Statistical inference practice

Joris LIMONIER

December 17, 2021

Contents

1 Homework for October 8, 2021

1.1 Chapter 6 - Exercise 2

Let $X_1, \dots, X_n$ be i.i.d. random variables with distribution $\mathcal{U}(0, \theta)$.
Let $\hat{\theta}_n = \max(X_1, \dots, X_n)$.

What is the bias of $\hat{\theta}_n$?
$\hat{\theta}_n$ is unbiased if $\mathbb{E}[\hat{\theta}_n] = \theta$.
We compute $F_{\hat{\theta}_n}$, the CDF of $\hat{\theta}_n$:

$$\begin{aligned}
F_{\hat{\theta}_n}(x) &= \mathbb{P}(\hat{\theta} \leq x) \\
&= \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) \\
&= \mathbb{P}(X_1 \leq x) \dots \mathbb{P}(X_n \leq x) && \textit{(independence)} \\
&= [\mathbb{P}(X_1 \leq x)]^n && \textit{(identity of distribution)} \\
&= \begin{cases} 0 & x < 0 \\ \left[\frac{x}{\theta}\right]^n & 0 \leq x \leq \theta \\ 1 & x > \theta \end{cases}
\end{aligned}$$

Then $f_{\hat{\theta}_n}$, the PDF of $\hat{\theta}_n$ is given by:

$$\begin{aligned}
f_{\hat{\theta}_n}(x) &= \frac{d}{dx} F_{\hat{\theta}_n}(x) \\
&= \begin{cases} \frac{nx^{n-1}}{\theta^n} & x \in [0, \theta] \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

Hence the expected value is given by:

$$\begin{aligned}
\mathbb{E}[\hat{\theta}_n] &= \int_{-\infty}^{+\infty} t f(t) dt \\
&= \frac{n}{\theta^n} \int_0^\theta t^n dt \\
&= \frac{n}{\theta^n (n+1)} \left[t^{n+1}\right]_0^\theta \\
&= \frac{n}{\theta^n (n+1)} \left[\theta^{n+1} - 0\right] \\
&= \frac{n\theta}{n+1}
\end{aligned}$$

Since $\mathbb{E}[\hat{\theta}_n] = n\theta \neq \theta$, we have that $\hat{\theta}_n$ is not an unbiased estimator for θ. However, $\mathbb{E}[\hat{\theta}_n] \xrightarrow[n\to\infty]{} \theta$, therefore $\hat{\theta}_n$ is asymptotically unbiased.

What is SE, the standard error of $\hat{\theta}_n$?

$$\begin{aligned}
SE &= SE(\hat{\theta}_n) \\
&= \sqrt{Var(\hat{\theta}_n)} \\
&= \sqrt{\mathbb{E}[\hat{\theta}_n^2] - \mathbb{E}[\hat{\theta}_n]^2} \qquad (1)
\end{aligned}$$

We need to find $\mathbb{E}[\hat{\theta}_n^2]$.

$$\begin{aligned}
\mathbb{E}[\hat{\theta}_n^2] &:= \int_{-\infty}^{+\infty} t^2 f(t) dt \\
&= \frac{n}{\theta^n} \int_0^\theta t^{n+1} dt \\
&= \frac{n}{\theta^n (n+2)} \left[t^{n+2}\right]_0^\theta \\
&= \frac{n}{\theta^n (n+2)} \left[\theta^{n+2} - 0\right] \\
&= \frac{n\theta^2}{n+2}
\end{aligned}$$

Then (1) becomes:

$$
\begin{aligned}
SE &= \sqrt{\mathbb{E}[\hat{\theta}_n^2] - \mathbb{E}[\hat{\theta}_n]^2} \\
&= \sqrt{\frac{n\theta^2}{n+2} - \left[\frac{n\theta}{n+1}\right]^2} \\
&= \sqrt{n\theta^2 \left[\frac{1}{n+2} - \frac{n}{(n+1)^2}\right]} \\
&= \sqrt{n\theta^2 \left[\frac{1}{n+2} - \frac{n}{n^2+2n+1}\right]} \\
&= \sqrt{n\theta^2 \left[\frac{n^2+2n+1}{(n+2)(n^2+2n+1)} - \frac{n(n+2)}{(n^2+2n+1)(n+2)}\right]} \\
&= \sqrt{n\theta^2 \frac{n^2+2n+1-n(n+2)}{(n^2+2n+1)(n+2)}} \\
&= \sqrt{\frac{n\theta^2}{(n^2+2n+1)(n+2)}} \\
&= \frac{\theta}{n+1}\sqrt{\frac{n}{n+2}}
\end{aligned}
$$

What is MSE, the Mean-Square Error of $\hat{\theta}_n$?

