Department of Mathematics	midterm exam Statistics
VU	October 25, 2019

- **You are not allowed to use calculators, phones, laptops, or other tools.**
- **Clearly put your name and student number on all sheets that you submit.**
- **The asterix (*) in problems 2(c) and 2(d) indicates that these are optional bonus questions. You can obtain a 10 without doing these problems.**

Good luck!

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables with density p_θ given by

$$p_\theta(x) = \frac{1}{2}\theta e^{-\theta|x|}, \qquad x \in \mathbb{R},$$

where $\theta > 0$ is an unknown parameter.

(a) Compute the first two moments $\mathbb{E}_\theta X_1$ and $\mathbb{E}_\theta X_1^2$ of X_1 and give a method of moments estimator for θ.

(b) Determine the maximum likelihood estimator for θ.

(c) Now we want to obtain a Bayesian estimator for the parameter θ. As prior distribution we choose a gamma distribution, with parameters $\alpha, \beta > 0$. This is the distribution with density

$$\pi(\theta) = \frac{\theta^{\alpha-1}\beta^\alpha e^{-\beta\theta}}{\Gamma(\alpha)}, \qquad \theta > 0,$$

where

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

is Euler's gamma function.

Determine the posterior distribution of θ.

(d) Determine the Bayes estimator for θ. You may use the fact that the expectation and variance of the gamma distribution with parameters α and β are given by α/β and α/β^2, respectively.

2. Let $X_1, \ldots, X_n$ be i.i.d. random variables with density given by

$$p_\lambda(x) = \frac{1}{\lambda}e^{-x/\lambda}, \quad x > 0,$$

where $\lambda > 0$ is an unknown parameter.

(a) Determine the Fisher information for λ in the whole vector $(X_1, \ldots, X_n)$.

(b) Determine the Cramér-Rao lower bound for the variance of an unbiased estimator for λ.

(c*) Show that for the maximum likelihood estimator $\hat{\lambda}_n$ we have the convergence in distribution

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow{\text{d}} N(0, 1/i_\lambda)$$

as $n \to \infty$, with i_λ the Fisher information in a single observation. Prove this statement yourself, without using the general theorem about the MLE.

(d*) Explain in which sense the maximum likelihood estimator $\hat{\lambda}_n$ is an (asymptotically) optimal estimator for λ.

3. Opinion poller Maurice is interested in the fraction p of Dutch people in favour of reducing the number of cattle in The Netherlands. He calls 200 people at random and asks them whether they are in favour of this. Let X be the number of people that answer that they agree.

   (a) If we assume that Dutch people are independent thinkers, what is a reasonable statistical model for this problem?

   (b) Maurice wants to know whether the fraction p is greater than 1/2. Formulate this as a statistical testing problem.

   (c) Give the derivation of the standard test of level $\alpha = 0.05$ for this problem. You may use the fact that for the distribution function F of the binomial distribution with parameters $n = 200$ and $p = 1/2$ it holds that

   $$F(110) = 0.931, \quad F(111) = 0.948, \quad F(112) = 0.961, \quad F(113) = 0.972.$$

   (d) Suppose that 111 of the 200 people say that they are in favour of reducing the number of cattle. Can Maurice then conclude that indeed $p > 1/2$?

4. Let $X_1, \ldots, X_n$ be i.i.d., $N(\mu, 1)$-distributed random variables, with unknown parameter $\mu \in \mathbb{R}$.

   (a) We want to show that $\mu > 0$. Formulate this as a testing problem and give an appropriate test statistic.

   (b) Suppose that we observe the realisation $x_1, \ldots, x_n$. Give an expression for the corresponding p-value for this testing problem, in terms of the distribution function of a known distribution. Motivate your answer by giving a derivation.

Norming:

1(a): 3	**2(a): 3**	**3(a): 1**	**4(a): 2**
1(b): 3	**2(b): 2**	**3(b): 1**	**4(b): 3**
1(c): 3	**2(c*): 3**	**3(c): 3**	
1(d): 2	**2(d*): 2**	**3(d): 1**	