

AG432 FINANCIAL QUANTITATIVE METHODS  
COURSEWORK TEST

LEWIS BRITTON

ACADEMIC YEAR 2020/2021

# Table of Contents

<b>1</b>	<b>Variance-Covariance</b>	<b>1</b>
<b>2</b>	<b>Errors</b>	<b>2</b>
2.1	Introduction to Errors . . . . .	2
2.2	Banks in 2008 . . . . .	3
2.2.1	Pre-Crisis . . . . .	3
2.2.2	Post-Crisis . . . . .	3
<b>3</b>	<b>Matrix OLS</b>	<b>5</b>
3.1	Preliminaries . . . . .	5
3.2	Assumptions . . . . .	6
3.3	Beta Estimations . . . . .	7
3.3.1	Gauss-Markov Theorem: Beta Estimations . . . . .	7
3.4	Variance Estimations . . . . .	9
3.4.1	Variance-Covariance Matrix for $\hat{\beta}$ . . . . .	9
3.5	Unbiasedness of OLS . . . . .	10

# 1 Variance-Covariance

Given a row vectors of portfolio weights for three companies  $v = [0.1 \ 0.3 \ 0.6]$  and the variance-covariance matrix for the three companies  $S = \begin{bmatrix} 1.5 & -0.1 & 1 \\ -0.1 & 1 & 2 \\ 1 & 2 & 1.5 \end{bmatrix}$ , calculate the variance of the portfolio.

$$\begin{aligned}\sigma_p^2 &= W_p \text{cov}_p W_p' \\ \therefore \sigma_p^2 &= W_p \text{cov}_p W_p' = [w_1 \ w_2 \ w_3] \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \\ &= [0.1 \ 0.3 \ 0.6] \begin{bmatrix} 1.5 & -0.1 & 1 \\ -0.1 & 1 & 2 \\ 1 & 2 & 1.5 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix} \\ &= [(0.15 - 0.03 + 0.6) \ (-0.1 + 0.3 + 1.2) \ (0.1 + 0.6 + 0.9)] \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix} \\ &= (0.15 - 0.03 + 0.6)0.1 \ (-0.1 + 0.3 + 1.2)0.3 \ (0.1 + 0.6 + 0.9)0.6 \\ &= 0.072 + 0.447 + 0.96 \\ \sigma_p^2 &= 1.479\end{aligned}$$

## 2 Errors

Use the concept of Type I and Type II errors to explain bank lending behaviour before and after the financial crisis in 2008

### 2.1 Introduction to Errors

Hypotheses		
<b>H<sub>0</sub></b> : Null Hypothesis <i>Aim to Reject; may Fail-to-Reject; rarely Accept</i>		
<b>H<sub>A</sub></b> : Alternative Hypothesis <i>Favour in the case of Rejection of the Null</i>		
<b>Type I Error</b> : Rejection of null hypothesis when it is true		
<b>Type II Error</b> : Failure-to-reject null hypothesis when it is false		
<b>Reduce Type I Error Risk</b> : Reduce significance level; harder to reject null		
<b>Reduce Type II Error Risk</b> : Use large sample; ensuring significant spread		
<b>Probability <math>\alpha</math></b> : Probability of making Type I Error		
<b>Probability <math>\beta</math></b> : Probability of making Type II Error		
Error Examples		
Decision	H <sub>0</sub> is True (Accused is Innocent)	H <sub>0</sub> is False (Accused is Guilty)
Reject H <sub>0</sub> (Accused Convicted)	<b>WRONG</b> Decision (Type I Error) Probability $\alpha$	<b>CORRECT</b> Decision
Fail-To-Reject H <sub>0</sub> (Accused Goes Free)	<b>CORRECT</b> Decision	<b>WRONG</b> Decision (Type II Error) Probability $\beta$
t-stat		
$t\text{-stat} = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{N}}\right)}$		
t-stat  > 1.96: reject the null hypothesis at the 5% significance level;		
t-stat  > 2.58: reject the null hypothesis at the 1% significance level		
Given 1000 degrees of freedom		
p-value		
p-value < 0.05: reject the null hypothesis at the 5% significance level;		
p-value < 0.01: reject the null hypothesis at the 1% significance level		
Given 1000 degrees of freedom		
Confidence Relevance		
5% and 1% significance levels are also referred to as the 95% and 99% confidence [in rejecting the null] intervals. Correct syntax: refer to 5% and 1% significance levels when referring to p-values; 95% and 99% confidence levels when referring to hypotheses.		
Error Relevance		
5% and 1% levels are used to ensure accuracy and reduced probability of making a Type I error. That is, “accepting a max 5%/1% chance that you are wrong when rejecting the null”; “you are min 95%/99% confident you are right when rejecting”.		

