ECON3206/5206: Financial Econometrics

Week-1: A Review of Random Variables and Distributions

Rachida Ouysse
School of Economics[1]

May 20, 2020

Random Variables: Definitions

- RV and probability distribution
  - Two types of RV
    - Discrete RV: it takes a finite or countable number of values.
    - Continuous RV: it may take any value in a interval.
  - Probability distribution (how likely the values occur)
    - Discrete RV
      - Probability mass (pmf):

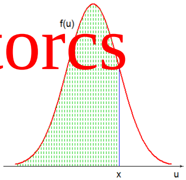
X	x_1	x_2	x_3	...
$P(X = x_i)$	p_1	p_2	p_3	...

      - Cumulative distribution (cdf): $F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i$

    - Continuous RV
      - Probability density (pdf): $f(u)$
      - Cumulative distribution (cdf):

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

Area under $f(u)$ curve up to x



Random Variables: Unconditional (marginal) Expectations (moments)

- Characterise RVs
  - Mean or expected value of X
    - The weighted average of all possible values, a measure of location
    - Discrete RV: $E(X) = \sum_i x_i P(X = x_i)$
    - Continuous RV: $E(X) = \int_{-\infty}^{\infty} u f(u) du$
  - Mean of $g(X)$ (eg. $X^2, e^X, \ldots$)
    - Discrete RV: $E[g(X)] = \sum_i g(x_i) P(X = x_i)$
    - Continuous RV: $E[g(X)] = \int_{-\infty}^{\infty} g(u) f(u) du$
  - Variance of X:
    - A measure of the amount of variation in all possible values

      $\text{Var}(X) = E\{[X - E(X)]^2\} = E(X^2) - [E(X)]^2$
  - Covariance between X and Y
    - A measure of association

      $\text{Cov}(X, Y) = E\{[X - E(X)][Y - E(Y)]\}$

Random Variables: Conditioning

– Conditional distribution of Y given X

- The distribution of Y when X is "known".
- It depends on the value of X.
- It is denoted as $Y|X$.

eg. the distribution of *wage* for *age* = 20,
the distribution of *wage* for *age* = 30, ...
wage|*age* depends on *age*.

bhp$_t$|***bhp***$_{t-1}$:
dist. of ***bhp***$_t$ when ***bhp***$_{t-1}$ is "fixed"

– Conditional expectation $E[g(Y)|X]$

- It is calculated with the conditional distribution, treating X as "known" or "fixed".
- It depends on X and, hence, is also a RV.

Random Variables

- Conditional mean of Y given X
  - $E(Y|X)$ generally depends on X.

eg. Linear regression $Y = a + bX + \varepsilon$.

$$E(Y|X) = a + bX + E(\varepsilon|X) = a + bX.$$

- Conditional variance of Y give X
  - $\mathrm{Var}(Y|X)$ generally depends on X.

eg. Linear regression $Y = a + bX + \varepsilon$.

$$\mathrm{Var}(Y|X) = \mathrm{Var}(\varepsilon|X).$$

Properties of Expectation operator

Let X and Y be RVs and a be a constant.

– Expectation is a linear operator

$E(a) = a, \quad \text{Var}(a) = 0;$

$E(aX) = aE(X);$

$E(X + Y) = E(X) + E(Y);$

$E[g(X)|X] = g(X)$ for any function $g(\cdot)$;

$E(Y) = E[E(Y|X)]$ (iterated expectations);

$E(Y|X) = E(Y)$ if Y is independent of X;

$$\boxed{E\left(\frac{X}{Y}\right) \neq \frac{E(X)}{E(Y)}}$$

– Variance in a nonlinear operator

$\text{Var}(aX) = a^2\text{Var}(X);$

$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y);$

$\text{Var}(Y) = E[\text{Var}(Y|X)] + \text{Var}[E(Y|X)].$

Sample moments

- Suppose we have (time series) observations
  $\{X_t\}_1^T = \{X_1, X_2, \ldots, X_T\}$ and $\{Y_t\}_1^T = \{Y_1, Y_2, \ldots, Y_T\}$
- Sample mean
  - $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t$
  - a measure of location (central tendency)
  - an estimator of population mean $E(X)$
- Sample variance and standard deviation
  - $\hat{\sigma}_X^2 = \frac{1}{T-1}\sum_{t=1}^{T}(X_t - \bar{X})^2$, $\hat{\sigma}_X = \sqrt{\hat{\sigma}_X^2}$
  - a measure of variation
  - an estimator of the population variance $\mathrm{Var}(X)$

Behold the summation operator

Linear operator:

$$\sum_{t=1}^{T}(a+bx_t+cy_t)=\sum_{t=1}^{T}a+b\sum_{t=1}^{T}x_t+c\sum_{t=1}^{T}y_t$$

Summation of a constant:

$$\sum_{t=1}^{T}a=Ta$$

Nonlinear relations:

$$\sum_{t=1}^{T}x_t y_t \neq \sum_{t=1}^{T}x_t \sum_{t=1}^{T}y_t$$

$$\sum_{t=1}^{T}\frac{x_t}{y_t} \neq \frac{\sum_{t=1}^{T}x_t}{\sum_{t=1}^{T}y_t}$$