Statistics (Ph. D. Version)

Instructions to the Student

a. Answer all six questions. Each will be graded from 0 to 10.

b. Use a different booklet for each question. Write the problem number and your code number (NOT YOUR NAME) on the outside cover.

c. Keep scratch work on separate pages in the same booklet.

d. If you use a "well known" theorem in your solution to any problem, it is your responsibility to make clear which theorem you are using and to justify its use.

1. Let \( \{p(x; \theta_1, \theta_2), (\theta_1, \theta_2) \in \Theta_1 \times \Theta_2\} \) be a family of pdf's on a measurable space \((\mathcal{X}, \mathcal{A})\) with \((\theta_1, \theta_2)\) as a parameter. Assume that there exists \(\theta_1^* \in \Theta_1\) such that \(T_1 = T_1(X)\) is sufficient for the family \(\{p(x; \theta_1^*, \theta_2), \theta_2 \in \Theta_2\}\). Assume also that for any \(\theta_2 \in \Theta_2\), a statistic \(T_2 = T_2(X)\) is sufficient for the family \(\{p(x; \theta_1, \theta_2), \theta_1 \in \Theta_1\}\).

(i) Write the versions of the factorization theorem expressing the above properties.

(ii) Assuming \(p(x; \theta_1, \theta_2) > 0\) for all \(x, \theta_1, \theta_2\), show that the statistic \(T = (T_1, T_2)\) is sufficient for \((\theta_1, \theta_2)\) (i.e., for the family \(\{p(x; \theta_1, \theta_2), (\theta_1, \theta_2) \in \Theta_1 \times \Theta_2\}\)).

2. Let \((x_1, \ldots, x_n)\) and \((y_1, \ldots, y_n)\) be independent samples from populations with pdf's \(f(x - \theta_1)\) and \(f(x - \theta_2)\), respectively, with \(f(x) = e^{-x}, x \geq 0\) and \(\theta_1, \theta_2\) as parameters.

(i) Find the MLE \(\hat{\Delta}_n\) of \(\Delta = \theta_1 - \theta_2\) and calculate \(E(\hat{\Delta}_n)\) and \(\text{var}(\hat{\Delta}_n)\).

(ii) Find the normalizing constants \(a_n > 0\) and the explicit formula of the pdf of the nondegenerate limiting distribution of \(a_n(\hat{\Delta}_n - \Delta)\) as \(n \to \infty\).

(Hint: Using the moment generating function may simplify the calculations.)
3. Let \((x_1, \ldots, x_n), (y_1, \ldots, y_n), (z_1, \ldots, z_n)\) be independent samples from exponential populations with densities 

\[ f(x; \lambda_1), f(y; \lambda_2), f(z; \lambda_3), \]

respectively, where \(f(u; \lambda) = (1/\lambda)e^{-u/\lambda}, u > 0\) and \(\lambda_i > 0, i = 1, 2, 3\) are parameters.

(i) Construct the LR (Likelihood Ratio) test of level \(\alpha\) for testing 

\[ H_0 : \lambda_1 = \lambda_2 = \lambda_3 \text{ vs } H_1 : \lambda_i \neq \lambda_j \text{ for some } i, j. \]

(ii) Show that the power function of the LR test depends only on \(\lambda_2/\lambda_1, \lambda_3/\lambda_1\).

4. Given \(\theta, (x_1, \ldots, x_n)\) is a sample from a population with pdf \(e^{x-\theta}, x \leq \theta\) (notice that the distribution is concentrated on \((-\infty, \theta))\)

(i) Construct the conjugate family of prior pdf's parameterized by a two-dimensional parameter.

(ii) Assuming the prior pdf \(\pi(\theta)\) belongs to the conjugate family, find the Bayes estimator of \(\theta\) for the quadratic loss function \(L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2\).

5. Let \((x_1, \ldots, x_m)\) and \((y_1, \ldots, y_n)\) be independent samples of sizes \(m\) and \(n\) from populations with pdf's 

\[ f(x; \lambda_1) = (1/2)\lambda_1^3x^2e^{-\lambda_1x}, x \geq 0 \]

and 

\[ f(y; \lambda_2) = (1/6)\lambda_2^3y^2e^{-\lambda_2y}, y \geq 0, \]

respectively, with \(\lambda_1, \lambda_2\) as parameters.

(i) Based on the sufficient statistics, construct a pivot for \(\lambda_1/\lambda_2\).

(ii) Express the distribution of the pivot in terms of the \(F\)-distribution and construct a level \(1 - \alpha\) confidence interval for \(\lambda_1/\lambda_2\).

6. Let \((x_1, \ldots, x_n)\) be a sample from a normal population \(N(\mu, \sigma^2)\) with parameters \(\mu, \sigma^2\). The \(100(1 - \alpha)\)-th population percentile \(\eta_{a} = \eta(\alpha)\) is a function of \(\mu, \sigma^2, \eta_{a} = \eta_{a}(\mu, \sigma^2)\)

(i) Find the MLE \(\hat{\eta}_{a}(\alpha)\) of \(\eta(\alpha)\) and show that it is biased.

(ii) Find the limiting distribution of \(\sqrt{n}(\hat{\eta}_{a}(\alpha) - \eta(\alpha))\) as \(n \to \infty\).