Booth Honor Code:  
*I pledge my honor that I have not violated the Honor Code during this examination.*

Signature: ___________________________  Name: ___________________________  ID: ___________________________

Notes:
- This is a 3-hour, open-book, and open-notes exam.
- Write your answer in the blank space provided for each question.
- There are 14 pages, including some R output.
- For simplicity, ALL tests use the 5% significance level, and all risk-measure calculations use 1% tail probability. Furthermore, unless specified, all VaR and expected shortfall are for the next trading day.
- Round your answer to 3 significant digits.
- You may bring a PC or calculator to the exam.

**Problem A: (32 pts)** Answer briefly the following questions.

1. 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 for the inverse price $\frac{1}{P_t}$?

2. Suppose that the price of a stock follows a geometric Brownian motion with drift 5% per annum and volatility 36% per annum. The stock pays no dividends and its current price is $70. Assume further that the risk-free interest rate is 4% per annum. (a) What is the price of a European call option contingent on the stock with a strike price of $71 that will expire in 3 months? (b) What is the corresponding put option price?

3. Give two weaknesses of using value at risk (VaR) to quantify financial risk.
4. Define a coherent risk measure.

5. Describe two approaches to overcome the market microstructural noises in computing realized volatility.

6. Give two weaknesses of realized volatility as an estimate of daily stock volatility.

7. Describe two main assumptions used by the RiskMetrics?

8. Describe two ways to apply extreme value theory in calculating VaR.

9. Give two main difficulties in modeling multivariate volatility of asset returns.

10. (For Questions 10 and 11). The log return \( r_t \) of a stock follows the model

\[
\begin{align*}
    r_t &= -0.016 + 0.1\sigma_t^2 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^*_{6.5} \\
    \sigma_t^2 &= 0.03 + 0.05a_{t-1}^2 + 0.9\sigma_{t-1}^2
\end{align*}
\]

where \( t^*_v \) denotes standardized Student-\( t \) distribution ith \( v \) degrees of freedom. Suppose further that \( r_{100} = 0.02 \) and \( \sigma_{100} = 0.4 \). Calculate the 1-step ahead mean and volatility predictions of \( r_t \) at the first origin \( T = 100 \).

11. Calculate the 2-step ahead predictions of the mean and volatility of \( r_t \) at the forecast origin \( T = 100 \).
12. (For Questions 12 and 13). The quarterly earnings of a company follows the model
\[(1 - B)(1 - B^4)x_t = (1 - 0.4B)(1 - 0.6B^4)a_t, \quad a_t \sim N(0, \sigma^2),\]
where \(\sigma^2 = 0.16\). Let \(w_t = (1 - B)(1 - B^4)x_t\). Give the lags for which \(w_t\) has non-zero ACF.

13. Obtain the values of all non-zero ACFs of the \(w_t\) series.

14. Let \(x_t = (x_{1t}, x_{2t})'\) be a bivariate stationary time series. Denote the mean vector of \(x_t\) as \(E(x_t) = \mu\). Define the lag-\(j\) autocovariance matrix of \(x_t\). What is the meaning of the \((1,2)\)th element of \(\Gamma_j\)? Here \((1,2)\)th element denotes the upper-right element of the 2-by-2 matrix \(\Gamma_j\).

15. Give two reasons for modeling multivariate time series jointly.

16. Describe two simple approaches discussed in the lecture to model multivariate asset volatility.

**Problem B.** (44 points) Consider the daily adjusted close prices of the stocks of Apple and ExxonMobil from January 3, 2002, to May 31, 2011. The tick symbols are AAPL and XOM, respectively. From the data, we obtain the 2368 daily log returns of the stocks. Consider a long position of $1 million on each of the two stocks. Use the attached output to answer the following questions.

1. (6 points) If RiskMetrics is used, what is the underlying model for the AAPL stock? (Write down the fitted model.) What are the VaR and expected shortfall for the position on AAPL stock?
2. (6 points) Again, if RiskMetrics is used, what is the VaR for the position on XOM stock? What is the corresponding VaR for the next 10 trading days? The sample correlation between the two daily log returns is 0.3723. What is the VaR for the joint position of the two stocks?

