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

1. Consider the log returns of an asset. Let $r_{m,t}^m$ be the monthly return and $r_{i,t}^d$ be the return on the $i$-th trading day of the month. Write down two assumptions used for justifying the following identity: $\text{Var}(r_{m,t}^m) = n \times \text{Var}(r_{i,t}^d)$, where $n$ is the number of trading days in month $t$.

   Answer: (a) daily returns $r_{i,t}^d$ are serially uncorrelated and (b) they have the same variance.

2. Assume that $x_t$ follows the stochastic diffusion equation $dx_t = -0.9dt + 2dw_t$, where $w_t$ is the Wiener process. Let $G(x_t) = \exp(x_t)$. Derive the stochastic diffusion equation for $G(x_t)$.

   Answer: $rac{\partial G(x_t)}{\partial x_t} = G(x_t)$, $rac{\partial G(x_t)}{\partial t} = 0$, and $rac{\partial^2 G(x_t)}{\partial x_t^2} = G(x_t)$. Applying the Ito’s lemma, we have
   
   $$dG(x_t) = 1.1G(x_t) + 2G(x_t)dw_t,$$
   
   which is a geometric Brownian motion.

3. Assume that the underlying true returns $\{r_t\}$ $(t = 1, \ldots, \infty)$ of an asset are independent and identically distributed with mean 0.1 and variance 0.04. Furthermore, assume that for each time interval $t$, the probability of no trade is 0.6. Let $r_t^o$ be the observed return of the asset. Is $r_t^o$ serially correlated? If yes, compute its lag-1 ACF.

   Answer: Yes, $r_t^o$ is serially correlated. Its lag-1 ACF is $\rho_1 = -\frac{(0.6)(0.4)(0.1)^2}{0.4 \times 0.04 + 2(0.6)(0.1)^2} = -0.0857$.

4. Give two weaknesses of the realized volatility constructed via the intraday high-frequency transactions data.

   Answer: (a) ignore over-night volatility and (b) impact of market micro-structure noises.

5. In volatility modeling, give two reasons for preferring the GARCH-M model over the usual GARCH models even though the former is harder to estimate.

   Answer: (a) Risk premium assessment and (b) continuous diffusion.

6. Suppose that $\mathbf{X}_t$ follows a 2-dimensional stationary vector AR(1) model,

   $\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.5 \end{bmatrix} + \begin{bmatrix} 0.8 & 0.2 \\ 0.0 & -0.7 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix},$
where $\mathbf{a}_t$ 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 shows Granger causality because $x_{2t}$ does not depend on any past values of $x_{1t}$. In this particular instance, $x_{1t}$ does not cause $x_{2t}$, but $x_{2t}$ causes $x_{1t}$.

7. Give two reasons for analyzing jointly multiple financial time series.

Answer: (a) Finding relationship between variables and (b) improving accuracy in forecasting.

8. (For Questions 8 to 10): In a case study of credit card default, a sample of 10,000 people were used. The explanatory variables used are (a) $x_1$: credit card balance (in dollars), (b) $x_2$: income (measured in thousands), and (c) $x_3$: an indicator for student, i.e. $x_3 = 1$ if and only if the card holder is a student. A logistic regression is employed and the results are given below:

<table>
<thead>
<tr>
<th></th>
<th>coefficient</th>
<th>Std. Error</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>$-10.869$</td>
<td>$0.492$</td>
<td>$-22.08$</td>
<td>$&lt;0.0001$</td>
</tr>
<tr>
<td>$x_1$</td>
<td>$0.0057$</td>
<td>$0.0002$</td>
<td>$24.74$</td>
<td>$&lt;0.0001$</td>
</tr>
<tr>
<td>$x_2$</td>
<td>$0.0030$</td>
<td>$0.0082$</td>
<td>$0.37$</td>
<td>$0.7115$</td>
</tr>
<tr>
<td>$x_3$</td>
<td>$-0.6468$</td>
<td>$0.2362$</td>
<td>$-2.74$</td>
<td>$0.0062$</td>
</tr>
</tbody>
</table>

Based on the fitted model, is being a student more likely to default conditioned on given $x_1$ and $x_2$? Why?

