Final Exam Probability and Statistics EE1M31
18 April 2018, 9.00 – 11.00

Last name, initials: ____________ Grade: ____

Student number: ____________

Open questions

ALWAYS MOTIVATE YOUR ANSWERS AND WRITE NEATLY.
YOU MAY ANSWER IN EITHER DUTCH OR ENGLISH.

1. The lifetime X of an electrical component is a continuous random variable having the probability density function $f_\theta(x) = \frac{2}{\theta} - \frac{2}{\theta^2}x$ for $0 \le x \le \theta$ and $f_\theta(x) = 0$ otherwise.

   Let $X_1, \ldots, X_n$ be a random sample from f_θ.

   (a) (3 points) Investigate whether $S = \sqrt{\frac{6}{n}(X_1^2 + \ldots + X_n^2)}$ is an unbiased estimator for θ.
   Hint: first investigate whether S^2 is an unbiased estimator for θ^2.

$$\mathbb{E}[X_1^2] = \int_0^\theta x^2\left(\frac{2}{\theta} - \frac{2}{\theta^2}x\right)dx = \frac{2}{\theta}\int_0^\theta x^2\,dx - \frac{2}{\theta^2}\int_0^\theta x^3\,dx$$

$$= \frac{2}{\theta}\left[\frac{1}{3}x^3\right]_0^\theta - \frac{2}{\theta^2}\left[\frac{1}{4}x^4\right]_0^\theta = \frac{2}{3}\theta^2 - \frac{2}{\theta^2}\frac{1}{4}\theta^4 =$$

$$= \left(\frac{2}{3} - \frac{1}{2}\right)\theta^2 = \left(\frac{4}{6} - \frac{3}{6}\right)\theta^2 = \frac{1}{6}\theta^2$$

$\mathbb{E}[S^2] = \frac{6}{n}\cdot n\cdot\frac{1}{6}\theta^2 = \theta^2 \Rightarrow$ S is unbiased

$(\mathbb{E}[S])^2 < \mathbb{E}[S^2]$ $\qquad$ $\mathbb{E}[S] < \sqrt{\mathbb{E}[S^2]} \le \sqrt{\theta^2} = \theta$

by Jensen's inequality

S has negative bias

   (b) (3 points) Let $T = \frac{3}{n}(X_1 + \ldots + X_n)$ be another estimator for θ: determine its MSE.

$$\mathbb{E}[X_1] = \int_0^\theta x\left(\frac{2}{\theta} - \frac{2}{\theta^2}x\right)dx = \frac{2}{\theta}\int_0^\theta x\,dx - \frac{2}{\theta^2}\int x^2\,dx$$

$$= \frac{2}{\theta}\frac{1}{2}\theta^2 - \frac{2}{\theta^2}\frac{1}{3}\theta^3 = \theta - \frac{2}{3}\theta = \frac{1}{3}\theta$$

$\mathbb{E}[T] = \frac{3}{n}\cdot n\cdot\frac{1}{3}\theta = \theta \Rightarrow$ T unbiased

$$MSE[T] = Var(T) = \frac{9}{n^2}\cdot n\cdot Var(X_1) = \frac{9}{n}\left(\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2\right)$$

$$= \frac{9}{n}\left(\frac{1}{6}\theta^2 - \frac{1}{9}\theta^2\right) = \frac{1}{n}\left(1.5\theta^2 - \theta^2\right) = \frac{\theta^2}{2n}$$

2. Consider the following ordered dataset:

3.14 3.22 3.43 3.73 3.94 3.96 4.06 4.22 4.22 5.60 7.49

(a) (2 points) Calculate the 20th empirical percentile of this dataset.

$p = 0.2$ $n = 11$ $p(n+1) = 0.2 \cdot 12 = 2.4$ $k = \lfloor p(n+1) \rfloor = 2$

$\alpha = p(n+1) - k = 0.4$ (see p. 235 in the blue book)

$q_{11}(0.2) = 3.22 + 0.4\,(3.43 - 3.22) = 3.22 + 0.4 \cdot 0.21 = 3.304$

Assume that the dataset above has been sampled from the following density:

$$f_\theta(x) = \begin{cases} 0 & x < 3, \\ \frac{1}{\theta} e^{-\frac{(x-3)}{\theta}} & x \geq 3. \end{cases}$$

(b) (4 points) Calculate the maximum likelihood estimate for θ.
You may use that $\sum x_i = 47.01, \sum \ln x_i = 15.63, \sum \frac{1}{x_i} = 2.73$ *or* $\sum e^{-x_i} = 0.23$.

3. A manufacturer produces two new types of solar panels: type A and type B. The peak power of type A has a normal distribution with unknown expectation μ_A, but the variance is known: $\sigma_A^2 = 100$. The manufacturer wishes to estimate μ_A and therefore measures the peak power of 20 solar panels of type A.

   (a) (3 points) Explain with a computation why $\bar{x}_{20} \pm 1.96\frac{10}{\sqrt{20}}$ is the two-sided 95% confidence interval for μ_A.

$\dfrac{\bar{X}_n - \mu}{\sigma_A/\sqrt{n}}$ is distributed as $N(0,1)$ since $\bar{X}_n$ is an average of normal random variables

$$0.95 = P\left(-z_{0.025} \le \frac{\bar{X}_n - \mu}{\sigma_A/\sqrt{n}} \le z_{0.025}\right)$$

$$= P\left(-z_{0.025}\frac{\sigma_A}{\sqrt{n}} \le \bar{X}_n - \mu \le z_{0.025}\frac{\sigma_A}{\sqrt{n}}\right)$$

$$= P\left(-z_{0.025}\frac{\sigma_A}{\sqrt{n}} \le \mu - \bar{X}_n \le z_{0.025}\frac{\sigma_A}{\sqrt{n}}\right)$$

$$= P\left(\bar{X}_n - z_{0.025}\frac{\sigma_A}{\sqrt{n}} \le \mu \le \bar{x}_n + z_{0.025}\frac{\sigma_A}{\sqrt{n}}\right)$$

We see that $\left(\bar{x}_{20} - 1.95\frac{10}{\sqrt{20}}, \bar{x}_{20} + 1.95\frac{10}{\sqrt{20}}\right)$

is a 95% confidence interval.

For type B, the manufacturer claims that $\mu_B = 300\,W$. However, the Consumers Association received some negative feedback that the actual peak power is lower so they decided to investigate this claim. They tested 15 panels of type B and found a sample mean of $290.3\,W$ and a sample variance of $353.8\,W^2$. You may assume a normal distribution as model distribution.

   (b) (3 points) Investigate the claim of the manufacturer at significance level $\alpha = 0.10$.

We test $H_0: \mu_B = 300$ against $H_1: \mu_B < 300$, so we have a one-sided test.

$$T = \frac{\bar{X}_{15} - 300}{\sqrt{353.8}/\sqrt{15}} = \frac{290.3 - 300}{\sqrt{353.8}/\sqrt{15}} = -1.997$$

The critical value is $-t_{14,0.1} = -1.345$.

$-1.997 < -1.345$, so we reject H_0.

The test yields that the peak power is lower than the claimed $300\,W$.

Answers multiple choice:

1 e.

2 c.

3 f.

4 a.

5 d.

6 a.

7 d.

8 a.

9 c.