Problem A: (40 points) Answer briefly the following questions.

1. Consider a stationary two-dimensional time series $X_t$. Let $\Gamma_1 = \text{Cov}(X_t, X_{t-1})$ be the lag-1 auto-covariance matrix of $X_t$. More specifically, we have

$$X_t = \begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix}, \quad \Gamma_1 = \begin{bmatrix} \Gamma_{11}(1) & \Gamma_{12}(1) \\ \Gamma_{21}(1) & \Gamma_{22}(1) \end{bmatrix}. $$

Write down the meanings of $\Gamma_{11}(1)$ and $\Gamma_{12}(1)$.

Answer: $\Gamma_{11}(1)$ measures the linear dependence of $x_{1t}$ on its own lag $x_{1,t-1}$ whereas $\Gamma_{12}(1)$ denotes the linear dependence of $x_{1t}$ on $x_{2,t-1}$.

2. Suppose that $X_t$ follows a 2-dimensional stationary vector AR(1) model,

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} 1.0 \\ 2.0 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix},$$

where $a_t$ is a sequence of independent and identically distributed Gaussian random vectors. What are the meanings of the coefficients $-0.6$ and $0.3$, respectively.

Answer: $-0.6$ measures the linear dependence of $x_{2t}$ on $1, t-1$ in the presence of $x_{2,t-1}$. That is, it shows the added contribution of $x_{1,t-1}$ to $x_{2t}$ over $x_{2,t-1}$. The coefficient $0.3$ measures the linear dependence of $x_{1t}$ on $x_{2,t-1}$ in the presence of $x_{1,t-1}$.

3. Consider the log prices $p_{1t}$ and $p_{2t}$ of two assets. Assume that both log price series have a unit root. Describe two approaches that can be used to verify the existence of a linear combination $w_t = p_{1t} + \gamma p_{2t}$, which has no unit root.

Answer: (1) Co-integration test and (2) linear regression of $p_{1t}$ on $p_{2t}$ and checking the unit root in the residuals of the regression.

4. Suppose the price $P_t$ of a stock follows the stochastic diffusion equation (SDE) $dP_t = 0.04P_t dt + 0.25P_t dw_t$, where $w_t$ denotes the standard Brownian motion. What are the drift and diffusion terms of the squared process $P_t^2$?

Answer: Applying the Ito’s lemma and letting $G_t = P_t^2$, we have

$$dG_t = \left[ 0.08P_t^2 + \left( \frac{P_t}{4} \right)^2 \right] dt + 0.5P_t^2 dw_t.$$ 

Thus, the drift and diffusion terms are $[0.08 + (1/16)]P_t^2$ and $0.5P_t^2$, respectively.
5. Describe two alternative approaches to compute the daily price volatilities of a stock using the intra-day transaction data.

Answer: (1) Use realized volatility, (2) use daily open, high, low, and close prices with various methods of Section 3.15.

6. Consider the quarterly earnings of a company. Assume that the moving average component of the earnings follows the model \( x_t = (1 - 0.4B - 0.5B^4 + 0.2B^5)a_t \), where \( a_t \) is a univariate white noise series and \( B \) denotes the back-shift operator. Let \( \rho_\ell \) be the lag-\( \ell \) autocorrelation of \( x_t \), where \( \ell > 0 \). How many non-zero autocorrelations \( x_t \) has? Write down the lags of those non-zero autocorrelations.

Answer: 4 lags. They are lag 1, lag 3, lag 4, and lag 5.

7. Given one advantage and one dis-advantage of using value at risk (VaR) as a risk measure for a financial position.

Answer: Advantage: simplicity. Disadvantage: not a coherent measure except under normality.

8. Let \( F(x) \) be the cumulative distribution function of a loss variable. Assume that \( F(x) \) satisfies

\[
F(x) = \begin{cases} 
0.8725 & \text{for } x = 85 \\
0.9025 & \text{for } x = 90 \\
0.95 & \text{for } x = 96 \\
0.99 & \text{for } x = 100.
\end{cases}
\]

Next, define a new loss variable \( Y \) by

\[
Y = \begin{cases} 
X & \text{if } X \leq 96 \\
0 & \text{otherwise.}
\end{cases}
\]

What is the VaR of the loss \( X \) if the upper tail probability is 5%? What is the VaR of \( Y \) if the upper tail probability is 5%?

