<u>WISB263 Mathematical Statistics</u>
Exam V1
25th June 2020
Lecturer: Dr. Wioletta Ruszel
Teaching assistants: Leandro Chiarini, Ben Balkenende, Esther Steenkamer

Exercise 1 (25 p.) The average energy expenditures for eight elderly subjects were estimated on the basis of information received from a battery-powered heart rate monitor that each subject wore. Two overall averages were calculated for each subject, one for the summer months and one for the winter months, as shown in the following table. Let μ_D denote the difference between the summer and winter energy expenditure populations.

V1
Average daily energy expediture

Subject	Summer x_i	Winter y_i
1	1458	1424
2	1353	1501
3	2209	1495
4	1804	1739
5	1912	2031
6	1366	934
7	1598	1401
8	1406	1339

1. (10 p.) Perform a hypothesis test "TEST 1" for $(H_0) : \mu_D = 0$ against $(H_1) : \mu_D > 0$ at $\alpha = 0.01$ and determine the exact rejection region. Is there a smallest level of significance to reject H_0?
2. (10 p.) Perform a hypothesis test "TEST 2" for $(H_0) : \mu_D = 0$ against $(H_1) : \mu_D > 0$ now based on the assumption that the samples are normally distributed. The sample of summer months expeditures $X_1, ..., X_8$ has distribution $N(\mu_X, \sigma_X^2)$ and the sample of winter months expeditures $Y_1, ..., Y_8$ has distribution $N(\mu_Y, \sigma_Y^2)$ with σ_X, σ_Y unknown. For which α do you reject/accept H_0? Compute the p-value. Determine the rejection region for $\alpha = 0.01$.
3. (5 p.) Compare the tests "TEST 1" and "TEST 2". Which one is more appropriate?

Exercise 2 (25 p.) Consider an i.i.d. sample $X_1, ..., X_n$ on some probability space with common density given by

$$h_\theta(x) = \frac{g(\theta)}{x^4}, \ x \geq \theta,$$

where $\theta > 0$.

1. (2 p.) Determine $g(\theta)$ as a function of θ.

2. (5 p.) Determine the MLE $\hat{\theta}_n$ of θ and $\hat{\theta}_{n,1}$ for θ^3.

3. (5 p.) Calculate the cumulative distribution function $F(x)$ of $\hat{\theta}_n$ for all $x \in \mathbb{R}$ and determine its density $f(x)$.

4. (4 p.) Is $\hat{\theta}_n$ biased? Show that $\hat{\theta}_n$ is asymptotically unbiased.

5. (4 p.) Find the MoM estimator θ_n^* of θ and compute the variance.

6. (5 p.) Which estimator is more efficient for $n \geq 2$? Can we apply Cramer-Rao Theorem in this case?

Exercise 3 (25 p.) Assume we have an estimator $\hat{X}_n$ for a parameter $\theta > 0$ appearing in the density of an i.i.d. sample $X_1, ..., X_n$. The density of the estimator is

$$f_\theta(x) = 3n\frac{\theta^{3n}}{x^{3n+1}},\ x > \theta,$$

which means that $\mathbb{P}(\hat{X}_n \in (a,b)) = \int_a^b f_\theta(x)dx$ for $a < b$. The expected value is equal to $\mathbb{E}(\hat{X}_n) = \frac{3n\theta}{3n-1}$ and variance $Var(\hat{X}_n) = \frac{3n\theta^2}{(3n-2)(3n-1)^2}$.

1. (4 p.) Show that $\hat{X}_n$ is consistent.

2. (4 p.) Prove that $\frac{\hat{X}_n - \mathbb{E}(\hat{X}_n)}{\sqrt{Var(\hat{X}_n)}} \xrightarrow{d} Y$ as $n \to \infty$, where Y has density $f_Y(y) = e^{-(1+y)}$. (Hint: You can use that $\lim_{n\to\infty}\left(1 + \frac{1+a}{n}\right)^{-n} = \lim_{n'\to\infty}\left(1 + \frac{1+a\sqrt{\frac{n'}{n'-2}}}{n'-1}\right)^{-n'}$, where $n' = c \cdot n$ for some $c \geq 1$ and $a \in \mathbb{R}$.)