Answer: No, students are less likely to default conditioned on $x_1$ and $x_2$ because the coefficient is negative and statistically significant.

9. Letting $x_1$ and $x_2$ be fixed, interpret the coefficient $-0.6468$ of $x_3$.

Answer: Conditional log odd ratio. Specifically, $\exp(-0.6468) = \frac{\text{[Default/(Not Default)] | student}}{\text{[Default/(Not Default)] | not student}}$.

10. What is the estimated default probability for a student with credit balance $\$1,500$ and income $\$40,000$? What is the corresponding default probability if the holder is not a student?

Answer: The probability of default for student is

$$p = \frac{\exp(-10.869 + 0.0057(1500) + 0.003(40) - 0.6468)}{1 + \exp(-10.869 + 0.0057(1500) + 0.003(40) - 0.6468)} = 0.0549.$$ For non-student, the probability is

$$p = \frac{\exp(-10.869 + 0.0057(1500) + 0.003(40) - 0.6468)}{1 + \exp(-10.869 + 0.0057(1500) + 0.003(40) - 0.6468)} = 0.0998.$$
11. (For Questions 11-15): To estimate the risk of holding $1 million in Hewlett-Packard stock, tick symbol (HPQ), we consider the daily log return of the stock starting from January 4, 2002 for 3126 observations. A GARCH(1,1) model with Student-$t$ innovations is entertained. Write down the fitted model.

Answer: The model is

\[ r_t = -6.332 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^*_4.16 \]

\[ \sigma_t^2 = 3.411 \times 10^{-6} + 0.0394a_{t-1}^2 + 0.954\sigma_{t-1}^2. \]

12. Based on the fitted model, perform the hypothesis \( H_0 : v = 5 \) vs \( H_a : v \neq 5 \), where \( v \) denotes the degrees of freedom of the Student-$t$ innovations. Draw your conclusion.

Answer: \( t\)-ratio = \( \frac{4.164-5}{0.297} \) = -2.83, which is greater than 1.96 in absolute value. Therefore, the null hypothesis of \( v = 5 \) is rejected.

13. The volatility obtained from the prior question and the VIX index are used to better estimate the risk. Specifically, we employ quantile regression with explanatory variables (a) the lag-1 volatility of the stock \( (v_{m1}) \) and (b) the lag-1 VIX index \( (v_{ixm1}) \). Write down the parameter estimates when the 95% quantiles are used. Are these estimates significantly different from zero? Why?

Answer: The estimates (standard errors) are 0.00201(0.00287), 0.545(0.149), and 0.00087(0.00014). The intercept is no significantly different from zero, but the other two coefficients are, because \( p\)-value of the intercept is 0.48 whereas those of the other two estimates are close to zero.

14. Given \( (v_{m1}, v_{ixm1}) = (0.01726, 12.08) \), what is the VaR of the position with tail probability 0.05?

Answer: The fitted value is 0.00201 + 0.545(0.01726) + 0.00087(12.08) = 0.021926. Therefore, \( \text{VaR}_{0.95} = $21,926. \)

15. A quantile regression with \( \tau = 0.99 \) is also considered. Based on the fitted model, compute the VaR of the position with tail probability 0.01 when \( (v_{m1}, v_{ixm1}) = (0.01726, 12.08) \).

Answer: The fitted value is 1.312(0.01726) + 0.00116(12.08) = 0.0366579. Therefore, \( \text{VaR}_{0.99} = $36,657.9. \)

16. Let \( p_{1t} \) and \( p_{2t} \) be the daily log prices of two stocks. Assume that both \( p_{it} \) series are unit-root nonstationary. Let \( w_t = p_{1t} - \gamma p_{2t} - \beta_0 \) be a linear combination of the two log price series, where \( \beta_0 \) and \( \gamma \) are constant. It is also known that \( w_t \) follows the model

\[ w_t = 0.702w_{t-1} + 0.122w_{t-2} + a_t, \quad a_t \sim \text{iid} N(0, 3.3 \times 10^{-4}). \]

Can the two stocks be used in pairs trading? Why?

