

STAT4005 - Time Series
Tutorial 0 - Review of Preliminary Material

Ma Ting Tin, Martin

Week 0

0 Preliminary

0.1 Standard operations

Common Operators

Definition 1. *Let $X \sim F_X$ and $Y \sim F_Y$ (Cumulative density function).*

1. (Expectation) *The expectation of X is defined as*

$$\mathrm{E}(X) = \int x \,\mathrm{d}F_X(x) \overset{(\star)}{=} \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x,$$

where $(\star)$ holds only if the probability density function $f_X(x) = F_X'(x)$ exists.

2. (Variance) *The variance of X is defined as*

$$\mathrm{Var}(X) = \mathrm{E}\big[\{X - \mathrm{E}(X)\}^2\big] = \mathrm{E}(X^2) - \mathrm{E}(X)^2.$$

3. (Covariance) *The covariance between X and Y is defined as*

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big[\{X - E(X)\}\{Y - E(Y)\}\big] = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y).$$

4. (Correlation) *The correlation between X and Y is defined as*

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}.$$

Remark 1. *Except for the identity $(\star)$, all the definitions above does NOT rely on the existence of f_X and f_Y. Recall that f_X exists if and only if X is a continuous random variable.*

Remark 2. *The above quantities may not exist for some random variable, commonly those with too heavy tail. For example, if X follows the Cauchy distribution, $\mathrm{E}(X)$ does not exist.*

Remark 3. *Let $c \in \mathbb{R}$ and $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be a series of random variable.*

1. $\mathrm{E}(\cdot)$ *satisfies the following:*

   (a) $\mathrm{E}(\cdot)$ *is a linear operator, i.e.,*

$$\mathrm{E}\left(c + \sum_{i=1}^{n} X_i\right) = c + \sum_{i=1}^{n} \mathrm{E}(X_i).$$

   (b) *When $X_1, \ldots, X_n$ are independent,*

$$\mathrm{E}\left(\prod_{i=1}^{n} X_i\right) = \prod_{i=1}^{n} \mathrm{E}(X_i) \tag{1}$$

2. $\text{Var}(\cdot)$ satisfies the following:
   (a) (Unaffected by location shift) $\text{Var}(X + c) = \text{Var}(X)$.
   (b) When $X_1, \ldots, X_n$ are independent,

$$\text{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \text{Var}(X_i). \tag{2}$$

3. $\text{Cov}(\cdot, \cdot)$ satisfies the following
   (a) (Bilinearity)

$$\text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n} \text{Cov}\left(X_i, \sum_{j=1}^{m} Y_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \text{Cov}(X_i, Y_j).$$

   (b) (Unaffected by location shift) $\text{Cov}(X + c, Y) = \text{Cov}(X, Y)$.
   (c) (Relationship with Variance) $\text{Var}(X) = \text{Cov}(X, X)$.

(★☆☆) Standard Computation II

✎ **Exercise 1.** *Let $X_1, \ldots, X_n$ be a series of random variables. Define the sample average as*

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

show that

$$\text{Var}(\bar{X}) = \frac{1}{n^2} \left\{ \sum_{k=1}^{n} \text{Var}(X_k) + 2 \sum_{1 \le i < j \le n} \text{Cov}(X_i, X_j) \right\}.$$

Solution

$$\begin{aligned}
\text{Var}(\bar{X}) &= \text{Cov}\left(\frac{1}{n} \sum_{i=1}^{n} X_i, \frac{1}{n} \sum_{j=1}^{n} X_j\right) \\
&= \frac{1}{n^2} \text{Cov}\left(\sum_{i=1}^{n} X_i, \sum_{j=1}^{n} X_j\right) \\
&= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{Cov}(X_i, X_j) \\
&= \frac{1}{n^2} \sum_{\{i,j:i=j\}} \text{Cov}(X_i, X_j) + \frac{1}{n^2} \sum_{\{i,j:i \neq j\}} \text{Cov}(X_i, X_j) \\
&= \frac{1}{n^2} \sum_{k=1}^{n} \text{Cov}(X_k, X_k) + \frac{1}{n^2} \sum_{1 \le i \neq j \le n} \text{Cov}(X_i, X_j) \\
&= \frac{1}{n^2} \left\{ \sum_{k=1}^{n} \text{Var}(X_k) + 2 \sum_{1 \le i < j \le n} \text{Cov}(X_i, X_j) \right\}.
\end{aligned}$$

0.2 Unconditional and conditional independence

Independence

Definition 2. *Let $X \sim F_X$, $Y \sim F_Y$ and $Z \sim F_Z$. Then*

1. *X and Y are **independent** ($X \perp\!\!\!\perp Y$) if one of the following holds:*
   (a) $\mathbb{P}(X \leq x, Y \leq y) = F_X(x)F_Y(y)$ *for all* x, y.
   (b) $f_{X,Y}(x, y) = f_X(x)f_Y(y)$ *for all* x, y.
2. *X and Y are **conditionally independent** given Z $[(X \perp\!\!\!\perp Y)|Z]$ if one of the following holds:*
   (a) $\mathbb{P}(X \leq x, Y \leq y|Z = z) = F_{X|Z=z}(x)F_{Y|Z=z}(y)$ *for all* x, y, z.
   (b) $f_{X,Y|Z=z}(x, y) = f_{X|Z=z}(x)f_{Y|Z=z}(y)$ *for all* x, y, z.

