ARCH/GARCH Models
Risk Management

- Risk: the quantifiable likelihood of loss or less-than-expected returns.
- In recent decades the field of financial risk management has undergone explosive development.
- Risk management has been described as “one of the most important innovations of the 20th century”.
- But risk management is not something new.
José (Joseph) De La Vega was a Jewish merchant and poet residing in 17\textsuperscript{th} century Amsterdam.

There was a discussion between a lawyer, a trader and a philosopher in his book \textit{Confusion of Confusions}.

Their discussion contains what we now recognize as European options and a description of their use for risk management.
Risk Management

- There are three major risk types:
  - **market risk**: the risk of a change in the value of a financial position due to changes in the value of the underlying assets.
  - **credit risk**: the risk of not receiving promised repayments on outstanding investments.
  - **operational risk**: the risk of losses resulting from inadequate or failed internal processes, people and systems, or from external events.
Risk Management

- VaR (Value at Risk), introduced by JPMorgan in the 1980s, is probably the most widely used risk measure in financial institutions.
Risk Management

- Given some confidence level $\alpha \in (0,1)$. The VaR of our portfolio at the confidence level $\alpha$ is given by the smallest value $l$ such that the probability that the loss $L$ exceeds $l$ is no larger than $\alpha$. 

![Value at Risk Diagram](image)
The steps to calculate VaR

- Market position
- Volatility \( \sigma \)
- Days to be forecasted
- Level of confidence
- Report of potential loss
The success of VaR

Is a result of the method used to estimate the risk

The certainty of the report depends upon the type of model used to compute the volatility on which these forecast is based
Volatility

"What a week!"
Modeling Volatility with ARCH & GARCH

- In 1982, Robert Engle developed the autoregressive conditional heteroskedasticity (ARCH) models to model the time-varying volatility often observed in economical time series data. For this contribution, he won the 2003 Nobel Prize in Economics (*Clive Granger shared the prize for co-integration [http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2003/press.html]).

- ARCH models assume the variance of the current error term or innovation to be a function of the actual sizes of the previous time periods' error terms: often the variance is related to the squares of the previous innovations.

- In 1986, his doctoral student Tim Bollerslev developed the generalized ARCH models abbreviated as GARCH.
Volatility of Merval Index modelling with Garch (1,1)
Time series

Robert F. Engle (1942 -), currently teaches at NYU, won Nobel Memorial Prize in Economic Sciences in 2003 for developing ARCH.

Tim Bollerslev (1958 -), currently teaches at Duke University, invented GARCH model, was a student of Robert at UCSD.
Time series-ARCH

(*Note, $Z_t$ can be other white noise, no need to be Gaussian)

- Let $Z_t$ be $N(0,1)$. The process $X_t$ is an ARCH(q) process if it is stationary and if it satisfies, for all $t$ and some strictly positive-valued process $\sigma_t$, the equations

$$X_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2$$

- Where $\alpha_0 > 0$ and $\alpha_i \geq 0$, $i = 1, \ldots, q$.

- Note: $X_t$ is usually the error term in a time series regression model!
Time series-ARCH

- ARCH(q) has some useful properties. For simplicity, we will show them in ARCH(1).

- Without loss of generality, let a ARCH(1) process be represented by

\[ X_t = Z_t \alpha_0 + \alpha_1 X_{t-1}^2 \]

- Conditional Mean

\[ E(X_t | I_{t-1}) = E \left( Z_t \alpha_0 + \alpha_1 X_{t-1}^2 | I_{t-1} \right) = E(Z_t | I_{t-1}) \alpha_0 + \alpha_1 X_{t-1}^2 = 0 \]

- Unconditional Mean

\[ E(X_t) = E \left( Z_t \alpha_0 + \alpha_1 X_{t-1}^2 \right) = E(Z_t) E \left( \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2} \right) = 0 \]

- So \( X_t \) have mean zero
Time series-ARCH

- $X_t$ has conditional variance given by
  $$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2$$

- **Proof:**
  $$\text{Var}(X_t|I_{t-1}) = E(X_t^2|I_{t-1}) - [E(X_t|I_{t-1})]^2$$
  $$= E(Z_t^2|I_{t-1})(\alpha_0 + \alpha_1 X_{t-1}^2) - 0$$
  $$= \alpha_0 + \alpha_1 X_{t-1}^2$$
Time series-ARCH

