

IIIT Vadodara
Autumn 2020-21
CS605 Data Analytics
Lab-3 | August 15, 2020.
Parameter estimation and confidence interval

Q. 1: Parameter estimation:

1. Estimate the unknown parameter θ from a sample $X = (3, 3, 3, 3, 3, 7, 7, 7)$ drawn from a discrete probability distribution with following probability mass function (PMF),

$$P(3) = \theta$$

$$P(7) = 1 - \theta$$

Compute two estimators of θ :

- (a) the method of moments (MoM) estimator
- (b) the maximum likelihood estimator (MLE)

Also, estimate the standard error of each estimator of θ .

2. The number of times a computer code is executed until it runs without errors has a Geometric distribution with unknown parameter p . For 5 independent computer projects, a student records the following number of runs: 3 7 5 3 2.

Estimate p

- (a) by the MoM
- (b) by the MLE

3. Construct generalized programs for parts 1 and 2, for a user specified samples. Note down your observations in each case.

Q. 2: Confidence interval:

1. In order to ensure efficient usage of a server, it is necessary to estimate the mean number of concurrent users. According to records, the average number of concurrent users at 100 randomly selected times is 37.7, with a standard deviation $\sigma = 9.2$. Construct a 90% confidence interval for the expectation of the number of concurrent users. Make generalized program for user-specified values of mean and standard deviation.
2. Salaries of entry-level computer engineers have Normal distribution with unknown mean and variance. Three randomly selected computer engineers have salaries (INR): 30000, 50000, 70000. Construct a 90% confidence interval for the average salary of an entry-level computer engineer. Present a generalized program for a user-specified salaries (both number and values).