Hochschule RheinMain SS 2023
Prof. Dr. D. Lehmann

# 11. Übungsblatt zur Vorlesung Stochastik II

**Aufgabe 1)** Es sei

$$p_\theta(x) \;=\; p_{\theta_1,\cdots,\theta_m}(x_1,\cdots,x_n)$$

eine Wahrscheinlichkeitsdichte und $I(\theta)$ sei die Fisher-Informationsmatrix aus der Cramer-Rao Abschätzung. Betrachten Sie den folgenden "Schätzer"

$$\hat\theta(x) \;=\; \theta \,+\, I^{-1}(\theta)\,\nabla_\theta \log p_\theta(x) \tag{1}$$

Zeigen Sie:

**a)** Das $\hat\theta$ ist erwartungstreu, es gilt

$$\mathsf{E}[\,\hat\theta_k\,] \;=\; \theta_k$$

**b)** Das $\hat\theta$ hat minimale Varianz in der Menge aller erwartungstreuen Schätzer, es gilt

$$\mathsf{V}[\,\hat\theta_k\,] \;=\; [I^{-1}(\theta)]_{k,k}$$

Erinnern Sie sich dazu an die Gleichungen (7) und (8) aus dem Beweis der Cramer-Rao Abschätzung, das war das Theorem 7.1.2 in dem week11b.pdf.

*Bemerkung:* Im allgemeinen hängt die rechte Seite von (1) von den gesuchten Modellparametern $\theta = (\theta_1,\cdots,\theta_m)$ ab, das kann dann also nicht wirklich als Schätzer benutzt werden, deshalb die Anführungszeichen "Schätzer" weiter oben. Man kommt auf den Ausdruck (1), wenn man sich überlegt, für welche $\hat\theta$ im Beweis der Cramer-Rao Abschätzung an der Stelle, wo die Cauchy-Schwarzsche Ungleichung benutzt wird, das Gleichheitszeichen gilt.

**Aufgabe 2)** In Aufgabe 1 haben wir gesehen, dass der Ausdruck (1),

$$\hat\theta(x) \;=\; \theta \,+\, I^{-1}(\theta)\,\nabla_\theta \log p_\theta(x)$$

ein erwartungstreuer Minimum-Varianz-"Schätzer" ist. Wir wollen noch einmal das Standardbeispiel 1 betrachten mit unabhängigen, normalverteilten Zufallszahlen und Wahrscheinlichkeitsdichte oder Likelihood-Funktion (wir setzen $\nu := \sigma^2$)

$$p_\theta(x) \;=\; p_{\mu,\nu}(x_1,\cdots,x_n) \;=\; \prod_{i=1}^n \left\{ \tfrac{1}{\sqrt{2\pi\nu}}\, e^{-\frac{(x_i-\mu)^2}{2\nu}} \right\}$$

und das $\hat\theta$ aus (1) berechnen. Zeigen Sie dazu mit $\hat\theta = (\hat\mu,\hat\nu)$:

**a)** Das $\hat{\mu}$ gegeben durch Gleichung (1) reduziert sich auf

$$\hat{\mu}(x_1, \cdots, x_n) \quad = \quad \frac{1}{n}\sum_{i=1}^{n} x_i$$

Das ist also identisch mit dem Maximum Likelihood Schätzer und insbesondere unabhängig von den Modellparametern $\theta = (\mu, \nu)$, die kürzen sich weg.

**b)** Das $\hat{\nu}$ gegeben durch Gleichung (1) reduziert sich auf

$$\hat{\nu}(x_1, \cdots, x_n) \quad = \quad \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{2}$$

und unterscheidet sich von dem Maximum Likelihood Schätzer nur dadurch, dass auf der rechten Seite von (2) der theoretische (unbekannte) Modellparameter $\mu$ in den Quadraten auftaucht, und nicht ein geschätztes $\hat{\mu}$ mit $\hat{\mu} = \bar{x} = \sum_{i=1}^{n} x_i/n$.

Das Berechnen von Schätzern mit Gleichung (1) kann also durchaus Sinn machen, wenn es darum geht, Ideen zu entwickeln, wie Modellparameter möglichst effizient geschätzt werden können.