3.3 Methoden zur Evaluierung von Schätzern

Bis jetzt haben wir nur glaubwürdige Techniken zur Konstruktion von Punktschätzern besprochen. Falls unterschiedliche Schätzer für einen Parameter resultieren, welchen sollen wir verwenden?

Wir werden nun einige Kriterien einführen, um die Qualität von Schätzern zu bewerten.

3.3.1 Der Mittlere Quadratische Fehler

Definition 3.3.1: Der **Mittlere Quadratische Fehler** (mean squared error – MSE) eines Schätzers W für einen Parameter θ ist eine Funktion in θ definiert durch $\mathsf{MSE}_\theta(W) = \mathsf{E}_\theta(W - \theta)^2$.

Dies ist eine monotone Funktion der Distanz.

Auch der mittlere absolute Fehler $\mathsf{E}_\theta(|W-\theta|)$ (mean absolute error – MAE) wäre ein Kandidat dafür, aber der MSE hat demgegenüber einige Vorteile:

1. Der MSE ist analytisch ziemlich gut handhabbar,

2. Der MSE lässt sich zerlegen in

$$\begin{aligned} \mathsf{E}_\theta(W-\theta)^2 &= \mathsf{E}_\theta\Big([W-\mathsf{E}_\theta(W)]+[\mathsf{E}_\theta(W)-\theta]\Big)^2 \\ &= \mathsf{E}_\theta\Big(W-\mathsf{E}_\theta(W)\Big)^2+\mathsf{E}_\theta\Big(\mathsf{E}_\theta(W)-\theta\Big)^2 \end{aligned}$$

da

$$\mathsf{E}_\theta\Big(W\mathsf{E}_\theta(W)-\mathsf{E}_\theta^2(W)-W\theta+\mathsf{E}_\theta(W)\theta\Big)=0\,.$$

Also folgt

$$\mathsf{MSE}=\mathsf{var}_\theta(W)+\big(\mathsf{bias}_\theta(W,\theta)\big)^2$$

Definition 3.3.2: Der **Bias** (Grad der Verzerrung) eines Punktschätzers W für einen Parameter θ ist die Differenz zwischen dem erwarteten Wert von W und θ, d.h. $\text{bias}_\theta(W, \theta) = \mathsf{E}_\theta(W - \theta)$. Ein Schätzer dessen Bias identisch Null ist heißt **unverzerrt** (unbiased) oder auch **erwartungstreu** und er genügt $\text{bias}_\theta(W, \theta) = 0$ für alle θ.

Der MSE beinhaltet 2 Komponenten: eine misst die Variabilität des Schätzers (Präzision) und die andere misst die Abweichung (Genauigkeit). Ein Schätzer mit geringem MSE hat gute Präzision kombiniert mit geringer Abweichung.

Für unverzerrte Schätzer gilt $\text{MSE} = \mathsf{E}_\theta(W - \theta)^2 = \text{var}_\theta(W)$, und daher ist für unverzerrte Schätzer der MSE gleich der Varianz.

Beispiel 3.3.1: Seien $X_1, \ldots, X_n$ iid Normal(μ, σ^2). Es wurde bereits gezeigt, dass

$$\mathsf{E}(\overline{X}) = \mu\,, \qquad \mathsf{E}(S^2) = \sigma^2\,, \qquad \text{für alle } \mu \text{ und } \sigma^2.$$

Also sind $\overline{X}$ und S^2 unverzerrte Schätzer für μ und σ^2. Dies hält auch ohne der Annahme der Normalverteilung. Die MSE's dieser Schätzer sind somit

$$\begin{aligned} \mathsf{E}(\overline{X} - \mu)^2 &= \mathsf{var}(\overline{X}) = \frac{\sigma^2}{n} \\ \mathsf{E}(S^2 - \sigma^2)^2 &= \mathsf{var}(S^2) = \frac{2\sigma^4}{n-1}\,. \end{aligned}$$

Letzteres ergibt sich aus der bekannten Eigenschaft $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$, also ist $\mathsf{var}((n-1)S^2/\sigma^2) = 2(n-1)$.

Der MSE von $\overline{X}$ ist σ^2/n auch ohne Annahme der Normalverteilung. Dies gilt aber nicht für S^2.

Viele unverzerrte Schätzer haben auch einen glaubwürdigen MSE. Jedoch garantiert die Kontrollieren über den Bias nicht die Kontrolle über den MSE. Manchmal impliziert eine leichte Zunahme im Bias eine große Abnahme der Varianz, was zu einer Verbesserung im MSE führt.

