

Solutions to the Sample Final Exam

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

1. Describe two approaches to estimate volatility of an asset returns.

Answer: Any two of (a) Econometric modelling, (b) High-frequency data, (c) Implied volatility of options data

2. If the price P_t follows a geometric Brownian motion, then what is the stochastic diffusion equation for $G(P_t) = P_t^2$?

Answer: Using the Ito's lemma with $\partial G(P_t)/\partial P_t = 2P_t$ and $\partial G(P_t)/\partial t = 0$ and $\partial^2 G(P_t)/\partial P_t^2 = 2$, we obtain $dP_t^2 = (2\mu + \sigma^2)P_t^2 dt + 2\sigma P_t^2 dw_t$.

3. Consider a European call option of a nondivided-paying stock. Suppose that $P_t = \$20$, $K = \$18$, $r = 6\%$ per annum, and $T - t = 0.5$ year. If the price of the call option is $\$2.10$, what opportunities are there for an arbitrageur?

Answer: The theoretical minimum of call option price is $c_t \geq P_t - Ke^{-r(T-t)}$. However, for the given information, $P_t - Ke^{-r(T-t)} = 20 - 18e^{-0.03} = \2.53 and $c_t = \$2.10$. Consequently, an arbitrageur can buy a call option and short one share of the stock to obtain $20 - 2.10 = \$17.90$. Deposit the cash in the risk-free interest rate of 0.06 per annum for 6 months resulting in a cash value of $\$17.90e^{0.03} = \18.45 . Exercise the option (pay the Strike price $\$18$) to settle the short position to net the difference $\$0.45$.

4. Consider a 4-3-1 neural network for a stock return r_t . Denote the input as $r_{t-1}, r_{t-2}, r_{t-3}$ and r_{t-4} , and the hidden nodes as h_1, h_2 , and h_3 . Assume that the network uses a linear activation function for the output r_t and uses the skip layer. Write down the function for the output node.

Answer: $o_t = \alpha_0 + \sum_{i=1}^4 \alpha_i r_{t-i} + \sum_{j=1}^3 \beta_j h_{jt}$.

5. Suppose that you hold a long position on IBM stock worth $\$1$ million dollars. Suppose further that the daily log returns of IBM stock follow a GARCH(1,1) model with Gaussian innovations

$$r_t = 0.001 + a_t, \quad a_t = \sigma_t \epsilon_t$$
$$\sigma_t^2 = 0.001 + 0.08a_{t-1}^2 + 0.91\sigma_{t-1}^2.$$

It is also known that $r_T = -0.013$ and $\sigma_T^2 = 0.00033$. Calculate the 1-day ahead VaR.

Answer: $a_T = r_T - 0.001 = -0.014$ and $\sigma_{T+1}^2 = 0.001 + 0.08(-0.014)^2 + 0.91(0.00033) = 0.00132$. Also, the prediction for the return is 0.001. The VaR of the log return is $(0.001 - 2.326 * \sqrt{0.00132}) = -0.0835078$, where the negative sign signifies a loss. The VaR of the position is $\$1000000 * 0.0835078 = \83507.8 .

6. Consider, again, the log return of IBM stock of the prior question. Assume that all the information remains the same except that the innovation is a standardized Student- t distribution with 5 degrees of freedom. What is the corresponding VaR for the next trading day?

Answer: The 1% quantile of the log return is $0.001 - 3.365/\text{sqrt}(5/3) \times \sqrt{0.00132} = -0.0936995$. Consequently, the VaR is $1000000 \times 0.0936995 = \$93,699.5$

7. Consider, again, the daily log returns of IBM stock. Suppose that you hold a short position on the stock worth \$1 million dollars. In addition, you are given $\alpha_n = 0.931$, $\beta_n = 2.184$, and $k_n = -0.168$ when the extreme value theory is applied with the subgroup size $n = 63$ days. What is the VaR of your position for the next trading day? What is the VaR of your position for the next 10 trading days?

