Exercise 1

Assume that $x_1, \ldots, x_n$ are independent realizations of a random variable $X \sim N(1, \sigma^2)$.

(a) Write the density function of X.

(b) Obtain the likelihood function $L(\sigma^2)$.

(c) Obtain the log-likelihood function $\ell(\sigma^2)$.

(d) Obtain the score function $\ell_*(\sigma^2)$.

(e) Compute the maximum likelihood estimate $\hat{\sigma}^2$ of σ^2 (it is not required to check the sign of the second derivative).

(f) Check whether $\hat{\sigma}^2$ is a biased or an unbiased estimator of σ^2.

(g) Obtain an asymptotic approximation for the distribution of $\hat{\sigma}^2$.

(h) Obtain the exact distribution of $\hat{\sigma}^2$.

(i) Compute $\hat{\sigma}^2$ with the data (see the R-script below),

```
2.84  2.66 -0.62 -1.81  2.18  2.10  2.84 -0.98  1.12 -0.89
```

(j) Use the asymptotic distribution of $\hat{\sigma}^2$ to compute an approximate 95% confidence interval for σ^2.

(k) Use the exact distribution of $\hat{\sigma}^2$ to compute a 95% confidence interval for σ^2.

This is some `R` code, potentially useful **for all the exercises**:

```
> x <- c(2.84,  2.66, -0.62, -1.81,  2.18,  2.10,  2.84, -0.98,  1.12, -0.89)
> mean(x)
[1] 0.944
> sd(x)
[1] 1.829549
> var(x)
[1] 3.347249
> sum(x^2)
[1] 39.0366
> sum((x-1))
[1] -0.56
> sum((x-1)^2)
[1] 30.1566
> alpha <- c(0.001, 0.01, 0.025, 0.05, 0.95, 0.975, 0.99, 0.999)
> round(qnorm(alpha),2)
[1] -3.09 -2.33 -1.96 -1.64  1.64  1.96  2.33  3.09
> round(qchisq(alpha, 9), 2)
[1]  1.15  2.09  2.70  3.33 16.92 19.02 21.67 27.88
> round(qchisq(alpha, 10), 2)
[1]  1.48  2.56  3.25  3.94 18.31 20.48 23.21 29.59
> round(qt(alpha, 9), 2)
[1] -4.30 -2.82 -2.26 -1.83  1.83  2.26  2.82  4.30
> round(qt(alpha, 10), 2)
[1] -4.14 -2.76 -2.23 -1.81  1.81  2.23  2.76  4.14
```

Sketch of the solution:

(a)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-1)^2\right\}$$

(b)

$$L(\sigma^2) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-1)^2\right\}$$

(c)

$$\begin{aligned} \ell(\sigma^2) &= \sum_{i=1}^{n} \left\{-\frac{1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(x_i-1)^2\right\} \\ &= -\frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-1)^2 \end{aligned}$$

(d)

$$\ell_*(\sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-1)^2$$

(e) $\ell_*(\hat{\sigma}^2) = 0$ gives

$$-n + \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-1)^2 = 0$$

and therefore

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(x_i-1)^2}{n}$$

(f)

$$\begin{aligned} E(\hat{\sigma}^2) &= E\left[\frac{\sum_{i=1}^{n}(X_i-1)^2}{n}\right] \\ &= \frac{\sum_{i=1}^{n} E\left[(X_i-1)^2\right]}{n} \\ &= \frac{nE\left[(X-1)^2\right]}{n} \\ &= E\left[(X-1)^2\right] \\ &= \mathrm{Var}[(X-1)] \\ &= \sigma^2 \end{aligned}$$

where the result follows because $(X_i - 1) \sim N(0, \sigma^2)$.