3. (2 p.) Show that the support of $f_Y(\cdot)$ is $(-1, \infty)$.

4. (3 p.) Let $\alpha \in (0,1)$. Find constants c_1, c_2 such that $\mathbb{P}(Y < c_1) = \frac{\alpha}{2}$ and $\mathbb{P}(Y > c_2) = \frac{\alpha}{2}$. What is the probability $\mathbb{P}(Y \in [c_1, c_2])$?

5. (4 p.) Use the asymptotic distribution of $\frac{\hat{X}_n - \mathbb{E}(\hat{X}_n)}{\sqrt{Var(\hat{X}_n)}}$ to construct an asymptotic random confidence interval $[L(\mathbf{X}), R(\mathbf{X})]$ for θ with confidence $1 - \alpha$.

6. (4 p.) Define a formal hypothesis test with significance α for testing $(H_0) : \theta = \theta_0$ against $(H_1) : \theta \neq \theta_0$. Describe the rejection region implicitely.

7. (1 p.) Determine now the asymptotic rejection region using knowledge from (2) and (5).

8. (3 p.) Assume we get from our data the value of the test statistic is $t(\mathbf{x}) = \ln(10) - 1$. What is the smallest level of significance to reject H_0 in this case? How do we call this value? (Hint: You can use the asymptotic distribution of the test statistic).

Exercise 4 (25 p.) Let $X_1, ..., X_n$ be an i.i.d. random sample on some probability space with common density

$$f_\theta(x_1) = \frac{x_1}{\theta^2}e^{-\frac{x_1^2}{2\theta^2}},\ x_1 \geq 0,$$

and $\theta > 0$. We want to compare two hypothesis tests for $(H_0) : \theta = 1$ against $(H_1) : \theta \neq 1$ using two different test statistics, namely

$$T_1(\mathbf{X}) = \sum_{i=1}^{n} X_i^2, \text{ and } T_2(\mathbf{X}) = 2\ln(R(\mathbf{X})),$$

where $R(\mathbf{X})$ is the generalized likelihood ratio. You can use that the MLE is equal to $\hat{\theta}_n = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}$ and is asymptotically normal.

1. (3 p.) Define a formal hypothesis test based on statistic $T_1(\mathbf{X})$ at significance α.
2. (2 p.) Show that the MGF of X_1^2 is equal to $\frac{1}{1-2t\theta^2}$ for $|t| < \frac{1}{2}$.
3. (3 p.) Determine the distribution of $T_1(\mathbf{X})$ under H_0 and show that it equal to $\frac{1}{\theta^2}T_1(\mathbf{X})$ under H_1.
4. (4 p.) Let $\alpha = 0.05$ and $n = 60$ determine the critical region for this test. Compute the power when $\theta = \sqrt{2}$. If you did not find the distribution before use the normal distribution.
5. (4 p.) Consider now the second test statistic $T_2(\mathbf{X})$. Show that $T_2(\mathbf{X})$ is of the form $c_1 \ln(c_2 \sum_{i=1}^{n} X_i^2) + c_3 \sum_{i=1}^{n}(X_i^2 - c_4)$ where c_1, c_2, c_3, c_4 are some constants.
6. (4 p.) What is the asymptotic distribution of $T_2(\mathbf{X})$? Determine the asymptotic rejection region for $\alpha = 0.05$.
7. (5 p.) Assume we observe $\sum_{i=1}^{60} x_i^2 = 85$. Perform both tests at $\alpha = 0.05$ and decide whether to reject H_0 or not (you can use the second asymptotic test). If the null hypothesis would be 1-sided, which test is more powerfull? Argue without computing the power of the test based on $T_2(\mathbf{X})$.