Advanced Econometrics (I).
Finance Management (F2&F3)

p1. 2023/2/21.

△ Statistical inference ⇒ We assumed that the observable random variables $Y_1, Y_2, \ldots, Y_n$ were independent and identically distributed.

implication of this assumption = the expected value of Y_i, $E(Y_i)$, is constant, i.e., $E(Y_i) = \mu$, does not depend on the value of any other variables.

⇒ this assumption is unrealistic in many inferential problems.

① the mean stopping distance for a particular type of automobile will depend on the speed that the automobile is travelling.

②. the mean potency of an antibiotic depends on the amount of time that the antibiotic has been stored.

⇒ We undertake a study of inferential procedures that can be used when a random variable Y, called the dependent variable, has a mean that is a function of one or more nonrandom variables $x_1, x_2, \ldots x_K$, called independent variables.

$\Rightarrow$ probabilistic model

$Y = \beta_0 + \beta_1 x + \varepsilon$: where ε is a random variable possessing a specified probability distribution with mean 0.

$E(Y) = \beta_0 + \beta_1 x$

$\Rightarrow$ We think of Y as the sum of a deterministic component $E(Y)$ and a random component.

$\Rightarrow$ The population has a distribution with mean $\beta_0 + \beta_1 x$ & variance σ^2.

* The expected value and variance of linear functions of Random variables.

$\Rightarrow$ If $a_1, a_2, \cdots, a_n$ are constants, we will need to find the expected value and variance of a linear

p 3.
2023/3/21.

function of the random variables $Y_1, Y_2, \ldots, Y_n$.

$$U_1 = a_1Y_1 + a_2Y_2 + a_3Y_3 + \cdots + a_nY_n = \underline{\sum_{i=1}^{n} a_iY_i}.$$

Theorem 5.12

Let $Y_1, Y_2, \ldots, Y_n$ and $X_1, X_2, \ldots, X_m$ be random variables with $\underline{E(Y_i)=\mu_i}$, and $\underline{E(X_j) = \xi_j}$. Define

$$U_1 = \sum_{i=1}^{n} a_iY_i, \quad U_2 = \sum_{j=1}^{m} b_jX_j \quad \text{for constants } \underline{a_1, a_2, \ldots, a_n}$$

and $\underline{b_1, b_2, \ldots, b_m}$. Then the following hold:

ⓐ $E(U_1) = \sum_{i=1}^{n} a_i\mu_i$

ⓑ $V(U_1) = \sum_{i=1}^{n} a_i^2 V(Y_i) + 2\sum\sum_{1\le i<j\le n} a_ia_j \operatorname{cov}(Y_i, Y_j)$,

where the double sum is over all ~~parist~~ pairs (i,j) with $i<j$.

ⓒ $\operatorname{Cov}(U_1, U_2) = \sum_{i=1}^{n}\sum_{j=1}^{m} a_ib_j \operatorname{cov}(Y_i, X_j)$

$$V(U_1) = E(U_1 - E(U_1))^2 = E\left[\sum_{i=1}^{n} a_iY_i - \sum_{i=1}^{n} a_i\mu_i\right]^2$$

$$= E\left[\sum_{i=1}^{n} a_i(Y_i - \mu_i)\right]^2 = E\left[\sum_{i=1}^{n} a_i^2(Y_i-\mu_i)^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\ne j}}^{n} a_ia_j(Y_i-\mu_i)\right.$$

$$\left.(Y_j-\mu_j)\right] = \sum_{i=1}^{n} a_i^2 E(Y_i-\mu_i)^2 + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\ne j}}^{n} a_ia_j E\left[(Y_i-\mu_i)(Y_j-\mu_j)\right].$$

2023/2/8

By the definition of variance and covariance, we have

$$V(U_1)=\sum_{i=1}^{n} a_i^2 V(Y_i)+\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n} a_i a_j \operatorname{cov}(Y_i, Y_j).$$

$$\operatorname{cov}(U_1,U_2)=E\{[U_1-E(U_1)][U_2-E(U_2)]\}$$

$$=E\left[\left(\sum_{i=1}^{n} a_i Y_i-\sum_{i=1}^{n} a_i \mu_i\right)\left(\sum_{j=1}^{m} b_j X_j-\sum_{j=1}^{m} b_j \xi_j\right)\right]$$

$$=E\left\{\left(\sum_{i=1}^{n} a_i (Y_i-\mu_i)\right)\left(\sum_{j=1}^{m} b_j (X_j-\xi_j)\right)\right\}.$$

$$=E\left[\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j (Y_i-\mu_i)(X_j-\xi_j)\right]$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j E\left[(Y_i-\mu_i)(X_j-\xi_j)\right]=\sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j \operatorname{cov}(Y_i, X_j)$$

Ex: Let Y_1, Y_2 & Y_3 be random variables, where $E(Y_1)=1$, $E(Y_2)=2$, $E(Y_3)=-1$, $V(Y_1)=1$, $V(Y_2)=3$, $V(Y_3)=5$,

$\operatorname{cov}(Y_1,Y_2)=-0.4$, $\operatorname{cov}(Y_1,Y_3)=1/2$, and $\operatorname{cov}(Y_2,Y_3)=2$.
Find the expected value and variance of $U=Y_1-2Y_2+Y_3$.
If $W=3Y_1+Y_2$, find $\operatorname{cov}(U,W)$.

