Indian Statistical Institute
B S D S Second Year,
Second Semester, 2025-26
Final Examination
Advanced Statistical Methods I: Part I

25.05.26 **Total score 50** **Duration: 90 minutes**

Name **Student ID**

---

1. Write your name and ID on each page.
2. Numbers in brackets denote total points allotted to each question.
3. You may use calculator.
4. Laptops and phones are not allowed.
5. You are allowed to bring four pages (two A4 sheets two sided OR four A4 sheets one sided) of notes.
6. Show all your work.

Qn no	1	2	3	Total
Marks	15	20	15	50
Obtained				

Name **Student ID**

1. (3+5+7=15) Data was recorded on 150 workplace accidents. The number of days since the previous accident is summarized in the following table

Days	0	1	2	3	4	5	6	7	8	9	10	11	12	13	15	20
Count	36	35	23	23	8	4	7	4	3	2	0	1	1	1	1	1

   (a) Use a nonparametric estimation method to estimate $\theta = P(X > 3)$.

   (b) Obtain a 95% confidence interval for θ based on the nonparametric estimator.

   (c) Test the hypothesis that the third quartile of the number of accidents is 3, that is, $P(X \leq 3) = 0.75$ at 5% level.

   You may use $z_{0.975} = 1.96$ for the normal quantile.

Name **Student ID**

Name **Student ID**

---

2. (10+5+5=20) A trial measured the relative efficacy of a treatment for a variety of doses.

Dose	0	3	3	4	5	6	8	8	9	10
Efficacy	2.56	0.78	4.79	2.44	1.01	1.05	-1.07	-1.37	-6.18	-12.03

We want to model the efficacy as the dependent variable in a nonparametric regression model. In particular, we'll use a Kernel Estimator with an Epanechnikov kernel with bandwidth 2.

$$K_2(t) = \frac{3}{8}\left(1 - \frac{t^2}{4}\right) \quad \text{for} \quad -2 \leq t \leq 2.$$

(a) Calculate the regression estimate this would give for efficacy when the dose is 8.

(b) Calculate the error (fitted minus actual) in fitting the efficacy at the point where the Dose= 8 if the true mean function is

$$E(Efficacy) = Dose - \frac{1}{6}Dose^2$$

(c) If, instead, I used a bandwidth of 3, what will be the Epanechnikov kernel?

Name **Student ID**

Name **Student ID**

Name **Student ID**

---

3. (8+7=15) Let $X_1, \cdots X_n$ be distinct observations and $n = 2m$ is even. We are interested in the Jackknife estimator of the median. Define the pseudo-observations

$$\tilde{\theta}_i = \hat{\theta}_n + (n-1)(\hat{\theta}_n - \hat{\theta}_{(-i)})$$

where $\hat{\theta}_n$ is the median of the full sample and $\hat{\theta}_{(-i)}$ is the median of the sample with the i-th observation deleted. The Jackknife estimator of median is the average of the pseudo observations.

   (a) Show that the pseudo observations take exactly two values. Find those values and their frequencies.

   (b) Show that the Jackknife estimator of median is same as $\hat{\theta}_n$.

Name **Student ID**