- $X_t$ have unconditional variance given by $Var(X_t) = \frac{\alpha_0}{1-\alpha_1}$

- Proof 1 (use the law of total variance):

$$Var(X_t) = E(Var(X_t | I_{t-1})) + Var(E(X_t | I_{t-1}))$$
$$= E(\alpha_0 + \alpha_1 X_{t-1}^2) + Var(0)$$
$$= \alpha_0 + \alpha_1 E(X_{t-1}^2)$$
$$= \alpha_0 + \alpha_1 Var(X_{t-1})$$

Because it is stationary, $Var(X_t) = Var(X_{t-1})$.

So $Var(X_t) = \frac{\alpha_0}{1-\alpha_1}$. 
Proof 2:

Lemma: Law of Iterated Expectations

Let $\Omega_1$ and $\Omega_2$ be two sets of random variables such that $\Omega_1 \subseteq \Omega_2$. Let $Y$ be a scalar random variable. Then

$$E(Y|\Omega_1) = E[E(Y|\Omega_2)|\Omega_1]$$

$$Var(X_t|I_{t-2}) = E(X_t^2|I_{t-2}) = E[E(X_t^2|I_{t-1})|I_{t-2}]$$

$$= \alpha_0 + \alpha_1 E(X_{t-1}^2|I_{t-2}) = \alpha_0 + \alpha_0 \alpha_1 + \alpha_1^2 X_{t-2}^2$$

$$Var(X_t|I_{t-3}) = \alpha_0 + \alpha_0 \alpha_1 + \alpha_0 \alpha_1^2 + \alpha_1^3 X_{t-3}^2$$

$$\vdots$$

$$Var(X_t) = E(X_t^2) = E(E \ldots E(X_t^2|I_{t-1}) \ldots |I_1)$$

$$= \alpha_0 (1 + \cdots + \alpha_1^{t-1}) + \alpha_1^t X_0^2$$

Since $\alpha_1 < 1$, as $t \to \infty$, $Var(X_t) = \frac{\alpha_0}{1-\alpha_1}$
Time series-ARCH

- The unconditional distribution of $X_t$ is leptokurtic, it is easy to show.

$$Kurt(X_t) = \frac{E(X_t^4)}{[E(X_t^2)]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3$$

- Proof:

$$E(X_t^4) = E(E(X_t^4|I_{t-1})) = E(\sigma_t^4E(Z_t^4|I_{t-1}))$$
$$= E(Z_t^4)E(\alpha_0 + \alpha_1X_{t-1}^2)^2$$
$$= 3(\alpha_0^2 + 2\alpha_0\alpha_1E(X_{t-1}^2) + \alpha_1^2E(X_{t-1}^4))$$

So $E(X_t^4) = 3 \frac{\alpha_0^2}{1-\alpha_1} \frac{1+\alpha_1}{1-3\alpha_1^2}$. Also, $E(X_t^2) = \frac{\alpha_0}{1-\alpha_1}$

$$Kurt(X_t) = \frac{E(X_t^4)}{[E(X_t^2)]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2}$$
Time series-ARCH

- The unconditional distribution of $X_t$ is leptokurtic,

$$Kurt(X_t) = \frac{E(X_t^4)}{[E(X_t^2)]^2} = 3 \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2} > 3$$

- The curvature is high in the middle of the distribution and tails are fatter than those of a normal distribution, which is frequently observed in financial markets.

![Graph showing mean, median, mode, and tails types](image)
Time series-ARCH

- For ARCH(1), we can rewrite it as
  \[ X_t^2 = \sigma_t^2 Z_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + V_t \]

  where
  \[ V_t = \sigma_t^2 (Z_t^2 - 1) \]

- \( E(X_t^4) \) is finite, then it is an AR(1) process for \( X_t^2 \).

- Another perspective: \( E(X_t^2 | I_{t-1}) = \alpha_0 + \alpha_1 X_{t-1}^2 \)

- The above is simply the optimal forecast of \( X_t^2 \) if it follows an AR(1) process.
Before introducing GARCH, we discuss the EWMA (exponentially weighted moving average) model

\[ \sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda)X_{t-1}^2 \]

where \( \lambda \) is a constant between 0 and 1.

