

Graduate School of Business, University of Chicago
Business 41202, Spring Quarter 2007, Mr. Ruey S. Tsay

Final Exam

GSB Honor Code:

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

Signature:

Name:

ID:

Notes:

- Open notes and books.
- Write your answer in the blank space provided for each question.
- **Manage** your time carefully and answer as many questions as you can.
- There are 14 pages, including output.
- For simplicity, **ALL** tests use the 5% significance level, and all VaR calculations use tail probability of 1%.
- Round your answer to 3 significant digits.

Problem A: (30 pts) Answer briefly the following questions.

1. Suppose that the price P_t of a stock follows the geometric Brownian motion

$$dP_t = \mu P_t dt + \sigma P_t dw_t,$$

where μ and σ are constant and w_t is the standard Brownian motion. What is the model for the square-root of the price $G(P_t) = P_t^{0.5}$?

2. Consider the daily log returns, in percentages, of the stock of Boeing Company from January 1997 to December 2006 for 2516 observations. Summary statistics of the returns are in the attached output. Let μ be the mean of the log return. Test $H_o : \mu = 0$ vs $H_a : \mu \neq 0$. What is the test statistic? Draw your conclusion.

3. Again, consider the log returns of Boeing stock and the output. Let ρ_i be the lag- i autocorrelation of the return. Test $H_o : \rho_1 = \rho_2 = \dots = \rho_{10} = 0$ vs $H_a : \rho_i \neq 0$ for some $1 \leq i \leq 10$. What is the test statistic? Draw your conclusion.

4. Again, consider the log returns of Boeing stock and the output. Is there any ARCH effect in the return? Why?

5. Consider a non-dividend paying stock. Suppose that the current price is $P_t = \$30$ and the risk-free interest rate is 6% per annum. If the price of a European call option of the stock is \$3.10 when the strike price is \$31 and the time to expiration is 3 months. What is the price of a European put option of the stock with the same duration and strike price?

6. Describe two methods to identify the order of an AR model.

7. Give two weaknesses of GARCH models in modeling asset volatility.

8. Suppose that the true price of a stock follows the random walk model

$$P_t^* = P_{t-1}^* + a_t, \quad a_t \sim N(0, 2).$$

Suppose also that the true price is the average of the bid and ask prices and the bid-ask spread is 0.25. What is the impact of the bid-ask bounce on the serial correlation of the simple return of the stock?

9. Consider the daily log returns, in percentages, of the FedEx stock from January 1997 to December 2006 for 2516 observations. A GARCH(1,1) model is fitted to the data. The result indicates that the following IGARCH(1,1) model seems reasonable for the log returns:

$$r_t = a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad \sigma_t^2 = 0.98\sigma_{t-1}^2 + 0.02a_{t-1}^2.$$

In addition, we have $a_{2516} = 0.544$ and $\sigma_{2516}^2 = 1.95$. What is the VaR for the next trading day for a position that long the stock valued at \$1 million dollars?

10. The computer output also shows that the correlation coefficient between the daily log returns of FedEx and Boeing stocks is 0.23. If you hold both stocks each valued at \$1 million dollars and you know that the VaR for the Boeing stock is \$28,000 for the next trading day. What is the VaR of your combined portfolio for the next trading day?

11. Describe two major difficulties in modeling the volatility series of multiple asset returns.

12. Suppose that the daily log returns, in percentages, of a stock follows the model

$$r_t = 0.2 + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1), \quad \sigma_t^2 = 0.94\sigma_{t-1}^2 + 0.06a_{t-1}^2.$$

Suppose that at the forecast origin T , $r_T = 0.54$ and $\sigma_T^2 = 2.05$. What is the 2-step ahead forecast $r_T(2)$ of the return? What is the 2-step ahead volatility forecast $\sigma_T(2)$?

13. Consider the growth rate of U.S. quarterly GDP from 1948 to 2004. An AR(3) model is fitted to the data. Write down the fitted model. Does the model imply the existence of business cycles? If yes, calculate the average length of business cycles.