3. (4 points) If ARMA-GARCH models with Gaussian innovations are used to model the daily stock returns, what are the VaR and expected shortfall for AAPL stock?

4. (4 points) Again, based on GARCH models with Gaussian innovations, what are the VaR for XOM stock position and the joint position of the two stocks?

5. (4 points) If GARCH models with Student-\(t\) innovations are used, write down the fitted model for the XOM stock?

6. (4 points) What are the VaR and expected shortfall for the XOM position based on the GARCH model with Student-\(t\) innovations?

7. (4 points) Next, apply the traditional extreme value theory with block size 21 to the Apple stock. What are the estimates of (\(x_i\), \(\sigma\), \(\mu\))? What is the VaR for the position on the Apple stock?

8. (4 points) Finally, consider the peaks over threshold (POT) approach. The thresholds of 4% and 2.5% are used for AAPL and XOM, respectively. Write down the parameter estimates for both stocks? Are these estimates significantly different from zero? Why?
9. (4 points) What are the VaR and expected shortfall for each of the two stock positions based on the POT method?

10. (4 points) The exponentially weighted moving average method is used to model the volatility matrix of AAPL and XOM returns. Let $\Sigma_t$ denote the volatility matrix. Write down the fitted volatility model. Also, based on the model, the correlation between the two stock on May 31, 2011 is 0.4108. What is the VaR of the joint position based on the RiskMetrics method if this new correlation is used?

Problem C. (12 points). Consider the intraday trading of the Allstate (ALL) stock on December 01, 2010. There are 14971 trades within the normal trading hours. Thus, we have 14970 price change points. Among those, only 3843 have non-zero price changes. 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. Simple logistic regression is used to model $A_i$ and $D_i$. The output is attached. Answer the following questions:

1. (2 points) Write down the fitted model for $A_i$. Are the estimates significantly different from zero? Why?

2. (2 points) Based on the fitted model, calculate $P(A_i = 1|A_{i-1} = 0)$ and $P(A_i = 1|A_{i-1} = 1)$.

3. (2 points) What is the meaning of the estimate 1.64958?

4. (2 points) Write down the fitted model for $D_i$ conditional on $A_i = 1$. Are the estimates significantly different from zero? Why?
5. (4 points) Based one the fitted 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 $P(D_i = 1 | A_i = 1, D_{i-1} = -1)$.

**Problem D.** (12 points). Consider the daily close prices, in log scale, of ExxonMobil (XOM) and Chevron (CVX) stocks from July 31, 2002 to December 31, 2008. We analyze the data to explore trading opportunities. A simple linear regression gives the model

$$x_{om_t} = 0.3004 + 0.9386cvx_t,$$

from which one can construct a spread process $w_t$. Figure 1 shows the $w_t$ series with three horizontal lines at $Y = (0.25, 0.30, 0.35)$, respectively. Based on the plot and the discussions of pairs trading in class, answer the following questions:

1. Describe the basic idea behind pairs trading.

2. Write down the linear combination that provides a stationary process of the two log prices.

3. If $w_{t_0} = 0.35$, which stock is overvalued? Why?

4. (4 points) If the total trading costs (including initiation and closing of a position) for pairs trading are 2%, is there any opportunity to conduct pairs trading between the two stocks? If yes, briefly describe a trading strategy.

5. Why is stationarity of $w_t$ important in pairs trading?
Figure 1: Spread between XOM and CVX log prices: 7/31/2002 to 12/31/2008

Computer output.

#### Problem B ####

```r
getSymbols("AAPL", from="2002-01-03", to="2011-05-31")
[1] "AAPL"
getSymbols("XOM", from="2002-01-03", to="2011-05-31")
[1] "XOM"
aapl=diff(log(as.numeric(AAPL$AAPL.Adjusted)))
xom=diff(log(as.numeric(XOM$XOM.Adjusted)))
aapl=-aapl
xom=-xom
source("Igarch.R")
m1=Igarch(naapl,include.mean=F,volcnt=F)
Coefficient(s):  
Estimate Std. Error t value Pr(>|t|)  
beta 0.96718043 0.00448209 215.788 < 2.22e-16 ***  

length(naapl)
[1] 2368
m1$volatility[2368]
[1] 0.01164631
naapl[2368]
[1] -0.03041505
m2=Igarch(nxom,include.mean=F,volcnt=F)
Coefficient(s):  
Estimate Std. Error t value Pr(>|t|)  
beta 0.94048352 0.00580152 162.11 < 2.22e-16 ***
```
> m2$volatility[2368]
[1] 0.01031144
> nxom[2368]
[1] -0.01011448

