3 Schätzer und ihre Eigenschaften

Es seien $(\mathfrak{X}, \mathcal{B}, \{P_\vartheta : \vartheta \in \Theta\})$ ein statistischer Raum, $\gamma : \Theta \to \Gamma$ ein Funktional, wobei $\Gamma \supset \gamma(\Theta)$, A_Γ eine σ-Algebra auf Γ.

3.1 Definition

Ein **Schätzer** für $\gamma(\vartheta)$ ist eine messbare Abbildung $S : (\mathfrak{X}, \mathcal{B}) \to (\Gamma, A_\Gamma)$.
$S(x)$ heißt **Schätzwert** für $\gamma(\vartheta)$ zur Beobachtung $x \in \mathfrak{X}$.

3.2 Beispiel

$(\mathfrak{X}, \mathcal{B}) = (\{0,1\}^n, \mathcal{P}(\{0,1\}^n))$, $\Theta = (0,1)$, $P_\vartheta = \bigotimes_{j=1}^n \mathrm{Bin}(1,\vartheta)$

$$\gamma(\vartheta) = \vartheta$$

$$S(x_1, \ldots, x_n) := \frac{1}{n} \sum_{j=1}^n X_j$$

(relative Trefferhäufigkeit)

$$\Gamma = [0,1] = \bar{\Theta},\ A_\Gamma = \mathcal{B}^1 \cap [0,1]$$

[Beachte: $\gamma(\Theta) = \Theta \subset \Gamma$]

Die Güte eines Schätzers wird über die Verteilung $P_\vartheta^{S(X)}$ von $S(X)$ unter ϑ beurteilt. Für jedes $\vartheta \in \Theta$ sollte $P_\vartheta^{S(x)}$ „stark um $\gamma(\vartheta)$ konzentriert“ sein.

3.3 Definition (Sei $\Gamma \subset \mathbb{R}^k$.)

a) S **erwartungstreu** (unbiased) für $\gamma(\vartheta)$ $:\Leftrightarrow$ $E_\vartheta S(X) = \gamma(\vartheta)\ \forall \vartheta \in \Theta$

b) $b_S(\vartheta) := E_\vartheta S(X) - \gamma(\vartheta)$ heißt Verzerrung (bias) von S an der Stelle ϑ.

c) Ist $S_n = S_n(X_1, \ldots, X_n)$, $n \geq 1$ eine Schätzfolge, so heißt (S_n) **asymptotisch erwartungstreu** für $\gamma(\vartheta)$ $:\Leftrightarrow$

$$\lim_{n\to\infty} E_\vartheta S_n = \gamma(\vartheta)\ \forall \vartheta \in \Theta$$

<u>Erwartungstreue:</u> $\forall \vartheta \in \Theta$: Schwerpunkt von $P_\vartheta^{S(X)}$ ist $\gamma(\vartheta)$

3.4 **Definition** (Sei $\Gamma \subset \mathbb{R}$.)

S mediantreu für $\gamma(\vartheta)$ $:\Leftrightarrow$ $\operatorname{med}_\vartheta S(X) = \gamma(\vartheta)\ \forall \vartheta \in \Theta$.

Dabei:
Sei Y Zufallsvariable mit Verteilungsfunktion F.

$$F^{-1}(q) := \inf\{x \in \mathbb{R} :\ F(x) \geq q\},\ 0 < q < 1$$

$$\operatorname{med} Y := \operatorname{med} F := \frac{1}{2}(F^{-1}(\frac{1}{2}) + \underbrace{F^{-1}(\frac{1}{2} + 0)}_{=\lim_{x \to \frac{1}{2}+} F^{-1}(x)})$$

$\to$ Median[8]

<u>Mediantreue:</u> $\forall \vartheta \in \Theta$:

$$P_\vartheta(S(X) \leq \gamma(\vartheta)) = P_\vartheta(S(X) \geq \gamma(\vartheta)) \geq \frac{1}{2}$$

(In jeweils 50% der Fälle Unter- bzw. Überschätzung.)

<u>Beispiele:</u>

a) $X_1, \ldots, X_n$ reellwertig, $X_1, \ldots, X_n \overset{uiv}{\sim} P_\vartheta$, $\mu(\vartheta) := E_\vartheta X_1$ $(E_\vartheta |X_1| < \infty)$.

   $\bar{X}_n$ ist erwartungstreu für $\mu(\vartheta)$
   X_1 ist erwartungstreu für $\mu(\vartheta)$

b) Wie a), $E_\vartheta X_1^2 < \infty$. $\sigma^2(\vartheta) := \operatorname{Var}_\vartheta(X_1)$.

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

   S_n^2 ist erwartungstreu für $\sigma^2(\vartheta)$.

[8] $\to$ Abbildung 3.1

Beweis:

$$\begin{aligned}
E_\vartheta S_n^2 &= \frac{1}{n-1} E_\vartheta \left[\sum_{i=1}^n ((X_i - \mu(\vartheta)) - (\bar{X}_n - \mu(\vartheta)))^2 \right] \\
&= \frac{1}{n-1} \left[\sum_{i=1}^n \underbrace{E_\vartheta (X_i - \mu(\vartheta))^2}_{=\operatorname{Var}_\vartheta(X_i)} - n \underbrace{E_\vartheta (\bar{X}_n - \mu(\vartheta))^2}_{=\operatorname{Var}_\vartheta(\bar{X}_n) = \frac{\sigma^2(\vartheta)}{n}} \right] \\
&= \frac{1}{n-1} (n\sigma^2(\vartheta) - \sigma^2(\vartheta)) \\
&= \sigma^2(\vartheta)
\end{aligned}$$

c) $X_1, \ldots, X_n \overset{uiv}{\sim} \mathcal{N}(\mu, \sigma^2)$, $\vartheta = (\mu, \sigma^2)$, $\gamma(\vartheta) = E_\vartheta X_1 = \mu$

$$\bar{X}_n \sim \mathcal{N}(\mu, \frac{\sigma^2}{n}) \;\Rightarrow\; \operatorname{med}_\vartheta \bar{X}_n = \mu$$

$\Rightarrow \bar{X}_n$ ist mediantreu für $\gamma(\vartheta)$.