The MSE is given by:

$$
\begin{aligned}
MSE &:= bias^2(\hat{\theta}_n) + Var(\hat{\theta}_n) \\
&= \left[\mathbb{E}[\hat{\theta}] - \theta\right]^2 + \frac{n\theta^2}{(n+1)^2(n+2)} \\
&= \left[\frac{n\theta}{n+1} - \theta\right]^2 + \frac{n\theta^2}{(n+1)^2(n+2)} \\
&= \left[\frac{-\theta}{n+1}\right]^2 + \frac{n\theta^2}{(n+1)^2(n+2)} \\
&= \frac{\theta^2}{(n+1)^2}\left[1 + \frac{n}{n+2}\right] \\
&= \frac{\theta^2}{(n+1)^2}\frac{2n+2}{n+2} \\
&= \frac{2\theta^2}{(n+1)(n+2)}
\end{aligned}
$$

1.2 Chapter 6 - Exercise 3

Let $X_1, \ldots, X_n$ be i.i.d. random variables with distribution $\mathcal{U}(0, \theta)$.
Let $\hat{\theta}_n := 2\bar{X}_n$.

What is the bias of $\hat{\theta}_n$?

$\hat{\theta}_n$ is unbiased if $\mathbb{E}[\hat{\theta}_n] = \theta$.
We compute $\mathbb{E}[\hat{\theta}_n]$:

$$\begin{aligned}
\mathbb{E}[\hat{\theta}_n] &= \mathbb{E}\left[2\bar{X}_n\right] \\
&= \mathbb{E}\left[2\frac{X_1 + \ldots + X_n}{n}\right] \\
&= \frac{2}{n}\mathbb{E}\left[X_1 + \ldots + X_n\right] \\
&= \frac{2}{n}\mathbb{E}\left[X_1\right] + \ldots + \mathbb{E}\left[X_n\right] \\
&= 2\mathbb{E}\left[X_1\right] \\
&= \theta
\end{aligned}$$

Therefore $\hat{\theta}$ is unbiased.

What is SE, the standard error of $\hat{\theta}_n$?

$$\begin{aligned}
SE &= SE(\hat{\theta}_n) \\
&= \sqrt{Var(\hat{\theta}_n)} \\
&= \sqrt{Var(2\bar{X}_n)} \\
&= \frac{2}{n}\sqrt{Var\left(X_1 + \ldots + X_n\right)} \\
&= \frac{2}{n}\sqrt{Var\left(X_1\right) + \ldots + Var\left(X_n\right)} && \textit{(The } X_i \textit{ are i.i.d.)} \\
&= \frac{2}{n}\sqrt{\frac{n\theta^2}{12}} \\
&= \frac{2\theta}{2\sqrt{3n}} \\
&= \frac{\theta}{\sqrt{3n}}
\end{aligned}$$

What is MSE, the Mean-Square Error of $\hat{\theta}_n$?
The MSE is given by:

$$\begin{aligned} MSE &:= bias^2(\hat{\theta}_n) + Var(\hat{\theta}_n) \\ &= \underbrace{\left[\mathbb{E}[\hat{\theta}] - \theta\right]^2}_{=0} + \frac{\theta^2}{3n} \\ &= \frac{\theta^2}{3n} \end{aligned}$$

2 Homework for October 20, 2021

2.1 Chapter 7 - Exercise 2

Let $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$ and let $Y_1, \ldots, Y_m \sim \text{Bernoulli}(q)$.

- Find the plug-in estimator and estimated standard error for p.
- Find an approximate 90 percent confidence interval for p.
- Find the plug-in estimator and estimated standard error for $p - q$.
- Find an approximate 90 percent confidence interval for $p - q$.

Find the plug-in estimator and estimated standard error for p.
Let ϕ be the plug-in estimator for p, it is given by:

$$\phi = \mathbb{E}[X_i], \ i = 1, \ldots, n$$

and

$$\hat{\phi} = \mathbb{E}[Z], \text{ with } \mathbb{P}(Z = X_i \mid X_1, \ldots, X_n) = \frac{1}{n}$$

$$\begin{aligned} \hat{\phi} &= \mathbb{E}[Z] \\ &= \sum_{i=1}^{n} X_i \mathbb{P}(Z = X_i) \\ &= \sum_{i=1}^{n} \frac{1}{n} X_i \\ &= \overline{X} \end{aligned}$$

and the standard error se is given by:

$$se(\phi) = \sqrt{Var(\phi)}$$

$$= \sqrt{Var\left(\overline{X}\right)}$$

$$= \sqrt{\frac{p(1-p)}{n}}$$

Find an approximate 90 percent confidence interval for p.

We know that 90% ($i.e.$ $\alpha = 0.05$) confidence intervals are of the following form:

$$\overline{X} \pm z_{\alpha/2} se(p_{pin})$$
$$= \overline{X} \pm 1.645 \sqrt{\frac{1}{n}\left[\sum_{i=1}^{n} X_i^2\right] - \overline{X}^2}$$

Find the plug-in estimator and estimated standard error for $p - q$.