> cor(xom,aapl)
[1] 0.3723159

> library(fGarch)
> m3=garchFit(~garch(1,1),data=naapl,trace=F)
> summary(m3)

Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 1), data = naapl, trace = F)

Mean and Variance Equation:
  data ~ garch(1, 1) [data = naapl]

Conditional Distribution: norm

Error Analysis:

|      | Estimate | Std. Error | t value | Pr(>|t|) |
|------|----------|------------|---------|----------|
| mu   | -2.257e-03 | 4.553e-04  | -4.956  | 7.18e-07 *** |
| omega| 7.342e-06  | 2.866e-06  | 2.562   | 0.0104 *   |
| alpha| 5.252e-02  | 8.982e-03  | 5.848   | 4.98e-09 *** |
| beta1| 9.370e-01  | 1.146e-02  | 81.739  | < 2e-16 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box Test R Q(10)</td>
<td>16.56879  0.084467</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>22.27602  0.3257231</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>5.821137  0.83006</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>10.91932  0.9482879</td>
</tr>
</tbody>
</table>

> predict(m3,2)

<table>
<thead>
<tr>
<th></th>
<th>meanForecast</th>
<th>meanError</th>
<th>standardDeviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.002256832</td>
<td>0.01591036</td>
<td>0.01591036</td>
</tr>
<tr>
<td>2</td>
<td>-0.002256832</td>
<td>0.01605746</td>
<td>0.01605746</td>
</tr>
</tbody>
</table>

> m4=garchFit(~arma(0,1)+garch(1,1),data=nxom,trace=F)
> summary(m4)

Title: GARCH Modelling

Call: garchFit(formula = ~arma(0, 1) + garch(1, 1), data = nxom, trace = F)

Mean and Variance Equation: data ~ arma(0, 1) + garch(1, 1) [data = nxom]
Conditional Distribution: norm

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | -7.945e-04 | 2.390e-04 | -3.324 0.000888 *** |
| ma1      | -8.464e-02 | 2.210e-02 | -3.831 0.000128 *** |
| omega    | 4.361e-06  | 1.163e-06 | 3.749 0.000177 *** |
| alpha1   | 7.473e-02  | 9.586e-03 | 7.796 6.44e-15 *** |
| beta1    | 9.057e-01  | 1.242e-02 | 72.942 < 2e-16 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box Test R Q(10)</td>
<td>5.734271 0.8370726</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>17.87614 0.595567</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>18.36097 0.04917135</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>22.34701 0.3219817</td>
</tr>
</tbody>
</table>

> predict(m4,2)

<table>
<thead>
<tr>
<th>meanForecast</th>
<th>meanError</th>
<th>standardDeviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 1.262777e-05</td>
<td>0.01108575</td>
<td>0.01108575</td>
</tr>
<tr>
<td>2 -7.944639e-04</td>
<td>0.01121277</td>
<td>0.01117344</td>
</tr>
</tbody>
</table>

> m5=garchFit(~garch(1,1),data=naapl,trace=F,cond.dist="std")
> summary(m5)

Title: GARCH Modelling

Call: garchFit(formula = ~garch(1, 1), data = naapl, cond.dist = "std", trace = F)

Mean and Variance Equation: data ~ garch(1, 1) [data = naapl]

Conditional Distribution: std

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | -1.841e-03 | 4.161e-04 | -4.423 9.72e-06 *** |
| omega    | 4.361e-06  | 2.370e-06 | 1.840 0.0657 . |
| alpha1   | 4.603e-02  | 1.042e-02 | 4.419 9.89e-06 *** |
| beta1    | 9.482e-01  | 1.183e-02 | 80.147 < 2e-16 *** |
| shape    | 5.582e+00  | 6.093e-01 | 9.161 < 2e-16 *** |

