INDIAN STATISTICAL INSTITUTE
CHENNAI CENTRE
M.STAT First Year
2015-16 Semester II

Nonparametric Inference and Sequential Analysis
Final Examination

Total Marks 50. *Duration: 3 hours*

1. Suppose that $X_1, ..., X_n$ are i.i.d. random variables having absolute continuous distribution function F. Define $X_{(n)} = max(X_1, ..., X_n)$. Find the asymptotic distribution of the random variable $Y = n(1 - F_{X_{(n)}}(x))$, where $F_{X_{(n)}}$ is the distribution function of $X_{(n)}$. 4

2. Suppose that $X_1, ..., X_n$ is an independent sample from cumulative distribution function F_1 and $Y_1, ..., Y_m$ is an independent sample from cumulative distribution function F_2. The parameter of interest is
$$\theta = \frac{\int x dF_1(x)}{\int y dF_2(y)} = \frac{\mu_1}{\mu_2}.$$
Let $\widehat{\theta} = \frac{\bar{X}}{\bar{Y}}$ be an estimator of θ, where $\bar{X}$ and $\bar{Y}$ are the sample means. Show that the influence functions are
$$L_{t;1}(x; F_1; F_2) = \frac{x - \mu_1}{\mu_2} \quad \text{and} \quad L_{t;2}(y; F_1; F_2) = -\frac{(y - \mu_2)\theta}{\mu_2}.$$
5

3. Gini mean difference (GMD) is defined as $E|X_1 - X_2|$, where X_1 and X_2 are two independent copies of X whose cumulative distribution function is F. Find a plug-in estimator of GMD. Write down the bootstrap algorithm to find the distribution of the estimator so obtained. 6

4. Define the consistency of a test statistic. Suppose F is a continuous symmetric cumulative distribution function with unique median θ. Prove that the Wilcoxon-signed rank test is consistent for tests on θ. 5

5. Define the Kendall's-τ. Find an estimator of the same. 4

6. Suppose that $\widehat{f}$ is a histogram estimator of the density f. Define
$$\bar{f} = E(\widehat{f}) = \sum_{j=1}^{m} \frac{p_j}{h} I(x \in B_j),$$
where $p_j = \int_{B_j} f(u)du$. Also define
$$l(x) = \left(max\{\sqrt{\widehat{f}(x)} - c, 0\}\right)^2, \quad u(x) = \left(\sqrt{\widehat{f}(x)} + c\right)^2,$$
where $c = \frac{Z_{\alpha/2m}}{2} \frac{m}{n}$ with bin size m and sample size n. Show that $(l(x), u(x))$ is an approximate $1 - \alpha$ confidence band for $\bar{f}$. 6

7. Suppose we observe two independent random samples $X_1, ..., X_n$ and $Y_1, ..., Y_n$ from distributions $F(x)$ and $G(x-\theta)$, respectively. We wish to test the null hypothesis $H_0 : \theta = 0$ versus the alternatives $H_1 : \theta > 0$. Under some assumption on F and G, show that the asymptotic relative efficiency (ARE) between the Mann-Whitney test and the independent t-test is $12Var(X)(\int f^2(y)dy)^2$. Obtain the ARE value when the parent population is
a) standard normal b) standard uniform. 8

8. Suppose that $X_1, ..., X_n, ...$ are i.i.d. normal random variables with mean μ and variance σ^2, both unknown. Discuss the sequential procedure for estimating μ as one needs to minimize $E(A|\bar{X}_n - \mu| + n)$, where $A > 0$ is known constant and $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. 4

9. Suppose that $X_1, ..., X_n, ...$ are i.i.d. random variables from a Poisson (λ) distribution with $0 < \lambda < \infty$. We assume that λ is unknown. Let us test a simple null hypothesis $H_0 : \lambda = \lambda_0$ versus a simple alternative hypothesis $H_1 : \lambda = \lambda_1 (> \lambda_0)$ with preassigned type I and type II error probabilities α, β, $0 < \alpha < \beta < 1$. Write down the SPRT. Also write down the expression for Wald's approximation of the OC function and the ASN function. 7

10. Suppose that $X_1, ..., X_n, ...$ are i.i.d. random variables with a $N(\mu, \mu)$ distribution involving an unknown parameter $\mu, 0 < \mu < \infty$. Find a $100(1-\alpha)\%$ fixed length confidence interval for μ. 5