Let Π be the plug-in estimator for $p - q$ and $\chi \in \{X_1, \dots, X_n, Y_1, \dots, Y_m\}$

$$\begin{aligned} \mathbb{P}(\Pi = \chi) &= \frac{1}{\#\{X_1, \dots, X_n\} + \#\{Y_1, \dots, Y_m\}} \\ &= \frac{1}{m+n} \end{aligned}$$

$$\begin{aligned} se(\Pi) &= \sqrt{Var(\Pi)} \\ &= \sqrt{\mathbb{E}[\Pi^2] - \mathbb{E}[\Pi]^2} \\ &= \sqrt{\left[\sum_{i=1}^{n} \chi^2 \mathbb{P}\{\Pi = \chi\}\right] - \left[\sum_{i=1}^{n} \chi \mathbb{P}\{\Pi = \chi\}\right]^2} \\ &= \sqrt{\left[\sum_{i=1}^{n} \chi^2 \frac{1}{m+n}\right] - \left[\sum_{i=1}^{n} \chi \frac{1}{m+n}\right]^2} \\ &= \sqrt{\frac{1}{m+n}\left[\sum_{i=1}^{n} \chi^2\right] - \overline{\chi}^2} \end{aligned}$$

Find an approximate 90 percent confidence interval for $p - q$.

We know that 90% (*i.e.* $\alpha = 0.05$) confidence intervals are of the following form:

$$\begin{aligned}
&\overline{\chi} \pm z_{\alpha/2} se(p_{pin}) \\
&= \overline{\chi} \pm 1.645 \sqrt{\frac{1}{m+n}\left[\sum_{i=1}^{n} \chi^2\right] - \overline{\chi}^2}
\end{aligned}$$

2.2 Chapter 7 - Exercise 5

Let x and y be two distinct points. Find $\text{Cov}\left(\hat{F}_n(x), \hat{F}_n(y)\right)$.

2.3 Chapter 7 - Exercise 6

3 Homework for October 29

3.1 Custom exercise

Let $N = 50$, $Y_1, \ldots, Y_n$ are i.i.d. $\mathcal{N}(0,1)$. Let $X_i = e^{Y_i}$.
Let $\theta = \text{skewness}(X) = (e+2)\sqrt{e-1}$ (X is log normal distributed)

Compute the 3 types of normal confidence intervals for θ.

Repeat the experiment to check how often θ belongs to the confidence intervals.

4 Homework for November 5, 2021

4.1 Chapter 9 - Exercise 1

Let $X_1, \ldots, X_n \sim \text{Gamma}(\alpha, \beta)$. Find the method of moments estimator for α and β. From tables, we have that:

$$\begin{cases}
\mathbb{E}[X_i] = \frac{\alpha}{\beta} \\
Var(X_i) = \frac{\alpha}{\beta^2}
\end{cases}$$

Equating with empirical expected value and empirical variance respectively.

$$\begin{cases} \frac{1}{n}\sum_{i=1}^n X_i = \mathbb{E}[X_i] \\ \frac{1}{n}\sum_{i=1}^n (\bar{X} - X_i)^2 = Var(X_i) \end{cases}$$

$$\implies \begin{cases} \frac{1}{n}\sum_{i=1}^n X_i = \frac{\hat{\alpha}}{\hat{\beta}} \\ \frac{1}{n}\sum_{i=1}^n (\bar{X} - X_i)^2 = \frac{\hat{\alpha}}{\hat{\beta}^2} \end{cases}$$

$$\implies \begin{cases} \hat{\alpha} = \frac{\hat{\beta}}{n}\sum_{i=1}^n X_i \\ \frac{1}{n}\sum_{i=1}^n (\bar{X} - X_i)^2 = \frac{\hat{\alpha}}{\hat{\beta}^2} \end{cases}$$

$$\implies \begin{cases} \hat{\alpha} = \frac{\hat{\beta}}{n}\sum_{i=1}^n X_i \\ \frac{1}{n}\sum_{i=1}^n (\bar{X} - X_i)^2 = \frac{1}{\hat{\beta}^2}\frac{\hat{\beta}}{n}\sum_{i=1}^n X_i \end{cases}$$

$$\implies \begin{cases} \hat{\alpha} = \frac{\hat{\beta}}{n}\sum_{i=1}^n X_i \\ \frac{1}{n}\sum_{i=1}^n (\bar{X} - X_i)^2 = \frac{1}{\hat{\beta} n}\sum_{i=1}^n X_i \end{cases}$$

$$\implies \begin{cases} \hat{\alpha} = \frac{\beta}{n}\sum_{i=1}^n X_i \\ \hat{\beta} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n (\bar{X} - X_i)^2} \end{cases}$$

$$\implies \begin{cases} \hat{\alpha} = \frac{1}{n}\frac{\left[\sum_{i=1}^n X_i\right]^2}{\sum_{i=1}^n (\bar{X} - X_i)^2} \\ \hat{\beta} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n (\bar{X} - X_i)^2} \end{cases}$$

4.2 Chapter 9 - Exercise 2

Let $X_1, \ldots, X_n \sim$ Uniform (a, b) where a and b are unknown parameters and $a < b$.

- (a) Find the method of moments estimators for a and b.
- (b) Find the MLE $\hat{a}$ and $\hat{b}$.
- (c) Let $\tau = \int x dF(x)$. Find the MLE of τ.
- (d) Let $\hat{\tau}$ be the MLE of τ. Let $\widetilde{\tau}$ be the nonparametric plug-in estimator of $\tau = \int x dF(x)$. Suppose that $a = 1, b = 3$, and $n = 10$. Find the MSE of $\hat{\tau}$ by simulation. Find the MSE of $\widetilde{\tau}$ analytically. Compare.

