Name : _____

SMU
SINGAPORE MANAGEMENT UNIVERSITY
MASTERS

# Master of Science in Quantitative Finance

# QF 603
# Quantitative Analysis of Financial Markets

## INSTRUCTIONS TO STUDENTS

1 The time allowed for this examination paper is 2 hours
2 You are REQUIRED to answer ALL questions. No marks will be deducted for wrong answers.
3 This is a closed-book examination. You are allowed to use calculators falling within previously notified categories
4 There are 12 printed pages including this instruction sheet.
5 You are required to return full set of question paper at the end the examination. Please write your name on the top right hand corner of this instruction sheet.
6 Write all answers (including MCQ answers) in exam booklet provided.
7 State workings clearly for non MCQ questions
8 Workings for MCQ questions will not be evaluated, only final answer.
9 For MCQs, write question number + number representing correct answer (1,2,3 or 4) in answer booklet. Select only 1 out of 4 options for each question as your answer.

**Section A - MCQ (20 points)**

Student (t-distribution; not normal) –> may make error (uncertainty)

Q1: What is intuition behind construction of a t-statistic?

Hypothesis testing

1. We compare an estimated value with its 'uncertainty', to see if estimated value is "large" relative to uncertainty band
2. Checks that distribution of underlying variable is normal
3. Checks if regression as a whole explains significant fraction of variation in dependent variable relative to null hypothesis where all slope coefficients are 1.0
4. Mitigates the effects of outliers in the system

line of best fit; correlation (perfect if +/- 1); slope coefficients (null –> zero slope coeff)

Q2: For a typical OLS regression ran in widely available software packages (e.g. python, r, matlab, excel, c++, etc), what is meaning of F-statistic reported?

comparing variances

1. Reported F-statistic measures collinearity amongst independent variables in the system
2. Checks if regression as a whole explains significant fraction of variation in dependent variable relative to null hypothesis where all slope coefficients are 1.0
3. Checks if regression as a whole explains significant fraction of variation in dependent variable relative to null hypothesis where all slope coefficients are 0.0
4. The reported F-statistic measures serial correlation amongst independent variables in the system

Q3: Consider following 3 statements on practical applications of ACF and PACF plots in time series modelling. Which are correct?

Statement 1: We can distinguish between AR and MA models because AR has an ACF that decays slowly as lags increase, while MA has an ACF that drops sharply beyond order of the MA model

Statement 2: We can establish order of AR model because AR model has PACF that drops sharply after the lag corresponding to order of AR model

Statement 3: We can establish order of MA model because MA model has PACF that drops sharply after the lag corresponding to order of MA model

1. Statement 1 & Statement 2
2. Statement 2 & Statement 3
3. Statement 1 & Statement 3
4. They are all correct

analog to SD ~252 days

Q4: The daily return volatility of a stock is 1.0%. What is the annual return volatility? [2 d.p]

V[Annual] = 252 * V[Daily]
SD[Annual] = sqrt(252) * SD[Daily]

1. 8.93%
2. 252.00%
3. None of these
4. 15.87%

Remember ACF –> most important: lag 1

Q5: What is purpose of Durbin Watson test?

1. To detect the presence of sample autocorrelation at arbitrary lag
2. To detect the presence of sample autocorrelation at lag 1 only
3. To detect the presence of heteroskedasticity in the dependent variable only
4. None of these

Cannot be equal to additive/multiplicative decomposition

Q6: Which of the following is **not** an effect of applying log transformation to all dependent and independent variables in an estimation?

1. Mitigate (but not eliminate) the effect of outliers
2. Potentially convert a multiplicative data generating process into a linear format which can be estimated using linear estimation
3. Potentially cause additional missing variable values because log of a non-positive number is undefined
4. Decompose time series data into trend and seasonal components

Addition, multiplication

Q7: Consider the following 4 steps for analysis of time series data. What is correct order to conduct these 4 steps in?

Model those with seasonality first

Step 2: A. Model cyclical / seasonal component using AR/MA/or another time series model
Step 1: B. Decompose time series into cyclical / seasonal component and trending component
Step 3: C. Model trend component includes unseasonal
Step 4: D. Recombine models for trend and cyclical components into an integrated model

1. A,B,C,D
2. D,C,B,A
3. B,A,C,D
4. C,A,B,D

Q8: Recall Bayes theorem states that:

P(A n B)

**P (A | B) = [P (B | A) * P(A)] / P(B)**

P(A n B) = 1/4

P(B) = 3/10

P(A|B) = (1/4)/(3/10) = 0.833

Assume we have two bonds, with following probability matrix on upgrades and downgrades

Probability Matrix

| | | **Bond A** | | |
|---|---|---|---|---|
| | | No Default | Default | |
| **Bond B** | No Default | 60% | 10% | 70% |
| | Default | 5% | 25% | 30% |
| | | | | |
| | | 65% | 35% | 100% |

What is probability Bond A defaults given Bond B defaults (nearest whole %)?