The EWMA approach has the attractive feature that relatively little data need to be stored.
Time series-EWMA

- We substitute for $\sigma_{t-1}^2$, and then keep doing it for $m$ steps

$$\sigma_t^2 = (1 - \lambda) \sum_{i=1}^{m} \lambda^{i-1} X_{t-i}^2 + \lambda^m \sigma_{t-m}^2$$

- For large $m$, the term $\lambda^m \sigma_{t-m}^2$ is sufficiently small to be ignored, so it decreases exponentially.
The EWMA approach is designed to track changes in the volatility.

The value of $\lambda$ governs how responsive the estimate of the daily volatility is to the most recent daily percentage change.

For example, the RiskMetrics database, which was invented and then published by JPMorgan in 1994, uses the EWMA model with $\lambda = 0.94$ for updating daily volatility estimates.
The GARCH processes are generalized ARCH processes in the sense that the squared volatility $\sigma_t^2$ is allowed to depend on previous squared volatilities, as well as previous squared values of the process.
Time series-GARCH

(*Note, Zt can be other white noise, no need to be Gaussian)

- Let $Z_t$ be $N(0,1)$. The process $X_t$ is a GARCH(p, q) process if it is stationary and if it satisfies, for all $t$ and some strictly positive-valued process $\sigma_t$, the equations

$$X_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i X_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

- Where $\alpha_0 > 0$ and $\alpha_i \geq 0$, $i = 1, \ldots, q$, $\beta_j \geq 0$, $j = 1, \ldots, p$.

- Note: $X_t$ is usually the error term in a time series regression model!
Time series-GARCH

- GARCH models also have some important properties. Like ARCH, we show them in GARCH(1,1).

- The GARCH(1, 1) process is a covariance-stationary white noise process if and only if $\alpha_1 + \beta < 1$. The variance of the covariance-stationary process is given by $\alpha_0 / (1 - \alpha_1 - \beta)$.

- In GARCH(1,1), the distribution of $X_t$ is also mostly leptokurtic – but can be normal.

$$Kurt(X_t) = 3 \frac{1 - (\alpha_1 + \beta)^2}{1 - (\alpha_1 + \beta)^2 - 2\alpha_1^2} \geq 3$$
Time series-GARCH

- We can rewrite the GARCH(1,1) as

\[ X_t^2 = \sigma_t^2 Z_t^2 = \alpha_0 + (\alpha_1 + \beta)X_{t-1}^2 - \beta V_{t-1} + V_t \]

where

\[ V_t = \sigma_t^2 (Z_t^2 - 1) \]

- \( E(X_t^4) \) is finite, then it is an ARMA(1,1) process for \( X_t^2 \).
Time series-GARCH

- The equation for GARCH(1,1) can be rewritten as 
  
  \[ \sigma_t^2 = \gamma V_L + \alpha_0 + \alpha_1 X_{t-1}^2 + \beta \sigma_{t-1}^2 \]

- where \( \gamma + \alpha_1 + \beta = 1 \).

- The EWMA model is a special case of GARCH(1,1) where
  \( \gamma = 0, \alpha_1 = 1 - \lambda, \beta = \lambda \)
The GARCH (1,1) model recognizes that over time the variance tends to get pulled back to a long-run average level of $V_L$.

Assume we have known $\sigma_t^2 = V$, if $V > V_L$, then this expectation is negative

$$E[\sigma_{t+1}^2 - V | I_t] = E[\gamma V_L + \alpha_1 X_t^2 + \beta V - V | I_t]$$
$$< E[\gamma V + \alpha_1 X_t^2 + \beta V - V | I_t]$$
$$= E[-\alpha_1 V + \alpha_1 X_t^2 | I_t]$$
$$= -\alpha_1 V + \alpha_1 V$$
$$= 0$$
Time series-GARCH

- If $V < V_L$, then this expectation is positive

\[
E[\sigma_{t+1}^2 - V | I_t] = E[\gamma V_L + \alpha_1 X_t^2 + \beta V - V | I_t] \\
> E[\gamma V + \alpha_1 X_t^2 + \beta V - V | I_t] \\
= E[-\alpha_1 V + \alpha_1 X_t^2 | I_t] \\
= -\alpha_1 V + \alpha_1 V \\
= 0
\]

- This is called mean reversion.
Time series-(ARMA-GARCH)

• In the real world, the return processes maybe stationary, so we combine the ARMA model and the GARCH model, where we use ARMA to fit the mean and GARCH to fit the variance.