## 2.2 Banks in 2008

Banks experienced a lot of cases of Type I Errors prior to the crisis in 2008 and they experienced many cases of making Type II Errors post-2008 crisis. These tended to be centred around the same hypothetical null and alternative hypotheses:

$H_0$ : Mortgage Seeker is Risky

$H_A$ : Mortgage Seeker is Not Risky

### 2.2.1 Pre-Crisis

Prior to the 2008 crisis, it could be argued that the hypothetical null hypothesis ( $H_0$ ) is as follows: “people seeking mortgage loans are risky”; and the corresponding alternative hypothesis ( $H_A$ ) is: “people seeking mortgage loans are not risky”. In other words, banks wanted to reject the null hypothesis in favour of the alternative hypothesis so they were able to give out more loans to non-worthy people. That is, to issue sub-prime mortgages, for their own self-interest. However, in practice, banks found this null hypothesis to in fact be true, therefore contradicting their desires. Furthermore, they rejected the null hypothesis even though it was true, making this a Type I Error [Reject  $H_0$  when  $H_0$  is true]. Therefore, loan application granted where it shouldn't be. Cost of Type I Error = loss of mortgage value plus remaining interest to be paid [assigned:  $\$500,000 + x$ ] upon default.

### 2.2.2 Post-Crisis

Post-crisis, the null hypothesis and alternative hypothesis remained the same: “people seeking mortgage loans are ( $H_0$ )/are not ( $H_A$ ) risky”. Banks would simply react in a different manner, with added precaution after observing the domino effect they had helped cause over the course of 2007 and 2008. This time round, banks may refuse to give a perfectly eligible citizen a mortgage loan. That is, they find  $H_0$  “people seeking mortgage loans are risky” to be false but fail to reject it. Meaning that they are acting as if it is true. In this case they should be favouring the alternative hypothesis but neglect this. This is therefore a Type II Error [Fail-to-reject  $H_0$  when  $H_0$  is false]. Therefore, loan application aren't granted where they should be. Cost of Type II Error = full potential interest [assigned:  $x = \$100,000$ ]. By doing this they turn away the opportunity to earn the interest however, it is safer to lose just potential interest (acting in a Type II Error – this case) than it is to lose the full loan in the case of

default, plus the remaining potential interest (acting in the Type I Error – pre-2008). But doing this over a long period will have significantly negative effects.

From a more, hypothetical, statistical point of view (and using the post-2008 reaction of banks): say the null hypothesis was as described, the t-test was run and the associated p-value of the t-test was 0.0003 (p-value < 0.01). This would state that “we should be willing to accept a 0.03% probability that we are wrong when rejecting  $H_0$ ”. Therefore we would reject the null hypothesis that mortgage seekers are risky. But despite this statistical evidence to the contrary, it was failed to be rejected anyway meaning banks could act in their own (safe) favour. Thus, rejecting mortgage loans when they should grant them; turning away safe customer when they should accept them.

## 3 Matrix OLS

### 3.1 Preliminaries

Recall basic OLS regression:

- $\hat{Y}_i = \hat{\alpha} + \hat{\beta}X_i + \varepsilon_i$
- $Y = \alpha + \beta_1X_1 + \beta_2X_2 + \dots + \beta_KX_K + \varepsilon_i$
- $Y = \alpha + \sum_{k=1}^K \beta_kX_k + \varepsilon_i$
- $Y$ : the dependent variable (what effects  $Y$ ?)
- $X_k$ : the explanatory variable  $k$  (does it effect  $Y$ ?)
- $\alpha$ : the *y-intercept* (level of  $Y$  when  $X = 0$ )
- $\beta$ : the slope of the line (degree of relationship)
- $\varepsilon$ : the residuals (random error term)
  - Captures explanation not contained within explanatory variables(s)
- For observation  $i$
- (Hat denotes ‘estimation’)

