NSM EndSem - B. Stat. III (2024 - 2025)

Date: 07.05.2025

Time: 180 Minutes Maximum Marks: 50

We follow usual notations from class lectures. Answer ALL questions.

1. (a) Define the univariate kernel density estimator (KDE).
(b) Derive an expression for the bias of KDE. Can we construct a kernel function so that the bias equals 0? Justify.
(c) Under appropriate assumptions, derive an upper bound for the MSE of KDE.
(d) State an upper bound for the MSE of the multivariate KDE. Interpret it.
[1+(1+2)+3+(2+1)]

2. Let $X_1, \ldots, X_n$ be i.i.d. from a continuous df F which is symmetric about θ. Consider the hypothesis $H_0 : \theta = 0$ against $H_1 : \theta > 0$, and the parameter $\gamma = P[(X_1 + X_2 + X_3)/3 > 0]$. Describe a U-statistic to estimate γ. Derive its asymptotic distribution. Compute all relevant constants under H_0.
Construct a test based on the proposed U-statistic. Prove its consistency under H_1.
[1+2+2+(1+1)]

3. Let $X_1, \ldots, X_n$ be an i.i.d. sample from a continuous df F.
(a) Define Walsh averages. Derive an expression for Wilcoxon's signed rank statistic using Walsh averages.
(b) Establish the asymptotic normality of this statistic. Give clear arguments, and state all necessary assumptions.
(c) Consider the median of all Walsh averages. Is this estimator better than the usual median? Justify. [(1+2)+3+2]

4. (a) Let $X_1, \ldots, X_n$ be i.i.d. from a continuous df F, with F symmetric about $\mu \in \mathbb{R}$. Prove that $E(\widetilde{X}_n) = \mu$ for any $n \geq 1$.
If F is *positively skewed* with $F(\mu) = 1/2$ (i.e., $F(\mu+x)+F(\mu-x) < 1$ for all $x \in \mathbb{R}$), then prove or disprove that (i) $E(\overline{X}_n) < \mu$ and (ii) $E(\widetilde{X}_n) > \mu$.
(b) State the run test for randomness of a binary sequence. How will you generalize this test for the case when the set is $\{0, 1, \ldots, 9\}$? [(2+3)+(1+1)]

5. (a) Fix $\Delta \in \mathbb{R}$. Assume $G(x) = F(x - \Delta)$ for all $x \in \mathbb{R}$ with F being a continuous df. Consider the two sample testing problem of equality of distributions when you have two i.i.d. samples of sizes m and n from F and G, respectively.
We want to test $H_0 : \Delta = 0$ against the alternative $H_1 : \Delta > 0$. State the usual two sample test based on means (assume appropriate moments exist), and the Mann-Whitney test. Prove that the tests are consistent under appropriate conditions, and hence, draw a comparison between the two tests.
(b) State Wald's SPRT. Consider the Bernoulli distribution with parameter θ. Derive the SPRT explicitly for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$. Show the different regions geometrically. [(2+3+1)+(2+1)]

6. Define Spearman's rank correlation (say, R_n) for bivariate data from a continuous distribution. Derive a simplified formula for R_n. Compute the expectation of R_n under the null hypothesis of independence.

Establish a mathematical relationship between R_n and Kendall's τ_n. Can we use this result to derive the asymptotic distribution of R_n? Justify.

[(1+2+2)+(3+2)]

All the best!