltsverzeichnis

1 Einleitung 1

2 Grundbegriffe 2

3 Maximum-Likelihood-Prinzip 5

4 Übungsaufgaben 8

1 Einleitung

- Unbekannter Parameter 1 ϑ soll auf Basis einer Stichprobe geschätzt werden.

- Schätzwert: 2 $\hat{\vartheta}$

- Vorgehen: Verwendung einer *Schätzfunktion* 3 $\hat{\Theta} = g(X_1, \ldots, X_n)$

$\Rightarrow$ Schätzwert 4 $\hat{\vartheta}$ ist Realisierung der Zufallsvariablen 5 $\hat{\Theta}$

Im Folgenden werden wir Kriterien kennenlernen, um Schätzfunktionen zu beurteilen und zu konstruieren. Wir gehen immer, falls nichts anderes gesagt wird, von einer einfachen Stichprobe aus, das heißt $X_1, \ldots, X_n$ sind i.i.d.

2 Grundbegriffe [Pap11, Teil III, Kapitel 3.2]

Erwartungstreue

- Eine Schätzfunktion

  6 $\hat{\Theta} = g(X_1, \ldots, X_n)$

  heißt *erwartungstreu* oder *unverzerrt* für ϑ, falls:

  7 $\mathrm{E}(\hat{\Theta}) = \vartheta$

- Falls nur gilt 8 $\lim_{n\to\infty} \mathrm{E}(\hat{\Theta}) = \vartheta$

  so heißt $\hat{\Theta}$ *asymptotisch erwartungstreu* für ϑ.

- **Beispiel** Wir untersuchen folgende Funktionen auf Erwartungstreue:

$$\hat{\Theta}_1 = \bar{X}, \quad \hat{\Theta}_2 = \frac{X_1 + X_n}{2}, \quad \hat{\Theta}_3 = \frac{1}{n-1}\sum_{i=1}^{n} X_i$$

1. 9 $\mathrm{E}(\hat{\Theta}_1) = \mathrm{E}(\bar{X}) = \mu \Rightarrow$ erwartungstreu
2. 10 $\mathrm{E}(\hat{\Theta}_2) = \mathrm{E}(\frac{X_1 + X_n}{2}) = \frac{1}{2}(\mathrm{E}(X_1) + \mathrm{E}(X_2)) = \frac{1}{2}(\mu + \mu) = \mu$
   $\Rightarrow$ erwartungstreu

11

3. $$E(\hat{\Theta}_3) = E(\frac{1}{n-1}\sum_{i=1}^{n} X_i) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i) = \frac{n}{n-1}\cdot\mu \neq \mu$$
$\Rightarrow$ nicht erwartungstreu, aber wegen
$$\lim_{n\to\infty}(\frac{n}{n-1}\cdot\mu) = \mu$$
asymptotisch erwartungstreu.

Wirksamkeit

- Welche der erwartungstreuen Schätzfunktionen ist „besser“?
- Von zwei erwartungstreuen Schätzfunktionen $\hat{\Theta}_1$ und $\hat{\Theta}_2$ für ϑ heißt $\hat{\Theta}_1$ *wirksamer* als $\hat{\Theta}_2$, falls:

12

$\mathrm{Var}(\hat{\Theta}_1) < \mathrm{Var}(\hat{\Theta}_2)$

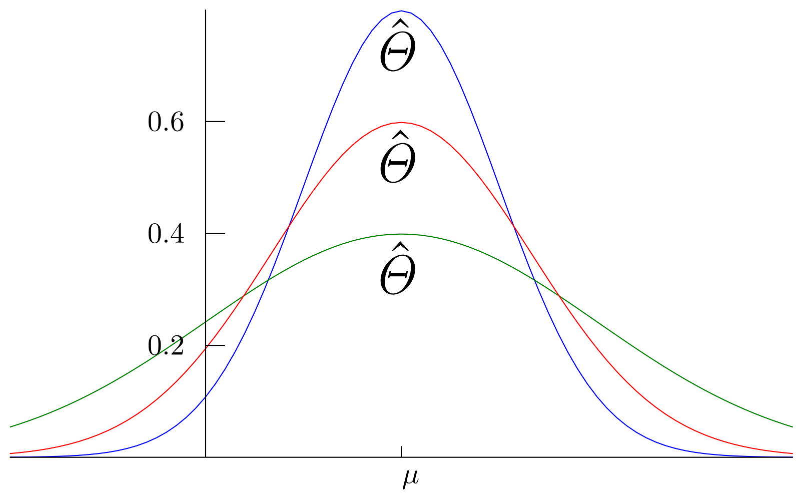
- **Beispiel:**

13

$$\mathrm{Var}(\hat{\Theta}_1) = \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \qquad \mathrm{Var}(\hat{\Theta}_3) = \frac{n}{n-1}\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n-1}$$
$$\mathrm{Var}(\hat{\Theta}_2) = \mathrm{Var}(\frac{X_1+X_n}{2}) = \frac{1}{4}(\sigma^2+\sigma^2) = \frac{\sigma^2}{2}$$

