week10b: Kapitel 6: Erwartungstreue und Verteilung von Schaetzern, Teil3

Bevor wir nächste Woche in einem ganz allgemeinen Setting die Cramer-Rao Abschätzung beweisen werden und dann als Folgerung erhalten werden, dass die Schätzer $\hat{\mu}_{\rm ML}$ und $\widehat{s^2}$ effizient oder asymptotisch effizient sind, wollen wir an dieser Stelle noch einige Folgerungen aus den Rechnungen vom letzten Mal festhalten. Dazu wollen wir uns zunächst an das 5. Übungsblatt erinnern, da gab es unter anderem die Aussagen (ii) und (iii), das waren die folgenden Sachen:

(ii) Es seien jetzt $\phi_1, \phi_2, ..., \phi_n$ eine Folge von unabhängigen, identisch normalverteilten Zufallsvariablen mit Erwartungswert $\mu = \mathsf{E}[\phi_i]$ und Varianz $\sigma^2 = \mathsf{V}[\phi_i]$ für alle $i = 1, 2, ..., n$. Dann ist die Zufallsvariable

$$Y_n \;:=\; \frac{\frac{1}{n}\sum_{i=1}^n \phi_i \,-\, \mu}{\sigma/\sqrt{n}}$$

für jedes n standard-normalverteilt (also nicht erst im Limes für grosse n).

Ok, das ist klar, das hatten wir schon in dem week3a im Theorem 2.2.1 bewiesen. Die Aussage (iii) auf dem Übungsblatt 5 war dann folgendes:

(iii) Bei unbekannter Varianz σ^2 der Zufallszahlen $\phi_1, \phi_2, ..., \phi_n$ kann man die Varianz schätzen mit dem Ausdruck

$$\hat{\sigma}^2 \;:=\; \tfrac{1}{n-1}\sum_{i=1}^n (\phi_i - \bar{\phi})^2 \qquad \text{wobei} \quad \bar{\phi} \;:=\; \tfrac{1}{n}\sum_{i=1}^n \phi_i \tag{1}$$

Nun gilt folgende Aussage $A(n)$ für $n \geq 2$ ($\hat{\sigma}$ ist für $n = 1$ nicht definiert): Die Zufallsvariable

$$Y_n \;:=\; \frac{\frac{1}{n}\sum_{i=1}^n \phi_i \,-\, \mu}{\hat{\sigma}/\sqrt{n}} \tag{2}$$

ist t-verteilt mit $n-1$ Freiheitsgraden oder kurz t_{n-1}-verteilt.

In dem Übungsblatt hatten wir dann diese Aussage durch eine R-Simulation verifiziert. Mit den Rechnungen vom letzten Mal ist es jetzt aber nicht mehr allzu schwierig, auch den entsprechenden theoretischen Beweis nachzuliefern. Machen wir vielleicht ein separates Theorem und benutzen wir die Notationen aus dem week10a.pdf:

Theorem 6.3.1: Es seien $x_1, \cdots, x_n$ normalverteilte Zufallszahlen mit Mittelwert μ und Varianz σ^2 und es sei

$$\widehat{s^2}(x_1, \cdots, x_n) \;=\; \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{3}$$

die erwartungstreue Version des Maximum-Likelihood-Schätzers für das σ^2. Dann gilt für alle $n \geq 2$: Die Zufallsgrösse

$$y_n \;=\; y_n(x_1, \cdots, x_n) \;:=\; \frac{\frac{1}{n}\sum_{i=1}^{n} x_i \;-\; \mu}{\sqrt{\widehat{s^2}/n}} \tag{4}$$

ist t-verteilt mit $n-1$ Freiheitsgraden oder kurz t_{n-1}-verteilt.

Beweis: Wir müssen zeigen: Für eine beliebige Funktion $F : \mathbb{R} \to \mathbb{R}$ gilt

$$\int_{\mathbb{R}^n} F\big(\, y_n(x_1, \cdots, x_n)\big) \; \prod_{i=1}^{n} \Big\{\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \;\frac{dx_i}{\sqrt{2\pi\sigma^2}} \,\Big\} \;=\; \int_{\mathbb{R}} F(y)\; p_{t_{n-1}}(y)\; dy \tag{5}$$

mit der Dichte $p_{t_{n-1}}(y)$ der t_{n-1}-Verteilung. Wie im week10a im Beweis von Teil (b) von Theorem 6.2.1 schreiben wir