(a) Find the method of moments estimators for a and b.

We compute the first order of moments:

$$\begin{aligned}
\mathbb{E}[X_1] &= \int_a^b x f(x) dx \\
&= \frac{1}{b-a}\int_a^b x dx \\
&= \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b \\
&= \frac{1}{b-a}\frac{b^2-a^2}{2} \\
&= \frac{a+b}{2} \qquad (2)
\end{aligned}$$

Now compute the second order of moments:

$$\begin{aligned}
\mathbb{E}[X_1^2] &= \int_a^b x^2 f(x) dx \\
&= \int_a^b x^2 \frac{1}{b-a} dx \\
&= \frac{1}{b-a}\left[\frac{x^3}{3}\right]_a^b \\
&= \frac{1}{b-a}\left[\frac{b^3-a^3}{3}\right] \\
&= \frac{1}{b-a}\left[\frac{(b-a)\left(a^2+ab+b^2\right)}{3}\right] \\
&= \frac{a^2+ab+b^2}{3} \qquad (3)
\end{aligned}$$

Now on the one hand we equate (2) with $\hat{\mu_1} := \frac{1}{n}\sum_{i=1}^n X_i$. On the other hand, we equate (3) with $\hat{\mu_2} := \frac{1}{n}\sum_{i=1}^n X_i^2$. Therefore we get a system of equations:

$$\begin{cases} \hat{\mu}_1 = \frac{a+b}{2} \\ \hat{\mu}_2 = \frac{a^2+ab+b^2}{3} \end{cases}$$

$$\begin{cases} b = 2\hat{\mu}_1 - a \\ 3\hat{\mu}_2 = a^2 + a\left[2\hat{\mu}_1 - a\right] + \left[2\hat{\mu}_1 - a\right]^2 \end{cases}$$

$$\begin{cases} b = 2\hat{\mu}_1 - a \\ 3\hat{\mu}_2 = a^2 + 2\hat{\mu}_1 a - a^2 + 4\hat{\mu}_1^2 - 4\hat{\mu}_1 a + a^2 \end{cases}$$

$$\begin{cases} b = 2\hat{\mu}_1 - a \\ a^2 - 2\hat{\mu}_1 a + (4\hat{\mu}_1^2 - 3\hat{\mu}_2) = 0 \end{cases} \qquad (4)$$

Then, the second equation of (4) yields:

$$\begin{aligned}
a_1 &= \frac{2\hat{\mu}_1 - \sqrt{(2\hat{\mu}_1)^2 - 4\left(4\hat{\mu}_1^2 - 3\hat{\mu}_2\right)}}{2} \\
&= \hat{\mu}_1 - \sqrt{\hat{\mu}_1^2 - 4\hat{\mu}_1^2 + 3\hat{\mu}_2} \\
&= \hat{\mu}_1 - \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)}
\end{aligned}$$

and

$$\begin{aligned}
a_1 &= \frac{2\hat{\mu}_1 + \sqrt{(2\hat{\mu}_1)^2 - 4\left(4\hat{\mu}_1^2 - 3\hat{\mu}_2\right)}}{2} \\
&= \hat{\mu}_1 + \sqrt{\hat{\mu}_1^2 - 4\hat{\mu}_1^2 + 3\hat{\mu}_2} \\
&= \hat{\mu}_1 + \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)}
\end{aligned}$$

Let b_1, b_2 be associated with a_1, a_2 respectively. Then the first equation of (4) becomes:

$$\begin{cases} b_1 := 2\hat{\mu}_1 - \left(\hat{\mu}_1 - \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)}\right) \\ b_2 := 2\hat{\mu}_1 - \left(\hat{\mu}_1 + \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)}\right) \end{cases}$$
$$\begin{cases} b_1 := \hat{\mu}_1 + \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)} \\ b_2 := \hat{\mu}_1 - \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)} \end{cases}$$

Since $b_2 > a_2$, which is impossible by the exercise, we have that the method of moment estimators are:

$$\begin{cases} a = \hat{\mu}_1 - \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)} \\ b = \hat{\mu}_1 + \sqrt{3\left(\hat{\mu}_2 - \hat{\mu}_1^2\right)} \end{cases}$$

(b) Find the MLE $\hat{a}$ and $\hat{b}$.