1. 10%
2. 15%
3. 67%
4. 83%

Q9. Looking at same matrix as for Q8, what is probability that Bond B defaults given that Bond A defaults?

P(A) = 7/20
P(A|B) = (1/4)/(7/20) = 0.714

1. 71%
2. 29%
3. 30%
4. 35%

Recall mean = 0 for t-4, t-5, ... specifically 3 periods back

Q10. Assume we have time series with true data generating process MA(3). Assume we know value of all constants in MA(3) model. You have information on realized values of time series from 10 periods ago. Is the 10 period ago data useful in forecasting time series for next period?

t-10
no correlation (10 > 3) –> not useful

1. Yes, there will be some small (but persistence) correlation between value from 10 periods ago and value in the next period.
2. No, there is 0 correlation between value from 10 periods ago and value in next period
3. Yes, realized value from 10 periods ago = expected value in next period
4. No, there is negative correlation between value from 10 periods ago and value in next period due to convergence back to mean value

we can have 2 periods ago, 3, 4, …

Q11. Assume we have a time series with true data generating process AR(1). Assume we know value of all constants in AR(1) model. You have information on realized values of the time series from 3 periods ago. Is the 3 period ago data useful in forecasting time series for next period?

iterative decrease in correlation with increasing backward, i.e., t-2, t-3, …

1. Yes, there will be some small (but persistence) correlation between value from 3 periods ago and value next period.
2. No, there is 0 correlation between value from 3 periods ago and value in next period
3. Yes, realized value from 3 periods ago = expected value in next period cannot hold true
4. No, there is always negative correlation between value from 3 periods ago and value in next period due to convergence back to mean value

For question 12 to 14, consider the following data series:

X1: [1.0, 1.0, -0.1, 2.5, -0.3, 0.1, 0.5]

Q12: What is median of X1?

1. 0.4
2. 0.2
3. 0.1
4. 0.5

Q13: What is mean of X1 [2 d.p.]?

Sum = 4.7
Size = 7
Mean = 4.7/7

1. 0.21
2. None of these
3. 0.67
4. -0.12

Q14: What is mode of X1 [1 d.p]?

most frequent value (in this case, 1.0)

1. 0.5
2. 0.8
3. 1.0
4. None of these

Q15: For an integer dependent variable representing customer satisfaction on a scale of 1 to 10, what estimation methodology should we use?

low to high –> ordered

1. Probit regression
2. Ordered probit regression
3. Multinomial probit regression
4. None of these

Q16: For survival analysis estimating the effect of the value of an income statement account on firm survival / bankruptcy in a right – censored dataset, what is the best analysis framework?

1. Principal Components Analysis
2. None of these
3. Proportional Hazards Cox model
4. Estimate parameters for a Poisson model

main –> largest

dimensions –> idea: vectors (eigenvalue, eigenvector)

Q17: How is first principal component ("PC1") found in principal components analysis?

1. It is the independent variable that explains the highest fraction of variation in the data
2. It is the eigen vector corresponding to the largest eigen value
3. None of these
4. It is the independent variable that has the highest correlation to the dependent variable

Q18: What is the effect of varying covariate values in the proportional hazards model?

1. Hazard rate is the product of a non-parametric baseline hazard and a function determined by a list of covariates. Varying covariate values have the impact of shifting the entire baseline hazard rate uniformly
2. Varying covariate values have no impact because the hazard rate is determined only by the baseline hazard function

3. Increasing a single covariate has the effect of increasing the gradient of the baseline hazard function
4. None of these

f(x) = 1

Q19: A variable takes a uniform distribution between 0 and 1. What is the unconditional mean of variable?

1. Undetermined, we are not given the distribution's pdf or cdf
2. 0
3. 1.41
4. 0.5

E[X] = int_{-\infty}^\infty x f(x) dx

Recall Bayes' theorem: P(A|B) = [P(B|A) * P(A)] / P(B) [prior probability]

Q20: What is a potential application of Bayesian updating?

