WISB263 Mathematical Statistics
Final Exam
29th June 2021, 15.15-18.15
Total amount of points: 100 + 10 bonus points
Lecturer: Wioletta Ruszel

Exercise 1 (25 p.) A number of coins from the reign of King Manuel I Komnenos (1143-1180) were discovered in Cyprus. They arise from four different coinages at intervals throughout his reign. The data below give the silver content ($\%A_g$) of coins. We want to test the hypothesis whether there is any significant difference in their silver content with the passage of time.

The silver content of coins is:

First	Second	Third	**Forth**
5.9	6.91	4.9	5.3
6.8	9.0	5.5	5.6
6.4	6.6	4.6	5.5
7.0	8.1	4.5	5.1
6.61	9.3		6.2
7.7	9.2		5.8
7.2	8.6		5.81
6.9			
6.21			

(i) (4 p.) Is the data from the table (first and forth reign) paired or unpaired? Give arguments to support your answer.

(ii) (8 p.) Perform an appropriate hypothesis test to test whether there is a difference between the silver content of the coins in the **first** and **forth reign** at significance at most $\alpha = 0.05$. Is there evidence that the silver content has been reduced or enhanced?

(iii) (8 p.) Perform the same test as in (ii) assuming that the data is normally distributed with known variance $\sigma^2 = 1$. Compute the p-value. (If you cannot compute the p-value exactly find an upper and lower bound.)

(iv) (5 p.) Compare the findings of the two tests. Which one is more appropriate and has stronger evidence against the null hypothesis? You can assume that the p-value of the first test in (ii) is equal to 0.00034.

Exercise 2 (25 p.) Let $\mathbf{X} = (X_1, ..., X_n)$ be a random sample such that $X_1 \sim U(\theta, c \cdot \theta)$, where $c > 1$ and $\Theta = \mathbb{R}_+$.

(i) (5 p.) Determine the MLE $\hat{\theta}_n$ of θ.

(ii) (5 p.) Show that the density g_θ of the MLE $\hat{\theta}_n$ is equal to

$$g_\theta(y) = \frac{cn}{(c-1)\theta} \left(\frac{cy - \theta}{(c-1)\theta} \right)^{n-1}$$

for some $y \in D \subset \mathbb{R}$ and 0 otherwise. Determine D.

(iii) (5 p.) Compute $\mathbb{E}_\theta(\hat{\theta}_n)$. Is the estimator unbiased? In case the estimator is biased, determine whether the estimator overestimates or underestimates the true parameter θ.

(iv) (5 p.) Let $c = 3$. We want to design a hypothesis test for testing

$$(H_0) : \theta = 5 \text{ against } (H_1) : \theta \neq 5$$

at significance level $\alpha = 0.05$ using the generalized likelihood ratio statistic $R(\mathbf{X})$. Determine $R(\mathbf{X})$ and the exact rejection region (not the asymptotic one).

(v) (3 p.) Assume that you observe $\{12, 6.1, 8, 5.3, 5, 4.7, 7.6, 11.4, 9, 10.7\}$. Perform the hypothesis test from (iv) in order to decide whether the data was likely coming from a $U(5, 15)$ distribution at significance $\alpha = 0.05$.

(vi) (2 p.) Determine the type-II error when $\theta = 6$ for the data from (v).

Exercise 3 (25 p.) Consider a random sample $\mathbf{X} = (X_1, ..., X_n)$ with common density

$$f_\theta(x_1) = \frac{\theta}{x_1^{\theta+1}}$$

where $\Theta = (2, \infty)$ and $x_1 > 1$.

(i) (4 p.) Determine the MLE $\hat{\theta}_n$ for θ.

(ii) (2 p.) Is the parametric family of exponential type?

(iii) (5 p.) Design a u.m.p. test for testing $(H_0) : \theta \leq \theta_0$ against $(H_1) : \theta > \theta_0$ at significance α. Describe the rejection region implicitly.

(iv) (2 p.) Show that the Fisher information $I(\theta)$ for the density f_θ is equal to $\frac{1}{\theta^2}$.

(v) (4 p.) Assume that $\hat{\theta}_n$ is asymptotically normal with the corresponding variance (satisfies Theorem 2.11). Determine the Wald random confidence interval for θ for general α.

(vi) (5 p.) Design an asymptotic hypothesis test for

$$(H_0) : \ \theta = \theta_0 \text{ against } (H_1) : \theta \neq \theta_0$$

at significance α and determine the rejection region using the information from (v).

(vii) (3 p.) Perform the hypothesis test defined in (vi) when $\theta_0 = 4$ for testing data $\{x_1, ..., x_{100}\}$ such that $\sum_{i=1}^{100} \ln(x_i) = 43.79$ at $\alpha = 0.01$ and determine the confidence interval.

Exercise 4 (25 p.) Consider independent random variables $Z_1, ..., Z_n$ such that

$$Z_i \sim N(\mu, \sigma^2(1 + x_i))$$

and $x_i > -1$ for all $i = 1, ..., n$.

(i) (5 p.) Write the random variables $Y_1, ..., Y_n$ in terms of $Z_1, ..., Z_n$ such that for all $i = 1, ..., n$

$$Y_i = a_i + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma^2)$. Determine a_i for all $i = 1, ..., n$.

(ii) (7 p.) What is the distribution of $\mathbf{Y} = (Y_1, ..., Y_n)$? Compute directly the unbiased MLE's $\hat{\mu}_n$ and $S^2_{\overline{Y}_n}$ for μ resp. σ^2 for the sample $\mathbf{Y}$.

(iii) (5 p.) Assume first that $\frac{n-1}{\sigma^2} S^2_{\overline{Y}_n}$ is χ^2_{n-1} distributed. Discuss the hypothesis test

$$(H_0) : \sigma = 2 \text{ against } (H_1) : \sigma < 2.$$

Which test statistic could you choose? Determine the rejection region for $\alpha = 0.05$ and $n = 10$. For which $s^2_{\overline{y}_{10}}$ do we accept H_0?

(iv) (3 p.) Find a matrix $\mathbf{X}$ and B such that we can write

$$\mathbf{Y} = \mathbf{X}B + \epsilon_n$$

where $\epsilon_n \sim N_n(0, \sigma^2 Id)$. How do you call $\mathbf{Y}$ written in this way?

(v) (5 p.) Show that $S^2_{\overline{Y}_n}$ is χ^2_{n-1} distributed.

BONUS (10 p.) Consider a two-sided test with critical region $\{S_n(\mathbf{X}) \leq c_1\} \cup \{S_n(\mathbf{X}) \geq c_2\}$. Prove that the p-value for this test is uniformly distributed over $[0, 1]$.