Let $\theta := (a, b) \in \Theta \subseteq \mathbb{R}^2$. We define $\mathcal{L}$ the likelihood function as follows:

$$\begin{aligned}
\mathcal{L}(\theta) &= f(X_1, \ldots, X_n \mid \theta) \\
&= \prod_{i=1}^{n} f(X_i \mid \theta)
\end{aligned}$$

Now we want to find $\hat{\theta} := (\hat{a}, \hat{b})$ the argument maximizing the likelihood function

$$\begin{aligned}
\hat{\theta} &:= (\hat{a}, \hat{b}) \\
&:= \arg\max_{\Theta} \mathcal{L}(\theta) \\
&= \arg\max_{\Theta} \log \mathcal{L}(\theta)
\end{aligned}$$

therefore we have

$$\begin{aligned}
\log \mathcal{L}(\theta) &= \log \prod_{i=1}^{n} f(X_i \mid \theta) \\
&= \sum_{i=1}^{n} \log f(X_i \mid (a, b)) \\
&= \sum_{i=1}^{n} \log \frac{1}{b-a} \\
&= -n \log(b-a)
\end{aligned}$$

hence

$$\begin{cases} \frac{\partial}{\partial a} \log \mathcal{L}(\theta) = \frac{n}{b-a} \\ \frac{\partial}{\partial b} \log \mathcal{L}(\theta) = -\frac{n}{b-a} \end{cases}$$

Now we have that

$$\begin{aligned}
&\begin{cases} \frac{\partial}{\partial a} \log \mathcal{L}(\theta) > 0 \\ \frac{\partial}{\partial b} \log \mathcal{L}(\theta) < 0 \end{cases} \\
\implies &\begin{cases} \log \mathcal{L}(\theta) \text{ is increasing with respect to } a \\ \log \mathcal{L}(\theta) \text{ is decreasing with respect to } b \end{cases} \\
\implies &\begin{cases} \hat{a} = \min\{X_1, \ldots, X_n\} \\ \hat{b} = \max\{X_1, \ldots, X_n\} \end{cases}
\end{aligned}$$

(c) Let $\tau = \int x dF(x)$. Find the MLE of τ.

Let $\theta := \tau \in \Theta \subseteq \mathbb{R}$. We have the following:

$$\begin{aligned}
\tau &= \int x dF(x) \\
&= \int x \frac{dF(x)}{dx} dx \\
&= \int x f(x) dx \\
&= \mathbb{E}[X_1] \\
&= \frac{a+b}{2}
\end{aligned}$$

Hence

$$\hat{\tau} = \frac{\hat{a} + \hat{b}}{2}$$

(d) Let $\hat{\tau}$ be the MLE of τ. Let $\widetilde{\tau}$ be the nonparametric plug-in estimator of $\tau = \int x dF(x)$. Suppose that $a = 1, b = 3$, and $n = 10$. Find the MSE of $\hat{\tau}$ by simulation. Find the MSE of $\widetilde{\tau}$ analytically. Compare.

The MSE is defined by:

$$
\begin{aligned}
MSE(\tilde{\tau}) &:= Var(\tilde{\tau}) + bias^2(\tilde{\tau}) \\
&= \mathbb{E}[\tilde{\tau}^2] - \mathbb{E}[\tilde{\tau}]^2 + [\mathbb{E}[\tilde{\tau}] - \tilde{\tau}]^2 \\
&= \mathbb{E}[\bar{X}^2] - \mathbb{E}[\bar{X}]^2 + \underbrace{[\mathbb{E}[\bar{X}] - \bar{X}]^2}_{=0} \\
&= \frac{1}{n}\left[\frac{a^2+ab+b^2}{3} - \left[\frac{a+b}{2}\right]^2\right] \\
&= \frac{1}{n}\left[\frac{a^2+ab+b^2}{3} - \frac{a^2+2ab+b^2}{4}\right] \\
&= \frac{1}{n}\left[\frac{a^2+ab+b^2-3a^2-6ab-3b^2}{12}\right] \\
&= \frac{1}{n}\left[\frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12}\right] \\
&= \frac{1}{n}\left[\frac{a^2-2ab+b^2}{12}\right] \\
&= \frac{1}{n}\left[\frac{(a-b)^2}{12}\right] \\
&= \frac{(a-b)^2}{12n}
\end{aligned}
$$

5 Preparation for mid-term

5.1 Chapter 9 - Exercises 5

Let $X_1, \ldots, X_n \sim$ Poisson (λ). Find the method of moments estimator, the maximum likelihood estimator and the Fisher information $I(\lambda)$.

We know that the first moment for a Poisson distribution $\mu_1 = \lambda$. We want to evaluate $\hat{\theta} := \lambda$. We set $\mu_1 = \bar{X}$, which gives:

$$\mu_1 = \bar{X} \implies \hat{\theta} = \bar{X}$$

Now we want to find the maximum likelihood estimator.

$$
\begin{aligned}
\mathcal{L}(\theta) &= f(X_1, \ldots, X_n \mid \theta) \\
&= \prod_{i=1}^{n} f(X_i \mid \theta) \\
&= \prod_{i=1}^{n} \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}
$$

[NOT FINISHED]

5.2 Chapter 9 - Exercises 6

Let $X_1, \ldots, X_n \sim N(\theta, 1)$. Define

$$Y_i = \begin{cases} 1 & \textbf{if } X_i > 0 \\ 0 & \textbf{if } X_i \leq 0 \end{cases}$$

Let $\psi = \mathbb{P}(Y_1 = 1)$.

(a) Find the maximum likelihood estimator $\widehat{\psi}$ of ψ.

(b) Find an approximate 95 percent confidence interval for ψ.

(c) Define $\widetilde{\psi} = (1/n)\sum_i Y_i$. Show that $\widetilde{\psi}$ is a consistent estimator of ψ.

