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

1. Describe the two key assumptions used by the RiskMetrics to justify the square root of time rule in VaR calculation.

   Answer: (a) The log return of the asset is normally distributed with mean zero. (b) The volatility of the log returns follows an IGARCH(1,1) model without drift.

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

\[
\begin{bmatrix}
  x_{1t} \\
  x_{2t}
\end{bmatrix} = \begin{bmatrix}
  1.5 \\
  0.5
\end{bmatrix} + \begin{bmatrix}
  0.2 & 0.0 \\
  0.6 & 0.8
\end{bmatrix} \begin{bmatrix}
  x_{1,t-1} \\
  x_{2,t-1}
\end{bmatrix} + \begin{bmatrix}
  a_{1t} \\
  a_{2t}
\end{bmatrix},
\]

where $a_t$ is a sequence of independent and identically distributed Gaussian random vectors with mean zero and covariance matrix $\Sigma > 0$. Does the model imply the existence of Granger causality? Why? If there is a Granger causality, state the relation.

   Answer: Yes, the model implies a Granger causality, because $\phi_{12} = 0$ and $\phi_{21} \neq 0$. [You may state that $x_{1t}$ does not depend on the past value of $x_{2t}$, but $x_{2t}$ depends on the past value of $x_{1t}$.] The relation is that $x_{1t}$ causes $x_{2t}$.

3. Describe two applications of seasonal time-series models in finance.

   Answer: Any two of (a) modeling quarterly earnings, (b) modeling energy related assets, and (c) handling intraday diurnal pattern.

4. Suppose the price $P_t$ of a stock follows the stochastic diffusion equation (SDE) $dP_t = P_t dt + 2P_t dw_t$, where $w_t$ denotes the standard Brownian motion. What is the stochastic diffusion equation for the square-root process $G_t = \sqrt{P_t}$? Justify briefly your answer.

   Answer: Applying the Ito’s lemma, we have $dG_t = G_t dw_t$. The drift term vanishes.

5. Provide two volatility models with leverage effect for the daily log return $r_t$ of a stock.

   Answer: (a) EGARCH model, (b) TGARCH or GJR model.

6. Provide two potential sources of serial correlations in observed log returns of an asset even though the true log returns are serially uncorrelated.

   Answer: Any two of (a) nonsynchronous trading, (b) bid-ask pounce, and (c) existence of risk premium.

7. The methods of empirical quantile and quantile regression can be used to estimate VaR. Give two reasons that the quantile regression approach is preferred.

   Answer: Any two of (a) Do not assume that the future returns and past returns have the same distribution, (b) flexibility in including explanatory variables, and (c) adaptive to changing economic environment.
8. Describe two assumptions used in pairs trading in the equity market.

**Answer:** (a) The log prices of the asset are unit-root nonstationary, (b) a linear combination of the log prices is stationary (mean reverting).

9. Consider an AR(2) time series \( (1 - 0.805B - 0.1215B^2)x_t = 0.134 + a_t \), where \( B \) is the backshift operator and \( \{a_t\} \) are iid \( N(0,1) \) variables. Is \( x_t \) stationary? Why? If it is stationary, calculate \( E(x_t) \), the mean of \( x_t \).

**Answer:** Yes, the process is stationary because the solutions of \( 1 - 0.805x - 0.1215x^2 = 0 \) are outside the unit circle. The mean is \( E(x_t) = 0.134/(1 - 0.805 - 0.1215) \approx 1.823 \).

10. Describe two methods that can mitigate the effects of market micro-structure noises in calculating daily realized volatility of a stock.

**Answer:** (a) Use the optimal sampling interval and (b) use the subsampling method.

11. Suppose that the true (unobserved) log price \( p^*_t \) of an asset follows the model \( p^*_t = p^*_{t-1} + a_t \), where \( \{a_t\} \) is sequence of iid \( N(0,1) \) random variables. Suppose that the observed log price \( p_t \) follows the model \( p_t = p^*_t + \epsilon_t \), where \( \{\epsilon_t\} \) is another sequence of iid \( N(0,0.25) \) random variables and \( \{\epsilon_t\} \) and \( \{a_t\} \) are independent. Let \( r_t = p_t - p_{t-1} \) be the observed log return. Calculate the lag-1 and lag-2 autocorrelations of \( r_t \).

**Answer:** \( r_t = a_t + \epsilon_t - \epsilon_{t-1} \). Therefore, \( \text{Var}(r_t) = 1 + 0.25 + 0.25 = 1.5 \), \( \gamma_1 = \text{Cov}(r_t, r_{t-1}) = -0.25 \) and \( \gamma_2 = \text{Cov}(r_t, r_{t-2}) = 0 \). Consequently, \( \rho_1 = -0.25/1.5 = -1.67 \) and \( \rho_2 = 0 \).