Remark 4. *(1) and (2) are some immediate implications of independence among random variables.*

In this course, you may not be asked to justify the independence or conditional independence among random variables according to the above definition. However, you have to be able to distinguish their differences and write your argument accordingly in the assessment. The following is a simple example to illustrate the difference between those two concepts.

(★☆☆) Difference in unconditional and conditional independence

Exercise 2. *Let X_1, X_2, X_3 be independent random variables. Choose the appropriate answer.*

Type of independence	$X_1 + X_2$ *and* $X_2 + X_3$	$X_1 + X_2$ *and* $2X_1$
Unconditional	✔/✘	✔/✘
Conditional on X_1	✔/✘	✔/✘
Conditional on X_2	✔/✘	✔/✘
Conditional on X_2, X_3	✔/✘	✔/✘

Solution

Type of independence	$X_1 + X_2$ and $X_2 + X_3$	$X_1 + X_2$ and $2X_1$
Unconditional	✘	✘
Conditional on X_1	✘	✔
Conditional on X_2	✔	✘
Conditional on X_1, X_3	✘	✔

Understanding conditional dependence is crucial in handling a more complicated time series model. For example, it can help to find the unconditional expectation and variance in the hierarchical model.

Tower Property & Law of total variance

Theorem 1. *Let X and Y be random variables. Then the following hold:*

1. (Tower Property) $\mathrm{E}(X) = \mathrm{E}\{\mathrm{E}(X|Y)\}$.
2. (Law of total variance) $\mathrm{Var}(X) = \mathrm{E}\{\mathrm{Var}(X|Y)\} + \mathrm{Var}\{\mathrm{E}(X|Y)\}$.

(★☆☆) Application of Theorem 1

✎ **Exercise 3.** *Suppose that*

$$X = a_0 + a_1 Y + \epsilon,$$

where $[\epsilon|Y] \sim \mathrm{N}(0, Y^2)$, *and* $Y \sim \mathrm{N}(\mu, \sigma^2)$. *Evaluate* $\mathrm{E}(X)$ *and* $\mathrm{Var}(X)$.

Solution

- Note that $[\epsilon|Y] \sim \mathrm{N}(0, Y^2)$ implies $\mathrm{E}(\epsilon|Y) = 0$. Hence,

$$\mathrm{E}(X|Y) = \mathrm{E}(a_0 + a_1 Y + \epsilon|Y) = a_0 + a_1 \mathrm{E}(Y|Y) + \mathrm{E}(\epsilon|Y) = a_0 + a_1 Y,$$

and we have

$$\mathrm{E}(X) = \mathrm{E}\{\mathrm{E}(X|Y)\} = \mathrm{E}(a_0 + a_1 Y) = a_0 + a_1 \mu.$$

- Notice that Y is non-random given the value of Y, i.e., $\mathrm{Var}(Y|Y) = 0$. Hence,

$$\mathrm{Var}(X|Y) = \mathrm{Var}(a_0 + a_1 Y + \epsilon|Y) = \mathrm{Var}(\epsilon|Y) = Y^2,$$

and we have

$$\mathrm{E}\{\mathrm{Var}(X|Y)\} = \mathrm{E}(Y^2) = \mathrm{Var}(Y) + \{\mathrm{E}(Y)\}^2 = \sigma^2 + \mu^2.$$

Recall that $\mathrm{E}(X|Y) = a_0 + a_1 Y$, then

$$\mathrm{Var}\{\mathrm{E}(X|Y)\} = \mathrm{Var}(a_0 + a_1 Y) = a_1^2 \sigma^2.$$

Hence,

$$\mathrm{Var}(X) = \mathrm{E}\big\{\mathrm{Var}(X|Y)\big\} + \mathrm{Var}\big\{\mathrm{E}(X|Y)\big\} = \sigma^2 + \mu^2 + a_1^2\sigma^2 = \mu^2 + (a_1^2 + 1)\sigma^2.$$

0.3 Least-square estimation

Suppose we have n-couples of observations, i.e., $(y_i, x_{1i}, \ldots, x_{pi})$ for $i = 1, \ldots, n$. The quantity of interest is the relationship between Y and $X_1, \ldots, X_p$. The multiple regression model is given by

$$Y = \sum_{k=1}^{p} \beta_k X_k + \epsilon, \qquad \text{where } \epsilon \sim [0, \sigma^2].$$

The least-square estimator $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \ldots, \hat{\beta}_p)$ is the minimizer of the error sum of squares (SSE), i.e.,

$$\hat{\boldsymbol{\beta}} = \underset{\beta_1,\ldots,\beta_p \in \mathbb{R}}{\arg\min} \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{p} \beta_k x_{ki} \right)^2$$

Write $Y = (y_1, \ldots, y_n)^T$ and

$$\boldsymbol{X} = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{p1} \\ x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix}.$$

Then the least-square estimator is given by

$$\hat{\boldsymbol{\beta}} = (X^T X)^{-1} X^T Y.$$

Email contact: martinmtt@link.cuhk.edu.hk ; Personal website: https://sites.google.com/view/martinmtt/home