$$\begin{aligned}
y_n \;&=\; \frac{\frac{1}{n}\sum_{i=1}^{n} x_i \;-\; \mu}{\{\, \widehat{s^2}(x_1, \cdots, x_n)/n \,\}^{1/2}} \\
&=\; \frac{\frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)}{\{\, \widehat{s^2}(x_1 - \mu, \cdots, x_n - \mu)/n \,\}^{1/2}} \\
&=\; \frac{\sigma \times \frac{1}{n}\sum_{i=1}^{n} \frac{x_i-\mu}{\sigma}}{\{\, \sigma^2 \times \widehat{s^2}(\frac{x_1-\mu}{\sigma}, \cdots, \frac{x_n-\mu}{\sigma})/n \,\}^{1/2}} \\
&=\; \frac{\frac{1}{n}\sum_{i=1}^{n} \frac{x_i-\mu}{\sigma}}{\{\, \widehat{s^2}(\frac{x_1-\mu}{\sigma}, \cdots, \frac{x_n-\mu}{\sigma})/n \,\}^{1/2}}
\end{aligned}$$

so dass

$$\begin{aligned}
&\int_{\mathbb{R}^n} F\big(\, y_n(x_1, \cdots, x_n)\big) \; \prod_{i=1}^{n} \Big\{\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \;\frac{dx_i}{\sqrt{2\pi\sigma^2}} \,\Big\} \;= \\
&\qquad \int_{\mathbb{R}^n} \frac{\frac{1}{n}\sum_{i=1}^{n} z_i}{\{\, \widehat{s^2}(z_1, \cdots, z_n)/n \,\}^{1/2}} \; \prod_{k=1}^{n} \Big\{\, e^{-\frac{z_k^2}{2}} \;\frac{dz_k}{\sqrt{2\pi}} \,\Big\}
\end{aligned} \tag{6}$$

mit

$$\widehat{s^2}(z_1, \cdots, z_n) \;=\; \frac{1}{n-1} \sum_{i=1}^{n} \Big\{\, z_i \;-\; \tfrac{1}{n}\textstyle\sum_{j=1}^{n} z_j \,\Big\}^2$$

Die folgende Rechnung können wir jetzt eins zu eins aus dem Beweis vom Theorem 6.2.1 übernehmen: Wir definieren den Vektor

$$v_1 \;:=\; \tfrac{1}{\sqrt{n}}\,(1, 1, \cdots, 1) \;\in\; \mathbb{R}^n$$

mit $\|v_1\|^2 = 1$. Wir haben dann

$$\langle z, v_1 \rangle \;=\; \tfrac{1}{\sqrt{n}} \textstyle\sum_{j=1}^n z_j \tag{7}$$

und können schreiben

$$z_i - \tfrac{1}{n} \textstyle\sum_{j=1}^n z_j \;=\; \Big[\, z - \langle z, v_1 \rangle v_1 \,\Big]_{i-\text{te Komponente}}$$

Weiterhin, da v_1 ein Einheitsvektor ist, können wir das v_1 mit geeigneten Vektoren $v_2, \cdots, v_n$ zu einer Orthonormalbasis

$$B \;=\; \{v_1, v_2, \cdots, v_n\}$$

des $\mathbb{R}^n$ erweitern. Wir haben dann also die Skalarprodukte

$$\langle v_j, v_k \rangle \;=\; \delta_{j,k}$$

und können das $z \in \mathbb{R}^n$ nach den $v_k \in \mathbb{R}^n$ entwickeln,

$$z \;=\; \sum_{k=1}^n \langle z, v_k \rangle \, v_k$$

Insbesondere,

$$\begin{aligned}
z_i - \tfrac{1}{n} \textstyle\sum_{j=1}^n z_j \;&=\; \Big[\, z - \langle z, v_1 \rangle v_1 \,\Big]_{i-\text{te Komponente}} \\
&=\; \left[\, \sum_{k=1}^n \langle z, v_k \rangle \, v_k - \langle z, v_1 \rangle v_1 \right]_{i-\text{te Komponente}} \\
&=\; \left[\, \sum_{k=2}^n \langle z, v_k \rangle \, v_k \right]_{i-\text{te Komponente}}
\end{aligned}$$

und wir können schreiben

$$\begin{aligned}
\widehat{s^2}(z_1, \cdots, z_n) \;&=\; \frac{1}{n-1} \sum_{i=1}^n \Big\{ z_i - \tfrac{1}{n} \textstyle\sum_{j=1}^n z_j \Big\}^2 \\
&=\; \frac{1}{n-1} \left\| \sum_{k=2}^n \langle z, v_k \rangle v_k \right\|^2 \\
&=\; \frac{1}{n-1} \left\langle \sum_{k=2}^n \langle z, v_k \rangle v_k \,, \sum_{\ell=2}^n \langle z, v_\ell \rangle v_\ell \right\rangle \\
&=\; \frac{1}{n-1} \sum_{k,\ell=2}^n \langle z, v_k \rangle \langle z, v_\ell \rangle \underbrace{\langle v_k, v_\ell \rangle}_{=\delta_{k,\ell}} \\
&=\; \frac{1}{n-1} \sum_{k=2}^n \langle z, v_k \rangle^2
\end{aligned} \tag{8}$$