12. Let \( C_i \) be the price change of the \( i \)th trade for a stock in a given day. For a heavily traded stock, describe two empirical characteristics of \( C_i \).

**Answer:** Any two of (a) assuming discrete values often, (b) having heavy tails, (c) approximately symmetric with respect to zero.

13. Let \( r_t \) and \( r_{m,t} \) be the monthly log returns of an asset and the market index, respectively. Suppose we like to extend the well-known Market model \( r_t = \alpha + \beta r_{m,t} + \epsilon_t \) so that the intercept \( \alpha \) or the slope \( \beta \) may depend on the performance of the market. Describe one extension that allows for the intercept, but not the slope, to change depending on the sign of the market. Also, describe another model that allows for the slope, but not the intercept, to change depending on the sign of the market.

**Answer:** Define the dummy variable \( M_t = 1 \) if \( r_{m,t} < 0 \) and \( M_t = 0 \), otherwise. The first extension is \( r_t = \alpha_1 + \alpha_2 M_t + \beta r_{m,t} + \epsilon_t \). The second extension if \( r_t = \alpha + \beta_1 r_{m,t} + \beta_2 (M_t \times r_{m,t}) + \epsilon_t \).

14. **(For Questions 14 to 18)** Consider the U.S. monthly growth rate of consumer price index, seasonally adjusted, from February 1947 to April 2013 for 795 observations. Analysis of the series is provided in the attached R output. Use the output to answer Questions 14 to 18. Write down the fitted pure time series model for the series \( x_t \), including the residual variance. (See model m1.)

**Answer:** \( (1 - 1.243B + 0.322B^2 - 0.063B^3)(1 - 0154B^{12})(r_t - 0.0029) = (1 - 0.825B)(1 - 0.383B^{12})a_t \), where \( \sigma_a^2 = 7.26 \times 10^{-6} \).
15. Based on the provided model checking statistics, is there any serial correlation in the residuals? Is there any ARCH effect in the residuals? Justify your answer briefly.

**Answer:** No, there are no significant serial correlations in the residuals, because the Ljung-Box statistics show $Q(12) = 14.85$ with $p$-value 0.25. On the other hand, the residuals show ARCH effect, because the Ljung-Box statistics of the squared residuals $Q(12) = 285.9$ with $p$-value close to zero.

16. A Gaussian GARCH(1,1) model is entertained, write down the fitted model for the residuals.

**Answer:** The residual is $a_t = \sigma_t \epsilon_t$, where $\epsilon_t \sim iid N(0,1)$ and $\sigma^2_t = 3.397 \times 10^{-7} + 0.216 \sigma^2_{t-1} + 0.749 \sigma^2_{t-1}$.

17. A GARCH(1,1) model with Student-$t$ innovations is also entertained. Let $v$ be the degrees of freedom of the Student-$t$ innovation. Consider the hypothesis $H_0 : v = 5$ versus $H_a : v \neq 5$. Calculate the test statistic and draw your conclusion.

**Answer:** $t = (5.252 - 5)/0.96 = 0.26$, which is insignificant at the 5% level. Thus, one cannot reject the null hypothesis that the degrees of freedom of the Student-$t$ innovations are 5.

18. One may combine the time series model with a volatility model to describe the dynamics of U.S. monthly CPI growth rates. Which volatility model is preferred? Why?

**Answer:** The GARCH(1,1) model with Student-$t$ innovations is preferred because it has smaller information criteria.

19. Give two types of nonlinear time series models discussed in the lectures that are useful in finance.

**Answer:** Any two of (a) threshold model, (b) Markov switching model, and (c) neural network.

20. Give two models that can be used to model price changes in high-frequency financial transactions.

**Answer:** Ordered probit model and the decomposition model with price change $C_i = A_i D_i S_i$.

**Problem B.** (26 points) Consider the daily adjusted close prices of Apple (AAPL) and McDonald’s (MCD) stocks, starting from some day in January 2002 for 2850 observations. We compute the daily log returns from the price series, resulting in 2849 returns. Consider a long position of $1 million on Apple stock and $1 million on McDonald’s stock. Use the attached output to answer the following questions.
1. (4 points) If RiskMetrics is used, calculate the VaR for Apple and McDonald’s positions separately.

**Answer:** For Apple stock: VaR = $49,378. For NCD, VaR = $19,404.

2. (3 points) A Gaussian GARCH(1,1) model is applied to the Apple log returns. Based on the model, compute 1-day and 5-day VaR of the Apple position.