Answer: For \( X \), VaR = 96 and for \( Y \), it is 90.

9. Consider a 2-2-1 feed-forward neural network. Let \( x_i \) denote the input nodes (\( i = 1, 2 \)), \( h_j \) denote the hidden nodes (\( j = 1, 2 \)) and \( o \) be the output node. Write down the logistic function for the hidden node \( h_2 \). Also, if the skip layer is used and the output is a continuous variable, write down the function for network.

Answer: \( h_2 = \frac{\exp[\omega_2 + \omega_1 x_1 + \omega_2 x_2]}{1 + \exp[\omega_2 + \omega_1 x_1 + \omega_2 x_2]} \). For the network, we have

\[
o_t = \alpha_0 + \alpha_1 h_1 + \alpha_2 h_2 + \alpha_3 x_1 + \alpha_4 x_2,
\]

where \( h_2 \) is given before and \( h_1 \) is defined in a similar manner with \( \omega_{j2} \) replaced by \( \omega_{j1} \), where \( j = 0, 1, 2 \).
10. Consider the monthly log returns of the S&P 500 index. Suppose we are interested in predicting the direction of the movement of the index, i.e., up or down. Describe two methods discussed in the lecture that can be used for such an analysis.

Answer: (1) Logistic regression, and (2) neural network.

11. Describe two advantages of using the peaks over threshold (POT) method of extreme value theory over the traditional block maximum method.

Answer: (1) Can handle volatility clustering, and (2) the result is less sensitive to the choice of threshold. [or It takes into consideration the frequency and exceedance of extreme returns.]

12. Give two characteristics of the intraday transaction data of U.S. stocks.

Answer: Any two of (1) large sample size, (2) non-synchronous trading, (3) diurnal pattern, and (4) discrete price values.

13. The GARCH-M models are harder to estimate for stock returns, compared with the conventional GARCH volatility models. Give two reasons that justify the use of GARCH-M models.

Answer: (1) GARCH-M model can handle risk premium, (2) the model has a diffusion limit.

14. (For Questions 14 to 16) Suppose that the quarterly log earnings $x_t$ of company A follows the model

$$(1 - B)(1 - B^4)x_t = (1 - 0.57B)(1 - 0.18B^4)a_t$$

$$a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1)$$

$$\sigma_t^2 = 8.09 \times 10^{-5} + 0.244a_{t-1}^2 + 0.711\sigma_{t-1}^2.$$ 

Suppose that the last 5 log earnings, residuals, and volatilities are

<table>
<thead>
<tr>
<th>time</th>
<th>96</th>
<th>97</th>
<th>98</th>
<th>99</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_t$</td>
<td>1.05</td>
<td>1.30</td>
<td>0.91</td>
<td>1.12</td>
<td>1.18</td>
</tr>
<tr>
<td>$a_t$</td>
<td>0.0221</td>
<td>-0.0318</td>
<td>0.0010</td>
<td>-0.0108</td>
<td>0.0113</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>0.03608</td>
<td>0.03416</td>
<td>0.03323</td>
<td>0.02951</td>
<td>0.02669</td>
</tr>
</tbody>
</table>

Consider the forecast origin $T = 100$. Calculate the 1-step and 2-step ahead predictions of the log earnings?

Answer: Use $x_{t+1} = x_t + x_{t-3} - x_{t-4} - 0.57a_t - 0.18a_{t-3} + 0.1026a_{t-4}$. For 1-step prediction, we have

$$x_{100}(1) = 1.18 + 1.30 - 1.05 - 0.57 \times (.0115) - 0.18(-.0318) + .1026(0.0221) = 1.43.$$ 

$$x_{100}(2) = 1.43 + 0.91 - 1.30 - 0.18(0.001) + 0.1026(-0.0318) = 1.037.$$
15. Again, consider the forecast origin $T = 100$. Compute the 1-step and 2-step ahead volatility forecasts (not the variances).