14. Give two objectives of analysis of high-frequency financial data that cannot be achieved by using daily data.

15. The square root of time rule of RiskMetrics is based on some critical assumptions. List two of these assumptions used.

Problem B. (25 points) Consider the daily log returns, in percentages, of the Boeing stock and the output. Suppose that you hold a long position of the stock valued at \$1 million dollars. Answer the following questions.

1. Apply the traditional extreme value theory to the negative log returns with block size 63. What are the estimates of the three parameters k , α , and β ? Are these estimates statistically significant? Why?
2. Based on the estimates of block size 63, what is the VaR of your position for the next trading day?
3. Turn to the approach of peaks over the threshold. Using threshold 2.0, we estimate the parameters of generalized Pareto distribution for the stock returns. Does the log returns have a heavy left tail? Why?
4. Based on estimated generalized Pareto distribution, what is the VaR of your financial position for the next trading day? What is the VaR of your financial position for the next 10 trading days?
5. (2 pts) Again, based on the fitted Pareto distribution with threshold 2.0, what is the expected shortfall when the 1% VaR is used?
6. (3 pts) If empirical quantiles are used, what is the VaR of your financial position for the next trading day?

Problem C. (25 pts) Again, consider the daily log returns, in percentages, of Boeing stock from January 1997 to December 2006. Suppose also that your position remains unchanged.

1. A GARCH(1,1) model with Gaussian distribution is fitted to the data. Based on the fitted model, what is the VaR of your financial position for the next trading day?
2. A GARCH(1,1) model with Student- t innovations is also fitted to the data. Write down the fitted model.
3. Based on the fitted GARCH(1,1) model with Student- t innovations, what is the VaR of your position for the next trading day? [The 99% quantile of t_6 is 3.14.]
4. A GJR model is also fitted to the daily log returns of Boeing stock. Write down the fitted model. Is the model adequate? Why?
5. Based on the fitted GJR model and ignoring the constant term of the volatility equation, compute the leverage impact $\frac{\sigma_t^2(\epsilon=-3)}{\sigma_t^2(\epsilon=3)}$, where ϵ is the standardized innovation.

Problem D. (20 points) To study the direction of price movement of FedEx stock, we define the dependent variable y_t as

$$y_t = \begin{cases} 1 & \text{if } r_t > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where r_t is the daily simple return of the stock. The independent variables are r_{t-2} and r_{t-3} . [We started with five lags, but only kept those that are statistical significant.] Two methods are discussed in class to model the direction of price movement. The first method is the neural network and the second method is linear logistic regression. For the neural network, we use a simple 2-2-1 network with input variables r_{t-2} and r_{t-3} . Computer output is attached. Answer the following questions.

1. Write down the model for the two nodes in the hidden layer.
2. Write down the model for the output node.
3. Write down the model for the linear logistic regression.
4. Suppose $x_{n-2} = -0.018$ and $x_{n-1} = 0.015$. Based on the fitted logistic regression, what is $P(y_{n+1} = 1)$?
5. Consider the fitted values of the neural network and the logistic regression. Let “dif” be the difference between the fitted values of the two methods. The summary statistics of “dif” are given. Is there any major difference between the two methods? Why?

Computer output. For most parts, R and S-Plus have the same output.

Problem A

**** BA log returns, in percentages, ****

```
> da=read.table("d-ba9706.txt")
> ba=log(da[,2]+1)*100
> basicStats(ba)
      round.ans..digits...6.
nobs          2516.000000
NAs            0.000000
Minimum       -19.388819
Maximum        11.000173
Mean           0.026282
Median         0.000000
Sum            66.125665
SE Mean        0.042731
LCL Mean       -0.057509
UCL Mean        0.110073
Variance        4.593997
Stdev           2.143361
Skewness       -0.624541
Kurtosis        7.543017
```

```
> Box.test(ba,lag=10,type='Ljung')
      Box-Ljung test
```

```
data:  ba
X-squared = 17.3332, df = 10, p-value = 0.06731
```

```
> Box.test(ba^2,lag=10,type='Ljung')
      Box-Ljung test
```

```
data:  ba^2
X-squared = 145.0247, df = 10, p-value < 2.2e-16
```