**Answer:** 1-day VaR is $49,054. For 5-day VaR, one need to compute $E(r_{t+1} + \cdots + r_{t+5}|F_t) = -0.01042$ and $\text{Var}(r_{t+1} + \cdots + r_{t+5}|F_t) = 0.0024371$. Therefore, 5-day VaR is $104,427$.

3. (3 points) A GARCH(1,1) model with Student-$t$ innovations is applied to the Apple stock returns. Based on the model, what are the VaR and expected shortfall for the Apple position?

**Answer:** VaR = $54,687$ and ES = $71,686$.

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

**Answer:** The estimates are $(x_i, \sigma, \mu) = (0.150, 0.0154, 0.032)$ and their standard errors are $(0.057, 0.0011, 0.00146)$. The estimates are significant at the 5% level because their $t$-ratios are greater than 1.96.

5. (2 points) What are the 1-day and 10-day VaRs for the Apple position based on the traditional block-maximum method?

**Answer:** 1-day VaR = $58,896$ and 10-day VaR = $10.150 \times 58896 = $83,193$.

6. (2 points) What are the VaR and expected shortfall for the Apple position based on the POT approach?

**Answer:** VaR = $64,583$ and ES = $86,236$.

7. (4 points) Consider the approach based on generalized Pareto distribution. The threshold of 3% is used for the AAPL log returns. Write down the parameter estimates for the stock? Are these estimates significantly different from zero? Why?

**Answer:** The estimates are $(x_i, \beta) = (0.114, 0.0153)$ and the standard errors are 0.064 and 0.0014, respectively. The estimate of $x_i$ is not significantly different from zero at the 5% level, because its $t$-ratio is $t = 0.114/0.064 = 1.78$.

8. (4 points) Consider, again, the RiskMetrics method. If constant correlation is used, what is the VaR of the combined position? If time-varying correlation is used, what is the VaR of the combined position?

**Answer:** Constant correlation is 0.297 so that VaR = $\sqrt{49378^2 + 19404^2 + 2 \times 0.297 \times 49378 \times 19404} = $58,170. For time-varying correlation, we have $p_t = 0.195$ so that VaR = $\sqrt{49378^2 + 19404^2 + 2 \times 0.195 \times 49378 \times 19404} = $56,466.
**Problem C.** (18 points). Caterpillar (CAT) transactions data on December 5, 2012. Focus on the trades during the regular trading hours only.

1. (2 points) Write down the fitted model \textbf{md1} for \( A_i \). Are the estimates significantly different from zero? Why?
   
   \textbf{Answer:} \( P(A_i = 1|A_{i-1}) = \frac{\exp(-0.024+1.243A_{i-1})}{1+\exp(-0.024+1.243A_{i-1})} \). The estimates are highly significant because their \( p \)-values are close to zero.

2. (2 points) Based on the fitted model, calculate the probabilities \( P(A_i = 1|A_{i-1} = 1) \) and \( P(A_i = 1|A_{i-1} = 0) \).
   
   \( \text{bf Answer: } P(A_i = 1|A_{i-1} = 1) = 0.555 \) and \( P(A_i = 1|A_{i-1} = 0) = 0.255 \).

3. (4 points) Write down the fitted model for \( D_i \) conditional on \( A_i = 1 \) when both \( D_{i-1} \) and \( C_{i-1} \) are used. [That is, \textbf{md2}] Are the estimates significantly different from zero? Why?
   
   \textbf{Answer:} The model is
   \[
   P(D_i = 1|A_i=1) = \frac{\exp(-0.0016 - 0.789D_{i-1} - 13.26C_{i-1})}{1 + \exp(-0.0016 - 0.789D_{i-1} - 13.26C_{i-1})}.
   \]
   The fit of the constant term is not significant because its \( p \)-value is 0.92, but the coefficients of \( D_{i-1} \) and \( C_{i-1} \) are highly significant.

4. (3 points) Based on the fitted model \textbf{md2}, compute the probabilities \( P(D_i = 1|A_i = 1, D_{i-1} = -1, C_{i-1} = -0.02) \) and \( P(D_i = 1|A_i = 1, D_{i-1} = 1, C_{i-1} = 0.01) \).
   
   \textbf{Answer:} \( P(D_i = 1|A_i = 1, D_{i-1} = -1, C_{i-1} = -0.02) = 0.741 \) and \( P(D_i = 1|A_i = 1, D_{i-1} = 1, C_{i-1} = 0.01) = 0.284 \).