1. Combine a prior hypothesis with realized data, where small size of dataset can be supplemented by prior hypothesis / view
2. Correct for conditional heteroskedasticity in the independent variables
3. Allows for clustering of standard errors by economically meaningful groups (e.g. industries, year)
4. None of these

**Section B – Estimations (20 points)**

Section B1 - Regression diagnostics (12 points)

For questions in this section, refer to following OLS estimation result

```
<class 'statsmodels.iolib.summary.Summary'>
"""
                            OLS Regression Results
==============================================================================
Dep. Variable:              lnepratio   R-squared:                       0.085
Model:                            OLS   Adj. R-squared:                  0.084
Method:                 Least Squares   F-statistic:                     44.65
Date:                Tue, 24 Nov 2020   Prob (F-statistic):           2.30e-80
Time:                        13:53:12   Log-Likelihood:            -1.5595e+05
No. Observations:              115861   AIC:                         3.119e+05
Df Residuals:                  115851   BIC:                         3.120e+05
Df Model:                           9
Covariance Type:              cluster
=================================================================================================================
                                                    coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------------------
const                                             -3.6332      0.034   -107.420      0.000      -3.700      -3.567
lnoperatingmargin                                  0.2969      0.016     18.678      0.000       0.266       0.328
Agriculture Forestry And Fishing                   0.6587      0.357      1.847      0.065      -0.040       1.358
Construction                                       0.3404      0.064      5.292      0.000       0.214       0.467
Finance Insurance And Real Estate                 -0.1784      0.127     -1.405      0.160      -0.427       0.070
Manufacturing                                     -0.0182      0.052     -0.351      0.726      -0.120       0.083
Mining                                            -0.0450      0.048     -0.947      0.344      -0.138       0.048
Retail Trade                                       0.1256      0.074      1.699      0.089      -0.019       0.270
Services                                          -0.2630      0.085     -3.105      0.002      -0.429      -0.097
Transportation Communications Electric Gas And Sanitary Service  -0.1285  0.058  -2.215  0.027  -0.242  -0.015
==============================================================================
Omnibus:                    18757.262   Durbin-Watson:                   1.031
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           159948.347
Skew:                           0.533   Prob(JB):                         0.00
Kurtosis:                       8.657   Cond. No.                         51.5
==============================================================================

Warnings:
[1] Standard Errors are robust to cluster correlation (cluster)
"""
```

B1. Comment on Durbin Watson statistic, and what this implies for OLS assumptions (2 points)

B2. Comment on Jarque-Bera statistic, and what this implies for the OLS assumptions (2 points)

B3. What coefficients are significant at 5% level of significance? List variable names (3 points)

B4. What does the p-value of F-statistic imply? (2 points)

B4 Correct Answer: A large F-statistic indicates that the OLS regression model is effective at explaining the variation in lnepratio. The observed differences between sample averages are statistically significant and could not be due to random chance alone.

B5. Why is adjusted R-square lower than R-square? (3 points)

Section B2: Time series (8 points)

B6. Recall our discussion on AR(1) processes:

# Autoregressive Process AR(1)

- Repeated substitution for $Y_t$ in the AR(1) process

$$Y_t = \theta + \lambda(\theta + \lambda Y_{t-2} + u_{t-1}) + u_t$$

Recall: lambda is the first-order autocorrelation by definition of AR(1)

leads to

$$Y_t = (1 + \lambda + \lambda^2 + \cdots)\theta + (u_t + \lambda u_{t-1} + \lambda^2 u_{t-2} + \cdots).$$

- For each $t$, provided $|\lambda| < 1$,

$$\begin{aligned}
\mathbb{E}(Y_t) &= (1 + \lambda + \lambda^2 + \cdots)\theta = \frac{\theta}{1-\lambda}; \\
\mathbb{V}(Y_t) &= \mathbb{V}(u_t + \lambda u_{t-1} + \lambda^2 u_{t-2} + \cdots) = \sigma_u^2(1 + \lambda^2 + \lambda^4 + \cdots) \\
&= \frac{\sigma_u^2}{1-\lambda^2}; \\
\mathbb{C}(Y_t, Y_{t-1}) &= \mathbb{C}(\theta + \lambda Y_{t-1} + u_t, Y_{t-1}) = \lambda \frac{\sigma_u^2}{1-\lambda^2}.
\end{aligned}$$

Given an AR(1) process Y(t) = 10 + 0.1Y(t-1) + $u_t$, where E($u_t$) = 0 and Variance($u_t$) = 2

i. What is mean of Y(t)? (2 points)
ii. What is first order autocorrelation of this process? (3 points)
iii. What is 10th order autocorrelation of this process? (3 points)

lambda

lambda^k

**Section C: Estimation design (20 points)**

C1. When should we use error correction model (ECM) rather than ARIMA model? (5 points)

When 2 variables are cointegrated, we can make use of spot deviations from LR spread to improve our forecast by predicting it will converge to LR spread

C2. What is purpose of principal components regression, and why do we discard some of the principal components? (5 points)

The purpose of principal components regression is to handle multicollinearity by transforming predictors into uncorrelated components, which stabilizes the regression estimates. We discard some components because they capture minimal variance and may represent noise, allowing the model to focus on the most informative components and improve generalization.