Answer: Using $\sigma^2_{t+1} = 8.09 \times 10^{-5} + 0.244a_t^2 + 0.711\sigma_t^2$, we have $\sigma^2_{100}(1) = 8.09 \times 10^{-5} + 0.244(0.0113)^2 + 0.711(0.02669)^2$. Therefore, $\sigma_{100}(1) = 0.02487$. Similarly, we have $\sigma_{100}(2) = 0.02489$.

16. Compute the unconditional standard error of $a_t$. Also, let $w_t = (1-B)(1-B^4)x_t$. What is the mean equation for $w_t$ conditional on information available at time $t-1$?

Answer: $\sqrt{8.09 \times 10^{-5} - 0.244 - 0.711} = 0.0424$. The mean equation of $w_t$ is $(1 - 0.57B)(1 - 0.18B^4)a_t$.

17. Describe two applications of VaR (value at risk) in making financial decision.

Answer: (1) Market risk management, (2) Compute risk-adjusted asset allocation.

18. Consider the AR(2) model $x_t = 1.1x_{t-1} - 0.5x_{t-2} + a_t$, $a_t \sim N(0,1)$. Does it exhibit business-cycle pattern? If yes, what is the average length of the business cycles?

Answer: Yes, the model implies business cycles. The average length is 9.25.

19. Describe two ways that a GARCH model can generate a heavy-tailed distribution.

Answer: (1) Use heavy-tailed innovations, (2) The dynamic dependence via the coefficient of $a^2_{t-1}$.

20. Describe two scenarios under which we observe an asset return series that follow a moving-average model.

Answer: (1) Bid-ask bounce, (2) Smoothing the returns, especially in hedge funds.

**Problem B.** (26 points) Consider the daily adjusted close prices of the stocks of IBM and Caterpillar (CAT), starting from January 4, 2005 for 1866 observations. We compute the daily log returns from the price series, resulting in 1865 returns. The tick symbols are IBM and CAT, respectively. Consider a short position of $2 million on IBM stock and $1 million on CAT stock. Use the attached output to answer the following questions.

1. (3 points) If RiskMetrics is used, calculate the VaR for the portfolio.

   Answer: For IBM, VAR = $2000000 \times 0.0199 = $39800$; For CAT, VAR = $1000000 \times 0.0455 = 45500$. The correlation is 0.544. Therefore, VaR for the portfolio is $\sqrt{39800^2 + 45500^2 + 2 \times 0.544 \times 39800 \times 45500} = $74997.

2. (3 points) If Gaussian GARCH(1,1) models are used for both log return series, what is the VaR for the portfolio?

   Answer: For IBM, VaR = $49000$. For CAT, VaR = $51100$. For the portfolio, we have VaR = $87957$.

3. (2 points) Focus on the Gaussian GARCH(1,1) model for the IBM stock return. What is the expected shortfall for the stock?

   Answer: ES = $2000000 \times 0.0279 = $55800$. 

4
4. (4 points) If a GARCH(1,1) model with standardized Student-t innovations is employed for the IBM log returns, what are the VaR and expected shortfall?

Answer: VaR = $65000 and ES=$84400. [With modified RMesure program, the answers are VaR = $51800 and ES = $67200.]

5. (4 points) If the traditional method of block maximum is used to study the extreme-value behavior of IBM log returns, what are the parameter estimates when block size is 21? Are the estimates statistically significant? Why?

Answer: The estimates are (xi,sigma,mu) = (0.259, 0.00812, 0.0189). Their standard errors are (0.090, 0.00071, 0.000971). The estimates are all statistically significant at the 5% level, because their t-ratios are greater than 2.0.

6. (2 points) What are the 1-day and 10-day VaRs for the IBM position based on the traditional block-maximum method?

Answer: 1-day VaR = $69000 and 10-day VaR = (10)^{0.259} \times 69000 = $125,271.

7. (4 points) Finally, consider the peaks over threshold (POT) approach. The threshold of 1% is used for both IBM and CAT log returns. Write down the parameter estimates for both stocks? Are these estimates significantly different from zero? Why?

Answer: For IBM, (xi,beta) = (0.144,0.00820) with standard errors (0.0608, 0.000633). For CAT, (xi,beta) = (0.0924, 0.0132) with standard errors (0.0441, 0.000799). These estimates are all significantly different from zero because their t-ratios are greater than 2.