Answer: Using (7.30) of the text, we have $\text{VaR} = 2.184 + \frac{0.931}{-0.168} \{1 - [-63 \ln(0.99)]^{-0.168}\} = 2.626241$. VaR is $1000000 \times 2.626241/100 = \$26,262.41$. The VaR for the next ten trading days is $(10)^{-0.168} \times 26262.41 = \$38,666.47$. **The estimates were obtained using percentage log returns so that the VaR 2.616241 is for percentage log returns.**

8. Consider a bi-variate time series $x_t = (x_{1t}, x_{2t})'$. Define the lag-1 autocovariance matrix as $\Gamma_1 = \text{Cov}(x_t, x_{t-1})$. What is the meaning of the (1, 2)th element Γ_{12} of the matrix Γ_1 ?

Answer: The linear dependence of x_{1t} on $x_{2,t-1}$.

9. Give a method to test for ARCH effect?

Answer: Let a_t be the residual of the mean equation. Use the linear regression $a_t^2 = \beta_0 + \sum_{i=1}^m \beta_i a_{t-i}^2 + e_t$ and consider the null hypothesis $H_o : \beta_1 = \beta_2 = \dots = \beta_m = 0$ vs $H_a : \beta_i \neq 0$ for some $1 \leq i \leq m$. The partial F test of the regression is the test statistic.

10. Describe two features of high-frequency financial data.

Answer: Any two of (a) irregular time intervals, (b) high excess kurtosis, (c) diurnal pattern, and (d) large sample size.

11. What is the impact of bid and ask bounce on the lag-2 autocorrelation function of a high-frequency stock return series?

Answer: None, because bid-ask bounce only affects lag-1 serial correlation.

12. Consider two asset price series P_{1t} and P_{2t} . Define the concept that the two price series are co-integrated of order 1.

Answer: Both P_{1t} and P_{2t} series have 1 unit root, but there is a non-trivial linear combination $w_t = \alpha_1 P_{1t} + \alpha_2 P_{2t}$ that is weakly stationary and invertible.

13. Give two weaknesses of using the empirical quantile to calculate value at risk?

Answer: (a) Assume future loss cannot exceed that of the past, (b) ignore the change in business and economic condition.

14. Suppose that the mean and standard deviation of the daily log returns, in percentages, of a stock are 0.04 and 1.65, respectively. What are the mean and standard deviation of the corresponding simple daily returns?

Answer: The mean and standard deviation of the log return are $\mu = .04/100$ and $\sigma = 1.65/100$. Using Eq. (1.17) of the text, $E(R_t) = \exp[(.04/100) + \frac{(1.65/100)^2}{2}] - 1 = 0.00054$ $\text{Var}(R_t) = \exp[(2 \times 0.04/100) + (1.65/100)^2][\exp(1.65/100)^2 - 1] = 0.000273$ so that the standard error is 0.0165.

15. Give a reason that indicates the U.S. monthly unemployment rate is not a linear time series.

Answer: The time series is not time reversible. That is, the series increases quickly and decays slowly.

Problem B. (30 pts) Consider the daily log returns, in percentages, of the Intel stock from January 1975 to December 2006. Suppose that you hold a long position of the stock valued at \$2 million dollars, and your friend, Jason, holds a short position of the stock valued at \$1 million dollars. Use the output (the same for both S-Plus and R) to answer the following questions.

1. For your position, write down the three parameter estimates of the generalized extreme value distribution when the block size of 63 is used, including their standard errors.

Answer: $-k = \xi = 0.1728(.0632)$, $\sigma = 2.197(0.167)$, and $\beta = \mu = 5.150(0.215)$, where k and β are the notation used in the text, and the number in parentheses is standard error.

2. What is the return level of your position if 12 blocks are used? What is the associated 95% confidence interval?

Answer: Based on the output, the return level for 12 blocks is 11.822, and the c.i. is (10.501, 13.827).

3. What is the VaR of your position for the next trading day? What is the VaR of your position for the next 10 trading days?

Answer: Since the output is based on negative returns and the program uses the right-hand tail, you can apply Eq. (7.30) directly to compute your VaR. For the percentage log returns, the VaR is $5.15 - (2.197/.1728)[1 - (-63 \ln(.99))^{-.1728}] = 6.195$. The VaR for your position is $\$2000000 \times 6.195/100 = \$123,895$. For 10 trading days, the VaR is $(10)^{.1728} \times 123,895 = \$184,440$.