• For example, ARMA(1,1)-GARCH(1,1)

\[
X_t = \mu + \phi X_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}
\]

\[
\varepsilon_t = \sigma_t Z_t
\]

\[
\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\]
Time Series in R

Data from Starbucks Corporation (SBUX)
Program Preparation

- Packages:

  >require(quantmod): specify, build, trade and analyze quantitative financial trading strategies

  >require(forecast): methods and tools for displaying and analyzing univariate time series forecasts

  >require(urca): unit root and cointegration tests encountered in applied econometric analysis are implemented

  >require(tseries): package for time series analysis and computational finance

  >require(fGarch): environment for teaching ‘Financial Engineering and Computational Finance’
Introduction

>getSymbols('SBUX')
>chartSeries(SBUX,subset='2009::2013')
Method of Modeling

> ret = na.omit(diff(log(SBUX$SBUX.Close)))

> plot(r, main='Time plot of the daily logged return of SBUX')
KPSS test

- KPSS tests are used for testing a null hypothesis that an observable time series is stationary around a deterministic trend.
- The series is expressed as the sum of deterministic trend, random walk, and stationary error, and the test is the Lagrange multiplier test of the hypothesis that the random walk has zero variance.
- KPSS tests are intended to complement unit root tests.
KPSS test

\[ x_t = D_t \beta + u_t + \omega_t \]
\[ \omega_t = \omega_{t-1} + \varepsilon_t \]
\[ \varepsilon_t \sim WN(0, \sigma^2) \]

where \( D_t \): contains deterministic components

\( u_t \): stationary time series

\( \omega_t \): pure random walk with innovation variance
Trend Check

- KPSS test: null hypothesis: $\sigma^2 = 0$

```r
> summary(ur.kpss(r,type='mu',lags='short'))
```

- Return is a stationary around a constant, has no linear trend

```
# KPSS Unit Root Test #

Test is of type: mu with 7 lags.

Value of test-statistic is: 0.1797

Critical value for a significance level of:

10pct  5pct  2.5pct  1pct
critical values 0.347  0.463  0.574  0.739
```
ADF test

- ADF test is a test for a unit root in a time series sample.
- ADF test:
  \[ \Delta X_t = \alpha + \beta t + \gamma X_{t-1} + \delta_1 \Delta X_{t-1} + \cdots + \delta_p \Delta X_{t-p-1} + \varepsilon_t \]
  
  null hypothesis: \( \gamma = 0 \Rightarrow X_t \) has unit root

- It is an augmented version of the Dickey-Fuller test for a larger and more complicated set of time series models.

- ADF used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence.
Trend Check

- ADF Test: null hypothesis: \( X_t \) has AR unit root (nonstationary)

\[
\text{summary(}\text{ur.df(r,type='trend',lags=20,selectlags='BIC')}\text{)}
\]

- Return is a stationary time series with a drift

```
Coefficients:
            Estimate  Std. Error   t value  Pr(>|t|)  
(Intercept)  3.225e-03   1.201e-03    2.685  0.00736 **
z.lag.1     -1.063e+00   4.173e-02  -25.475  < 2e-16 ***
tt          -2.200e-06   1.634e-06   -1.346  0.17856
z.diff.lag  -1.060e-02   2.846e-02   -0.372  0.70966
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.0205 on 1233 degrees of freedom
Multiple R-squared:  0.5372,    Adjusted R-squared:  0.5361
F-statistic: 477.1 on 3 and 1233 DF,  p-value: < 2.2e-16
```
Check Seasonality

- `>par(mfrow=c(3,1))`
- `>acf(r)`
- `>pacf(r)`
- `>spec.pgram(r)`
Random Component

- Demean data \( r_1 = r - \text{mean}(r) \)

\( \texttt{acf}(r_1); \texttt{pacf}(r_1); \)
Random Component

> fit = arima(r, order = c(1, 0, 0))

