ECON21030 (S25): Econometrics - Honors

Lecturer: Joseph Hardwick
Notes by: Aden Chen

Tuesday 25^{th} March, 2025

Contents

1 Introduction

- The "small bin" problem, dimension reduction, and linearity.
- Given the model $y_i = \beta x_i + \epsilon_i$, ϵ_i is the **error**, and $\hat{\epsilon}_i = y_i - \hat{y}_i$ is the **residual**. The residual is sample-dependent.
-
  - $\min_b \sum |x_i - a - bx_i|$ gives an estimate of the conditional median of y given x. This is called the "quantile regression.'
  - $\min_b \sum |x_i - a - bx_i|^2$ gives the conditional expectation function $\mathrm{E}[Y|X]$. This is called the "ordinary least squares.'

2 Probability

Definition 2.1. A random variable X is **absolutely continuous** if there exists a density function f_X such that

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, \mathrm{d}t.$$

Remark 2.2. Absolutely continuous distributions assign probability 0 to any finite set of points.

2.1 Expectation

Proposition 2.3.

- E *is linear.*
- *If* $X \leq Y$ *with probability* 1*, then* $\mathrm{E}\, X \leq \mathrm{E}\, Y$.

Theorem 2.4 (Jensen's Inequality)**.** *If X is such that* $\mathrm{E}\, X$ *and* $\mathrm{E}\, g(X)$ *exist and g is convex, then*

$$g(\mathrm{E}\, X) \leq \mathrm{E}\, g(X)$$

where the inequality is strict if g is strictly convex and X is not constant.

Proof. From the convexity of g we know $g(x) \geq g(y) + g'(y)(x - y)$ for any x and y. Setting $y = \mu =: \mathrm{E}\, X$ gives

$$g(X) \geq g(\mu) + g'(\mu)(X - \mu), \quad \forall x, y.$$

Taking expectation on both sides gives the desired result. □

Example 2.5. Wages are often modeled using a log-normal distribution: $\log w \sim \mathcal{N}(\mu, \sigma^2)$. Then, $\mathrm{E} \log w = \mu$, but $\mathrm{E}\, w = \mathrm{E}(\exp \log w) \geq e^{\mu}$ (the inequality is strict when $\sigma^2 > 0$). It turns out that $\mathrm{E}\, w = \exp\left(\mu + \sigma^2/2\right)$.

2.2 Moments

Definition 2.6. 'If $\mathrm{E}(X^k)$ exists, then

- $\mathrm{E}(X^k)$ is the ***k*-th moment of *X***.

- $\mathrm{E}[(X - \mathrm{E}\, X)^k]$ is the ***k*-th central moment of *X***. The case $k = 2$ gives the variance of X.

Theorem 2.7 (Markov's Inequality)**.** *If $X \geq 0$ and $c > 0$, then*

$$\mathbb{P}(X \geq c) \leq \frac{\mathrm{E}(X)}{c}.$$

(Equality is attained when $\mathbb{P}(X = 0 \text{ or } X = c) = 1$.)

Proof. Construct

$$Y \coloneqq c \cdot \mathbb{1}_{\{x \geq c\}}(X).$$

Then $Y \leq X$ and

$$\mathrm{E}(Y) = c \cdot \mathbb{P}(X \geq c) \leq \mathrm{E}(X).$$

□

Theorem 2.8 (Chebychev's Inequality)**.** *If $c > 0$, then for any μ we have*

$$\mathbb{P}(|X - \mu| \geq c) \leq \frac{\mathrm{E}[(X - \mu)^2]}{c^2}.$$

Proof. Apply Markov's inequality to $(X - \mu)^2$. □