(d) Compute the asymptotic relative efficiency of $\widetilde{\psi}$ to $\widehat{\psi}$. Hint: Use the delta method to get the standard error of the MLE. Then compute the standard error (i.e. the standard deviation) of $\widetilde{\psi}$.

(e) Suppose that the data are not really normal. Show that $\widehat{\psi}$ is not consistent. What, if anything, does $\widehat{\psi}$ converge to?

5.3 Chapter 9 - Examples 20

5.4 Chapter 9 - Examples 21

Let $X_1, \ldots, X_n \sim N\left(\theta, \sigma^2\right)$ where σ^2 is known. The score function is $s(X;\theta) = (X - \theta)/\sigma^2$ and $s'(X;\theta) = -1/\sigma^2$ so that $I_1(\theta) = 1/\sigma^2$. The MLE is $\widehat{\theta}_n = \bar{X}_n$. According to Theorem 9.18, $\bar{X}_n \approx N\left(\theta, \sigma^2/n\right)$. In this case, the Normal approximation is actually exact.

5.5 Chapter 9 - Examples 22

Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$. Then $\widehat{\lambda}_n = \bar{X}_n$ and some calculations show that $I_1(\lambda) = 1/\lambda$, so

$$\widehat{\text{se}} = \frac{1}{\sqrt{nI\left(\widehat{\lambda}_n\right)}} = \sqrt{\frac{\widehat{\lambda}_n}{n}}$$

Therefore, an approximate $1 - \alpha$ confidence interval for λ is $\widehat{\lambda}_n \pm z_{\alpha/2}\sqrt{\widehat{\lambda}_n/n}$.

6 In class exercise December 3, 2021

6.1 Exercise 1

Let $X_1, \dots, X_N \sim \mathcal{N}(\mu, \sigma^2)$ with μ given.

i) Compute $\hat{\sigma}_{ML}$ and estimator $se(\hat{\sigma}_{ML})$

$$
\begin{aligned}
\log \mathcal{L}(\sigma^2) &= \sum_{i=1}^{N} \log f_\sigma(X_i) \\
&= \sum_{i=1}^{N} \left[\log \frac{1}{\sigma\sqrt{2\pi}} \right] e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \\
&= \left[N \log \frac{1}{\sigma\sqrt{2\pi}} \right] - \frac{1}{2} \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^2 \\
&= -N(\log \sigma\sqrt{2\pi}) - \frac{1}{2} \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^2 \\
&= -N(\log \sigma + \log \sqrt{2\pi}) - \frac{1}{2} \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^2
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial}{\partial \sigma} \log \mathcal{L}(\sigma^2) &= \frac{\partial}{\partial \sigma} \left[-N(\log \sigma + \log \sqrt{2\pi}) - \frac{1}{2} \sum_{i=1}^{N} \left(\frac{X_i - \mu}{\sigma} \right)^2 \right] \\
&= -\frac{N}{\sigma} - \frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^2 \frac{\partial}{\partial \sigma} \sigma^{-2} \\
&= -\frac{N}{\sigma} - \frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^2 (-2)\sigma^{-3} \\
&= -\frac{N}{\sigma} + \sum_{i=1}^{n} (X_i - \mu)^2 \sigma^{-3}
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial}{\partial \hat{\sigma}_{ML}} \log \mathcal{L}(\hat{\sigma}^2_{ML}) = 0 \\
\implies & -\frac{N}{\hat{\sigma}_{ML}} + \sum_{i=1}^{n} (X_i - \mu)^2 \, \hat{\sigma}_{ML}^{-3} = 0 \\
\implies & \hat{\sigma}_{ML}^{-2} = \frac{N}{\sum_{i=1}^{n} (X_i - \mu)^2} \\
\implies & \hat{\sigma}_{ML}^{2} = \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{N} \\
\implies & \hat{\sigma}_{ML} = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{N}}
\end{aligned}
$$

$$
se(\hat{\sigma}_{ML}) = \frac{1}{I_N(\sigma)}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial \sigma^2} \left[\log \mathcal{L}(\sigma)\right] &= \frac{\partial}{\partial \sigma} \left[-\frac{N}{\sigma} + \sum_{i=1}^{n} (X_i - \mu)^2 \, \sigma^{-3} \right] \\
&= \frac{N}{\sigma^2} - 3 \sum_{i=1}^{n} (X_i - \mu)^2 \, \sigma^{-4}
\end{aligned}
$$

$$
\begin{aligned}
I_N(\sigma) &= -\mathbb{E}\left[\frac{\partial^2}{\partial \sigma^2} \log \mathcal{L}(\sigma) \right] \\
&= -\mathbb{E}\left[\frac{N}{\sigma^2} - 3 \sum_{i=1}^{n} (X_i - \mu)^2 \, \sigma^{-4} \right] \\
&= -\frac{N}{\sigma^2} - \frac{-3}{\sigma^4} \sum_{i=1}^{n} \mathbb{E}\left[(X_i - \mu)^2 \right] \\
&= -\frac{N}{\sigma^2} - \frac{-3}{\sigma^4} \sum_{i=1}^{n} Var\left[X_i^2 \right] \\
&= -\frac{N}{\sigma^2} - \frac{-3}{\sigma^4} \sum_{i=1}^{n} \sigma^2 \\
&= -\frac{N}{\sigma^2} - \frac{-3N}{\sigma^2} \\
&= \frac{2N}{\sigma^2}
\end{aligned}
$$