(g)

$$
\begin{aligned}
j(\sigma^2) &= -\frac{d\ell_*(\sigma^2)}{d\sigma^2} \\
&= -\left[\frac{n}{2(\sigma^2)^2} - \frac{2}{2(\sigma^2)^3}\sum_{i=1}^{n}(x_i-1)^2\right] \\
&= -\frac{n}{2(\sigma^2)^2} + \frac{n\hat{\sigma}^2}{(\sigma^2)^3}
\end{aligned}
$$

so that

$$
\begin{aligned}
j(\hat{\sigma}^2) &= -\frac{n}{2(\hat{\sigma}^2)^2} + \frac{n\hat{\sigma}^2}{(\hat{\sigma}^2)^3} \\
&= \frac{1}{(\hat{\sigma}^2)^2}\left[-\frac{n}{2} + n\right] \\
&= \frac{n}{2\hat{\sigma}^4}
\end{aligned}
$$

and

$$
j(\hat{\sigma}^2)^{-1} = \frac{2\hat{\sigma}^4}{n} \quad \text{so that} \quad \hat{\sigma}^2 \dot{\sim} N\left(\sigma^2, \frac{2\hat{\sigma}^4}{n}\right)
$$

(h) Because $\sum_{i=1}^{n}[(X_i-1)/\sigma]^2 \sim \chi^2_n$ it follows that

$$
\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_n \quad \text{or, equivalently,} \quad \hat{\sigma}^2 \sim \frac{\sigma^2}{n}\chi^2_n
$$

(i) $\hat{\sigma}^2 = 3.01566$

(j) $j(\hat{\sigma}^2)^{-1} = 2 \times 3.01566^2/10 = 1.818$ and, therefore, $\sqrt{j(\hat{\sigma}^2)^{-1}} = 1.3486$ and an approximate 95% confidence interval is

$$
3.01566 \pm 1.96 \times 1.3486 \quad \Rightarrow \quad (0.372;\ 5.65)
$$

(k)

$$
\Pr\left(3.25 \le \frac{n\hat{\sigma}^2}{\sigma^2} \le 20.48\right) = 0.95
$$

so that

$$
\Pr\left(\frac{10 \times \hat{\sigma}^2}{20.48} \le \sigma^2 \le \frac{10 \times \hat{\sigma}^2}{3.25}\right) = 0.95
$$

and a 95% exact confidence interval is (1.47; 9.27)

Exercise 2

Let $X_1, \ldots, X_n$ a random sample from a random variable X with density $f(x) = e^{-\frac{x}{\theta}}/\theta$ where $x > 0$ and $\theta > 0$. Compute the mean square error of the following estimators of θ.
(Hint: $E(X) = \theta$ and $\mathrm{Var}(X) = \theta^2$.)

(a) $\hat{\theta}_1 = X_1$.

(b) $\hat{\theta}_2 = (X_1 + 2X_2)/3$.

(c) $\hat{\theta}_3 = \bar{X}$.

(d) $\hat{\theta}_4 = 5$.

Sketch of the solution:

(a) $MSE(\hat{\theta}_1) = \theta^2$

(b) $MSE(\hat{\theta}_2) = \frac{5\theta^2}{9}$

(c) $MSE(\hat{\theta}_3) = \frac{\theta^2}{n}$

(d) $MSE(\hat{\theta}_4) = (\theta - 5)^2$

Exercise 3

For a certain disease, there exists a treatment that cures 70% of the cases. A laboratory proposes a new treatment claiming that it is better than the previous one. Out of 200 patients having received the new treatment, 148 of them have been cured. As the expert in charge of deciding whether the new treatment should be authorized, what are your conclusions? **Provide all the following points: system of hypotheses, statistical test and its distribution under H_0, (approximate) p-value and discuss the conclusions.**

Sketch of the solution:

I we denote by π the probability of recover under the new treatment, we want to test the following system of hypotheses

$$\begin{cases} H_0, & \pi = 0.7 \\ H_1 & \pi > 0.7 \end{cases}$$

Furthermore,

$$SE_0 = \sqrt{\frac{0.7 \times 0.3}{200}} = 0.0324 \quad \text{and} \quad \hat{\pi} = \frac{148}{200} = 0.74$$

and we use the test statistic,

$$T = \frac{\hat{\pi} - 0.7}{SE_0}$$

which, under H_0, is asymptotically normally distributed. The observed value of the test statistic is $t^{obs} = (0.7 - 0.74)/0.0324 = 1.2344$ and we can se that p-value> 0.05 and therefore the information provided by the data does not allow us to reject H_0.