8. (4 points) What are the VaR and expected shortfall for each of the two stock positions based on the POT method?

Answer: For IBM, VaR = $80200 and ES = $109400. For CAT, VaR = $62600 and ES = $82500.

**Problem C. (18 points).** Consider the intraday trading of the Caterpillar (CAT) stock from December 04 to December 08, 2010. There were 155,072 trades in the data. Among them, there were 27,859 price increases and 27,510 price decreases. Let $A_i$ and $D_i$ be the action and direction of the price change for the $i$th trade. That is, $A_i = 1$ if and only if the $i$th trade results in a non-zero price change, and $D_i = 1$ if $C_i > 0$, $D_i = -1$ if $C_i < 0$, and $D_i = 0$, otherwise, where $C_i$ denotes the price change of the $i$th trade. The data for the last four trades are given below:

<table>
<thead>
<tr>
<th>ith trade</th>
<th>155069</th>
<th>155070</th>
<th>155071</th>
<th>155072</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$D_i$</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>$A_{i-1}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$D_{i-1}$</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>
Logistic regressions are used to model $A_i$ and $D_i$. The output is attached. Answer the following questions:

1. (3 points) Write down the fitted model $m_1$ for $A_i$. Are the estimates significantly different from zero? Why?
   
   Answer: $P(A_i = 1|F_{i-1}) = \frac{\exp[-1.054+1.192A_{i-1}+0.0079D_{i-1}]}{1+\exp[-1.054+1.192A_{i-1}+0.0079D_{i-1}]}$. The coefficient of $D_{i-1}$ is not significant at the 5% level, because it z-value is only 0.932 with $p$-value 0.351.

2. (2 points) Based on the refined model, called $m_2$, and the data, obtain the probability that the next trade has a non-zero price change at the origin $i = 155072$. That is, calculate $P(A_{155073} = 1|F_{155072})$, where $F_i$ denotes information available at the $i$-th trade (inclusive).
   
   Answer: $P(A_i = 1|A_{i-1}) = \frac{\exp[-0.0224 - 1.147D_{i-1} - 0.0027A_{i-1}]}{1 + \exp[-0.0224 - 1.147D_{i-1} - 0.0027A_{i-1}]}$. Since $A_{155072} = 0$, we have $P(A_{155073} = 1|A_{155072} = 0) = 0.258$.

3. (3 points) Write down the fitted model for $D_i$ conditional on $A_i = 1$ when both $D_{i-1}$ and $A_{i-1}$ are used. [That is, $m_3$.] Are the estimates significantly different from zero? Why?
   
   Answer: The fitted model is $P(D_i = 1|A_i = 1, F_{i-1}) = \frac{\exp[0.0212 - 1.147D_{i-1} - 0.0027A_{i-1}]}{1 + \exp[0.0212 - 1.147D_{i-1} - 0.0027A_{i-1}]}$.

   In this particular instance, the coefficient of $A_{i-1}$ is insignificant, that of $D_{i-1}$ is significant, whereas that of the constant is marginal with $p$-value 0.0718.

4. (2 points) Write down the refined model for $D_i$. That is, the model $m_4$.
   
   Answer: The refined model is $P(D_i = 1|A_i = 1, F_{i-1}) = \frac{\exp[0.0212 - 1.147D_{i-1}]}{1 + \exp[0.0212 - 1.147D_{i-1}]}$.

   where both coefficients are significant at the 5% level.

5. (6 points) Based one the refined model, calculate $P(D_i = 1|A_i = 1, D_{i-1} = 1)$, $P(D_i = 1|A_i = 1, D_{i-1} = 0)$ and $P(D_i = 1|A_i = 1, D_{i-1} = -1)$.
   
   Answer: $P(D_i = 1|A_i = 1, D_{i-1} = 1) = 0.245$. $P(D_i = 1|A_i = 1, D_{i-1} = 0) = 0.505$ and $P(D_i = 1|A_i = 1, D_{i-1} = -1) = 0.763$.

6. (2 points) Interpret the implications of the probabilities calculated in the prior question #5.
   
