Fachbereich Mathematik
Prof. Dr. J. Lehn
A. Neuenkirch
B. Niese
A. Rößler

TECHNISCHE UNIVERSITÄT DARMSTADT

SS 2006 05.07.2006

Einführung in die Mathematische Statistik

6. Tutorium

Optimale Schätzverfahren

Es seien $X_1, \ldots, X_n$ unabhängige identisch $B(1, \theta)$, $\theta \in]0, 1[$, verteilte Zufallsvariablen. Betrachtet man "lineare" Schätzer der Form

$$T_n^{\text{lin}}(X_1, \ldots, X_n) = a_1 X_1 + \ldots + a_n X_n, \qquad a_1, \ldots, a_n \in \mathbb{R},$$

für $\tau(\theta) = \theta$, dann ist aus der Vorlesung (Beispiel 3.8) bekannt, daß das arithmetische Mittel

$$T_n^*(X_1, \ldots, X_n) = \frac{1}{n} X_1 + \ldots + \frac{1}{n} X_n$$

ein optimales Schätzverfahren in folgendem Sinn ist:

$$\begin{aligned} &\mathrm{Var}_\theta(T_n^*(X_1, \ldots, X_n)) \\ &\qquad = \inf \left\{ \mathrm{Var}_\theta(T_n^{\text{lin}}(X_1, \ldots, X_n))) : \; T_n^{\text{lin}} \text{ erwartungstreuer Schätzer für } \tau(\theta) = \theta \right\}. \end{aligned}$$

Also ist T_n^* ist unter allen erwartungstreuen linearen Schätzern T_n^{lin} für $\tau(\theta) = \theta$ dasjenige Schätzverfahren mit minimaler Varianz.

In diesem Tutorium werden zeigen, daß das arithmetische Mittel in dieser Situation sogar der optimale erwartungstreue Schätzer schlechthin für $\tau(\theta) = \theta$ ist.

Satz 1 *Es seien $X_1, \ldots, X_n$ unabhängige identisch $B(1, \theta)$, $\theta \in]0, 1[$, verteilte Zufallsvariablen. Dann gilt*

$$\begin{aligned} &\mathrm{Var}_\theta(T_n^*(X_1, \ldots, X_n)) \\ &\qquad = \inf \left\{ \mathrm{Var}_\theta(T_n(X_1, \ldots, X_n)) : \; T_n : \mathbb{R}^n \to \mathbb{R} \textit{ erwartungstreuer Schätzer für } \tau(\theta) = \theta \right\}. \end{aligned}$$

Entscheidend für den Beweis dieser Aussage ist folgender Satz.

Satz 2 (Frechét-Rao-Cramer) *Es seien $X_1, \ldots, X_n$ unabhängige identisch $B(1, \theta)$, $\theta \in]0, 1[$, verteilte Zufallsvariablen. Weiterhin sei*

$$L_\theta(x_1, \ldots, x_n) = \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{1 - x_i}, \qquad x_1, \ldots, x_n \in \{0, 1\}.$$

Dann gilt für jedes Schätzverfahren $T_n : \mathbb{R}^n \to \mathbb{R}$, das erwartungstreu für $\tau(\theta) = \theta$ ist:

$$\forall \theta \in]0, 1[: \quad \mathrm{Var}_\theta(T_n(X_1, \ldots, X_n)) \geq \frac{1}{\mathrm{E}_\theta \left| \frac{\partial}{\partial \theta} \ln(L_\theta(X_1, \ldots, X_n)) \right|^2}.$$

Für den Beweis der Frechét-Rao-Cramer Ungleichung, die in ähnlicher Form auch unter viel allgemeineren Voraussetzungen noch gültig ist, benötigen wir die Cauchy-Schwarz Ungleichung, die Sie ohne Beweis verwenden können.

Satz 3 (Cauchy-Schwarz Ungleichung) *Es seien* X *und* Y *zwei Zufallsvariablen mit* $\mathrm{E}\,|X|^2 < \infty$ *und* $\mathrm{E}\,|Y|^2 < \infty$. *Dann gilt*

$$|\mathrm{E}(XY)| \leq \sqrt{\mathrm{E}\,|X|^2} \cdot \sqrt{\mathrm{E}\,|Y|^2}.$$

Aufgabe 1 Zeigen Sie:

$$(a) \qquad 1 = \sum_{(x_1,\ldots,x_n)\in\{0,1\}^n} L_\theta(x_1,\ldots,x_n),$$

$$(b) \qquad \theta = \mathrm{E}_\theta(T_n(X_1,\ldots,X_n)) = \sum_{(x_1,\ldots,x_n)\in\{0,1\}^n} T_n(x_1,\ldots,x_n)\cdot L_\theta(x_1,\ldots,x_n),$$

$$(c) \qquad 0 = \mathrm{E}\left(\frac{\partial}{\partial\theta}\ln(L_\theta(X_1,\ldots,X_n)\right) = \sum_{(x_1,\ldots,x_n)\in\{0,1\}^n} \frac{\partial}{\partial\theta}L_\theta(x_1,\ldots,x_n).$$

Bestimmen Sie dazu zuerst die Wahrscheinlichkeit

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n), \qquad x_1,\ldots,x_n \in \{0,1\}.$$

Aufgabe 2 Beweisen Sie Satz 2. Gehen Sie dabei wie folgt vor:

- Berechnen Sie
$$\mathrm{E}_\theta\left(T_n(X_1,\ldots,X_n)\cdot\frac{\partial}{\partial\theta}\ln(L_\theta(X_1,\ldots,X_n)\right),$$
indem Sie zuerst $\mathrm{E}_\theta(T_n(X_1,\ldots,X_n))$ nach θ differenzieren. (Vgl. Aufgabe 1.)
- Folgern Sie aus dem ersten Schritt und Aufgabe 1
$$\mathrm{E}_\theta\left((T_n(X_1,\ldots,X_n)-\tau(\theta))\cdot\frac{\partial}{\partial\theta}\ln(L_\theta(X_1,\ldots,X_n))\right) = 1,$$
und wenden Sie nun Satz 3 an.

Aufgabe 3 Zeigen Sie jetzt Satz 1, indem Sie

$$\mathrm{E}_\theta\left|\frac{\partial}{\partial\theta}\ln(L_\theta(X_1,\ldots,X_n))\right|^2$$

und

$$\mathrm{Var}_\theta(T_n^*(X_1,\ldots,X_n))$$

berechnen.