

MG221: Applied Probability & Statistics

Midterm 2019

1. In the first knock-out round of a certain T20 cricket tournament, 10 teams *viz.* Afghanistan, Australia, Bangladesh, England, India, New Zealand, Pakistan, South Africa, Sri Lanka and West Indies are randomly paired up to play each other in 5 matches. What is the probability that none of the five South Asian teams *viz.* Afghanistan, Bangladesh, India, Pakistan and Sri Lanka play one another in these 5 matches? [10]
2. Two friends X and Y go to school by taking the metro from the same station. Everyday, both of them independently arrive at the metro station at a random time (uniformly distributed) between 7:00 and 8:00 AM, wait for each other for a maximum of 15 minutes, and if the other person arrives within this time only then they take the metro together to school. Assuming that there is no waiting time for the metro, answer the following:
 - a. What is the probability that they take the metro together? [5]
 - b. Given that they took the metro together, what is the probability that at least one of them arrived before 7:30 AM? [5]
3. A sample of 3 balls are randomly drawn without replacement from an urn containing 5 red, 3 green and 2 blue balls. The balls are identical in all respect except their colours *i.e.* the 5 red balls are indistinguishable from one another and so are the 3 green and the 2 blue balls. Answer the following:
 - a. Given that we are only interested in the composition in the sample of 3 balls rather than what coloured ball appeared in which draw, write down the sample space and the corresponding probabilities of the individual outcomes of this experiment. [12]
 - b. Let B denote the number of blue balls and N denote the number of colours present in the sample of 3 balls. Give the joint distribution of (B, N) . [8]
 - c. Find the mean, median and mode of B . [5]
 - d. What is the “best” predictor of B , if one knows N ? What is this predictor called? [4+1=5]
 - e. What is the “best” linear predictor of N , if one knows B ? [5]
 - f. Find $\rho_{B,N}$, the correlation coefficient between B and N , and interpret its value. [2+3=5]
- 4 a. If $X_1, \dots, X_n$ are independent with $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, \dots, n$ then show that given real numbers $c_1, \dots, c_n$, $\sum_{i=1}^n c_i X_i \sim N(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \sigma_i^2)$. [5]
 - b. Weights of female and male employees working in an office building are both Normally distributed with respective means of 70 and 75 Kg and standard deviations of 15 and 10 Kg. If 6 male and 6 female employees board an elevator in that office building, which has a recommended maximum capacity of 1000 Kg, then what is the probability of the elevator being overloaded? [5]
5. The probability of an airline passenger, flying in a particular route, showing up for the scheduled flight even after buying a seat is 0.9. An airline operates a 100-seater plane in that route. Answer the following:
 - a. If the airlines sells 110 seats for one flight in that route, what is the probability of its plane flying full? [5]

- b. At least how many seats the airline must sell for one flight, so that it can be at least 99% certain of its plane flying full? [7]

6. A certain reserve forest is claimed to have at least 250 tigers residing in it. To test this claim, a wild-life activist group laid camera-traps all over the forest. In the first week of observation, they captured images of 40 distinct tigers. In the following week, they captured 50 distinct images, but 15 of them were of the same tigers, which were also photographed in the previous week. Answer the following:

- a. If indeed there are 250 tigers in the reserve forest, what is the probability of the second week's 50 images containing *exactly* 15 of the first week's images? [5]
- b. Since an increasing number of repeated images in the second week does not strengthen the claim of the reserve forest, and the probability of this random variable assuming a single value (like 15) is low, find the probability of the second week's 50 images containing *at least* 15 of the first week's images, assuming that there are 250 tigers in the reserve forest. [5]
- c. In light of the probabilities obtained in **a** and **b** above, how tenable do you think is the reserve-forest's claim? [3]

7. Car insurance policies are sold by a certain web-portal at a homogeneous Poisson rate of 2.4 per hour. The premium amounts of the sold policies are i.i.d. Normal with a mean of Rs.10,000 and an SD of Rs.2,000, and are independent of the times of sales of the policies. Answer the following:

- a. What is the probability of the portal selling between 10 and 15 policies (both inclusive) during a 5 hour period? [4]
- b. Starting counting from any point of time, what is probability that the 30-th policy sold, becomes the 5-th one exceeding a premium amount of Rs.12,000? How many policies is the portal expected to sell, to have 5 exceeding premium amounts of Rs.12,000 each? What is its most likely value? [4+1+1=6]
- c. Starting counting from 8:00AM, what is the probability that the 10-th policy is sold before (that day's) 12:00 noon? Give the mean, median, mode and standard deviation of the time of the sale of the 10-th policy. [3+1+2+1+1=8]
- d. Find the mean and variance of the total premium amount sold in a 10 hour period. What is the (approximate) probability of this amount exceeding Rs.200,000? [2+2+3=7]