Answer: Yes, because the \( w_t \) series is mean-reverting. The characteristic roots of \( w_t \) are approximately 0.846 and 0.144, respectively, indicating that \( w_t \) is stationary with half-life 4.14.
17. The airline model \((1 - B)(1 - B^4)x_t = (1 - \theta_1 B)(1 - \theta_4 B^4)a_t\) is widely applicable in modeling quarterly earnings of a company, where \(0 < \theta_i < 1\). Why? In addition, write down the lags of non-zero ACF of the corresponding MA process \(w_t = (1 - B)(1 - B^4)x_t\). [No calculation of ACF is needed.]

Answer: It is a double exponential smoothing model with one regular exponential smoothing and one seasonal exponential smoothing. The non-zero lags are 1, 3, 4, and 5. [Lag 0 is 1, which is obviously not zero.]

18. Describe two possible applications of moving-average (MA) models in finance.

Answer: (a) Modeling bid-ask bounce and (b) detecting data smooth.

19. Describe two possible applications of ordered Probit models in finance.

Answer: (a) Modeling price change in tick-by-tick data analysis and (b) classifying asset returns into categories.

20. Give two multivariate volatility models that produce positive-definite volatility matrices with probability 1.

Answer: Any two of (a) exponentially weighted moving average, (b) BEKK, and (c) DCC model.

**Problem B.** (10 points) Consider again the daily log return of Hewlett-Packard stock, starting from January 4, 2002 for 3126 observations. Let \(r_t\) be the log return and define

\[
D_t = \begin{cases} 
1 & \text{if } r_t > 0 \\
0 & \text{if } r_t \leq 0.
\end{cases}
\]

Let \(z_t = \text{VIX}_t - \text{VIX}_{t-1}\) be the change series of the VIX index, and define

\[
\text{DVIX}_t = \begin{cases} 
1 & \text{if } z_t > 0 \\
0 & \text{if } z_t \leq 0.
\end{cases}
\]

We are interested in modeling \(D_t\) with the following four explanatory variables: (a) \(r_{t-1}\), (b) \(r_{t-2}\), (c) \(\text{DVIX}_{t-1}\) and (d) \(\text{DVIX}_{t-2}\).

1. A logistic linear regression is employed. Write down the fitted model.

Answer: \(\logit = -2.864r_{t-1} - 0.994r_{t-2} + 0.0288\text{DVIX}_{t-1} + 0.0828\text{DVIX}_{t-2}\). Alternatively,

\[
P(D_t = 1) = \frac{\exp(-2.864r_{t-1} - 0.994r_{t-2} + 0.0288\text{DVIX}_{t-1} + 0.0828\text{DVIX}_{t-2})}{1 + \exp(-2.864r_{t-1} - 0.994r_{t-2} + 0.0288\text{DVIX}_{t-1} + 0.0828\text{DVIX}_{t-2})}.
\]

2. Based on the fitted model, is past return \(r_{t-1}\) helpful in predicting \(D_t\) when \(\alpha = 0.1\) is used as the Type-I error? Why?

Answer: Yes, because the \(p\)-value of the coefficient of \(r_{t-1}\) is 0.0878, which is less than 0.1.
3. Given \((r_{t-1}, r_{t-2}, \text{DVIX}_{t-1}, \text{DVIX}_{t-2}) = (0.002076, 0.00775, 1.0, 1.0)\), compute the probability \(P(D_t = 1)\).

Answer: \(p = 0.524\).

4. A 4-2-1 neural network with skip layer is entertained. Write down the models for the two hidden nodes \(h_1\) and \(h_2\).

