Universität Bielefeld
Tanja Pasurek
Ehsan Abedi WiSe 2024/25

Foundations of Statistics

Homework 10

Parameter estimation (Chapter 3)

Exercise 1. (Based on Ch. 3.3-4.)

Let $X_1, ..., X_n$ be an i.i.d. random sample from a uniform distribution $\mathrm{Unif}(a, b)$ with unknown parameters $a, b \in \mathbb{R}$ ($a < b$).

(a) Apply the method of moments using sample mean and sample standard deviation to find estimators $\tilde{a}_n$ and $\tilde{b}_n$ for a and b.
(*Hint:* proceed as described on p. 10 in Ch. 3.3.)

(b) Apply the method of the maximum likelihhod to show that

$$\widehat{a}_n := \min\{X_1, ..., X_n\} \text{ and } \widehat{b}_n := \max\{X_1, ..., X_n\}$$

are the maximum likelihhod estimators for a and b.
(*Hint:* proceed analogously to Exp. 3 in Ch. 3.4.)

(c) Check that $\widehat{a}_n$ and $\widehat{b}_n$ are asymptotically unbiased, that is,

$$\mathbb{E}[\widehat{a}_n] \to a \quad \text{and} \quad \mathbb{E}[\widehat{b}_n] \to b \quad \text{as} \quad n \to \infty.$$

(*Hint:* In HW 6, Ex 3(c), we have already found the distribution of the smallest and largest order statistics of a uniformly distributed random sample. Here you need to compute the corresponding expectations.)

(d) Let $\tau := \int_{-\infty}^{\infty} x^2 f(x)\,\mathrm{d}x$, where $f(x)$ is the PDF of $\mathrm{Unif}(a, b)$. Find the MLE $\hat{\tau}_n$ of τ using the invariance property of MLE.

Exercise 2. (Based on Ch. 3.3-5.)

Let $X_1, ..., X_n$ be an i.i.d. random sample from a Gamma distribution $\Gamma(\alpha, \beta)$ with unknown parameters $\alpha, \beta > 0$ (see the Def. 10 in Ch. 1.6.)

(a) Find the method of moments estimator $\hat{\theta} = (\hat{\alpha}, \hat{\beta})$ of the parameter vector $\theta = (\alpha, \beta)$ (*Hint:* proceed as described on p. 10 in Ch. 3.3.)

(b) Define the log-likelihood function $\ell(\alpha, \beta)$ and write down the system of equations to find its maximum. Could you solve it explicitly?

(c) The following dataset with size $n = 16$ comes from a Gamma distribution:

0.8666	1.3620	1.4664	1.0032	2.0429	2.9602	0.7447	1.0066
1.4548	2.0756	1.0561	1.5033	1.8870	3.9882	1.3323	1.1682

Define the log-likelihood in `R` and use numerical optimization to find ML-estimators. (*Hint:* proceed analogously to p. 14 in Ch. 3.5.) Compare the results with MM-estimators found in **(a)**. (Use `R`'s command `var()` without changing its default denominator).

Exercise 3. (Based on Ch. 3.9.)

Consider a Bernoulli distribution $\mathrm{Ber}(\pi)$ with parameter $\pi \in (0, 1)$. Its PMF can be represented by the following formula

$$f_\pi(x) = \mathbb{P}(X = x) = \begin{cases} \pi^x(1-\pi)^{1-x}, & x \in \{0, 1\}, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Let $x_1, ..., x_n$ be a realization of a random sample $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathrm{Ber}(\pi)$. Calculate the observed Fisher information for this dataset.

(b) Calculate the expected Fisher information for $X_1, ..., X_n \overset{\text{iid}}{\sim} \mathrm{Ber}(\pi)$.

(c) Show that the estimator $\hat{\pi}_n = \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ attains the explicit Cramér–Rao bound.

Exercise 4. (Based on Ch. 3.9.) The goal of this exercise is to study the sampling distribution of ML-estimators for large n.

Let $X_1, ..., X_n$ be an i.i.d. random sample from the density

$$f_\lambda(x) = \frac{\lambda}{2\sqrt{x}} e^{-\lambda\sqrt{x}}, \quad \forall x > 0,$$

for some unknown $\lambda > 0$,

(a) Compute the ML-estimator $\widehat{\lambda}_n$.