Beispiel 3.3.1: Fortsetzung Ein alternativer Schätzer für σ^2 ist der MLE

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{n-1}{n}S^2\,.$$

Dafür folgt

$$\mathsf{E}(\hat{\sigma}^2) = \mathsf{E}\left(\frac{n-1}{n}S^2\right) = \frac{n-1}{n}\sigma^2 < \sigma^2\,, \qquad \mathsf{bias}(\hat{\sigma}^2, \sigma^2) = -\frac{\sigma^2}{n}\,.$$

Also ist $\hat{\sigma}^2$ verzerrt für σ^2.

Als Varianz resultiert

$$\mathsf{var}(\hat{\sigma}^2) = \mathsf{var}\left(\frac{n-1}{n}S^2\right) = \left(\frac{n-1}{n}\right)^2 \mathsf{var}\left(S^2\right) = 2\frac{n-1}{n^2}\sigma^4$$

und als MSE

$$\mathsf{E}(\hat{\sigma}^2 - \sigma^2)^2 = \mathsf{var}(\hat{\sigma}^2) + \mathsf{bias}^2(\hat{\sigma}^2, \sigma^2) = 2\frac{n-1}{n^2}\sigma^4 + \left(-\frac{\sigma^2}{n}\right)^2 = \frac{\sigma^4}{n^2}(2n-1)\,.$$

Zusammen liefert dies

$$\mathsf{E}(\hat{\sigma}^2 - \sigma^2)^2 = \sigma^4\frac{2n-1}{n^2} < \sigma^4\frac{2}{n-1} = \mathsf{E}(S^2 - \sigma^2)^2\,, \qquad \text{für alle } \sigma^2.$$

Somit hat $\hat{\sigma}^2$ einen geringeren MSE als das unverzerrte S^2.

Da der MSE eine Funktion in θ ist, gibt es auch nicht den *besten Schätzer*. Oft überschneiden sich die MSE's zweier Schätzer, also ist der eine Schätzer besser in einem Teil des Parameterraums als der andere, und umgekehrt im anderen Teil.

3.3.2 Beste unverzerrte Schätzer

Ein Vergleich von Schätzern basierend auf deren MSE's zeigt häufig **keinen** klaren Favoriten. Eigentlich gibt es gar keinen *besten MSE Schätzer*. Der Grund dafür liegt darin, dass die Klasse der betrachteten Schätzer zu groß ist ($\hat{\theta} = 17$ ist unschlagbar gut falls $\theta = 17$ gilt, ist aber auch sehr schlecht für andere Werte von θ.) Eine Möglichkeit, um dieses Problem zu behandeln, besteht darin, die Klasse der Schätzer einzuschränken. Wir betrachten nur noch **unverzerrte** Schätzer.

Falls W_1 und W_2 unverzerrt sind, $\mathsf{E}_\theta(W_1) = \mathsf{E}_\theta(W_2) = \theta$, dann sind deren MSE's gleich deren Varianzen, und wir können den Schätzer mit der geringeren Varianz nehmen. Finden wir einen Schätzer mit gleichmäßig kleinster Varianz — einen **besten unverzerrten Schätzer** — haben wir es geschafft.

Eigentlich ist die betrachtete Klasse sogar größer. Sei W^* ein Schätzer für θ mit $\mathsf{E}_\theta(W^*) = \tau(\theta) \neq \theta$. Betrachte die Klasse von Schätzern

$$\mathcal{C}_\tau = \{W : \mathsf{E}_\theta(W) = \tau(\theta)\} .$$

Für beliebige $W_1, W_2 \in \mathcal{C}_\tau$ gilt

$$\mathsf{bias}_\theta(W_1, \theta) = \mathsf{bias}_\theta(W_2, \theta) = \tau(\theta) - \theta$$

und wir haben

$$\mathsf{E}_\theta(W_1 - \theta)^2 - \mathsf{E}_\theta(W_2 - \theta)^2 = \mathsf{var}_\theta(W_1) - \mathsf{var}_\theta(W_2) .$$

Daher basiert der MSE Vergleich auch hier nur auf den Vergleich der Varianzen.