Also bekommen wir mit (7) und (8)

$$\begin{aligned}
&\int_{\mathbb{R}^n} F\big(\, y_n(x_1, \cdots, x_n)\big) \ \prod_{i=1}^{n} \Big\{ \ e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \ \frac{dx_i}{\sqrt{2\pi\sigma^2}} \ \Big\} \\
&\quad = \ \int_{\mathbb{R}^n} \frac{\frac{1}{n} \sum_{i=1}^{n} z_i}{\{\ \widehat{s^2}(z_1, \cdots, z_n)/n \ \}^{1/2}} \ \prod_{k=1}^{n} \Big\{ \ e^{-\frac{z_k^2}{2}} \ \frac{dz_k}{\sqrt{2\pi}} \ \Big\} \\
&\quad = \ \int_{\mathbb{R}^n} \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} z_i}{\{\ \widehat{s^2}(z_1, \cdots, z_n) \ \}^{1/2}} \ \prod_{k=1}^{n} \Big\{ \ e^{-\frac{z_k^2}{2}} \ \frac{dz_k}{\sqrt{2\pi}} \ \Big\} \\
&\quad = \ \int_{\mathbb{R}^n} \frac{\langle z, v_1 \rangle}{\{\ \frac{1}{n-1} \ \sum_{k=2}^{n} \langle z, v_k \rangle^2 \ \}^{1/2}} \ \prod_{k=1}^{n} \Big\{ \ e^{-\frac{z_k^2}{2}} \ \frac{dz_k}{\sqrt{2\pi}} \ \Big\} \qquad (9)
\end{aligned}$$

Wir machen wieder die orthogonale Variablen-Substitution

$$\begin{aligned}
\phi_1 \ &= \ \langle v_1, z \rangle \\
\phi_2 \ &= \ \langle v_2, z \rangle \\
&\ \vdots \\
\phi_n \ &= \ \langle v_n, z \rangle
\end{aligned}$$

oder in Vektor-Matrix-Schreibweise

$$\phi \ = \ V z$$

mit der orthogonalen Matrix ($V^{-1} = V^T$)

$$V \ := \ \begin{pmatrix} - & v_1 & - \\ & \vdots & \\ - & v_n & - \end{pmatrix} \ \in \ \mathbb{R}^{n \times n}$$

und haben wieder

$$\begin{aligned}
\langle z, z \rangle \ &= \ \langle \phi, \phi \rangle \\
d^n z \ &= \ d^n \phi
\end{aligned}$$

und bekommen dann die folgende Integraldarstellung:

$$\begin{aligned}
&\int_{\mathbb{R}^n} F\big(\, y_n(x_1, \cdots, x_n)\big) \ \prod_{i=1}^{n} \Big\{ \ e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \ \frac{dx_i}{\sqrt{2\pi\sigma^2}} \ \Big\} \\
&\quad = \ \int_{\mathbb{R}^n} F\Bigg(\ \frac{\langle z, v_1 \rangle}{\{\ \frac{1}{n-1} \ \sum_{k=2}^{n} \langle z, v_k \rangle^2 \ \}^{1/2}} \ \Bigg) \ \prod_{k=1}^{n} \Big\{ \ e^{-\frac{z_k^2}{2}} \ \frac{dz_k}{\sqrt{2\pi}} \ \Big\} \\
&\quad = \ \int_{\mathbb{R}^n} F\Bigg(\ \frac{\phi_1}{\{\ \frac{1}{n-1} \ \sum_{k=2}^{n} \phi_k^2 \ \}^{1/2}} \ \Bigg) \ \prod_{k=1}^{n} \Big\{ \ e^{-\frac{\phi_k^2}{2}} \ \frac{d\phi_k}{\sqrt{2\pi}} \ \Big\} \qquad (10)
\end{aligned}$$