---

> predict(m5,2)

<table>
<thead>
<tr>
<th>meanForecast</th>
<th>meanError</th>
<th>standardDeviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 -0.001840778</td>
<td>0.01507109</td>
<td>0.01507109</td>
</tr>
<tr>
<td>2 -0.001840778</td>
<td>0.01517161</td>
<td>0.01517161</td>
</tr>
</tbody>
</table>

> qstd(.99,nu=5.582)
Title: GARCH Modelling

Call: garchFit(formula = ~arma(0, 1) + garch(1, 1), data = nxom, cond.dist = "std")

Mean and Variance Equation: data ~ arma(0, 1) + garch(1, 1)  
[Data = nxom]

Conditional Distribution: std

Error Analysis:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| mu       | -9.508e-04 | 2.286e-04 | -4.159 3.20e-05 *** |
| ma1      | -9.450e-02 | 2.142e-02 | -4.411 1.03e-05 *** |
| omega    | 4.309e-06  | 1.326e-06 | 3.249 0.00116 ** |
| alpha1   | 7.700e-02  | 1.266e-02 | 6.082 1.19e-09 *** |
| beta1    | 9.041e-01  | 1.529e-02 | 59.147 < 2e-16 *** |
| shape    | 8.378e+00  | 1.314e+00 | 6.377 1.81e-10 *** |

---

Standardised Residuals Tests:

<table>
<thead>
<tr>
<th>Statistic</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ljung-Box Test R Q(10)</td>
<td>6.080044 0.808494</td>
</tr>
<tr>
<td>Ljung-Box Test R Q(20)</td>
<td>18.20665 0.5737976</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(10)</td>
<td>17.16920 0.0707029</td>
</tr>
<tr>
<td>Ljung-Box Test R^2 Q(20)</td>
<td>21.30077 0.3796236</td>
</tr>
</tbody>
</table>

> predict(m6, 2)

```
> predict(m6, 2)
   meanForecast meanError standardDeviation
1  -6.315802e-05  0.01107138     0.01107138
2  -9.508449e-04  0.01120999     0.01116106
```

> qstd(.99, nu = 8.378)

```
[1] 2.500269
```

> library(evir)

```
> library(evir)
```

> m7 = gev(naapl, block = 21)

```
> m7
$n.all
 [1] 2368
$n
 [1] 113
$data
 [1] 0.044291017 0.075408193 0.031186683 0.046677969 0.053124524 0.162460517
```
$block
[1] 21

$par.est

\[
\begin{array}{ccc}
\text{xi} & \sigma & \mu \\
0.12562658 & 0.01583669 & 0.03335727 \\
\end{array}
\]

$par.se

\[
\begin{array}{ccc}
\text{xi} & \sigma & \mu \\
0.057799359 & 0.001192286 & 0.001637734 \\
\end{array}
\]

> source("evtVaR.R")
> evt VaR(0.12563,0.01584,0.03336)
[1] 0.06057381
>
> m8=gpd(naapl,threshold=0.04)
> m8
$n
[1] 2368
$data

\[
\begin{array}{c}
0.04429102 0.04332127 0.05504399 0.07264511 0.07540819 0.04667797 \\
\end{array}
\]

threshold: 0.04
$p.less.thresh: 0.9522804
$n.exceed: 113

$par.est

\[
\begin{array}{c}
0.32113620 0.01120458 \\
\end{array}
\]

$par.se

\[
\begin{array}{c}
0.128110955 0.001716158 \\
\end{array}
\]

> riskmeasures(m8,c(0.95,0.99,0.999))

\[
\begin{array}{ccc}
p & \text{quantile} & \text{sfall} \\
0.950 & 0.03948086 & 0.05574018 \\
0.990 & 0.06274113 & 0.09000372 \\
0.999 & 0.12583440 & 0.18294322 \\
\end{array}
\]

> m9=gpd(nxom,threshold=0.025)
> m9
$n
[1] 2368
$data

\[
\begin{array}{c}
0.02604230 0.02613568 0.02588947 0.03355710 0.04009306 0.03017261 \\
\end{array}
\]
$threshold: 0.025
$p.less.thresh: 0.9518581
$n.exceed: 114