Definition 3.3.3: Ein Schätzer W^* heißt **bester unverzerrter Schätzer** für $\tau(\theta)$, falls dieser $\mathsf{E}_\theta(W^*) = \tau(\theta)$ für alle θ genügt, und für alle anderen Schätzer W mit $\mathsf{E}_\theta(W) = \tau(\theta)$ gilt, dass $\mathrm{var}_\theta(W^*) \leq \mathrm{var}_\theta(W)$ für alle θ. Den Schätzer W^* nennt man auch **uniform minimum variance unbiased estimator** (UMVUE) für $\tau(\theta)$.

Den UMVUE zu finden, falls dieser existiert, ist nicht immer einfach.

Beispiel 3.3.2 (unverzerrte Poisson Schätzer): Seien $X_1, \ldots, X_n$ iid Poisson(λ), also $\mathsf{E}(X) = \mathrm{var}(X) = \lambda$. Deshalb folgt

$$\begin{aligned} \mathsf{E}_\lambda(\overline{X}) &= \lambda\,, \quad \text{für alle } \lambda, \\ \mathsf{E}_\lambda(S^2) &= \lambda\,, \quad \text{für alle } \lambda, \end{aligned}$$

und beide Statistiken sind erwartungstreue Schätzer für λ. Nun gilt

$$\mathrm{var}_\lambda(\overline{X}) = \lambda/n\,, \quad \text{für alle } \lambda,$$

aber die Herleitung von $\mathsf{var}_\lambda(S^2)$ ist ein enormer Aufwand. Jedenfalls gilt $\mathsf{var}_\lambda(\overline{X}) \le \mathsf{var}_\lambda(S^2)$.

Sogar wenn wir dies geschafft haben zu zeigen, ist noch die gesamte Klasse von Schätzern vorhanden

$$W_a(\overline{X}, S^2) = a\overline{X} + (1-a)S^2$$

mit $0 \le a \le 1$. Für jedes a gilt

$$\mathsf{E}_\lambda(W_a(\overline{X}, S^2)) = \lambda\,, \quad \text{für alle } \lambda.$$

Damit haben wir unendlich viele unverzerrte Schätzer für λ. Falls $\overline{X}$ besser ist als S^2, ist $\overline{X}$ auch besser als jedes $W_a(\overline{X}, S^2)$?

Angenommen wir könnten eine untere Schranke $B(\theta)$ für die Varianz jedes unverzerrten Schätzers für $\tau(\theta)$ spezifizieren. Finden wir den Schätzer W^* mit $\mathsf{var}_\theta(W^*) = B(\theta)$, dann haben wir einen **besten unverzerrten Schätzer**.

Satz 3.3.1: (Cramér-Rao Ungleichung) Sei $X_1, \ldots, X_n$ eine (nicht notwendigerweise iid) Stichprobe mit Dichte $f(\mathbf{x}|\theta)$, und sei $W(\mathbf{X}) = W(X_1, \ldots, X_n)$ ein beliebiger Schätzer, der

$$\frac{d}{d\theta}\mathsf{E}_\theta(W(\mathbf{X})) = \int \frac{\partial}{\partial\theta}\left[W(\mathbf{x})f(\mathbf{x}|\theta)\right]\, d\mathbf{x}$$

und

$$\mathsf{var}(W(\mathbf{X})) < \infty$$

genügt. Dann gilt

$$\mathsf{var}_\theta(W(\mathbf{X})) \geq \frac{\left(\frac{\partial}{\partial\theta}\mathsf{E}_\theta(W(\mathbf{X}))\right)^2}{\mathsf{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta)\right)^2}.$$

Satz 3.3.1 gilt für beliebige Stichproben (nicht notwendigerweise iid). Für iid Stichproben (Zufallsstichproben) ergibt sich folgende Vereinfachung.

Korollar 3.3.1: (Cramér-Rao Ungleichung für den iid Fall) Falls die Voraussetzungen des Satzes 3.3.1 erfüllt sind und zusätzlich $X_1, \ldots, X_n$ eine iid Stichprobe mit Dichte $f(x|\theta)$ ist, dann gilt

$$\mathsf{var}_\theta(W(\mathbf{X})) \geq \frac{\left(\frac{\partial}{\partial\theta}\mathsf{E}_\theta(W(\mathbf{X}))\right)^2}{n\mathsf{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right)^2}.$$

Noch einige Bemerkungen:

Die Cramér-Rao Schranke wurde hier nur für stetig verteilte Zufallsvariablen angeführt, sie hält aber auch für diskrete Populationen. Die Hauptbedingung erlaubt das Vertauschen von Differentiation und Integration und wird für den diskreten Fall zur Bedingung der Vertauschbarkeit von Differentiation und Summation.