Das Integral (10) ist jetzt aber genau von der Form, wie sie auch schon im Kapitel 3.1 im Theorem 3.1.3 Teil (b) im week5a.pdf betrachtet wurde. Wir können also das Theorem 3.1.3 anwenden und bekommen dann sofort

$$\begin{aligned}
&\int_{\mathbb{R}^n} F\big(\,y_n(x_1,\cdots,x_n)\,\big)\;\prod_{i=1}^n \Big\{\, e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\;\frac{dx_i}{\sqrt{2\pi\sigma^2}}\,\Big\} \\
&= \quad \int_{\mathbb{R}^n} F\left(\frac{\phi_1}{\big\{\,\frac{1}{n-1}\,\sum_{k=2}^n \phi_k^2\,\big\}^{1/2}}\right)\;\prod_{k=1}^n \Big\{\, e^{-\frac{\phi_k^2}{2}}\;\frac{d\phi_k}{\sqrt{2\pi}}\,\Big\} \\
&\overset{\text{Theorem 3.1.3 (b)}}{=} \quad \int_{\mathbb{R}} F(y)\; p_{t_{n-1}}(y)\; dy
\end{aligned} \tag{11}$$

Damit ist das Theorem bewiesen. $\blacksquare$

Die Verteilungen der Schätzer

Die Verteilungen von $\hat{\mu}_{\mathrm{ML}}$ und $\widehat{s^2}$ können wir jetzt sofort angeben: Das

$$\hat{\mu}_{\mathrm{ML}} \;=\; \frac{1}{n}\sum_{i=1}^n x_i$$

ist als Summe von normalverteilten Zufallsvariablen wieder normalverteilt, mit Mittelwert μ und Varianz

$$\mathsf{V}\big[\,\hat{\mu}_{\mathrm{ML}}(x_1,\cdots,x_n)\,\big] \;=\; \frac{\sigma^2}{n}$$

Für das $\widehat{s^2}$ hatten wir im Beweis von Theorem 6.2.1 (oder auch eben nochmal) die Darstellung

$$\widehat{s^2} \;=\; \frac{\sigma^2}{n-1}\sum_{k=2}^n \phi_k^2$$

hergeleitet, mit standard-normalverteilten ϕ_k's. Die Kombination

$$\sum_{k=2}^n \phi_k^2 \;\in\; \chi^2_{n-1}$$

ist chi-quadrat verteilt mit $n-1$ Freiheitsgraden, die Dichte dazu hatten wir im Kapitel 3.1 hergeleitet. Also können wir schreiben, wieder mit einem beliebigen $F:\mathbb{R}\to\mathbb{R}$

$$\mathsf{E}\big[\,F(\,\widehat{s^2}\,)\,\big] \;=\; \int_{\mathbb{R}^n} F\left(\frac{\sigma^2}{n-1}\sum_{k=2}^n \phi_k^2\right)\;\prod_{k=1}^n \Big\{\, e^{-\frac{\phi_k^2}{2}}\;\frac{d\phi_k}{\sqrt{2\pi}}\,\Big\}$$

Das ϕ_1 kommt in dem F nicht vor und wir können es ausintegrieren. Dann bleibt folgendes übrig:

$$
\begin{aligned}
\mathsf{E}\big[\, F(\,\widehat{s^2}\,)\,\big] \quad &= \quad \int_{\mathbb{R}^{n-1}} F\left(\frac{\sigma^2}{n-1} \sum_{k=2}^{n} \phi_k^2 \right) \prod_{k=2}^{n} \left\{ e^{-\frac{\phi_k^2}{2}} \, \frac{d\phi_k}{\sqrt{2\pi}} \right\} \\
&\overset{\text{Theorem 3.1.3 (a)}}{=} \quad \int_0^\infty F\big(\tfrac{\sigma^2}{n-1}\, y \big) \, p_{\chi^2_{n-1}}(y) \, dy \\
&= \quad \int_0^\infty F(\, z\,) \, \tfrac{n-1}{\sigma^2} \, p_{\chi^2_{n-1}}\big(\tfrac{n-1}{\sigma^2}\, z \big) \, dz
\end{aligned}
$$

Also ist die Dichte von dem $\widehat{s^2}$ gegeben durch

$$
\mathsf{Prob}\Big[\; \widehat{s^2} \in [z, z+dz) \;\Big] \quad = \quad \tfrac{n-1}{\sigma^2} \; \times \; p_{\chi^2_{n-1}}\big(\tfrac{n-1}{\sigma^2}\, z \big) \, dz
$$

wobei das $p_{\chi^2_{n-1}}$ also die Dichte der χ^2-Verteilung mit $n-1$ Freiheitsgraden ist, eine explizite Formel ist in dem Theorem 3.1.3 im Teil (a) angegeben.

Diese Resultate könnte man natürlich noch einmal mit einer R-Simulation verifizieren.