   Answer: The probabilities indicate that (a) if there was no price change at the $(i-1)$th trade, then the price change at the $i$th trade, if any, would be 50-50 in either direction; (b) if there was a price drop at the $(i-1)$th trade, then the price change at the $i$th trade, if any, would be positive with probability about 75%; and (c) if there was a price drop...
increase at the \((i - 1)\)th trade, then the price change at the \(i\)th trade, if any, would be positive only 25%.

In short, the probabilities are consistent with the bid-ask bounce.
Problem D. (16 points). This problem applies CreditMetrics to compute VaR of a loan for the next year. Consider a five-year fixed-rate loan of $100 million made to a borrower rated A at 5% annual interest rate. Based on the historical data, we have the following one-year transition probability for A-rated borrower

<table>
<thead>
<tr>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Defaulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.28</td>
<td>3.32</td>
<td>88.4</td>
<td>5.52</td>
<td>1.65</td>
<td>0.52</td>
<td>0.12</td>
<td>0.15</td>
</tr>
</tbody>
</table>

One-year forward zero-coupon curves plus credit spreads by credit rating category:

<table>
<thead>
<tr>
<th>Category</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3.60</td>
<td>4.17</td>
<td>4.73</td>
<td>5.12</td>
</tr>
<tr>
<td>AA</td>
<td>3.65</td>
<td>4.22</td>
<td>4.78</td>
<td>5.17</td>
</tr>
<tr>
<td>A</td>
<td>3.72</td>
<td>4.32</td>
<td>4.93</td>
<td>5.32</td>
</tr>
<tr>
<td>BBB</td>
<td>4.10</td>
<td>4.67</td>
<td>5.25</td>
<td>5.63</td>
</tr>
<tr>
<td>BB</td>
<td>5.55</td>
<td>6.02</td>
<td>6.78</td>
<td>7.27</td>
</tr>
<tr>
<td>B</td>
<td>6.05</td>
<td>7.02</td>
<td>8.03</td>
<td>8.52</td>
</tr>
<tr>
<td>CCC</td>
<td>15.05</td>
<td>15.02</td>
<td>14.03</td>
<td>13.52</td>
</tr>
</tbody>
</table>

Value of the loan at the end of Year 1, under different rating changes (including first-year coupon):

<table>
<thead>
<tr>
<th>Rating</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Defaulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>104.78</td>
<td>XXXX</td>
<td>XXXX</td>
<td>103.00</td>
<td>97.59</td>
<td>93.76</td>
<td>XXXX</td>
<td>51.13</td>
</tr>
</tbody>
</table>

1. (3 points) Calculate the value of the loan at the end of Year 1 if the borrower is upgraded to AA, remains at A, or down-graded to CCC. That is, fill in the missing parts of the value of the loan at the end of Year 1.

Answer: \( P_2 = 5 + 5/(1.0365) + 5/(1.0422)^2 + 5/(1.0478)^3 + 105/(1.0517)^4 = 104.6 \) million; \( P_4 = 5 + 5/(1.0410) + 5/(1.0467)^2 + 5/(1.0525)^3 + 105/(1.0563)^4 = 104.1 \) million; and \( P_7 = 5 + 5/(1.1505) + 5/(1.1502)^2 + 5/(1.1403)^3 + 105/(1.1352)^4 = 79.7 \) million.

2. (3 points) What is the mean value of the loan at the end of Year 1?

Answer: The mean is $103.77 million.

3. (3 points) What is the standard error of the value of the loan at the end of Year 1?

Answer: The standard error is $2.48 million.

4. (4 points) Based on CreditMetrics, the loss on the loan is assumed to be normally distributed with mean zero and variance given by that of the value of the loan. What is the 5% VaR of the loan for Year 1? What is the corresponding expected shortfall?

Answer: The VaR = $4.08 million. The expected shortfall is \( ES = f(x) \sigma_t = 0.10311 \times 2.48 = 5.11 \) million, where \( f(x) \) is the density function of the standard normal distribution. (See Lecture 9.)
5. (3 points) If the actual distribution is used, what is the 2.44% VaR of the loan for Year 1?
   Answer: VaR = $(103.77 - 97.59) = $6.18 million.