Die Größe

$$\mathsf{E}_\theta \left(\frac{\partial}{\partial \theta} \log f(\mathbf{X}|\theta) \right)^2$$

nennt man **Informationszahl** (oder Fisher-Information) der Stichprobe $\mathbf{X}$. Steigt die Informationszahl, so wächst die Information über θ und wir bekommen eine kleinere Schranke für die Varianz des besten unverzerrten Schätzers.

Für jede differenzierbare Funktion $\tau(\theta)$ haben wir somit eine untere Schranke für die Varianz eines beliebigen erwartungstreuen Schätzers W für $\tau(\theta)$. Diese Schranke hängt nur von $\tau(\theta)$ und $f(\mathbf{x}|\theta)$ ab und ist eine gleichmäßige untere Schranke für die Varianz. Ein Schätzer, für den $\mathsf{E}_\theta(W) = \tau(\theta)$ gilt, und der diese untere Varianzschranke erreicht, ist der beste unverzerrte Schätzer für $\tau(\theta)$.

Das folgende Resultat erlaubt oft eine vereinfachte Berechnung der Cramér-Rao Varianzschranke.

Lemma 3.3.1: Falls $f(x|\theta)$ der Identität

$$\frac{d}{d\theta}\mathsf{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) = \int \frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right) f(x|\theta)\right] dx$$

genügt (gilt für die Exponentialfamilie), dann ist

$$\mathsf{E}_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2 = -\mathsf{E}_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right) .$$

(Beweis Übung)

Beispiel 3.3.2 (Fortsetzung): Für den Poissonfall haben $\overline{X}$, S^2 und $W_a(\overline{X}, S^2)$ alle Erwartungswert λ, d.h. hier ist $\tau(\lambda) = \lambda$ und $\partial\tau(\lambda)/\partial\lambda = 1$.

Wir sind in der Exponentialfamilie und es folgt

$$\begin{aligned}
\mathsf{E}_\lambda \left(\frac{\partial}{\partial\lambda} \log f(\mathbf{X}|\lambda) \right)^2 &= \mathsf{E}_\lambda \left(\frac{\partial}{\partial\lambda} \log \prod_{i=1}^{n} f(X_i|\lambda) \right)^2 \\
&= -n\mathsf{E}_\lambda \left(\frac{\partial^2}{\partial\lambda^2} \log f(X_1|\lambda) \right) \\
&= -n\mathsf{E}_\lambda \left(\frac{\partial^2}{\partial\lambda^2} \log \left(\tfrac{1}{X_1!} e^{-\lambda} \lambda^{X_1} \right) \right) \\
&= -n\mathsf{E}_\lambda \left(\frac{\partial^2}{\partial\lambda^2} \big(-\log(X_1!) - \lambda + X_1 \log\lambda \big) \right) \\
&= -n\mathsf{E}_\lambda \left(-\frac{X_1}{\lambda^2} \right) = n\frac{\lambda}{\lambda^2} = \frac{n}{\lambda}\,.
\end{aligned}$$

Korollar 3.3.1 liefert für beliebige Schätzer $W(\mathbf{X})$ mit $\mathsf{E}_\lambda(W(\mathbf{X})) = \lambda$

$$\mathsf{var}_\lambda(W(\mathbf{X})) \geq \frac{1}{n/\lambda} = \frac{\lambda}{n}\,.$$

Da $\mathsf{var}_\lambda(\overline{X}) = \lambda/n$ ist $\overline{X}$ der gleichmäßig beste unverzerrte Schätzer für λ.

Bemerkungen:

Eine kritische Annahme im Cramér-Rao Satz 3.3.1 ist die Möglichkeit im Integral zu differenzieren. Für die Exponentialfamilie ist dies möglich. Generell muss man aber prüfen, ob dies auch gewährleistet ist.

Im Allgemeinen ist Satz 3.3.1 **nicht** anwendbar wenn der Träger der Dichte oder der Wahrscheinlichkeitsfunktion vom Parameter θ abhängt.

Beispiel 3.3.3 (Gleichverteilung): Seien $X_1, \ldots, X_n$ iid stetig Uniform$(0, \theta)$, d.h. mit Dichte $f(x|\theta) = \frac{1}{\theta} I_{[0,\theta]}(x)$. Diese ist **nicht differenzierbar** in θ und der Cramér-Rao Satz ist **nicht anwendbar**.