Know that:

1.  $\beta = \frac{\Delta Y}{\Delta X} = \frac{\partial Y}{\partial X}$ 
  - “For each (+) unit on the  $x$  axis, expect the  $y$  value to  $\Delta$  by  $\beta$ ”
2.  $\varepsilon_i = (Y_i - \hat{Y}_i)$ 
  - Each observation has a predicted  $\hat{Y}_i$  on the line of best fit, directly above or below its real  $Y_i$
  - Distance between each of these value is the *error*
3. This is a *Linear Regression* which uses a straight ‘line of best fit’ (hyperplane)
  - For non-linear data, a *Polynomial Model* can be used to account for concave/convex data
  - Hence:  $Y_i = \alpha + \beta_1X_1 + \beta_2X_2^2 + \dots + \beta_KX_K^K + \varepsilon_i$

## 3.2 Assumptions

1.  $\sum \varepsilon_i = 0$ 
  - “The sum of all residuals in the model are equal to zero”
2.  $\text{var}_{\varepsilon_i} = E(\varepsilon_i^2) = \sigma^2$ ; or  $E(\varepsilon\varepsilon') = \sigma^2 I$  in vector format
  - “Observations have constant errors”
  - *Homoskedasticity*: constant errors
  - *Heteroskedasticity*: non-constant errors (adjust to *Robust Standard Errors*)
    - $\text{var}_{\varepsilon} = \sigma^2 \omega_i^2$
    - For  $i \in \{1, \dots, N\}$
    - The *White Test* is used for heteroskedasticity testing ( $H_0$ : Homoskedasticity is present  $\therefore$  you wish to fail-to-reject  $H_0$ ; associated p-value > critical value)
3.  $\text{cov}_{\varepsilon_i, \varepsilon_j} = 0$  (for  $i \neq j$ ) or  $E(X\varepsilon') = 0$  in vector format
  - “Error terms should be uncorrelated”
  - Expected observations should be uncorrelated with errors
  - *Instrumental Variable Approach* (IV): where explanatories are correlated with error terms (endogeneity problem), don’t use OLS
    - $X$  may be *exogenous*, where factors within the model cause changes in  $X$ ;  $\therefore \Delta X$  associated w/  $\Delta \varepsilon$
    - *Endogenous* context: variables correlated with error term
    - *Exogenous* context: variables uncorrelated with error term
    - If two explanatories have high collinearity, omit one
4. (Extra for matrix OLS)  $(X'X)^{-1}$  exists
  - *Multicollinearity* is not present Under multicollinearity, two variables may have high correlation (e.g. 1, -1)
  - Model would struggle to understand which one explains  $Y$
  - Hence, w/o multicollinearity, each explanatory gives unique information



### 3.3 Beta Estimations

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i + \varepsilon_i$$

$$\therefore \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_N \end{bmatrix} \begin{bmatrix} 1 & X_{1,1} & X_{2,1} & \dots & X_{k,1} \\ 1 & X_{1,2} & X_{2,2} & \dots & X_{k,2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & X_{1,N} & X_{2,N} & \dots & X_{k,N} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

- Going forward, the following are unknown:  $\beta, \sigma^2$
- We wish to maximize the total variation of  $\beta_i X_i$  in order to explain  $Y$
- In result, this will minimise variation of the random error value
- In other words, choose  $\alpha$  and  $\beta$  to  $\min\{\sum(Y_i - \hat{Y}_i)^2\}$
- Optimization is required  $[E(\varepsilon\varepsilon')] \text{ with respect to } \beta]$ , to the first order
- Again, choose  $\alpha$  and  $\beta$  to:
  - $\min\{\sum(Y_i - \hat{Y}_i)^2\} = \min\{E(\varepsilon_i^2)\} = \min\{E(\varepsilon\varepsilon')\}$
  - Thus, minimizing for  $\sim 0$  @ stationary point
  - The following uses relevant matrix algebraic notation
  - Hence, the use of  $(Y - \beta X)$  to represent the (error) distance of observations from the line of best fit
  - As  $\varepsilon = (Y - \beta X)$  in a vector context:

