

INFO 7374 MACHINE LEARNING IN FINANCE

ASSIGNMENT II

1. In making non-parametric kernel density estimation, you found that the estimated kernel density has a lot of zigzags, what could you do with it?
 - A. change a kernel function
 - B. decrease bandwidth
 - C. increase bandwidth
 - D. change the design base
2. Given an Ornstein-Uhlenbeck process, i.e. AR (1): $y_t = \alpha y_{t-1} + u_t$, when $\alpha \rightarrow 1$, the OLS estimates would follow distribution that is asymptotically
 - A. skewed to the right
 - B. normal
 - C. students-t
 - D. skewed to the left
3. Which of the following tests could help you test whether a time series has a unit root?
 - A. Granger Test
 - B. Dicky-Fuller Test
 - C. Hausman Test
 - D. Durbin-Watson Test
4. An analyst at an investment firm has estimated a regression of the firm's fixed income portfolio using a 3-factor model. The results are shown below.

Regression Statistics	
R Square	0.98
Adjusted R Square	0.97
Standard Error	1.02
Observations	600

Regression Output	Coefficients	Standard Error	t Stat	P-value
Intercept	0.56	1.01	0.51	0.31
Factor 1	2.53	2.29	0.50	0.12
Factor 2	2.05	2.60	0.79	0.21
Factor 3	1.78	3.69	0.44	0.33

Correlation Matrix	Portfolio Returns	Factor 1	Factor 2	Factor 3
Portfolio Returns	1.0000	0.9312	0.9136	0.9629
Factor 1	0.9312	1.0000	0.7605	0.8863
Factor 2	0.9136	0.7605	1.0000	0.9144
Factor 3	0.9629	0.8863	0.9144	1.0000

Based on the regression results, what problem(s) does the factor model have?

- A. Heteroskedasticity
- B. Multicollinearity
- C. Autocorrelation
- D. Endogeneity

5. Which of the following statement correctly defines Cointegration?

- A. If X and Y are $I(1)$ processes, then there exists a variable U that is a linear combination of X and Y , but also an $I(0)$ process.
- B. If X and Y are $I(0)$ processes, then there exists a variable U that is a linear combination of X and Y , but also an $I(1)$ process.
- C. If X and Y are $I(1)$ processes, then there exists a variable U that is a linear combination of X and Y , but also an $I(1)$ process.
- D. If X and Y are $I(0)$ processes, then there exists a variable U that is a linear combination of X and Y , but also an $I(0)$ process.

6. Which of the following statement about pairs trading is incorrect?

- A. We select trading pairs by comparing their β s from the CAPM regression.
- B. A candidate pair has their β s as close as possible.
- C. A candidate pair has their β s as disperse as possible.
- D. We test whether the two stocks are cointegrated or not.

7. What could you infer about the kurtosis of the autoregressive conditional heteroskedasticity (ARCH) process

$$\begin{aligned} r_t &= \mu + \sigma_t \varepsilon_t \\ \sigma_t^2 &= \omega + \alpha(r_{t-1} - \mu)^2 \\ \sigma_1^2 &= \omega / (1 - \alpha) \end{aligned}$$

where ε_t is iid standard normal

- A. Kurtosis > 12
- B. Kurtosis > 6
- C. Kurtosis > 4
- D. Kurtosis > 3

8. When you are working on fitting a curve using a piece-wise spline function (fitting a curve with small pieces of straight lines), which of the following method would you consider using?

- A. Non-parametric kernel density
- B. Support Vector Machine
- C. Random Forest
- D. Least Angle Regression

9. When you measure the predictive accuracy of a point forecast you made for the mid-price in a high frequency trading algorithm, which of the following tests would you use?
- A. Dicky-Fully Test
 - B. Kolmogorov-Smirnov Test
 - C. Amisano-Giacomini Test
 - D. Diebold-Mariano Test
10. When we say that an estimator is a BLUE estimator, we mean that the following statements except?
- A. $\hat{\beta}$ is linear.
 - B. $\hat{\beta}$ is unbiased.
 - C. $\hat{\beta}$ is the most efficient estimator when compared with other estimators.
 - D. $\hat{\beta}$ is normally distributed.