Ignorieren des Indikators führt **fälschlicherweise** zu

$$\frac{\partial}{\partial \theta} \log f(x|\theta) = -\frac{\partial}{\partial \theta} \log \theta = -\frac{1}{\theta} \quad \Longrightarrow \quad \mathsf{E}_\theta \left(\frac{\partial}{\partial \theta} \log f(x|\theta) \right)^2 = \frac{1}{\theta^2} .$$

Cramér-Rao weist darauf hin, dass für unverzerrte Schätzer $W(\mathbf{X})$ für θ gilt

$$\mathsf{var}_\theta(W(\mathbf{X})) \geq \frac{\theta^2}{n} .$$

Wir suchen nun einen Schätzer mit kleiner Varianz. Die **suffiziente** Statistik für dieses Problem ist $Y = \max_i(X_i)$, die größte Ordnungsstatistik.

Zur Verteilung des Maximums: für $0 \le y \le \theta$ gilt

$$F_Y(y|\theta) = P_\theta(Y \le y) = P_\theta(X_1 \le y, \dots, X_n \le y) = \prod_{i=1}^n P_\theta(X_i \le y) = \left(\frac{y}{\theta}\right)^n,$$

$$f_Y(y|\theta) = \frac{\partial}{\partial y} F_Y(y|\theta) = n\left(\frac{y}{\theta}\right)^{n-1} \frac{1}{\theta} = n y^{n-1} \frac{1}{\theta^n}$$

$$\mathsf{E}_\theta(Y) = \int_0^\theta n y^n \frac{1}{\theta^n} dy = \frac{n}{\theta^n} \frac{1}{n+1} y^{n+1} \Big|_0^\theta = \frac{n}{n+1}\theta\,.$$

Ein unverzerrter Schätzer (der auf die suffiziente Statistik basiert) ist somit $\frac{n+1}{n}Y$. Weiters gilt

$$\mathsf{E}_\theta(Y^2) = \int_0^\theta n y^{n+1} \frac{1}{\theta^n} dy = \frac{n}{n+2}\theta^2$$

und

$$
\begin{aligned}
\mathsf{var}_\theta\left(\tfrac{n+1}{n}Y\right) &= \frac{(n+1)^2}{n^2}\left[\mathsf{E}_\theta(Y^2)-\mathsf{E}_\theta^2(Y)\right] = \frac{(n+1)^2}{n^2}\left[\frac{n}{n+2}\theta^2-\frac{n^2}{(n+1)^2}\theta^2\right] \\
&= \frac{\theta^2}{n(n+2)} \le \frac{\theta^2}{n},
\end{aligned}
$$

was gleichmäßig kleiner ist als die Cramér-Rao Schranke.

Natürlich war hier der Satz von Cramér-Rao für diese Dichte gar nicht anwendbar!

Auch für den Fall, dass der Satz von Cramér-Rao verwendbar ist, gibt es keine Garantie dafür, dass die Grenze **scharf** ist. Es könnte der Wert der Grenze ja auch grundsätzlich kleiner sein als die Varianzen aller unverzerrter Schätzer.

Für die einparametrige Exponentialfamilie kann man beispielsweise nur sagen, dass ein Parameter $\tau(\theta)$ existiert für den die Grenze scharf ist. In anderen Situationen ist die Schranke gar nicht **erreichbar**.

Beispiel 3.3.4 Seien $X_1, \ldots, X_n$ iid Normal(μ, σ^2). Schätze σ^2 bei μ unbekannt. Die Dichte genügt allen Voraussetzungen womit folgt

$$
\begin{aligned}
\log f(X_1|\mu,\sigma^2) &= -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log\sigma^2 - \frac{1}{2}\frac{(X_1-\mu)^2}{\sigma^2} \\
\frac{\partial}{\partial(\sigma^2)}\log f(X_1|\mu,\sigma^2) &= -\frac{1}{2\sigma^2} + \frac{1}{2}\frac{(X_1-\mu)^2}{\sigma^4} \\
\frac{\partial^2}{\partial(\sigma^2)^2}\log f(X_1|\mu,\sigma^2) &= \frac{1}{2\sigma^4} - \frac{(X_1-\mu)^2}{\sigma^6} \\
-\mathsf{E}\left(\frac{\partial^2}{\partial(\sigma^2)^2}\log f(X_1|\mu,\sigma^2)\right) &= -\frac{1}{2\sigma^4} + \frac{\mathsf{E}(X_1-\mu)^2}{\sigma^6} = -\frac{1}{2\sigma^4} + \frac{1}{\sigma^4} = \frac{1}{2\sigma^4}
\end{aligned}
$$