(b) Compute the observed Fisher information $J_n(\lambda; x_1, \cdots x_n)$ and the (expected) Fisher information $\mathcal{I}_n(\lambda)$.

(c) Our goal now is to study the empirical distribution of the ML-estimator $\widehat{\lambda}_n$ for large n. Remember that $\widehat{\lambda}_n$ is a random variable. For this task, you will need to create <u>three plots</u> and <u>one table</u> as the final output.

Assume that we know the unknown parameter $\lambda = 0.3$. Apply the inverse transform method (recall Ch. 1.8) to simulate n random samples from the density f_λ and compute $\widehat{\lambda}_n$. Repeat this experiment $K = 10^6$ times for $n \in \{8, 64, 128\}$. Make three histograms with `xlim=c(0.1,0.5)`. To each plot, add the empirical density curve using `density()`, and a vertical line representing the sample mean of $\widehat{\lambda}_n$, both with `col="blue",lty=2`.

We will see in Ch. 3.10 that, for large n, the distrubution of ML-estimators can be approximated by a normal distribution:

$$\hat{\lambda}_n \overset{\text{approx}}{\sim} \mathcal{N}(\lambda, \mathcal{I}_n(\lambda)^{-1})$$

To each plot, add the theoretical normal density, and a vertical line at the theoretical value λ, both with `col="red",lty=1`.

Finally make a table, comparing the sample variance of $\widehat{\lambda}_n$ using `var()` and the theoretical approximation $\mathcal{I}_n(\lambda)^{-1}$.

Exercise 5. (Based on Ch. 3.4 and 3.9.) The goal of this exercise is to study the change in Fisher information under reparametrization.

Let X_i, $i \in \{1, 2, \cdots n\}$ be an i.i.d. random sample from a normal distribution $\mathcal{N}(0, \sigma^2)$ with mean 0 and unknown variance $\sigma^2 > 0$.

(a) Find the Fisher information $\mathcal{I}(\sigma)$ for a single variable X_i considering the standard deviation $\sigma > 0$ as the parameter.

(b) Find the Fisher information $\widetilde{\mathcal{I}}(\theta)$ for a single variable X_i considering the variance $\theta := \sigma^2$ as the parameter. Compare the result with **(a)**.

(c) Let σ be the unknown parameter. Find the ML-estimator $\hat{\sigma}_n$.

(d) Let θ be the unknown parameter. Find the ML-estimator $\hat{\theta}_n$ directly and compare the result with the estimator obtained by the invariance principle.

(e) Show that $\hat{\sigma}_n$ is biased, whereas $\hat{\theta}_n$ is unbiased.

(f) For which of the estimators $\hat{\sigma}_n$, $\hat{\theta}_n$, CLT can be immediately applied to approximate its distribution for large n. Identify it, and apply CLT accordingly.

(g) We will see in Ch. 3.10 that for large enough n, we can approximate

$$\hat{\sigma}_n \overset{\text{approx}}{\sim} \mathcal{N}(\sigma, \mathcal{I}_n(\sigma)^{-1})$$
$$\hat{\theta}_n \overset{\text{approx}}{\sim} \mathcal{N}(\theta, \widetilde{\mathcal{I}}_n(\theta)^{-1})$$

Compute $\mathcal{I}_n(\sigma)^{-1}$ and $\widetilde{\mathcal{I}}_n(\theta)^{-1}$. Compare the results with **(f)**.

(h) Now let X be a random variable for which the PDF or the PMF is given by $f_\phi(x)$, depending on a parameter $\phi \in \mathbb{R}$. Let $\mathcal{I}(\phi)$ denote the Fisher information of f_ϕ. Suppose now that the parameter ϕ is replaced by a new parameter θ, where $\phi = g(\theta)$ for some differentiable function $g : \mathbb{R} \to \mathbb{R}$. Let $\widetilde{\mathcal{I}}(\theta)$ denote the new Fisher information with respect to the parameter θ. Show that

$$\widetilde{\mathcal{I}}(\theta) = \left[g'(\theta)\right]^2 \mathcal{I}\left[g(\theta)\right].$$

(i) Apply the general result above to **(a)** and compare it with **(b)**.

Have a wonderful holiday season!