Indian Institute of Technology Dharwad, Karnataka, India

EE 201 / EE 227: Data Analysis / Data Analysis (Second Half)
Quiz (Autumn 2021)

Date: 29/10/2021
Time Limit: 5:00 p.m. to 6:30 p.m., 1.5 Hours + 10 more minutes for submission

---

Instructions

1. Make suitable assumptions when necessary and state them clearly in your solutions.

2. Please write your own solutions. Any kind of copying will fetch zero marks for the entire quiz.

3. Submission is on Google Classroom. Late submissions will fetch zero marks regardless of how much late they are. So, plan your submissions properly.

Questions

1. (10 points) **History Repeats with Corrections:** Suppose you are taking a course where you are graded +1 and -1 for "pass" and "fail", respectively and the grades are equally likely. Let $X_1, X_2, \cdots X_n$ denote the n measurements of your grade in the class. You have designed two estimators for "pass" - a) $\hat{\theta_1} = X_1 + 1$ (yes, just the first measurement + 1!!) and b) $\hat{\theta_2} = \max\{X_1, X_2, \cdots X_n\}$. Determine whether these estimators are unbiased and consistent.

2. (10 points) **Queues:** Suppose you are at a movie theatre waiting for your friends to arrive. The number of people in the ticket queue is following a Poisson distribution with parameter λ. You count the number of people in the queue every 15 seconds for a total time of 20 minutes. Assume that each counting is independent. Come up with two estimators for λ. Both of them should be unbiased as well as consistent.

3. (10 points) **Highly likely:** Show that the maximum likelihood estimate of the parameter of an exponential distribution as well as that of a Geometric distribution is the same. Prove that they are asymptotically unbiased and consistent.

4. (10 points) **Over Confidence:**
   (a) (5 points) If the percentage of bowlers in the set of cricket players is 30%, how many players' details you have to collect so that, with a confidence level of 95%, the error in the estimate is $\leq 3.1\%$.
   (b) (5 points) If you have the details of 64 players and the percentage of bowlers in it is 35%, determine the corresponding confidence interval for the proportion of bowlers using a significance level of 1%.

Monday	62
Tuesday	47
Wednesday	44
Thursday	45
Friday	52

5. (10 points) **As cool as ice:** To determine whether the ice-creams are more likely to be sold on weekdays, data has been collected on all the ice-creams sold at an ice-cream parlour in Dharwad. The data (as shown in the table above) gives the numbers on each weekday.

   Design a test at a 5% level of significance to check if the ice-cream is equally likely to be sold on any weekday.

6. (10 points) **Dental Problems:** A medical report indicated that 75% of the children in India aged between 2 to 17 were affected by cavities in the past year. Government wants to assess whether the impact was similar in children living in your city. Samples from 125 children aged 2 to 17 living in your city were surveyed and 64 were reported with cavity over the past year. Is there a significant difference between children living in your city and the nation-wide data?