C3. "If the F-statistic or any T-statistic for a variable we are interested in improves, it is okay to remove y-intercept from an estimation". Discuss (5 points)

No. This will result in a biased estimate.

C4. Why do we cluster standard errors in our OLS estimations? (5 points)

To mitigate (but in practice, not completely remove) the effects of positive autocorrelation in financial data in terms of inflating our t-statistics. By clustering std errors, we are adding back a positive covariance term

**Section D: Capital Markets Theory (20 points)**

D1. What is the relationship between beta (slope coefficient in CAPM) and market efficiency? (4 points)

If markets are completely efficient, excess returns of an instrument should be completely predicted by its market beta

D2. Describe the applicability of cointegration and error correction models in pairs trading (4 points)

Valid pairs need to be cointegrated for pairs trading. If they are, we can use a (V)ECM to generate price predictions for trading

D3. What is the difference between the EWMA (Exponential Weighted Moving Average) model for volatility, and GARCH? Discuss main features and advantages of GARCH for modelling volatility (8 points)

D4. Does volatility have a long term directional time trend? Why/why not? (4 points)

D3 Correct Answer:
GARCH takes into account a target LR volatility level which EWMA does not.
GARCH has 2 features which are closely aligned to practical time series:
(i) volatility clustering due to AR component
(ii) mean reversion in volatility

D4 Correct Answer:
No. It is mean reverting due to fiscal and monetary policies, as well as the effect of valuations on (eventually) cushioning downside volatility.

fiscal/monetary: govts spend more, have lower i/rs (interest rate), looser discount lending, etc during downside volatility

**Section E: Probability theory and distributions (20 points)**

Part A: Bayesian Analysis (6 points)

Suppose that there are 3 types of portfolio managers.

1. Underperformers beat the market 5% of the time
2. In-line performers beat the market 50% of the time
3. Outperformers beat the market 70% of the time

Initially, our prior beliefs are that any given manager is more likely to be an inline performer. This is because we do not have any additional information. The prior beliefs can be given as follows:

1. P(p = 0.1) = 10%
2. P(p = 0.5) = 80%
3. P(p = 0.55) = 10%

A new manager beats market for **2 years in a row**.

[Note: formula for Bayes theorem is given in MCQ question A8]

P(B|A)

E1: What is the likelihood of each of the 3 types of managers beating the market 2 years in a row? (2 points)

P(B)

E2: What is the unconditional probability of observing the manager beat the market 2 years in a row given prior beliefs? (2 points)

P(A|B)

E3: What are the updated beliefs about the manager? (2 points)

Distributions (14 points)

2 sides (head, tail) –> binomial distribution

E4: A biased coin is flipped 1000 times, and the probability of heads each time is p. What is the variance of the total number of heads? (2 points)

E5: A random variable X follows a normal distribution.

V[N] = chi-squared

i. What distribution does the variance of the random variable follow? (3 points)

ii. If we wanted to test a hypothesis that the variance of the random variable is the same as the variance of another random variable (i.e. the ratio of both variances is 1), propose a test statistic and state its distribution. You can neglect degrees of freedom parameters in your answer (3 points)

F

sigma^2

E6. Stock A has annual expected return variance of 10%, while stock B has annual expected return variance of 50%. What is the expected return variance of a portfolio invested equally in stock A and stock B? Assume that the returns of both stocks have zero covariance (2 points) 50%

E[X] = sum(x P(X=x))

V[aX+bY] = a^2 V[X] + b^2 V[Y] + 2 ab Cov[X, Y]

E7. Using the same setup as E6, now assume that the two stocks have a covariance ~~of 0.7~~. Compute the new portfolio variance [2 d.p] (2 points)

correlation (not covariance)

E8. Neglecting the risk free rate, what is the annualized sharpe ratio of the portfolio constructed in (i) E6 and (ii) E7 if the expected returns of stock A is 20% per year, and 10% for stock B (to 2 d.p). Note that you need to output 2 numbers for this question (one for the E6 situation and another for E7). (2 points)

E[R] = (10% + 20%)/2 = 15%

Sharpe = E[R]/sqrt(E[Var])