4. Consider Jason's position. Perform a hypothesis testing to justify that the tail distribution of the daily log returns of Intel stock is heavier than that of a normal distribution.

Answer: For Jason's position, use the log returns. The estimate of the shape parameter is 0.024 with standard error 0.044. The t-ratio is $(0.024/0.044) = 0.55$, which is insignificant at the 5% level. Thus, for Jason's position, the data fail to support the heavy tail assumption.

5. Turn to the concept of peaks over threshold. What are the parameter estimates and their standard errors for Jason's position?

Answer: (Sorry, the detail of the output was not shown in the attached output.) With threshold 5.0, the estimates(standard error) are $\xi = 0.093(0.058)$ and $\beta = 1.564(0.130)$.

6. What is the VaR for Jason's position for the next trading day based on the generalized Pareto distribution? What is the expected loss?

Answer: The VaR is $1000000 \times 7.068/100 = \70680 and the expected loss is $\$90060$.

Problem C. (20 pts) Again, consider the daily log returns, in percentages, of Intel stock from January 1975 to December 2006. In addition, your position remains the same.

1. Write down the fitted mean and volatility equations when a Gaussian GARCH(1,1) model is fitted.

Answer: The mean equation is

$$r_t = 0.105 + 0.048r_{t-1} - 0.027r_{t-2} - 0.021r_{t-3} + a_t, \quad a_t = \sigma_t \epsilon_t.$$

The volatility equation is

$$\sigma_t^2 = 0.108 + 0.056a_{t-1}^2 + 0.930\sigma_{t-1}^2.$$

2. The fitted AR(3)-GARCH(1,1) model has some weaknesses, but, for simplicity, we use it to calculate the VaR. What is the VaR of your position for the next trading day?

Answer: The prediction gives mean 0.0368 and variance 2.497. Thus, the VaR for the percentage log return is $\text{VaR} = 0.0368 - 2.326\sqrt{2.497} = -3.6387$, where the negative sign denotes loss. The VaR for your position is $\text{VaR} = \$2000000 \times 3.6387/100 = \72774 .

3. Write down the fitted GJR model for the log return series. Is the leverage effect significant? Why?

Answer: The fitted GJR model is as follows:

Mean equation:

$$r_t = 0.085 + 0.045r_{t-1} - 0.036r_{t-2} - 0.017r_{t-3} + a_t, \quad a_t = \sigma_t \epsilon_t.$$

Volatility equation:

$$\sigma_t^2 = 0.050 + (0.031 + 0.015N_{t-1})a_{t-1}^2 + 0.954\sigma_{t-1}^2,$$

where $N_{t-1} = 1$ if $a_{t-1} < 0$, and $= 0$, otherwise. The leverage parameter is significant at the 5% level because its p-value is 0.031.

4. Based on the fitted GJR model, what is the VaR of your position for the next trading day?

Answer: From the forecast, the 1% quantile of the log return is $0.0246 - 2.326 \times \sqrt{2.398} = -3.577$ so that the VaR for the percentage log return is 3.577. The VaR for your position is $2000000 \times 3.577/100 = \71540 .

Problem D. (20 pts) Consider the TAQ trade data for the IBM in January 2007. There are more than 588,000 transactions within the regular trading hours. For simplicity, we only analyze the first 3000 transactions and employ the ADS model for the price change. In particular, we use logistic regression to estimate the following models

$$\text{logit}(p_i) = \gamma_0 + \gamma_1 A_{i-1} + \gamma_2 D_{i-1},$$

$$\text{logit}(w_i) = \beta_0 + \beta_1 D_{i-1},$$

where $p_i = P(A_i = 1 | F_{i-1})$, $w_i = P(D_i = 1 | A_i = 1, F_{i-1})$ with F_{i-1} denotes the information available at the $(i - 1)$ th transaction (inclusive). Answer the following questions.

1. What is the fitted logistic regression model for $p_i = P(A_i = 1 | F_{i-1})$?

Answer: $p_i = \frac{\exp[-0.157 + 0.845A_{i-1} - 0.119D_{i-1}]}{1 + \exp[-0.157 + 0.845A_{i-1} - 0.119D_{i-1}]}$. The estimates are all significant at the 5% level.

