Probability & Statistics. Assignment 10

1. Let $X_1, \ldots, X_6$ be a sample from a uniform distribution on $[0, \theta]$, where $\theta \in [1, 2]$ is an unknown parameter. Find an unbiased estimator for θ of variance less than 1/10.

2. Consider the linear regression model

   $Y_i = \alpha + \beta x_i + \epsilon_i$, $\epsilon_i \sim N(0, \sigma^2)$, $i = 1, \ldots, n$, where $x_1, \ldots, x_n$ are known, with $\sum_{i=1}^{n} x_i = 0$, and where $\alpha, \beta \in \mathbb{R}$ and $\sigma^2 \in (0, +\infty)$ are unknown. Find the maximum likelihood estimators $\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2$ and write down their distributions.

   Consider the following data:

x_i	-3	-2	-1	0	1	2	3
Y_i	-5	0	3	4	3	0	-5

   Fit the linear regression model and comment on its appropriateness.

3. The independent observations X_1, X_2 are distributed as Poisson random variables, with means μ_1, μ_2 respectively, where

   $ln\ \mu_1 = \alpha$
   $ln\ \mu_2 = \alpha + \beta$,
   with α and β unknown parameters. Write down $l(\alpha, \beta)$, the log-likelihood function, and hence find the following:

   (a) $\frac{\partial^2 l}{\partial \alpha^2}, \frac{\partial^2 l}{\partial \alpha \partial \beta}, \frac{\partial^2 l}{\partial \beta^2}$;

   (b) $\widehat{\beta}$, the maximum likelihood estimator of β.

4. Consider the model

   $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \epsilon_i$, for $1 \leq i \leq n$, where $x_1, \ldots, x_n$ are given values, with $\sum_i x_i = 0$, and where $\epsilon_1, \ldots, \epsilon_n$ are independent normal errors, each with zero mean and unknown variance σ^2.

   (a) Obtain equations for $(\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, the MLE of $(\beta_0, \beta_1, \beta_2)$. Do not attempt to solve these equations.

   (b) Obtain an expression for $\widehat{\beta_1^*}$, the MLE of β_1 in the reduced model $H_0:\ y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $1 \leq i \leq n$ with $\sum_i x_i = 0$, and $\epsilon_1, \ldots, \epsilon_n$ distributed as above.

5. Let $(x_1, \ldots, x_n)$ be a random sample from the normal PDF with mean μ and variance σ^2. (a) Write down the log-likelihood function $l(\mu, \sigma^2)$.

   (b) Find a pair of sufficient statistics, for the unknown parameters (μ, σ^2), carefully quoting the relevant heorem.

   (c) Find $(\widehat{\mu}, \widehat{\sigma}^2)$, the MLE of (μ, σ^2). Quoting carefully any standard distributional results required, show how to construct a 95% confidence interval for μ.

6. $X_1, \ldots, X_n$ form a random sample from a uniform distribution on the interval $(-\theta, 2\theta)$, where the value of the positive parameter θ is unknown. Determine the maximum likelihood estimator of the parameter θ.

7. (a) Aerial observations x_1, x_2, x_3, x_4 are made of the interior angles $\theta_1, \theta_2, \theta_3, \theta_4$ of a quadrilateral on the ground. If these observations are subject to small independent errors with zero means and common variance σ^2, determine the least squares estimator of $\theta_1, \theta_2, \theta_3, \theta_4$.

   (b) Obtain an unbiased estimator of σ^2 in the situation described in part (a).

   (c) Suppose now that the quadrilateral is known to be a parallelogram with $\theta_1 = \theta_3$ and $\theta_2 = \theta_4$. What now are the least squares estimates of its angles? Obtain an unbiased estimator of σ^2 in this case.

8. Let $X_1, X_2, \ldots, X_n$ be a random sample from the $N(\mu, \sigma^2)$-distribution. Prove that the random variables $\overline{X}$ (sample mean) and $\sum_{i=1}^n (X_i - \overline{X})^2$ (sample variance times (n-1)) are independent and determine their distributions.

9. (a) Explain what is meant by constructing a confidence interval for an unknown parameter θ from a given sample $x_1, \ldots, x_n$. Let a family of PDFs $f(x; \theta)$, $-\infty < \theta < \infty$, be given by

   $$f(x;\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta, \\ 0, & x < \theta. \end{cases}$$

   Suppose that $n = 4$ and $x_1 = -1,\ x_2 = 1.5,\ x_3 = 0.5,\ x_4 = 1$. Construct a 95% confidence interval for θ.

   (b) Let $f(x; \mu, \sigma^2)$ be a family of normal PDFs with an unknown mean μ and an unknown variance $\sigma^2 > 0$. Explain how to construct a 95% confidence interval for μ from a sample $x_1, \ldots, x_n$. Justify the claims about the distributions you use in your construction.

10. (a) State and prove the factorisation criterion for sufficient statistics, in the case of discrete random variables.

    (b) A linear function $y = Ax + B$ with unknown coefficients A and B is repeatedly measured at distinct points $x_1, \ldots, x_k$: first n_1 times at x_1, then n_2 times at x_2 and so on; and finally n_k times at x_k. The result of the i-th measurement series is a sample $y_{i1}, \ldots, y_{in_i}$, $i = 1, \ldots, k$. The errors of all measurements are independent normal variables, with mean zero and variance 1. You are asked to estimate A and B from the whole sample y_{ij}, $1 \leq j \leq n_i$, $1 \leq i \leq k$. Prove that the maximum likelihood and the least squares estimators of (A, B) coincide and find these.

    Denote by $\widehat{A}$ the maximum likelihood estimator of A and by $\widehat{B}$ the maximum likelihood estimator of B. Find the distribution of $(\widehat{A}, \widehat{B})$.

11. (a) Let $x_1, \ldots, x_n$ be a random sample from the PDF $f(x; \theta)$. What is meant by saying that $t(x_1, \ldots, x_n)$ is sufficient for θ? Let

    $$f(x;\theta) = \begin{cases} e^{-(x-\theta)}, & x \geq \theta, \\ 0, & x < \theta. \end{cases}$$

    and suppose $n = 3$. Let $y_1 < y_2 < y_3$ be ordered values of x_1, x_2, x_3. Show that y_1 is sufficient for θ.

    (b) Show that the distribution of $Y_1 - \theta$ is exponential of parameter 3. Your client suggests the following possibilities as estimators of θ:

$$\begin{aligned}\overline{\theta_1} &= x_3 - 1,\\ \overline{\theta_2} &= y_1,\\ \overline{\theta_1} &= \tfrac{1}{3}(x_1 + x_2 + x_3) - 1.\end{aligned}$$

How would you advise him?

12. (a) Derive the form of the MLEs of α, β and σ^2 in the linear model

    $Y_i = \alpha + \beta x_i + \epsilon_i$, $1 \leq i \leq n$,

    where $\epsilon_i \sim N(0, \sigma^2)$ and $\sum_{i=1}^n x_i = 0$.

    (b) What is the joint distribution of the maximum likelihood estimators $\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2$? Construct 95% confidence intervals for:

    - σ^2,
    - $\alpha + \beta$.