$\Rightarrow$ Falls $n > 2$ ist: $\mathrm{Var}(\hat{\Theta}_1) < \mathrm{Var}(\hat{\Theta}_2)$
$\Rightarrow$ $\hat{\Theta}_1$ wirksamer als $\hat{\Theta}_2$

Erwartungstreue und Wirksamkeit



14

Verteilung von G	ϑ	wirksamste erwartungstreue Schätzfunktion
unbekannt	μ	$\bar{X}$
$B(1, p)$	$p\ (= \mu)$	$\bar{X}$
$N(\mu, \sigma)$, σ bekannt / unbekannt	μ	$\bar{X}$
$N(\mu, \sigma)$, μ bekannt	σ^2	$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$
$N(\mu, \sigma)$, μ unbekannt	σ^2	S^2

Konsistente Schätzfunktionen

- Eine Folge von Schätzfunktionen $\hat{\Theta}_1, \hat{\Theta}_2, \ldots, \hat{\Theta}_n$ gemäß

  15

  $$\begin{aligned} \hat{\Theta}_1 &= g_1(X_1) \\ \hat{\Theta}_2 &= g_1(X_1, X_2) \\ &\vdots \\ \hat{\Theta}_3 &= g_1(X_1, \ldots, X_n) \end{aligned}$$

  heißt *konsistent* für ϑ, wenn für alle $c > 0$ gilt:

  16

  $$P\left(\left|\hat{\Theta}_n - \vartheta\right| \geq c\right) \xrightarrow[n\to\infty]{} 0$$

  Das heißt, die Wahrscheinlichkeit, ϑ deutlich zu verfehlen, geht gegen 0.
- Mit der Tschebyscheff-Ungleichung

  $$P\left(\left|X - \mathrm{E}(X)\right| \geq c\right) \leq \frac{\mathrm{Var}(X)}{c^2}$$

  ergibt sich die hinreichende Konsistenzbedingung:

  17

  $$\mathrm{E}(\hat{\Theta}_n) = \vartheta \text{ und } \lim_{n\to\infty} \mathrm{Var}(\hat{\Theta}_n) = 0$$

- Beispiel: Ist $\bar{X}_n$ konsistent für μ?

  18

  $$\begin{aligned} \mathrm{E}(\bar{X}_n) &= \mu \\ \mathrm{Var}(\bar{X}_n) &= \frac{\sigma^2}{n} \xrightarrow[n\to\infty]{} 0 \Rightarrow \text{konsistent} \end{aligned}$$

3 Maximum-Likelihood-Prinzip [Pap11, Teil III, Kapitel 3.3]

Gegeben:

- Ergebnis einer einfachen Stichprobe 19 $(x_1, \ldots, x_n)$
- Likelihoodfunktion 20 $f(x_1, \ldots, x_n|\vartheta)$

Beispiel:

- $G \sim B(1,p) \quad \Rightarrow \quad f_i(x_i) =$ 21 $p^{x_i}(1-p)^{1-x_i}$
- p ist unbekannt
- Einfache Stichprobe mit $n = 2$

$\Rightarrow$ Likelihoodfunktion:

22

$$\begin{aligned} f(x_1, x_2|p) &= f_1(x_1) \cdot f_2(x_2) \quad \text{(wegen Unabhängigkeit)} \\ &= p^{x_1}(1-p)^{1-x_1} \cdot p^{x_2}(1-p)^{1-x_2} \\ &= p^{x_1+x_2}(1-p)^{2-x_1-x_2} \end{aligned}$$

- Stichprobenergebnis $(0,1) \Rightarrow$ 23 $f(0,1|p) = p(1-p) = p - p^2$

Gesucht:

Schätzwert $\hat{\vartheta}$, der „am besten zu $(x_1, \ldots, x_n)$ passt“.

Maximum-Likelihood-Prinzip (ML-Prinzip):

- Wähle $\hat{\vartheta}$ so, dass für alle möglichen ϑ-Werte gilt:

  24 $f(x_1, \ldots, x_n|\hat{\vartheta}) \geq f(x_1, \ldots, x_n|\vartheta)$
- Maximierung: 1. Ableitung zu Null setzen und 2. Ableitung < 0 prüfen.
- Maximierung für:
  - konkretes Stichprobenergebnis (zum Beispiel $(0,1)$)

    25 $\Rightarrow$ ML-Schätzwert

- allgemeines Stichprobenergebnis (zum Beispiel (x_1, x_2))

26 $\Rightarrow$ ML-Schätzfunktion

- Maximierung der 27 logarithmierten Likelihoodfunktion liefert das selbe Ergebnis, ist aber oft einfacher:

28 $\ln f(x_1, \ldots, x_n|\hat{\vartheta}) \geq \ln f(x_1, \ldots, x_n|\vartheta)$

Das ist möglich, weil 29 $\ln x$ streng monoton mit x wächst.