2. Calculate the probability p_i if $A_{i-1} = 0$? What is the probability p_i if $A_{i-1} = D_{i-1} = 1$? What is the probability p_i if $A_{i-1} = 1$, but $D_{i-1} = -1$? fitted model?

Answer: When $A_{i-1} = 0$, $D_{i-1} = 0$ as well because there is no direction. In this case, $p_i = \exp(-.157)/(1 + \exp(-.157)) = 0.460$. When $A_{i-1} = D_{i-1} = 1$, $p_i = 0.639$. When $A_{i-1} = 1$, but $D_{i-1} = -1$, we have $p_i = 0.691$.

3. What is the fitted logistic regression for $w_i = P(D_i = 1 | A_i = 1, F_{i-1})$?

Answer: $w_i = \frac{\exp(0.203 - 0.596D_{i-1})}{1 + \exp(0.203 - 0.596D_{i-1})}$. The estimates are highly significant.

4. What is the probability w_i when $D_{i-1} = 1$ and $A_{i-1} = 1$? What is the probability w_i when $D_{i-1} = -1$ and $A_{i-1} = 1$? Interpret the results.

Answer: When $D_{i-1} = 1$ and $A_{i-1} = 1$, $w_i = \exp(0.203 - .596)/(1 + \exp(0.203 - .596)) = .403$. When $D_{i-1} = -1$ and $A_{i-1} = 1$, $w_i = 0.690$. The result states that when two consecutive price changes occur, the bid and ask bounce effect continues to apply.

R and S-Plus output. Note that for Problems B and D, the command and output are the same for both R and S-Plus.

