DEPARTMENT OF ACCOUNTING & FINANCE

University of Strathclyde Business School

Department of Accounting and Finance

Class code AG432

Class Test – Friday 16th October 2020

90 minutes

This test is worth 30% of your final mark for the class.

You may detach the question sheet from this answer sheet.

Make sure your name and registration number are clearly *printed*.

Name : Lewis Britton

Registration Number: 201724452

Start Time: 11:16

Finish Time: 12:14

Question 1 (33%)

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:

Answer

$$\sigma_P^2 = W_p cov_p W_p'$$

$$\therefore \sigma_P^2 = W_p cov W'_p = [w_1 \quad w_2 \quad w_3] \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \sigma_{1,3} \\ \sigma_{1,2} & \sigma_2^2 & \sigma_{2,3} \\ \sigma_{1,3} & \sigma_{2,3} & \sigma_3^2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$= [0.1 \quad 0.3 \quad 0.6] \begin{bmatrix} 1.5 & -0.1 & 1 \\ -0.1 & 1 & 2 \\ 1 & 2 & 1.5 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix}$$

$$= [(0.15 - 0.03 + 0.6) \quad (-0.01 + 0.3 + 1.2) \quad (0.1 + 0.6 + 0.9)\,] \begin{bmatrix} 0.1 \\ 0.3 \\ 0.6 \end{bmatrix}$$

$$= (0.15 - 0.03 + 0.6)0.1 + (-0.01 + 0.3 + 1.2)0.3 + (0.1 + 0.6 + 0.9)0.6$$

$$= 0.072 + 0.447 + 0.96$$

$$\sigma_P^2 = 1.479$$

"Variance of Portfolio is a [scalar] 1.479"

Question 2 (33%)

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

Answer

1 Intro To Errors

H_0: Null Hypothesis – assumed true unless **rejected** – true if **fail to reject**
H_A: Alternative Hypothesis – favoured if null is rejected

Recall the criteria to reject a null hypothesis:

- **p-value** < 0.01 (for the 1% significance level; 99% confidence interval);
- **p-value** < 0.05 (for the 5% significance level; 95% confidence interval);
- **p-value** < 0.10 (for the 10% significance level; 90% confidence interval)
- We like to use a 1% level
    - "Should be willing to accept 1% probability that we are wrong when rejecting H_0" therefore:
    - **Probability $\boldsymbol{\alpha}$**: probability of making a Type I Error
        - Reduce risk of making this error by reducing the value of α so that it is harder to obtain rejection of the null – especially when true
    - **Probability $\boldsymbol{\beta}$**: probability of making a Type II Error
        - Reduce risk of making this error by using a large enough sample size, ensuring a significantly consistent difference

Decision	H_0 is True (Accused is Innocent)	H_0 is False (Accused is Guilty)
Reject H_0 (Accused Convicted)	**WRONG** Decision (Type I Error) Probability α	**CORRECT** Decision
Fail-To-Reject H_0 (Accused Goes Free)	**CORRECT** Decision	**WRONG** Decision (Type II Error) Probability β

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

***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.

***Post-2008**:* 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.

Question 3 (34%)

Use matrix notation to derive the estimates of OLS regressions and briefly discuss the procedures of conducting OLS regressions.

Answer

Preliminaries

Recall basic regression:

- $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i + \varepsilon_i$
- $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon_i$
- Y is the dependent variable (what effects Y?)
- X is the explanatory variable (does it affect Y?)
- α y-intercept (level of Y when X = 0)
- β slope of the line (severity of the relationship)
- ε residuals (random error term) (outliers from the best fit)
  - Captures any explanation not contained within the explanatory variables
- For observation i
- *Recall also, when a variable has a 'hat' accent it is an **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 change by β

2: $\varepsilon_i = \left(Y_i - \hat{Y}_i\right)$

- Each observation will have a predicted $\hat{Y}_i$ on the best fit line, directly above or below their real Y_i
- Distance between these values, for each observation, is its **error**

3: This is a Linear Regression and uses a straight best-fit (hyperplane)

- For non-linear data, a Polynomial Model can be used to account for concave/convex data using X^2 values
- Hence: $Y_i = \alpha + \beta_1 X_1 + \beta_2 X_2^2 + \cdots + \beta_k X_k^k + \varepsilon_i$

Assumptions

1: $\sum \boldsymbol{\varepsilon_i} = \mathbf{0}$
- The sum of all residuals in the model are equal to zero

2: $\boldsymbol{var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma^2}$; or $\boldsymbol{E(\varepsilon\varepsilon') = \sigma^2 I}$ in vector format
- "Observations have constant errors"
- Homoscedasticity: constant errors
- Heteroscedasticity: non constant errors (must adjust → Robust Model)
    - $var(u) = \sigma^2\omega_i^2$
    - Where for: $i = \{1, \dots N\}$; i denotes that variance of the error can be different for each observation
    - The **White Test** can be used to hypothesise and test for this

3: $\boldsymbol{cov(\varepsilon_i, \varepsilon_j) = 0}$ **(for** $\boldsymbol{i \neq j}$**)**; or $\boldsymbol{E(X\varepsilon') = 0}$
- Error terms should be uncorrelated
- Expected observations should be uncorrelated with errors
- **Instrumental Variable Approach**: where explanatories are correlated with error terms
    - X may be **Endogenous**, where factors within the model cause changes in X therefore, ΔX associated with $\Delta\varepsilon$
    - Endogenous context: variables correlated with error term ε
    - Exogenous context: variables uncorrelated with error term ε
    - Use the IV (Instrumental Variable) approach, not OLS (Ordinary Least Squares)
    - If two explanatory variables have high collinearity, omit one