If Y_1 and Y_2 are random variables with μ_1 & μ_2, respectively, the covariance of Y_1 and Y_2 is

$$cov(Y_1, Y_2) = E[(Y_1 - \mu_1)(Y_2 - \mu_2)]$$

⇒ The larger the absolute value of the covariance of Y_1 & Y_2, the greater the linear dependence between Y_1 and Y_2

⇒ positive values indicate that Y_1 increases as Y_2 increases, negative values indicate that Y_1 decreases as Y_2 increases.

⇒ A zero value of the covariance indicates that the variables are uncorrelated and that is no linear dependence between Y_1 & Y_2.

⇒ correlation coefficient, ρ, a quantity related to the covariance and defined as

$$\rho = \frac{cov(Y_1, Y_2)}{\sigma_1 \sigma_2}$$

Ex: Let Y_1 and Y_2 be jointly distributed random variables with finite variances.

(a) Show that $[E(Y_1 Y_2)]^2 \le E(Y_1^2) E(Y_2^2)$

(b). Let ρ denote the correlation coefficient of Y_1 and Y_2. Using the inequality in part (a), show that $\rho^2 \le 1$.

p6 2023/2/21.

Hint: Observe ~~observe~~ that $E[(tY_1 - Y_2)^2] \geq 0$ for any real number t or, equivalently,

$$t^2 E(Y_1^2) - 2t E(Y_1 Y_2) + E(Y_2^2) \geq 0$$

This is a quaratic expression of the form $At^2 + Bt + C$; because it is nonnegative, we must have $B^2 - 4AC \leq 0$

In practice, (a). represents (Cauchy-Schwartz inequality)

$$E[g(X)h(Y)] \leq \{E[g^2(X)]E[h^2(Y)]\}^{1/2}.$$

Theorem Assume $Y = a + bX$, $b \neq 0$, $\sigma_X^2 = Var(X)$ exists,

when $b > 0$, $\rho_{XY} = 1$; when $b < 1$, $\rho_{XY} = -1$

Proof: $\because \mu_Y = a + b\mu_X$, & $\sigma_Y^2 = b^2 \sigma_X^2$,

$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$

$= E[(X - \mu_X)(a + bX - a - b\mu_X)] = bE(X - \mu_X)^2$

$= b\sigma_X^2$

$\rightarrow \rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{b\sigma_X^2}{|b|\sigma_X^2} = \frac{b}{|b|} = \begin{cases} 1, & b>0 \\ -1, & b<0. \end{cases}$ PJ.

Ex: Assume $Y = a + bX + \xi$ ξ, r.v (random variable)

$E(\xi)=0$, $Var(\xi) = \sigma_\xi^2 > 0$ & $E(\xi X)=0$,

$\Rightarrow \sigma_Y^2 = b^2\sigma_X^2 + \sigma_\xi^2$

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{b}{\sqrt{b^2 + \sigma_\xi^2/\sigma_X^2}}$$

$\therefore |\rho_{XY}| \longrightarrow$ noise-to-signal ratio.

Ex: Best Linear Least Square Prediction

$Y = a + bX + \xi$

Mean squared Error (MSE),

$MSE(\alpha, \beta) = E[Y - (\alpha + \beta X)]^2$, α^*, β^*

$\uparrow$

Minimize

$$\begin{cases} \alpha^* = \mu_Y - \frac{cov(X,Y)}{Var(X)}\mu_X \\ \beta^* = \frac{cov(X,Y)}{Var(X)} = \rho_{XY}\sqrt{\frac{Var(Y)}{Var(X)}} \end{cases}$$

P8.
2023/2/24

proof: F.O.C.

$$\frac{\partial MSE(\alpha,\beta)}{\partial\alpha}\Big|_{(\alpha^*,\beta^*)} = -2E[Y-(\alpha^*+\beta^* X)]=0.$$

$$\frac{\partial MSE(\alpha,\beta)}{\partial\beta}\Big|_{(\alpha^*,\beta^*)} = -2E[X(Y-(\alpha^*+\beta^* X))]=0.$$

$$\xi = Y-(\alpha^*+\beta^* X) \Rightarrow Y=\alpha^*+\beta^* X+\xi.$$

$$\Rightarrow E(\xi X)=0,$$

$\Rightarrow$ F.O.C $\Rightarrow E(\xi X)=0. \Rightarrow$ $\underline{\xi}$ does not include any information which is related to X,

$\Rightarrow$ <u>phillips Curve</u> : <u>inflation rate</u> and <u>unemployment rate</u>.

<u>Portfolio plan</u> Assume $Z=\alpha+\beta X+cY$, $\beta=c=1$, $\alpha=0$,

$\cancel{Var(Z)=Var(X+Y)}$

$$Var(X+Y)=\sigma_X^2+\sigma_Y^2+2cov(X,Y)$$

$$\Rightarrow cov(X,Y)>0 \Rightarrow Var(X+Y)>\sigma_X^2+\sigma_Y^2,$$

$$\Rightarrow \underline{Var(Z)=Var(\alpha+\beta X+cY)=\beta^2\sigma_X^2+c^2\sigma_Y^2+2\beta c\, cov(X,Y)}.$$