Thus

$$se(\hat{\sigma}_{ML}) = \frac{1}{\sqrt{I_N(\sigma)}} = \frac{\sigma}{\sqrt{2N}}$$

ii) Compute $\hat{\sigma}_{ML}$ and estimator $se(\hat{\sigma}_{ML})$

7 Homework for December 10, 2021

7.1 Chapter 9 - Exercise 3

3. Let $X_1, \ldots, X_n \sim N\left(\mu, \sigma^2\right)$.
Let τ be the .95 percentile, i.e. $\mathbb{P}(X < \tau) = 0.95$
(a) Find the MLE of τ.
(b) Find an expression for an approximate $1-\alpha$ confidence interval for τ.
(c) Suppose the data are:

3.23	−2.50	1.88	−0.68	4.43	0.17
1.03	−0.07	−0.01	0.76	1.76	3.18
0.33	−0.31	0.30	−0.61	1.52	5.43
1.54	2.28	0.42	2.33	−1.03	4.00
0.39					

Find the MLE $\widehat{\tau}$. Find the standard error using the delta method. Find the standard error using the parametric bootstrap.

7.1.1 (a)

$$\begin{aligned}
&\mathbb{P}\left(X_i < \tau\right) = 0.95 \\
&\Longrightarrow \mathbb{P}\left(\underbrace{\frac{X_i - \mu}{\sigma}}_{:=Z\sim\mathcal{N}(0,1)} < \frac{\tau - \mu}{\sigma}\right) = 0.95 \\
&\Longrightarrow \underbrace{\mathbb{P}\left(Z < \frac{\tau-\mu}{\sigma}\right)}_{:=\Phi\left(\frac{\tau-\mu}{\sigma}\right)} = 0.95 \\
&\Longrightarrow \Phi\left(\frac{\tau-\mu}{\sigma}\right) = 0.95 \\
&\Longrightarrow \frac{\tau-\mu}{\sigma} = \Phi^{-1}\left(0.95\right) \\
&\Longrightarrow \tau = \sigma\Phi^{-1}\left(0.95\right) + \mu \\
&\Longrightarrow \hat{\tau}_{ML} = \hat{\sigma}_{ML}\Phi^{-1}\left(0.95\right) + \hat{\mu}_{ML}
\end{aligned}$$

7.1.2 (b)

Let $\tau = \sigma\Phi^{-1}(0.95) + \mu := g(\mu, \sigma)$, so $\hat{\tau}_{ML} = g(\hat{\mu}_{ML}, \hat{\sigma}_{ML})$, by equivariance of ML.

Therefore:

$$\begin{aligned}\hat{\tau}_{ML} - \tau &= g(\hat{\mu}_{ML}, \hat{\sigma}_{ML}) - g(\mu, \sigma) \\ &\approx \nabla g(\mu, \sigma)\left[(\hat{\mu}_{ML}, \hat{\sigma}_{ML}) - (\mu, \sigma)\right]\end{aligned}$$

7.1.3 (c)

8 Simple Linear Regression Exercises

8.1 Exercise 1

8.1.1 (a)

Recall how the OLS estimator

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$$

is computed and show that the fitted values $\hat{Y}$ are obtained as

$$\hat{Y} = X\left(X^T X\right)^{-1} X^T Y$$

8.1.2 (b)

(b) Show that that $\hat{\beta}$ is an unbiased estimator of β and compute its variance.

$$\begin{aligned}\mathbb{E}[\hat{\beta}] &= \mathbb{E}\left[\left(X^T X\right)^{-1} X^T Y\right] \\ &= \left(X^T X\right)^{-1} X^T \mathbb{E}\left[Y\right] \\ &= \left(X^T X\right)^{-1} X^T \mathbb{E}\left[X\beta + \varepsilon\right] \\ &= \left(X^T X\right)^{-1} X^T X \mathbb{E}\left[\beta\right] \\ &= \left(X^T X\right)^{-1} X^T X \mathbb{E}\left[\beta\right] \\ &= \beta\end{aligned}$$

So $\hat{\beta}$ is unbiased.

$$
\begin{aligned}
Var(\hat{\beta}) &= \mathbb{E}\left[\left(\hat{\beta} - \mathbb{E}[\hat{\beta}]\right)\left(\hat{\beta} - \mathbb{E}[\hat{\beta}]\right)^T\right] \\
&= \mathbb{E}\left[\left(\hat{\beta} - \beta\right)\left(\hat{\beta} - \beta\right)^T\right] \\
&= \mathbb{E}\left[\left(\left(X^T X\right)^{-1} X^T \varepsilon\right)\left(\left(X^T X\right)^{-1} X^T \varepsilon\right)^T\right] \\
&= \mathbb{E}\left[\left(X^T X\right)^{-1} X^T \varepsilon \varepsilon^T X \left(\left(X^T X\right)^{-1}\right)^T\right] \\
&= \left(X^T X\right)^{-1} X^T \cdot \mathbb{E}\left[\varepsilon \varepsilon^T\right] \cdot X \left(\left(X^T X\right)^{-1}\right)^T \\
&= \left(X^T X\right)^{-1} X^T \cdot \sigma^2 I_n \cdot X \left(X^T X\right)^{-1} \\
&= \sigma^2 \left(X^T X\right)^{-1}
\end{aligned}
$$

