<u>Possible Topics to focus for Midterm examination</u>

<u>Topics from Lecture 1 (Read lecture notes on "Matrix Information- Lec1-A" and "Statistics-Sampling Part-1- lec-1B)</u>

<u>From Lec1-A</u>

1) Define followings: Square Matrix, Diagonal Matrix, Identity matrix, Symmetric Matrix. Show examples
2) Learn how to calculate determinant of 2*2 matrix and 3*3 Matrix
3) Learn about the formula to calculate the Inverse of a Matrix (again for 2*2 matrix and for 3*3 Matrix)

<u>From Lec1-B on Statistical Sampling</u>

Definition/Description

i) Population, Sample, Statistic, Estimator, Random Sampling, Sampling Distribution, Sampling Error

ii) Sampling distribution of Sample Mean (say $\bar{x}$, mean and variance of $\bar{x}$,)

(i) Probability Distribution,

(ii) Statistical Inference and various steps for statistical inference

(iii) <u>Interval</u> Estimations of Population parameters, Confidence Interval

Some more ..you just have to know definition of many things described in class.

4) What are the important characteristics of a good estimator of model parameters? (such **as Unbiased ness, Efficiency or minimum variance and Consistency**) Explain the meaning of each of the characteristics. **You can also describe them using appropriate diagrams**.

5) If the population from which the sample of X's are taken has normal probability density function with mean μ_x and 6_x^2 as variance, then what will be the sampling distribution of the random variable $\bar{X}$ in the following 3 cases:
   i) 6_x^2 is known and sample size n > 30
   ii) 6_x^2 is unknown and sample size n > 30
   iii) 6_x^2 is unknown and sample size n < 30

6) How you will derive the confidence interval of the population parameters such as population mean μ_x based on its estimates $\bar{X}$ from sample. **Provide the answer using both mathematical expression as well as diagrams**

<u>Topics from Lecture 2 (Read lecture notes: Probability -Lec 2 A and B: Fundamentals of Probability theory – Part I and Part II and Additional Notes on Probability)</u>

- What is the definition of probability based on (i) **Classical theory** (ii) **Frequency based theory?**
- Suppose X is a random variable (Discrete or Continuous). **Write the formula for the expected value (or mean) and variance of the random variable X when X is a discrete random variable and also the formula for expected value and variance of X when X is a continuous random variable.**

- Suppose Z is a random variable which is a function of another random variable X which has normal probability distribution. So, if

- $$Z = \frac{\bar{X} - \mu_x}{6_x}$$

- X has normal probability distribution X ~ N (μ_x , 6_x^2),

Then write the mathematical expression for probability density functions f(x) and f(z)

- If X and Y are jointly distributed random variables, then write the join density function of X & Y when both are discrete random variables and also for the case when both are continuous random variables.
- Describe and write the expression of Marginal Distribution of X and Y when both are discrete variables as well as when both are continuous variables.
-

<u>Topics from Lecture 3 & 4 (Read lecture notes: Linear Regression Model Part 1A and Part IB)</u>

1) Consider the following Linear Regression Model

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

In this model, derive the expression for estimated value of β_0 $and \beta_1$ by

(i) applying the method of moments and using the following two population moment restrictions;

$$E[X\,\varepsilon] = 0 \text{ and } E[\varepsilon] = 0$$

And

(ii) using the OLS (Ordinary Least Square) method which minimizes the sum of squares of residuals to derive the normal equations for estimating parameters $\beta_0\ and \beta_1$

Show that under both methodologies, normal equations for estimating the parameters are same.

Read about Goodness of Fit.

How do you evaluate how well the regression line fits the data?

Read notes on Linear Regression Part1-B, Read specially notes in the file ***Least Square Estimator and properties***

List the 6 assumptions of Classical Linear Regression Model which are

(i) Linearity, (ii) Full Rank (iii) Exogeneity of independent variables (iv)Assumptions regarding disturbance term- Homoscedasticity and nonauto correlation (v) Data Generation process: Stochastic vs non-stochastic(vi) Normal Distribution: The disturbance terms are normally distributed

Study all the assumption so that you can state them clearly and can explain them separately

2) In the model

Y = Xβ + ε (in vector-matrix form)

The formula for least square estimator of parameter vector **β is as follows:**

$\hat{\beta}$ (or b) = $[X'X]^{-1} X'Y$

Describe all the **small sample properties of the Least Square Estimator of parameter vector β.**

Prove that Least Square estimator is unbiased. Mention that which assumption you are using to prove this unbiasedness i.e., to prove $E(\hat{\boldsymbol{\beta}}$ (or b)|X) = $\boldsymbol{\beta}$

Define the standard error of parameters and explain each term in the following expression

$$Se(b_k) = ([S^2 (X'X)^{-1}]_{kk})^{1/2}$$

<u>Topics from Lecture 5 (Read lecture notes: Linear Regression Model Part 2 which will be covered in class on Oct 12)</u>

<u>Know in detail following topics as will be discussed in class on Oct 12</u>

- Model Specification
  - Omitting relevant variables
  - Including irrelevant variables
- Multicollinearity
- Errors of Measurement