Answer: The functions are
\[
h_1 = \frac{\exp(-11.06 + 39.43r_{t-1} + 16.82r_{t-2} + 17.25\text{DVIX}_{t-1} + 5.19\text{DVIX}_{t-2})}{1 + \exp(-11.06 + 39.43r_{t-1} + 16.82r_{t-2} + 17.25\text{DVIX}_{t-1} + 5.19\text{DVIX}_{t-2})}
\]
\[
h_2 = \frac{\exp(3.09 + 21.12r_{t-1} - 14.62r_{t-2} + 14.5\text{DVIX}_{t-1} - 21.81\text{DVIX}_{t-2})}{1 + \exp(3.09 + 21.12r_{t-1} - 14.62r_{t-2} + 14.5\text{DVIX}_{t-1} - 21.81\text{DVIX}_{t-2})}.
\]

5. Write down the model for the output node.

Answer: Let \(z = 9.63 - 14.68h_1 - 10.09h_2 + 3.97r_{t-1} - 2.15r_{t-2} + 15.08\text{DVIX}_{t-1} - 9.62\text{DVIX}_{t-2}\). The output function is
\[
h(z) = \begin{cases} 
1 & \text{if } z > 0 \\
0 & \text{if } z \leq 0.
\end{cases}
\]

**Problem C.** (28 points) Consider a portfolio consisting of two stocks: (a) Johnson and Johnson (JNJ) and (b) Hewlett-Packard (HPQ) each with $1 million. To estimate the risk of the portfolio, we use daily log returns of the stocks, starting from January 4, 2002 for 3126 observations.

1. (3 points) Focus on the JNJ stock. If RiskMetrics is used, write down the fitted model. What are the VaR and Expected shortfall?

   Answer: The model is
   \[
   r_t = a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1)
   \]
   \[
   \sigma_t^1 = 0.073r_{t-1}^2 + 0.927\sigma_{t-1}^2.
   \]
   The VaR is $13,647.4 and ES is $15,635.3.

2. (2 points) If constant correlation coefficient is used, what is the VaR for the portfolio under RiskMetrics?

   For HPQ, VaR is $36,673. Therefore, for the portfolio the VaR is
   \[
   \$ \sqrt{(13647.4)^2 + (36673)^2 + 2 \times 0.33 \times (13647.4)(36673)} = $43,144.9.
   \]

3. (4 points) Focus again on the JNJ stock. If GARCH(1,1) model with Gaussian innovations is used, what are the VaR and Expected shortfall? Also, what are the VaR and Expected shortfall for the next 10-trading days?

   Answer: For the next trading day: VaR = $13,802.3 and ES = $15,896.2. For the next 10 trading days, VaR = $44,892.6 and ES = $52,265.4.
4. (2 points) Focus on the JNJ stock. If GARCH(1,1) model with Student-t innovations is used, what are the VaR and Expected shortfall?

Answer: VaR = $15,763.9 and ES = $20,586.3.

5. (4 points) Focus on the HPQ stock. Consider the traditional extreme value theory with block size 21, write down the three parameter estimates and the associated VaR.

Answer: The estimates and their standard errors (in parentheses) are 0.349(0.0786), 0.0143(0.00113), and 0.0271(0.00135), respectively, for $\xi$, $\sigma$ and $\mu$. The VaR is $56,722.

6. (5 points) Focus on the HPQ stock and apply the generalized Pareto distribution with threshold 3.5%. Write down the parameter estimates. Are these estimates statistically significant? What are the corresponding VaR and Expected shortfall?

Answer: The estimates(standard errors) are 0.232(0.0934) and 0.0149(0.00181), respectively, for $\xi$ and $\beta$. The estimates are statistically significant as their $t$-ratio are greater than 1.96. Based on the estimates, VaR = $62,540 and ES is $90,209.4.

7. (3 points) Focus on the HPQ stock and apply the generalized Pareto distribution with threshold 4.5%. What are the VaR and Expected shortfall implied by the model? What is the VaR for the next 10-trading days?

Answer: VaR is $61,021.6 and ES is $91,118.6. For the next 10-trading day VaR is $ (61.021.6) * 10^{0.344} = $134,736.

8. (2 points) Are the VaR and Expected shortfall sensitive to the choice of the threshold when generalized Pareto distribution is used? Why?