Somit muss für jeden unverzerrten Schätzer $W(\mathbf{X})$ für σ^2 gelten, dass

$$
\mathsf{var}_{\mu,\sigma^2}(W(\mathbf{X})) \geq \frac{2\sigma^4}{n}\,.
$$

Wir haben bereits gezeigt, dass $\mathsf{E}_{\mu,\sigma^2}(S^2) = \sigma^2$ ($\hat{\sigma}^2$ ist nicht unverzerrt) und

$$\mathsf{var}_{\mu,\sigma^2}(S^2) = \frac{2\sigma^4}{n-1}\,.$$

Also wird die Cramér-Rao Schranke von S^2 **nicht erreicht**!

Bemerkungen:

Gibt es nun einen besseren Schätzer oder ist die Schranke gar nicht erreichbar?

Die Schranke ergab sich aus der Verwendung der Cauchy-Schwarz Ungleichung. Somit sind die Bedingungen für die Erreichbarkeit der Schranke die Bedingungen für die Gleichheit in der Cauchy-Schwarz Ungleichung.

Korollar 3.3.2: (Erreichbarkeit) Sei $\mathbf{X} = (X_1, \ldots, X_n)$ eine Zufallsstichprobe mit Dichte $f(x|\theta)$, die den Bedingungen im Satz von Cramér-Rao genügt, und $L(\theta|\mathbf{x})$ die Likelihood Funktion. Falls $W(\mathbf{X})$ ein beliebiger unverzerrter Schätzer für $\tau(\theta)$ ist, also falls $\mathsf{E}_\theta(W(\mathbf{X})) = \tau(\theta)$ gilt, dann erreicht $W(\mathbf{X})$ die Cramér-Rao Schranke genau dann wenn es eine Funktion $a(\theta)$ gibt, für die gilt

$$\frac{\partial}{\partial\theta} \log L(\theta|\mathbf{x}) = a(\theta)\Big(W(\mathbf{x}) - \tau(\theta)\Big)\,.$$

Beispiel 3.3.4 (Fortsetzung): Seien $X_1, \ldots, X_n$ iid Normal(μ, σ^2); schätze σ^2.

$$\begin{aligned}
\log L(\mu, \sigma^2|\mathbf{x}) &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 \\
\frac{\partial}{\partial(\sigma^2)}\log L(\mu, \sigma^2|\mathbf{x}) &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\mu)^2 \\
&= \frac{n}{2\sigma^4}\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2 - \sigma^2\right) .
\end{aligned}$$

Für $W(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$ gilt $nW(\mathbf{X})/\sigma^2 \sim \chi^2_n \quad \Rightarrow \quad \mathsf{E}(W(\mathbf{X})) = \sigma^2$.

Wähle $a(\sigma^2) = n/2\sigma^4$ womit gezeigt ist, dass der Schätzer $W(\mathbf{X})$ die untere Varianzschranke erreicht. Dieser lässt sich jedoch nur berechnen, wenn μ bekannt ist. Bei unbekanntem μ kann diese Schranke nicht erreicht werden.

Offen gebliebene Fragen:

1. Was tun falls $f(x|\theta)$ den Annahmen nicht genügt
   z.B. Beispiel 3.3.3: ist $Y(n+1)/n$ bester unverzerrter Schätzer für θ?

2. Falls die Schranke für zulässige Schätzer nicht erreichbar ist
   z.B. Beispiel 3.3.4: ist S^2 bester unverzerrter Schätzer für σ^2?

Mögliche Antwort: Verwende weiterhin das Konzept der Suffizienz!

Satz 3.3.2: (Rao-Blackwell) Sei W ein beliebiger unverzerrter Schätzer für $\tau(\theta)$, und sei T eine suffiziente Statistik für θ. Definiere $\phi(T) = \mathsf{E}_\theta(W|T)$. Dann gilt $\mathsf{E}_\theta(\phi(T)) = \tau(\theta)$ und $\mathsf{var}_\theta(\phi(T)) \leq \mathsf{var}_\theta(W)$ für alle θ, d.h. $\phi(T)$ ist ein gleichmäßig besserer unverzerrter Schätzer für $\tau(\theta)$.

Satz 3.3.3: (Eindeutigkeit) Falls W ein bester unverzerrter Schätzer für $\tau(\theta)$ ist, dann ist W eindeutig.