Typische Vorgehensweise bei der ML-Schätzung

1. Likelihoodfunktion aufstellen: 30 $f(x_1, \ldots, x_n|\vartheta)$
2. gegebenenfalls 31 logarithmieren: $\ln f(x_1, \ldots, x_n|\vartheta)$
3. Erste Ableitung nullsetzen: 32 $\frac{\partial}{\partial \vartheta} [\ln] f(x_1, \ldots, x_n|\vartheta) \stackrel{!}{=} 0$
4. Vorzeichen der zweiten Ableitung prüfen: 33 $\frac{\partial^2}{\partial \vartheta^2} [\ln] f(x_1, \ldots, x_n|\vartheta) \stackrel{?}{<} 0$

Achtung: Lösung *meist* per Ableitung, es gibt aber Ausnahmen!

Beispiel Seite 5: $f(x_1, x_2|p) = p^{x_1+x_2}(1-p)^{2-x_1-x_2}$ bzw. $f(0,1|p) = p - p^2$

- Konkreter Schätzwert:

34

$$\frac{\partial}{\partial p} f(0,1|p) = 1 - 2p \stackrel{!}{=} 0 \Rightarrow \hat{p} = \frac{1}{2}$$

$$\frac{\partial^2}{\partial p^2} f(0,1|p) = -2 < 0 \Rightarrow \hat{p} = \frac{1}{2} \text{ist ML-Schätzwert}$$

- Schätzfunktion: Logarithmieren sinnvoll

– Logarithmieren:

35

$$\ln f(x_1, x_2|p) = (x_1 + x_2)\ln(p) + (2 - x_1 - x_2)\ln(1-p)$$

– Ableiten und nullsetzen:

36

$$\begin{aligned} \frac{\partial}{\partial p} \ln f(x_1, x_2|p) &= \frac{x_1 + x_2}{p} - \frac{2 - x_1 - x_2}{1-p} \overset{!}{=} 0 \\ &\Leftrightarrow (x_1 + x_2)(1-p) = (2 - x_1 - x_2)p \\ &\Leftrightarrow x_1 + x_2 = 2p \Rightarrow \hat{p} = \frac{x_1 + x_2}{2} \end{aligned}$$

– 2. Ableitung prüfen:

37

$$\frac{\partial^2}{\partial p^2} \ln f(x_1, x_2|p) = -\frac{x_1 + x_2}{p^2} - \frac{2 - x_1 - x_2}{(1-p)^2} < 0$$

$$\Rightarrow \hat{p} = \frac{1}{2}(x_1 + x_2) = \bar{x} \text{ ist ML-Schätzfunktion}$$

Maximum-Likelihood-Schätzfunktionen

38

Verteilung von G	ϑ	ML-Schätzfunktion
$B(1, p)$	$p\ (= \mu)$	$\bar{X}$
$Exp(\lambda)$	μ	$\bar{X}$
$Exp(\lambda)$	σ^2	$\bar{X}^2$
$P(\mu)$	μ	$\bar{X}$
$N(\mu, \sigma)$, σ bekannt / unbekannt	μ	$\bar{X}$
$N(\mu, \sigma)$, μ bekannt	σ^2	$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$
$N(\mu, \sigma)$, μ unbekannt	σ^2	$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$

Hilfreicher Satz für ML-Schätzung:

> Ist h eine *streng monotone* Funktion und ist $\hat{\Theta}$ eine Maximum-Likelihood-Schätzfunktion für den Parameter ϑ, so ist die Stichprobenfunktion $h(\hat{\Theta})$ eine Maximum-Likelihood-Schätzfunktion für den transformierten Parameter $h(\vartheta)$.

- Hinweis: h kann streng monoton wachsend oder fallend sein.
- Beispiele:
  - $\hat{\Theta}$ ML-Schätzfunktion für $\sigma^2 \Rightarrow$ 39 $\sqrt{\hat{\Theta}}$ ML-Schätzfunktion für σ
  - Ist $G \sim Exp(\lambda)$, so ist $\frac{1}{\bar{X}}$ ML-Schätzfunktion für λ. Grund:

40

$\lambda = \frac{1}{\mu} = h(\mu)$ mit h streng monoton fallend

$\Rightarrow \hat{\lambda} = \frac{1}{\hat{\mu}} = \frac{1}{\bar{X}}$

4 Übungsaufgaben

Aufgabenblatt und [Pap11, Seite 641f: Zu Abschnitt 2 alle Aufgaben, zu Abschnitt 3 Aufgaben 1),2),3)].

Literatur

[Pap11] Lothar Papula. *Mathematik für Ingenieure und Naturwissenschaftler*. 6. Auflage. Bd. 3. Vieweg + Teubner, 2011.