Problem B

```
> library(evir)
> da=read.table("d-intc7506.txt")
> dim(da)
[1] 8078    2
```

```
> intc=log(da[,2]+1)*100
```

```
*** Analysis of the negative daily log returns.
```

```
> nintc=-intc
```

```
> m1=gev(nintc,block=63)
```

```
> m1
```

```
$n.all
```

```
[1] 8078
```

```
$n
```

```
[1] 129
```

```
$data
```

```
[1] 6.001543 5.841332 9.289410 8.557968 5.615774 7.203320 4.098868
[8] 6.197115 5.430836 9.824147 6.155631 4.520659 4.571938 4.465224
[15] 3.687147 4.854969 2.703210 5.188294 3.652914 4.000982 5.284186
[22] 3.905273 3.932312 10.627204 10.202165 4.094700 6.197115 6.284396
[29] 6.351505 5.469909 6.160948 8.003937 6.341915 3.921911 7.311923
[36] 5.716312 8.369688 10.008556 7.695664 8.700774 5.959078 6.700556
[43] 5.358012 5.588276 5.716312 5.406555 8.700774 4.879116 5.518507
[50] 7.218367 4.609630 21.357696 9.531118 5.909205 4.989427 15.776554
[57] 5.556557 7.006845 8.303384 6.478388 3.174870 3.298818 5.653861
[64] 7.941124 11.838947 7.809128 7.165713 14.581221 5.437171 11.315750
[71] 3.732811 3.089229 4.726982 11.806307 4.358623 7.001482 7.324834
[78] 6.270554 3.306053 3.805499 3.858492 4.750050 11.565780 6.471987
[85] 10.636102 5.690903 4.741661 3.393948 4.658859 7.785344 3.989534
[92] 6.787207 13.445754 5.156702 7.849141 4.029089 6.393071 6.975745
[99] 5.424501 6.630007 6.006853 9.761283 5.919814 24.888453 12.563125
[106] 12.178353 7.181829 9.996399 6.436787 20.480032 9.173159 19.892666
[113] 5.363287 5.827551 4.004105 4.492415 4.413999 2.644665 11.159100
[120] 3.302952 2.650826 2.633371 4.524844 3.684035 12.155765 2.482562
[127] 7.791830 3.341202 1.434236
```

```
$block
```

```
[1] 63
```

```
$par.ests
```

```
      xi      sigma      mu
0.1727636 2.1966596 5.1501344
```

```
$par.ses
```

```
      xi      sigma      mu
0.06318266 0.16721612 0.21539919
```

```
$nllh.final
```

```
[1] 317.9296
```

```
> rlnintc=rlevel.gev(m1,k.blocks=12)
```

```
> rlnintc
```

```
[1] 10.50057 11.82242 13.82748
```

```
>
```

```
> m3=gpd(nintc,threshold=5.0)
```

```
> m3var=riskmeasures(m3,c(0.95,0.99,0.999))
```

```
> m3var
```

```
      p  quantile  sfall
[1,] 0.950  4.238603  6.111836
[2,] 0.990  6.979415  9.924178
[3,] 0.999 13.834940 19.459890
```

```
>
```

```
*** Percentage log returns
```

```
> m2=gev(intc,block=63)
```

```
> m2
```

```
$n.all
```

```
[1] 8078
```

```
$n
```

```
[1] 129
```

```
$data
```

```
[1] 12.107539  6.227972 11.739185  7.979960  7.313893  7.336198  4.938998
[8]  6.196960  3.620656  4.908536  4.951371  4.233126  4.908536  7.371504
[15]  5.984314  6.798570  5.855189  4.256128  7.174375  3.681397  5.092123
[22]  5.835381  6.130204  6.879819  5.862733  5.696617  7.232066  7.243228
[29]  9.776172  6.769603  5.513192  7.898677  6.025750 13.306128  6.771472
[36]  5.481010  8.576477  6.560046  6.595627  8.536085  7.480130  9.266122
[43]  5.941919  5.587926  5.623855  7.410512  8.223415  5.989965 10.616020
[50]  6.569410  6.379137 23.410723  6.454165  4.839961  5.505621  6.297480
```

```

[57] 6.252399 4.406473 4.899014 5.715508 5.781598 5.160525 7.946716
[64] 6.062462 6.794833 9.051689 5.599273 8.701221 8.552611 5.551039
[71] 6.584392 5.317103 8.046416 6.303114 7.456001 4.016258 4.000886
[78] 4.672136 3.491337 3.363785 5.775934 4.182309 6.627451 6.062462
[85] 5.315207 6.878885 4.429436 6.153715 5.955110 5.387247 8.829472
[92] 12.849875 5.085470 6.289968 6.541314 7.167860 7.472707 5.336066
[99] 7.190197 7.129688 12.379161 10.716688 7.955028 9.367248 9.722652
[106] 18.334603 7.017883 9.907581 8.168136 10.275512 9.608261 8.987727
[113] 6.946076 6.858345 5.329429 3.954758 3.500028 4.639683 3.274789
[120] 5.149127 3.420817 3.042251 2.674903 3.540575 2.414612 3.028669
[127] 3.983591 4.104597 1.233363

```

```
$block
```

```
[1] 63
```

```
$par.ests
```

```

      xi      sigma      mu
0.02351354 2.01664562 5.46856444

```

```
$par.ses
```

```

      xi      sigma      mu
0.04406089 0.13575162 0.19314159

```

```
$converged
```

```
[1] 0
```

```
$nllh.final
```

```
[1] 294.0007
```

```
> rlintc=rlevel.gev(m2,k.blocks=12)
```

```
> rlintc
```

```
[1] 9.70545 10.53674 11.66984
```

```
> m4=gpd(intc,threshold=5)
```

```
> riskmeasures(m4,c(0.95,0.99,0.999))
```

```

      p  quantile  sfall
[1,] 0.950  4.441932  6.109701
[2,] 0.990  7.067958  9.005676
[3,] 0.999 11.580814 13.982443

```

```
>
```

Problem C

```
> g1=garch0xFit(formula.mean=~arma(3,0),formula.var=~garch(1,1),series=intc)
```

```
*****
```

```
** SPECIFICATIONS **
```

```
*****
```

```
Dependent variable : X
```

```
Mean Equation : ARMA (3, 0) model.
```

```
No regressor in the mean
```

```
Variance Equation : GARCH (1, 1) model.
```

```
No regressor in the variance
```

```
The distribution is a Gauss distribution.
```

```
Strong convergence using numerical derivatives
```

```
Log-likelihood = -19000.2
```