4: [Extra] $\boldsymbol{(X'X)^{-1}}$ Exists
- Multicollinearity is not present
- Under multicollinearity, two variables may have high correlation $(1, -1)$
- The model would struggle to understand which one explains Y
- Hence, w/o multicollinearity, each explanatory gives unique information

Beta Estimations [1]

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

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

- Going forward, the following are unknown: $\boldsymbol{\beta, \sigma^2}$
- We wish to maximise the total variation of $\boldsymbol{\beta_i X_i}$ in order to explain $\boldsymbol{Y}$
- In result, this will minimise variation of random error value ε_t: $\boldsymbol{var(\varepsilon_i)}$ or $\boldsymbol{E(\varepsilon\varepsilon')}$
- In other words: choose $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to $\mathbf{min} \sum\left(\boldsymbol{Y_i - \hat{Y}_i}\right)^2$
- Optimisation is required [$\boldsymbol{E(\varepsilon\varepsilon')}$ with respect to $\boldsymbol{\beta}$], to the first-order:
  - **Must**: choose $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ to:
  - $\mathbf{min} \sum\left(\boldsymbol{Y_i - \hat{Y}_i}\right)^2 = \mathbf{min}\left(\boldsymbol{E}\left(\boldsymbol{\varepsilon_i^2}\right)\right) = \mathbf{min}\left(\boldsymbol{E(\varepsilon\varepsilon')}\right)$
  - Thus, minimising for {= 0} @ stationary point
  - The following uses relevant matrix notation, not pure mathematical
  - Hence, the use of $(\mathrm{Y} - \beta X)$ to represent the (error) distance of observations from the best fit line
  - As $\boldsymbol{\varepsilon} = (\boldsymbol{Y - \beta X})$ in a vector context:

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

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

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

Gauss-Markov Theorem – Beta Estimations

This states that, given the aforementioned assumptions, there are no alternative linear estimators of β's with a smaller variance. That is, the OLS model is the **B**est **L**inear **U**nbiased **E**fficient (BLUE) estimator

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

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

- 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 ε:

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

or

$$E(\hat{\beta}) = E(\beta) + (X'X)^{-1}X'\varepsilon = \beta + (X'X)^{-1}E(X'\varepsilon) \text{ [2]}$$

- Therefore, $\hat{\beta}$ **is** a linear estimator of β
- Where β has minimal variance such that:

$$var(\beta) - var(\hat{\beta}) > 0$$

- Meaning that $\hat{\beta}$ estimations will never exceed β

Variance Estimations [2]

- **Recall**:
  - $\hat{\beta} = (X'X)^{-1}X'Y$
  - $\varepsilon = (Y - \beta X)$
- Estimate σ^2 with $\hat{\sigma}^2$ in a similar manner to β;
- For N observations and k explanatory variables

$$\hat{\sigma}^2 = \frac{\varepsilon'\varepsilon}{N-k}$$

$$\boxed{\therefore \hat{\sigma}^2 = \frac{(Y-\beta X)'(Y-\beta X)}{N-k}}$$

- Thus, like β, $\hat{\sigma}^2$ is an unbiased estimator of σ^2

Variance-Covariance Matrix for $\hat{\boldsymbol{\beta}}$

$$E\left((\hat{\beta}-\beta)(\hat{\beta}-\beta)'\right)$$
$$= E\big(((X'X)^{-1}X'\varepsilon)((X'X)^{-1}X'\varepsilon)'\big)$$
$$= E((X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1})$$

- Assume X is stochastic & rearrange:

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

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

$$= (X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1}$$
$$= \sigma^2 I(X'X)^{-1}X'X(X'X)^{-1}$$
$$\underline{= \sigma^2 I(X'X)^{-1} \text{ As Required}}$$

Unbiasedness of OLS

- **Meaning**: applying $\hat{\beta}$ to all observations will provide a **central** True Value β on the distribution of β values
  - Hence, $E(\hat{\beta}) = \bar{\beta}_i$
- **Recall:** $\hat{\beta} = (X'X)^{-1}X'Y$
- In matrix terms $(X'X)^{-1}X$ is referred to as a Projection Matrix which is a vector that captures the error (distance between observations and best-fit)
  - Denoted by $\boldsymbol{P_x}$
- Mathematics and matrix mathematics is shown:
  - $E(\hat{\beta}) = \bar{\beta}_i$
  - $\therefore E(\hat{\beta}) = E((X'X)^{-1}X'Y)$
  - $= E\left((X'X)^{-1}X'(X\beta + \varepsilon)\right)$
  - $= E((X'X)^{-1}X'X\beta + (X'X)^{-1}X'\varepsilon)$
  - $= E((X'X)^{-1}X'\varepsilon) + \beta$
  - $= \beta$
  - $\therefore$ requirements **satisfied**

t-testing

- Recalling **Beta** and **Variance** (headings [1] and [2]) and projection matrix;
- t-values are generated for each parameter using:

$$\mathrm{t-value} = \frac{\beta}{\sqrt{\hat{\sigma}^2(X'X)^{-1}}}$$

- Knowing that $Y \sim N(\mu, \sigma^2)$, we are working with normally distributed data in this scenario (t-distribution);
- the t-test shows how significant mean differences are;
- and, whether differences are by chance;
- t-value shows size of the difference compared to variation in the dataset;
- thus, we establish the level of significance (usually 1%, 5%, 10%);
- and furthermore, we know the relative critical value
    - 1% Significance Level (Example):
        - Call H_0: X **does not** explain Y
        - And H_A: X **does** explain Y
        - **Rejecting** Null Hypothesis at the 1% Significance Level;
        - It can be said with 99% certainty that X explains Y;
        - The result is **not** down to chance
    - This can be established by:
        - t-value (t-test) > t-value (t-distribution) [relative]
        - p-value (t-test) < 0.01