8.1.3 (c)

We now introduce the matrix $P_1 := X\left(X^T X\right)^{-1} X^T$ and show that it is a projection matrix [1] In order to do that

1. Show that P_1 is symmetric.

$$
\begin{aligned}
P_1^T &= \left[X\left(X^T X\right)^{-1} X^T\right]^T \\
&= X\left(X^T X\right)^{-1^T} X^T \\
&= X\left(X^T X\right)^{-1} X^T \\
&= P_1
\end{aligned}
$$

2. Show that $P_1^2 = P_1$.

$$
\begin{aligned}
P_1^2 &= \left[X\left(X^T X\right)^{-1} X^T\right]^2 \\
&= X\left(X^T X\right)^{-1} X^T X \left(X^T X\right)^{-1} X^T \\
&= X\left(X^T X\right)^{-1} X^T \\
&= P_1
\end{aligned}
$$

8.1.4 (d)

Prove that $X\hat{\beta} = X\beta + P_1\varepsilon$

$$\begin{aligned} X\hat{\beta} &= X\left(X^TX\right)^{-1}X^TY \\ &= X\left(X^TX\right)^{-1}X^T\left(X\beta+\varepsilon\right) \\ &= X\beta + X\left(X^TX\right)^{-1}X^T\varepsilon \\ &= X\beta + P_1\varepsilon \end{aligned}$$

8.1.5 (e)

Show that the OLS residuals $\hat{\varepsilon} := Y - X\hat{\beta}$ are obtained as

$$\hat{\varepsilon} = \left(I_n - P_1\right)Y$$

$$\hat{\varepsilon} := Y - X\hat{\beta}$$

8.1.6 (f)

Prove that the matrix $P_2 := (I_n - P_1)$ is also a projection matrix and show that $P_1P_2 = 0_n$.

8.1.7 (g)

Show that $\hat{\varepsilon} = P_2\varepsilon$.

$$\begin{aligned} \hat{\varepsilon} &= y - \hat{y} \\ &= y - X\left(X^TX\right)^{-1}X^Ty \\ &= \left[I - P_1\right]y \\ &= P_2y \\ &= P_2\left(X\beta+\varepsilon\right) \\ &= P_2\varepsilon \end{aligned}$$

Show that $\mathbb{E}[\hat{\varepsilon}] = 0$

$$\begin{aligned} \mathbb{E}[\hat{\varepsilon}] &= \mathbb{E}\left[P_2\varepsilon\right] \\ &= P_2\mathbb{E}\left[\varepsilon\right] \\ &= 0 \end{aligned}$$

Compute $Var(\hat{\varepsilon})$

$$\begin{aligned}
Var(\hat{\varepsilon}) &= \mathbb{E}\left[\hat{\varepsilon}\hat{\varepsilon}^T\right] \\
&= P_2\mathbb{E}\left[\varepsilon\varepsilon^T\right] P_2^T \\
&= P_2\sigma^2 P_2^T \\
&= \sigma^2 P_2 P_2^T \\
&= \sigma^2 P_2
\end{aligned}$$

$$\begin{aligned}
Cov(\hat{\beta},\hat{\varepsilon}) &= \mathbb{E}\left[\left(\hat{\beta}-\mathbb{E}\left[\hat{\beta}\right]\right)\left(\hat{\varepsilon}-\mathbb{E}\left[\hat{\varepsilon}\right]\right)^T\right] \\
&= \mathbb{E}\left[\left(\hat{\beta}-\beta\right)\left(P_2\varepsilon\right)^T\right] \\
&= \mathbb{E}\left[\left(XX\right)^{-1}X^T\left(P_2\varepsilon\right)^T\right] \\
&\vdots \\
&= 0
\end{aligned}$$

8.2 Exercise 2

Let

$$\sum_{i=1}^{n}\hat{\varepsilon}_i^2 = \|\hat{\varepsilon}\|^2 = \sum_{k=1}^{n-p} Z_k^2, \tag{5}$$

where $Z_k \sim \mathcal{N}\left(0,\sigma^2\right)$ and are all independent.

8.2.1 (a)

Modify (5) is such a way to obtain a sum of independent standard (unit variance) Gaussian random variables.

$$\begin{aligned}
\|\hat{\varepsilon}\|^2 &= \sum_{k=1}^{n-p} Z_k^2 \\
&= \sigma^2 \underbrace{\sum_{k=1}^{n-p}\frac{Z_k^2}{\sigma^2}}_{\sim\mathcal{N}(0,1)}
\end{aligned}$$

Hence $\|\varepsilon\|^2 \sim \sigma\chi^2_{n-p}$