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

| | Coefficient | Std.Error | t-value | t-prob |
|--------------|-------------|-----------|---------|--------|
| Cst(M) | 0.105329 | 0.026402 | 3.989 | 0.0001 |
| AR(1) | 0.048030 | 0.011823 | 4.062 | 0.0000 |
| AR(2) | -0.027159 | 0.011747 | -2.312 | 0.0208 |
| AR(3) | -0.020822 | 0.011605 | -1.794 | 0.0728 |
| Cst(V) | 0.108072 | 0.022343 | 4.837 | 0.0000 |
| ARCH(Alpha1) | 0.056183 | 0.0063166 | 8.895 | 0.0000 |
| GARCH(Beta1) | 0.929879 | 0.0083602 | 111.2 | 0.0000 |

| | | | |
|--------------------|------------|-------------------|---------|
| No. Observations : | 8078 | No. Parameters : | 7 |
| Mean (Y) : | 0.07774 | Variance (Y) : | 7.42717 |
| Skewness (Y) : | -0.24997 | Kurtosis (Y) : | 8.00171 |
| Log Likelihood : | -19000.171 | Alpha[1]+Beta[1]: | 0.98586 |

```
*****
```

```
** FORECASTS **
```

```
*****
```

```
Number of Forecasts: 15
```

| Horizon | Mean | Variance |
|---------|---------|----------|
| 1 | 0.03681 | 2.497 |
| 2 | 0.1278 | 2.57 |
| 3 | 0.1279 | 2.642 |
| 4 | 0.1072 | 2.712 |

```
*****
```

```
** TESTS **
```

```
*****
```

| Statistic | t-Test | P-Value |
|-----------|--------|---------|
|-----------|--------|---------|

| | | | |
|-----------------|----------|--------|-------------|
| Skewness | -0.32142 | 11.796 | 4.1080e-032 |
| Excess Kurtosis | 3.0273 | 55.557 | 0.00000 |
| Jarque-Bera | 3223.8 | .NaN | 0.00000 |

Information Criterium (to be minimized)

| | | | |
|---------|----------|--------------|----------|
| Akaike | 4.705910 | Shibata | 4.705909 |
| Schwarz | 4.711973 | Hannan-Quinn | 4.707984 |

Q-Statistics on Standardized Residuals

--> P-values adjusted by 3 degree(s) of freedom

Q(10) = 9.69811 [0.2063363]

Q(15) = 27.5535 [0.0064264]

Q(20) = 36.5015 [0.0039292]

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) = 16.4355 [0.0365538]

Q(15) = 19.6972 [0.1030206]

Q(20) = 21.0376 [0.2775194]

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

> g2=garch0xFit(formula.mean=~arma(3,0),formula.var=~gjr(1,1),series=intc,cond.dist="t

** SPECIFICATIONS **

Dependent variable : X

Mean Equation : ARMA (3, 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 7.16981 degrees of freedom.

Strong convergence using numerical derivatives

Log-likelihood = -18776.3

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

| | Coefficient | Std.Error | t-value | t-prob |
|--------|-------------|-----------|---------|--------|
| Cst(M) | 0.085370 | 0.024758 | 3.448 | 0.0006 |
| AR(1) | 0.044724 | 0.011089 | 4.033 | 0.0001 |
| AR(2) | -0.035831 | 0.011076 | -3.235 | 0.0012 |
| AR(3) | -0.016662 | 0.011080 | -1.504 | 0.1327 |

| | | | | |
|--------------|----------|-----------|-------|--------|
| Cst(V) | 0.050049 | 0.015980 | 3.132 | 0.0017 |
| ARCH(Alpha1) | 0.031275 | 0.0051149 | 6.114 | 0.0000 |
| GARCH(Beta1) | 0.954126 | 0.0072294 | 132.0 | 0.0000 |
| GJR(Gamma1) | 0.015316 | 0.0070982 | 2.158 | 0.0310 |
| Student(DF) | 7.169807 | 0.50377 | 14.23 | 0.0000 |

No. Observations : 8078 No. Parameters : 9
Mean (Y) : 0.07774 Variance (Y) : 7.42717
Skewness (Y) : -0.24997 Kurtosis (Y) : 8.00171
Log Likelihood :-18776.302

