Felix Kießner, M.Sc.

Advanced Statistics II – Summer Term 2017

Problem Set 2

1. **Home Assignment 1**

   Let $(X_i, i : 1 \to n)$ be a random sample from a Uniform distribution on the interval $(0, 1]$ with the respective order statistics $X_{[1]} \leq X_{[2]} \leq \cdots \leq X_{[n]}$.

   a) Find the density of the sample maximum as well as of the sample minimum

   b) Find the probability limits of the maximum and the minimum

   (as done in the PP-tutorial) and verify these results with Monte-Carlo Simulation methods. Choose $n = 10,\ 100,\ 1000$, $MC = 100,\ 500,\ 1000$. Comment (shortly) on the dependency of your results on n and MC. Plot your results in one figure using a 3x3 grid. Save this plot in a pdf file. You are only allowed to use functions (and packages) that are part of basic R.

2. Consider the following estimators $\hat{\mu}_X^{(j)}$ for the mean of a normally distributed population with $X \sim N(\mu, \sigma^2)$ and a random sample $\{X_i, i : 1 \to n\}$, where $3 < n < \infty$

$$\hat{\mu}_X^{(1)} = \frac{1}{n-1} \sum_{i=1}^{n} X_i; \quad \hat{\mu}_X^{(2)} = \frac{1}{3} \sum_{i=1}^{3} X_i; \quad \hat{\mu}_X^{(3)} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

   (a) Write a function for each of the estimators.

   (b) Write a function that computes the MSE for an estimator.

   (c) Discuss consistency in terms of MSE for each estimator. Therefore conduct a simulation study for the following parameters: $MC = 100$, $n = 3, \cdots, 100$, $\mu = 0,\ 1,\ 5$, $\sigma = 1,\ 2,\ 5$. Plot the MSE for each of the three estimators as a function of n as well as for each possible parameter combination. Arrange all graphs on one single figure. Save this figure as a pdf. Name all axes appropriately, add a meaningful legend and indicate in each graph to which parameter combination it refers. Which of the estimators is consistent? How does the MSE depend on n, μ, σ?

Home Assignment 2 (2 points)

The random sample $\{X_i, i : 1 \to n\}$ is drawn from an exponential distribution with pdf

$$f(x;\theta) = \frac{1}{\theta}\exp\{-\frac{x}{\theta}\} \cdot I_{(0,\infty)}(x)\ , \quad \theta > 0\,.$$

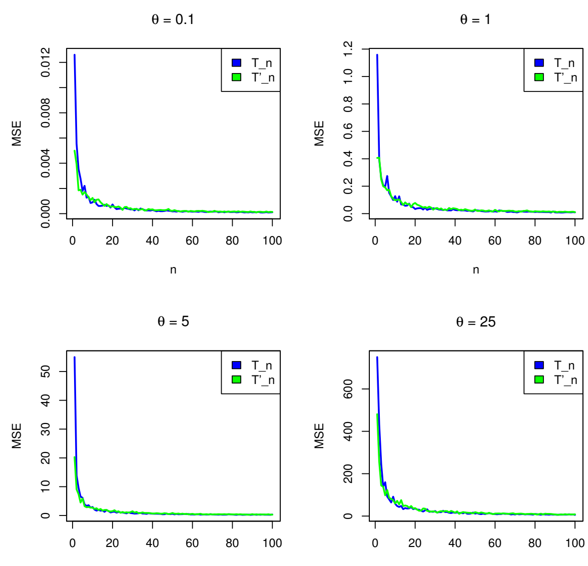
The following statistics are considered as estimators for θ:

$$T_n = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \text{and} \quad {T'}_n = \left[\frac{1}{2n}\sum_{i=1}^{n} X_i^2\right]^{1/2}\,.$$

(a) Write a function for each of the estimators.

(b) Write a function that computes bias, variance and MSE for an given estimator. (Hint: This function should return all of the three quantities!)

(c) **Use your function from part b)** to discuss consistency in terms of MSE for each estimator. Therefore conduct a simulation study for the following parameters: $MC = 100$, $n = 1, \cdots, 100$, $\theta = 0.1,\ 1,\ 5,\ 25$. Plot the MSE for each θ and estimator as a function of n. Arrange all graphs in one single figure. Save this figure as a pdf. Name all axes appropriately, add a meaningful legend and indicate in each graph to which parameter combination it refers. Which of the estimators is consistent? How does the MSE depend on n, θ? Overall your plot should look like the following:



(d) Compare now the variances of both estimators for each θ. Plot the individual variances against n. Arrange all graphs in one single figure. Save this figure as a pdf. Name all axes appropriately, add a meaningful legend and indicate in each graph to which parameter combination it refers. Which estimator has the lower variance for small ($n < 10$) and/or for large $n > 75$?

You are only allowed to use functions (and packages) that are part of basic R. Please use the provided script file for the first home assignment and fill out the header!

Hand in your script file (with the necessary comments) until 16.5.17 8pm via mail.