$par.ests

<table>
<thead>
<tr>
<th>xi</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.24801082</td>
<td>0.01132185</td>
</tr>
</tbody>
</table>

$par.ses

<table>
<thead>
<tr>
<th>xi</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.117737811</td>
<td>0.001651101</td>
</tr>
</tbody>
</table>

> riskmeasures(m9,c(0.95,0.99,0.999))

<table>
<thead>
<tr>
<th>p</th>
<th>quantile</th>
<th>sfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,]</td>
<td>0.950</td>
<td>0.02457325</td>
</tr>
<tr>
<td>[2,]</td>
<td>0.990</td>
<td>0.04675867</td>
</tr>
<tr>
<td>[3,]</td>
<td>0.999</td>
<td>0.09867414</td>
</tr>
</tbody>
</table>

> source("EWMAvol.R")
> rtn=cbind(naapl,nxom)
> mm3=EWMAvol(rtn,-0.1)

Coefficient(s):

|               | Estimate | Std. Error | t value | Pr(>|t|) |
|---------------|----------|------------|---------|----------|
| lambda        | 0.962889 | 0.002848   | 338.0   | <2e-16   *** |

---

> names(mm3)

[1] "Sigma.t" "return" "lambda"
> Sigt=mm3$Sigma.t
> cor2=Sigt[,2]/sqrt(Sigt[,1]*Sigt[,4])
> cor2[2368]

0.4108287

#### Problem C ######

> da=read.table("all-taq-trade-12012010.txt")
> head(da)

V1 V2 V3 V4 V5 V6
1 20101201 9 30 0 29.53 500
6 20101201 9 30 10 29.65 100

> m1=hfchg(da) ### Compute price change series
> names(m1)

[1] "pchange" "duration"
> chg=m1$pchange
> length(chg)

[1] 14970
> idx=c(1:14970)[chg != 0]
> Ai=rep(0,14970)
> Ai[idx]=1
> Di=rep(0,14970)
> jdx=c(1:14970)[chg > 0]
> njdx=c(1:14970)[chg < 0]
> Di[jdx]=1
> Di[njdx]=-1

> ai=Ai[2:14970]
> aim1=Ai[1:14969]
> m2=glm(ai~aim1,family=binomial)
> summary(m2)

Call: glm(formula = ai ~ aim1, family = binomial)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -1.59805   0.02534  -63.06  <2e-16 ***
aim1          1.64958   0.04103   40.20  <2e-16 ***
---
> di=Di[2:14970]
> dim1=Di[1:14969]

> ddi=di[ai==1]
> ddim1=dim1[ai==1]
> ddi=(ddi+abs(ddi))/2
> m3=glm(ddi~ddim1,family=binomial)
> summary(m3)

Call: glm(formula = ddi ~ ddim1, family = binomial)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)   0.02930   0.03583   0.818  0.414
ddim1        -1.40442   0.05664  -24.795  <2e-16 ***
---

#### Problem D #######
> getSymbols("CVX",from="2002-7-31",to="2008-12-31")
[1] "CVX"
> getSymbols("XOM",from="2002-7-31",to="2008-12-31")
[1] "XOM"
> cvx=log(as.numeric(CVX$CVX.Adjusted))
> xom=log(as.numeric(XOM$XOM.Adjusted))
> b1=lm(xom~cvx)
> acf(b1$residuals)
> pacf(b1$residuals)
> plot(b1$residuals,type='l')
> summary(b1)
Call: lm(formula = xom ~ cvx)

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.300446   0.011487   26.16  <2e-16 ***
cvx          0.938645   0.002965  316.63  <2e-16 ***
---
Residual standard error: 0.04619 on 1616 degrees of freedom
Multiple R-squared: 0.9841, Adjusted R-squared: 0.9841
F-statistic: 1.003e+05 on 1 and 1616 DF, p-value: < 2.2e-16

## create a spread variable w_t.
> plot(wt,type='l')
> abline(h=c(0.25,0.3,0.35))