3.5 **Definition** (Sei $\Gamma \subset \mathbb{R}^k$.)

Schätzfolge $S_n = S_n(X_1, \ldots, X_n)$, $n \geq 1$, heißt (schwach) **konsistent** für $\gamma(\vartheta)$ $:\Leftrightarrow$

$$P_\vartheta(\|S_n(X_1, \ldots, X_n) - \gamma(\vartheta)\| \geq \varepsilon) \to 0 \quad \forall \varepsilon > 0\ \forall \vartheta \in \Theta$$

$(\forall \vartheta \in \Theta: \; S_n \overset{P_\vartheta}{\to} \gamma(\vartheta))$

3.6 **Bemerkung** (Sei $\Gamma \subset \mathbb{R}$.)

(S_n) asymptotisch erwartungstreu für $\gamma(\vartheta)$ und $\operatorname{Var}_\vartheta S_n \to 0 (n \to \infty) \Rightarrow$ (S_n) konsistent für $\gamma(\vartheta)$.

[Beweis:]

$$\begin{aligned}
P_\vartheta(|S_n - \gamma| \geq \varepsilon) &\leq P_\vartheta(|S_n - ES_n| + |ES_n - \gamma| \geq \varepsilon) \\
&\leq P_\vartheta(|S_n - ES_n| \geq \frac{\varepsilon}{2} \text{ oder } |ES_n - \gamma| \geq \frac{\varepsilon}{2}) \\
&\leq P_\vartheta(|S_n - ES_n| \geq \frac{\varepsilon}{2}) + P_\vartheta(|ES_n - \gamma| \geq \frac{\varepsilon}{2}) \\
&\overset{(*)}{\leq} \frac{\operatorname{Var}(S_n)}{(\frac{\varepsilon}{2})^2} + P_\vartheta(|ES_n - \gamma| \geq \frac{\varepsilon}{2}) \\
&\to 0 \ (n \to \infty)
\end{aligned}$$

$(*)$: Tschebyscheff

Kurz:

$$|S_n - \gamma| \leq \underbrace{|S_n - E_\vartheta S_n|}_{\stackrel{P_\vartheta}{\to} 0 \ (1)} + \underbrace{|E_\vartheta S_n - \gamma|}_{\stackrel{P_\vartheta}{\to} 0 \ (2)} \stackrel{P_\vartheta}{\to} 0$$

(1): Tschebyscheff, (2): asymptotisch erwartungstreu

Obiges Beispiel a):
$\bar{X}_n$ konsistent für $\mu(\vartheta)$, falls $E_\vartheta X_1^2 < \infty$ nach 3.6.
Starkes Gesetz der großen Zahlen (SGGZ):
$\bar{X}_n \stackrel{P_\vartheta - f.s.}{\to} E_\vartheta X_1$ (ohne weitere Voraussetzung) $\Rightarrow \bar{X}_n \stackrel{P_\vartheta}{\to} E_\vartheta X_1$

3.7 Bemerkung und Definition (Sei $\Gamma \subset \mathbb{R}$.)

$$\mathrm{MQA}_S(\vartheta) := E_\vartheta(S(X) - \gamma(\vartheta))^2$$

heißt **mittlere quadratische Abweichung** von S an der Stelle ϑ.
Es gilt:

$$\mathrm{MQA}_S(\vartheta) = E_\vartheta(S(X) - E_\vartheta S(X))^2 + (E_\vartheta S(X) - \gamma(\vartheta))^2 = \mathrm{Var}_\vartheta\, S(X) + b_s^2(\vartheta)$$

3.8 Beispiel

$X_1, \ldots, X_n \stackrel{uiv}{\sim} \mathcal{N}(\mu, \sigma^2)$, μ, σ^2 unbekannt
Schätzer von σ^2:

$$\tilde{\sigma}_n^2(c) = c \cdot \sum_{j=1}^n (X_j - \bar{X}_n)^2$$

Ziel: c so wählen, dass MQA von $\tilde{\sigma}_n^2(c)$ minimal wird.

$$E_\vartheta(\tilde{\sigma}_n^2(c)) = c\sigma^2 \cdot E_\vartheta[\underbrace{\frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \bar{X}_n)^2}_{\sim \chi^2_{n-1}}] = c\sigma^2 \cdot (n-1)$$

$$\mathrm{Var}_\vartheta(\tilde{\sigma}_n^2(c)) = c^2\sigma^4 \cdot 2(n-1)$$

Damit:

$$\begin{aligned}
\mathrm{MQA}_{\tilde{\sigma}^2_n(c)}(\vartheta) &= 2(n-1)c^2\sigma^4 + (c\sigma^2(n-1)-\sigma^2)^2 \\
&= \dots \\
&= \sigma^4(n^2-1)\left[(c-\frac{1}{n+1})^2 + \frac{2}{(n-1)(n+1)^2}\right] \\
&\stackrel{!}{=} \min
\end{aligned}$$

Dies führt offensichtlich auf $c = \frac{1}{n+1}$.