$$\frac{d(\min\{E(\varepsilon\varepsilon')\})}{d\beta} = \frac{d((Y - \beta X)(Y - \beta X)')}{d\beta}$$

$$\therefore \frac{d(Y Y' - Y \beta' X' - X \beta Y' + X \beta \beta' X')}{d\beta} = 0$$

$$\therefore \hat{\beta} = (X' X)^{-1} X' Y$$

#### 3.3.1 Gauss-Markov Theorem: Beta Estimations

This states that, given the aforementioned assumptions, there are no alternative linear estimators of  $\beta$ 's with a smaller variance. That is, the OLS model is the Best Linear Unbiased Efficient (BLUE) estimator.

Recall:  $\hat{\beta} = (X'X)^{-1}X'Y$ , therefore:

$$\hat{\beta} = (X'X)^{-1}X'(X\beta + \varepsilon)$$

$$\therefore \hat{\beta} = \beta + (X'X)^{-1}X'\varepsilon$$

- As  $(X'X)^{-1}X'X = I$ ;
- OLS is unbiased given that either [1]  $X$  is not stochastic or [2]  $X$  is stochastic but independent of  $\varepsilon$
- Therefore:
- $E(\hat{\beta}) = E(\beta) + (X'X)^{-1}X'\varepsilon = \beta + (X'X)^{-1}X'E(\varepsilon)$   
or
- $E(\hat{\beta}) = E(\beta) + (X'X)^{-1}X'\varepsilon = \beta + (X'X)^{-1}E(X'\varepsilon)$
- Therefore,  $\hat{\beta}$  is a linear estimator of  $\beta$
- Where  $\beta$  has minimal variance such that:  $var_{\beta} - var_{\hat{\beta}} > 0$
- Meaning that  $\hat{\beta}$  estimations should never exceed  $\beta$

### 3.4 Variance Estimations

- Recall:

$$- \hat{\beta} = (X'X)^{-1}X'Y$$

$$- \varepsilon = (Y - \beta Y)$$

- Estimate  $\sigma^2$  with  $\hat{\sigma}^2$  in a similar manner to  $\beta$
- For  $N$  observations and  $k$  explanatory variables
- $\hat{\sigma}^2 = \frac{\varepsilon'\varepsilon}{N-k}$
- $\therefore \hat{\sigma}^2 = \frac{(Y-\beta X)'(Y-\beta X)}{N-K}$
- Thus, like  $\beta$ ,  $\hat{\sigma}^2$  is an unbiased estimator of  $\sigma^2$

#### 3.4.1 Variance-Covariance Matrix for $\hat{\beta}$

$$\begin{aligned} & E\left((\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right) \\ &= E\left(((X'X)^{-1}X'\varepsilon)((X'X)^{-1}X'\varepsilon)'\right) \\ &= E\left((X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1}\right) \end{aligned}$$

Assume  $X$  is stochastic and rearrange:

$$= (X'X)^{-1}X'E(\varepsilon\varepsilon')X(X'X)^{-1}$$

Recall  $E(\varepsilon\varepsilon') = \sigma^2 I$ :

$$\begin{aligned} &= (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1} \\ &= \sigma^2 I(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2 I(X'X)^{-1} \text{ as req.} \end{aligned}$$

### 3.5 Unbiasedness of OLS

- Meaning applying  $\hat{\beta}$  to all observations will provide a central *true value*  $\beta$  on the distribution of  $\beta$  values
  - Hence  $E(\hat{\beta}) = \bar{\beta}_i$
- $\hat{\beta} = (X'X)^{-1}X'$  is referred to as a *projection matrix* which is a vector that captures the error, denoted by  $P_x$

Matrix algebra shown below:

$$\begin{aligned}E(\hat{\beta}) &= \bar{\beta}_i \\ \therefore E(\hat{\beta}) &= E((X'X)^{-1}X'Y) \\ &= E((X'X)^{-1}X'(X\beta + \varepsilon)) \\ &= E((X'X)^{-1}X'X\beta + (X'X)^{-1}X'\varepsilon) \\ &= E((X'X)^{-1}X'\varepsilon) + \beta \\ &= \beta \text{ as req.}\end{aligned}$$