***** Growth Rate of U.S. Quarterly GDP *****

```
> da=read.table("q-gdpun.txt")
> gdp=da[,4]
> x=diff(gdp)
> m2=arima(x,order=c(3,0,0))
> m2
arima(x = x, order = c(3, 0, 0))
```

Coefficients:

```

          ar1      ar2      ar3  intercept
    0.3111  0.1223  -0.1131    0.0085
s.e.  0.0659  0.0685   0.0657    0.0009

```

sigma^2 estimated as 8.7e-05: log likelihood = 738.99, aic = -1467.99

```

> p1=c(1,-m2$coef[1:3])
> mm=polyroot(p1)
> mm
[1] 1.609803+1.242145i -2.138323-0.000000i 1.609803-1.242145i
> Mod(mm)
[1] 2.033320 2.138323 2.033320

```

```

*****
***** Problems B & C *****
*****

```

```

> nba=-ba
> m1=gev(nba,block=63)
> m1
$n.all
[1] 2516

```

```

$n
[1] 40

```

```

$data
[1] 3.688185 6.715528 5.046206 11.907637 4.703920 3.744230 14.783897
.....
[36] 2.831717 2.480512 3.309154 4.706016 3.320524
$block
[1] 63

```

```

$par.ests
      xi      sigma      mu
0.6198668 1.4191684 3.6360118
$par.ses
      xi      sigma      mu
0.2107701 0.2792803 0.2736889

```

```

> m2=gpd(nba,threshold=2.0)
> m2
$n
[1] 2516

```

```
$data
  [1] 2.495890 3.598994 2.076409 2.185714 2.226606 2.020271 2.307417
  .....
[316] 2.709374
```

```
$threshold
[1] 2
```

```
$n.exceed
[1] 316
```

```
$par.ests
      xi      beta
0.2146411 1.1285121
```

```
$par.ses
      xi      beta
0.06297248 0.09431986
```

```
> riskmeasures(m2,c(0.95,0.99))
      p quantile      sfall
[1,] 0.95 3.149277 4.900316
[2,] 0.99 5.792961 8.266527
>
> sba=sort(ba) % sorting
> length(ba)
[1] 2516
> 2516*0.01
[1] 25.16
> sba[24:27]
[1] -5.654919 -5.614716 -5.530132 -5.268373
>
> m3=garch0xFit(formula.mean=~arma(0,0),formula.var=~garch(1,1),series=ba)
*****
** SPECIFICATIONS **
*****
Dependent variable : X
Mean Equation : ARMA (0, 0) model.
No regressor in the mean
Variance Equation : GARCH (1, 1) model.
No regressor in the variance
The distribution is a Gauss distribution.
```

Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

	Coefficient	Std.Error	t-value	t-prob
Cst(M)	0.089082	0.036234	2.459	0.0140
Cst(V)	0.044535	0.016291	2.734	0.0063
ARCH(Alpha1)	0.058817	0.0089055	6.605	0.0000
GARCH(Beta1)	0.934182	0.010077	92.71	0.0000

No. Observations : 2516 No. Parameters : 4
Mean (Y) : 0.02628 Variance (Y) : 4.59217
Skewness (Y) : -0.62491 Kurtosis (Y) : 10.55140
Log Likelihood : -5313.728 Alpha[1]+Beta[1]: 0.99280

** FORECASTS **

Number of Forecasts: 15

Horizon	Mean	Variance
1	0.08908	1.454
2	0.08908	1.488

....

> m4=garch0xFit(formula.mean=~arma(0,0),formula.var=~garch(1,1),series=ba,cond.dist="t

** SPECIFICATIONS **

Dependent variable : X

Mean Equation : ARMA (0, 0) model.

No regressor in the mean

Variance Equation : GARCH (1, 1) model.

No regressor in the variance

The distribution is a Student distribution, with 6.02138 degrees of freedom.

Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

	Coefficient	Std.Error	t-value	t-prob
Cst(M)	0.059900	0.034174	1.753	0.0798
Cst(V)	0.053963	0.025599	2.108	0.0351
ARCH(Alpha1)	0.045461	0.012919	3.519	0.0004
GARCH(Beta1)	0.941718	0.017135	54.96	0.0000
Student(DF)	6.021384	0.64644	9.315	0.0000

Log Likelihood : -5204.023 Alpha[1]+Beta[1]: 0.98718

** FORECASTS **

Number of Forecasts: 15

Horizon	Mean	Variance
1	0.0599	1.674
2	0.0599	1.706

.....

> m5=garch0xFit(formula.mean=~arma(0,0),formula.var=~gjr(1,1),series=ba,cond.dist="t")

** SPECIFICATIONS **

Dependent variable : X

Mean Equation : ARMA (0, 0) model.

No regressor in the mean

Variance Equation : GJR (1, 1) model.

No regressor in the variance

The distribution is a Student distribution, with 6.14234 degrees of freedom.

Maximum Likelihood Estimation (Std.Errors based on Second derivatives)

	Coefficient	Std.Error	t-value	t-prob
Cst(M)	0.051306	0.034105	1.504	0.1326
Cst(V)	0.061139	0.026365	2.319	0.0205
ARCH(Alpha1)	0.027723	0.011760	2.357	0.0185
GARCH(Beta1)	0.938946	0.017489	53.69	0.0000
GJR(Gamma1)	0.037132	0.015838	2.344	0.0191
Student(DF)	6.142341	0.67201	9.140	0.0000

Log Likelihood : -5200.466

** FORECASTS **

Number of Forecasts: 15

Horizon	Mean	Variance
1	0.05131	1.688
2	0.05131	1.724

.....

** TESTS **

	Statistic	t-Test	P-Value
Skewness	-0.40616	8.3222	8.6329e-017
Excess Kurtosis	4.8449	49.655	0.00000
Jarque-Bera	2529.9	.NaN	0.00000

Information Criterium (to be minimized)

Akaike	4.138685	Shibata	4.138674
Schwarz	4.152589	Hannan-Quinn	4.143732

Q-Statistics on Standardized Residuals

Q(10) = 10.7315 [0.3788158]
Q(15) = 11.9655 [0.6816375]
Q(20) = 13.4701 [0.8563154]

H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

Q-Statistics on Squared Standardized Residuals

--> P-values adjusted by 2 degree(s) of freedom

Q(10) = 13.0545 [0.1099896]
Q(15) = 16.6970 [0.2135289]
Q(20) = 20.0996 [0.3272395]

H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

**** Problem D ****

**** FDX direction ****

> da=read.table("d-fdx9706.txt")

> fdx=da[,2]

> y=fdx[6:2516]

> x2=fdx[4:2514]

> x3=fdx[3:2513]

> ydir=ifelse(y>0,1,0)

> fdxin=cbind(x2,x3)

> m2=nnet(fdxin,ydir,skip=T,linout=F,size=2)

weights: 11

....

converged

> summary(m2)

a 2-2-1 network with 11 weights

options were - skip-layer connections

b->h1 i1->h1 i2->h1

1.34 -1.13 -0.75

b->h2 i1->h2 i2->h2

3.10 0.12 0.07

b->o h1->o h2->o i1->o i2->o

-0.71 1.06 -0.17 -5.61 -3.94

> pfit=predict(m2,fdxin)

>

```

> mm=glm(ydir~x2+x3,family=binomial)
> summary(mm)
Call:
glm(formula = ydir ~ x2 + x3, family = binomial)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.03074    0.04009  -0.767  0.44318
          x2    -5.79652    1.91677  -3.024  0.00249 **
          x3    -4.10553    1.90499  -2.155  0.03115 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> fit=mm$fitted.values

> dif=fit-pfit
> basicStats(dif)
      round.ans..digits...6.
nobs          2511.000000
NAs            0.000000
Minimum        -0.001046
Maximum         0.001156
Mean           -0.000040
Median         -0.000039
Sum            -0.100122
SE Mean         0.000003
LCL Mean       -0.000046
UCL Mean       -0.000033
Variance        0.000000
Stdev           0.000165
Skewness        0.040515
Kurtosis        3.693464

```