5. (3 points) Another model is employed for \( D_i \) conditional on \( A_i = 1 \) and using \( D_{i-1} \) and \( S_{i-1} \) as explanatory variables. [That is, \textbf{md3}] Based on the model, compute \( P(D_i = 1|A_i = 1, D_{i-1} = -1, S_{i-1} = 2) \) and \( P(D_i = 1|A_i = 1, D_{i-1} = 1, S_{i-1} = 5) \).
   
   \textbf{Answer:} \( P(D_i = 1|A_i = 1, D_{i-1} = -1, S_{i-1} = 2) = \exp(0.393 - 0.796 * (-1) - 0.099 * 2)/(1 + \exp(0.393 - 0.796 * (-1) - 0.099 * 2)) = 0.729 \) and \( P(D_i = 1|A_i = 1, D_{i-1} = 1, S_{i-1} = 5) = \exp(0.393 - 0.796 - 0.099*5)/(1 + \exp(0.393 - 0.796 - 0.099*5)) = 0.289 \).

6. (2 points) Consider the probabilities in questions 4 and 5. Since \( C_{i-1} = -0.02 \) corresponds to \( S_{i-1} = 2 \), why the two models give different probabilities?
   
   \textbf{Answer:} The two models \textbf{md2} and \textbf{md3} look similar and give similar probabilities, but the probabilities are not identical because the model \textbf{md3} uses categories, not the actual price changes.

7. (2 points) Compare the two models \textbf{md2} and \textbf{md3}. Which model is preferred? Why?
   
   \textbf{Answer:} The model \textbf{md2} is preferred for two reasons. First, it has a smaller AIC. Second, strictly speaking, \( S_i \) is a categorical variable, not numeric variable. In this particular, \( S_i \) is ordinal. [One reason is sufficient.]
Problem D. (16 points). Consider two financial assets. The first asset is the Vanguard energy ETF (VDE) and the second is energy select sector SPDR (XLE) of State Street Global Advisors. The sample period is from July 1, 2009 to June 30, 2011 for 505 observations. Denote the log price (namely, daily adjusted close price) of the asset by vde\(_t\) and xle\(_t\), respectively. Some analysis of the data is given. Answer the following questions:

1. (2 points) Are the two log price series co-integrated? Why?
   Answer: Yes, the two log prices are co-integrated, because their difference \(w_t\) series is stationary.

2. (2 points) Let \(w_t = vde_t - xle_t\). Write down the fitted model for \(w_t\).
   Answer: \((1 - 0.647B - 0.287B^2)(w_t - 0.381) = a_t\) with \(\sigma_a^2 = 1.529 \times 10^{-6}\).

3. (2 points) What is the approximate half-life for the \(w_t\) series?
   Answer: \(h = \log(0.5)/\log(0.95) = 13.51\).

4. (2 points) The time plot of \(w_t\) is shown in Figure 1. The horizontal line denotes the sample mean of \(w_t\). From the plot and the information provided, is there any pairs trading opportunities? Why?
   Answer: Yes, the plot and the mean-reverting of \(w_t\) indicate the possibilities for pairs trading.

5. (2 points) Suppose that the transaction cost is 2%. Are there sufficient opportunities to conduct pairs trading? Why?
   Answer: No, with cost at 2%, there is essentially no pairs trading opportunity. With \(E(w_t) = 0.381\), a horizontal line of 0.391 (0.381+1% of cost) is above all data points. Similarly, a horizontal line of 0.371 would below most of the data points.

6. (2 points) Suppose that the transaction cost is reduced to 1%. Are there sufficient opportunities to conduct pairs trading? Why?
   Answer: With trading cost of 1%, there are a few trading opportunities, because one can construct two horizontal lines separated by 1% that intersect with the \(w_t\) series.

7. (2 points) We extend the sample period so that the data span is from July 1, 2009 to December 31, 2011. Again, let \(w_t = vde_t - xle_t\). Write down the model for \(w_t\) and calculate the approximate half-life of this new \(w_t\) series.
   Answer: The model for \(w_t\) becomes \((1 - 0.625B - 0.246B^2 - 0.076B^3)(w_t - 0.38) = a_t\) with \(\sigma_a^2 = 1.604 \times 10^{-6}\). The half-life increases to 17.89, which is relatively high.

8. (2 points) Compare the two models for \(w_t\) and discuss the implication, if any, on pairs trading.
   Answer: By expanding the sampling period for 6 months, the \(w_t\) series moves closer to being unit-root nonstationary. This example demonstrates the necessity to monitor the behavior of \(w_t\). It also indicates that co-integration is a necessary, but not sufficient condition, for pairs trading.