Key Topics

Introduction to Probability and Statistics for Econometrics

Mathematical Concepts
Understand and verify the following properties

- \( \sum_{j=1}^{J} x_j p_j = x_1 p_1 + x_2 p_2 + \cdots + x_J p_J \)
- \( \sum_{j=1}^{J} (x_j + z_j) p_j = \sum_{j=1}^{J} x_j p_j + \sum_{j=1}^{J} z_j p_j \)
- \( \left( \sum_{j=1}^{J} x_j p_j \right)^2 = \sum_{j=1}^{J} \sum_{k=1}^{K} x_j x_k p_j p_k = \sum_{j=1}^{J} x_j^2 p_j^2 + \sum_{j=1}^{J} \sum_{k=1}^{K} x_j x_k p_j p_k \)

You should know this from prior course work, but the first two points are covered in the Random Variables and Distributions lecture and the third point is covered in the Estimator Variance lecture.

Probability Concepts
Understand that for events \( A \) and \( B \)

- If \( A \) and \( B \) are mutually exclusive, then \( \mathbb{P}(A \text{ or } B \text{ or both}) = \mathbb{P}(A) + \mathbb{P}(B) \)
- If \( A \) and \( B \) are independent, then \( \mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \cdot \mathbb{P}(B) \)
- \( \mathbb{P}(B|A) = \frac{\mathbb{P}(A \text{ and } B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B) \mathbb{P}(B)}{\mathbb{P}(A)} \)

The first two points are covered in the Probability Assignment lecture, the third point is covered in the Conditional Probability and Bayes Rule lecture.

Understand that for discrete random variables \( X \), which takes \( J \) distinct values, and \( U \), which takes \( K \) distinct values,

- \( \mathbb{E}(X) = \sum_{j=1}^{J} x_j \cdot \mathbb{P}(X = x_j) := \mu_X \)
- \( \text{Var}(X) := \mathbb{E}(X - \mu_X)^2 = \sum_{j=1}^{J} (x_j - \mu_X)^2 \cdot \mathbb{P}(X = x_j) \)
- \( \text{Cov}(X,U) := \mathbb{E}[(X - \mu_X)(U - \mu_U)] \)
  \[ = \sum_{j=1}^{J} \sum_{k=1}^{K} (x_j - \mu_X) (u_k - \mu_U) \cdot \mathbb{P}(X = x_j, U = u_k) \]
\[
\mathbb{E}(U|X = x_j) = \sum_{k=1}^{K} u_k \cdot \mathbb{P}(U = u_k|X = x_j)
\]

\[
\text{Var}(U|X = x_j) := \sum_{k=1}^{K} [u_k - \mathbb{E}(U|X = x_j)]^2 \cdot \mathbb{P}(U = u_k|X = x_j)
\]

The first point is covered in the Random Variables and Distributions lecture, the second and third points are covered in the Variance and Covariance lecture. The fourth point is covered in the Stronger Exogeneity Assumptions lecture.

Understand and verify the following implications

- \( \mathbb{E}(U|X) = 0 \) means \( \mathbb{E}(U|X = x_j) = 0 \) for all \((x_1, \ldots, x_J)\)

- \( \mathbb{E}(U|X) = 0 \Rightarrow \left\{ \begin{array}{l}
\mathbb{E}(U) = 0 \\
\text{Cov}(X, U) = 0
\end{array} \right. \)

- \( \mathbb{E}(U^2|X) = \sigma^2 \Rightarrow \mathbb{E}(U^2) = \sigma^2 \)

The first two points are covered in the Stronger Exogeneity Assumptions lecture. The second point is also addressed in the Exogeneity lecture.

### Statistical Concepts

Understand how to determine if

- an unknown parameter is identified

Understand

- For \( B \) an estimator of \( \beta \)

\[(\text{Mean}) \text{ Bias}(B) = \mathbb{E}(B) - \beta\]

This is covered in the Statistical Concepts lecture

### Linear Regression

For

\[ y_t = x'_t \beta + u_t, \quad t = 1, \ldots, n \]

### Interpretation

Understand how to interpret the elements of \( \beta \)

This concept is addressed in the Stochastic Models lecture

### Identification

Understand the concept of identification of \( \beta \)

This concept is addressed in the Identification lecture

Understand the role of each of these assumptions in identification of \( \beta \):
• $E(x_t u_t) = 0$
• $E(u_t|x_t) = 0$

The first point is addressed in the Estimation: Method of Moments lecture and further motivated in the Exogeneity lecture. The second bullet point is addressed in the Stronger Exogeneity Assumptions lecture.

Estimation
Understand how to construct Method-of-Moments estimators of $\beta$

This concept is addressed in the Estimation: Method-of-Moments lecture

Understand how to construct Ordinary Least Squares estimators of $\beta$

This concept is addressed in the Estimation: Least Squares lecture

Estimator Bias
Be able to establish that, under $E(u_t|x_1, \ldots, x_n) = 0$, the OLS estimators satisfy

$$E(\hat{\beta}|x_1, \ldots, x_n) = \beta.$$  

Error Variance
Understand the following definitions

• $u_t$ is unconditionally homoskedastic: $E(u_t^2) = \sigma^2$
• $u_t$ is conditionally homoskedastic: $E(u_t^2|x_1, \ldots, x_n) = \sigma^2$
• $u_t$ is unconditionally heteroskedastic: $E(u_t^2) = \sigma_t^2$
• $u_t$ is conditionally heteroskedastic: $E(u_t^2|x_1, \ldots, x_n) = \sigma_t^2$

Understand and verify the following implications

• $E(u_t^2|x_1, \ldots, x_n) = \sigma^2 \Rightarrow E(u_t^2) = \sigma^2$
• $E(u_t^2|x_1, \ldots, x_n) = \sigma_t^2 \Rightarrow$ either $E(u_t^2) = \sigma^2$ or $E(u_t^2) = \sigma_t^2$
  – example 1: $E(u_t^2|x_1, \ldots, x_n) = \omega^2 x_{1t}^2$ and $E(u_t^2) = \omega^2 E(x_{1t}^2) = \sigma^2$
  – example 2: $E(u_t^2|x_1, \ldots, x_n) = \begin{cases} \omega^2 x_{1t}^2 & \text{for } t < n' \\ \omega^2 x_{2t}^2 & \text{for } t \geq n' \end{cases}$ and $E(u_t^2) = \begin{cases} \omega^2 E(x_{1t}^2) & t < n' \\ \omega^2 E(x_{2t}^2) & t \geq n' \end{cases}$

These concepts are covered in the Error: Location and Scale Variation lecture