** FORECASTS **

Number of Forecasts: 15

| Horizon | Mean | Variance |
|---------|---------|----------|
| 1 | 0.02456 | 2.398 |
| 2 | 0.1155 | 2.431 |
| 3 | 0.1043 | 2.464 |
| 4 | 0.08615 | 2.497 |

** TESTS **

| | Statistic | t-Test | P-Value |
|-----------------|-----------|--------|-------------|
| Skewness | -0.32243 | 11.833 | 2.6326e-032 |
| Excess Kurtosis | 3.2594 | 59.817 | 0.00000 |
| Jarque-Bera | 3715.8 | .NaN | 0.00000 |

Information Criterium (to be minimized)

| | | | |
|---------|----------|--------------|----------|
| Akaike | 4.650979 | Shibata | 4.650976 |
| Schwarz | 4.658774 | Hannan-Quinn | 4.653645 |

Q-Statistics on Standardized Residuals

--> P-values adjusted by 3 degree(s) of freedom

Q(10) = 10.8521 [0.1452023]

Q(15) = 30.0892 [0.0027073]

Q(20) = 39.3867 [0.0015783]

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) = 21.2746 [0.0064526]

```

Q( 15) = 23.7361 [0.0336462]
Q( 20) = 25.2054 [0.1193622]
H0 : No serial correlation ==> Accept H0 when prob. is High [Q < Chisq(lag)]

```

Problem D

```

> da=read.table("taq-ibmjan07.txt")
> dim(da)
[1] 3000    3
> a=da[2:3000,1]
> am1=da[1:2999,1]
> dm1=da[1:2999,2]
> m1=glm(a~am1+dm1,family=binomial)
> m1

```

Call: glm(formula = a ~ am1 + dm1, family = binomial)

Coefficients:

| | | |
|-------------|--------|---------|
| (Intercept) | am1 | dm1 |
| -0.1566 | 0.8451 | -0.1188 |

Degrees of Freedom: 2998 Total (i.e. Null); 2996 Residual

Null Deviance: 4085

Residual Deviance: 3957 AIC: 3963

> summary(m1)

Call:

glm(formula = a ~ am1 + dm1, family = binomial)

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|---------|---------|--------|--------|--------|
| -1.5337 | -1.1117 | 0.8589 | 0.9469 | 1.2446 |

Coefficients:

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|----------|------------|---------|-------------|
| (Intercept) | -0.15659 | 0.05636 | -2.778 | 0.00546 ** |
| am1 | 0.84506 | 0.07622 | 11.087 | < 2e-16 *** |
| dm1 | -0.11882 | 0.05131 | -2.316 | 0.02058 * |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 4085.1 on 2998 degrees of freedom
Residual deviance: 3957.3 on 2996 degrees of freedom
AIC: 3963.3

Number of Fisher Scoring iterations: 4

```
> cnt=0                                ** The following steps pick up D(i) when A(i).ne.0.
> for (i in 2:3000){
+ if(da[i,1]==1){
+ cnt=cnt+1
+ d[cnt]=da[i,2]
+ dm1[cnt]=da[(i-1),2]
+ }
+ }
> cnt                                  *** Sample size for D(i) logistic regression.
[1] 1732
> d=d[1:1732]
> dm1=dm1[1:1732]
> d=ifelse(d>0.5,1,0)  ** Code -1 to zero for negative D(i) values.

> m2=glm(d~dm1,family=binomial)
> summary(m2)
```

Call:

```
glm(formula = d ~ dm1, family = binomial)
```

Deviance Residuals:

| Min | 1Q | Median | 3Q | Max |
|--------|--------|--------|-------|-------|
| -1.530 | -1.265 | 0.862 | 1.093 | 1.348 |

Coefficients:

| | Estimate | Std. Error | z value | Pr(> z) |
|-------------|----------|------------|---------|--------------|
| (Intercept) | 0.20289 | 0.04980 | 4.074 | 4.61e-05 *** |
| dm1 | -0.59561 | 0.06204 | -9.600 | < 2e-16 *** |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 2388.1 on 1731 degrees of freedom
Residual deviance: 2291.6 on 1730 degrees of freedom
AIC: 2295.6