Answer: No, the VAR and ES are not sensitive to the choice of threshold as both thresholds give close results.

9. (3 points) Revisit the RiskMetrics method, but consider a new portfolio that holds a short position on HPQ with $1 million and a long position on JNJ with $2 million. What is the VaR for the portfolio if the sample correlation is used?

Answer: The VaR for HPQ is $36,673 and that for JNJ is $27,294.8. The correlation becomes −0.33. Therefore, VaR is $37,805.6 for the portfolio.

Problem D. (22 points) Consider the daily log returns of IBM stock and the Bank of America Merrill Lynch AAA bond index as used in the lecture. Let $r_t$ be the 2-dimensional log returns, in percentages. The lag-1 sample cross-correlation matrix of $r_t$ is

$$
\rho_1 = \begin{bmatrix}
-0.0125 & 0.0263 \\
0.0376 & 0.0546
\end{bmatrix}.
$$
1. (2 points) Explain the meaning of each element of $\rho_1$.

Answer: Elements $-0.0125$ and $0.0546$ are the lag-1 ACF of IBM return and bond return, respectively. The element $0.0263$ measures the linear dependence of IBM return on the past lag of bond return. Finally, the element $0.0376$ quantifies the linear dependence of bond returns on the past lag of IBM return.

2. (2 points) Test the hypothesis $H_0 : \rho_1 = \rho_2 = \cdots = \rho_{10} = 0$ versus $H_a : \rho_i \neq 0$ for some $i$. Write down the test statistic and draw your conclusion.

Answer: The Ljung-Box statistics give $Q(10) = 73.7$ with $p$-value close to zero. Therefore, the null hypothesis is rejected, implying that there are serial correlations in the bivariate returns.

3. (3 points) The AIC criterion selects a VAR(2) model. Write down the fitted AR(1) coefficient matrix. Identify the significant coefficient estimates.

Answer: The coefficient estimates and their standard errors (in parentheses) are

$$\phi_1 = \begin{bmatrix} -0.00679(0.0203) & 0.0935(0.0933) \\ 0.0117(0.00445) & 0.0680(0.0204) \end{bmatrix}.$$  

The estimates $0.0117$ and $0.068$ are the only two significant estimates.

4. (2 points) Are the residuals of the VAR(2) model serially correlated? Why?

Answer: No, the Ljung-Box statistics of the residuals give $Q(10) = 42.394$ with $p$-value 0.37. Therefore, the null hypothesis of zero cross-correlation matrices cannot be rejected at the 5% level.

5. (2 points) The exponentially weighted moving average method is used to estimate the volatility matrices of $a_t$, the residual series of the VAR(2) model. Write down the fitted model.

Answer: Let $\Sigma_t$ be the volatility matrix. The model is

$$\Sigma_t = 0.962\Sigma_{t-1} + 0.038a_{t-1}a'_{t-1}.$$  

6. (3 points) Write down the range and mean of the resulting time-varying correlations.

Answer: The range is $(-0.702, 0.286)$ and the mean value is $-0.2$.

7. (3 points) A DCC model with bivariate Student-$t$ innovations is applied to the residual series $a_t$. Write down the estimates of $\theta_1$, $\theta_2$ and degrees of freedom of the DCC model.

Answer: The estimates (standard errors) are $0.977(0.01)$ and $0.0115(-.00483)$ for $\theta_1$ and $\theta_2$, respectively. The estimated degrees of freedom are $7.46$ with standard error $0.587$.

8. (3 points) Write down the range and mean of the time-varying correlations based on the fitted DCC model.

Answer: Range is $(-0.346, 0.0324)$ and the mean is $-0.175$.  

7
9. (2 points) Figure 1 shows the time plots of the time-varying correlation coefficients. The solid line is based on the EWMA method and the "*" is based on the DCC model. Comments of the two approaches to model the time-varying correlations.

Answer: The DCC model, as expected, gives low variability in the time-varying correlations. The patterns of the estimated correlations, however, are similar.