> tsdiag(fit) AR(1)
Random Component

- First difference > `diffr = na.omit(diff(r))`

```r
> plot(diffr); acf(diffr); pacf(diffr);
```
Random Component

> fit1 = arima(r, order=c(0,1,1)); tsdiag(fit1);
Random Component

>fit2 = arima(r, order = c(1, 1, 1)); tsdiag(fit2); ARIMA(1, 1, 1)
Model Selection

\[ AIC = 2k - 2\ln(L) \]
where \( k \) is the number of parameters, \( L \) is likelihood

Final model:
\[ X_t = 0.0017 - 0.0724X_{t-1} + \varepsilon_t \]
Shapiro-Wilk normality test

- The Shapiro-Wilk test, proposed in 1965, calculates a W statistic that tests whether a random sample comes from a normal distribution.

- Small values of W are evidence of departure from normality and percentage points for the W statistic, obtained via Monte Carlo simulations.
Residual Test

- \( \texttt{res} = \texttt{residuals(fit)} \)

- \( \texttt{shapiro.test(res)} \)  
  *not normally distributed*

```
Shapiro-Wilk normality test

data:  res
W = 0.9356, p-value < 2.2e-16
```
Residual Test

- `par(mfrow=c(2,1))`

- `hist(res); lines(density(res))`

- `qqnorm(res); qqline(res)`
ARMA+GARCH

- ARMA(1,0)+GARCH(1,1)
  
  >summary(garchFit(~arma(1,0)+garch(1,1),r,trace=F))

- ARIMA(1,1,1)+GARCH(1,1)
  
  >summary(garchFit(~arma(1,1)+garch(1,1),data=diffr,trace=F))
Model Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(1)+intercept</td>
<td>3098</td>
</tr>
<tr>
<td>ARIMA(1,1,1)</td>
<td>3091.4</td>
</tr>
<tr>
<td>AR(1)+GARCH(1,1)</td>
<td>3209.128</td>
</tr>
<tr>
<td>ARIMA(1,1,1)+GARCH(1,1)</td>
<td>3200.077</td>
</tr>
</tbody>
</table>

\[
x_t = 0.00185 - 0.0512x_{t-1} + \sigma_t \epsilon_t
\]

\[
\sigma_t^2 = 0.0805x_{t-1}^2 - 0.9033\sigma_{t-1}^2
\]
Forecasting

Forecasts of logged return with lag of 1

- Forecasting return
- Real return
- 95% CI

logged return

Dec 02 Dec 09 Dec 16 Dec 23 Dec 30 Jan 06 Jan 13

date
Time Series in SAS

Data from Starbucks Corporation (SBUX)
SAS Procedures for Time Series

- **PROC ARIMA**

This procedure do model identification, parameter estimation and forecasting for model ARIMA(p,d,q)

\[(1 - B)^d X_t = \mu + \frac{1 - \theta_1 B - \cdots - \theta_q B^q}{1 - \phi_1 B - \cdots - \phi_p B^p} Z_t\]

where \(\mu\) is ground mean of \(X_t\) and

\[\mu(1 - \phi_1 B - \cdots - \phi_p B^p) = \mu(1 - \phi_1 - \cdots - \phi_p)\]

is the usually intercept (drift).

- does **NOT** do ARCH/GARCH
SAS Procedures for Time Series

- PROC AUTOREG

This procedure estimates and forecasts the linear regression model for time series data with an autocorrelated error or a heteroscedastic error.

\[ Y_t = X_t \beta + \nu_t \quad \text{Linear regression model} \]

\[ \nu_t = -\phi_1 \nu_{t-1} - \cdots - \phi_p \nu_{t-p} + \varepsilon_t \quad \text{Autocorrelated error} \]

\[ \varepsilon_t = \sigma_t Z_t \quad \text{Heteroscedastic error} \]

\[ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{Q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{P} \gamma_j \sigma_{t-j}^2 \]

- Independent assumption invalid \quad \text{Autocorrelated error}
- Homoscedasticity assumption invalid \quad \text{Heteroscedastic error}
import data

DATA st.return;
  INFILE "\SBUXreturn.txt"
    firstobs=2;
  INPUT Date YYMMDD10. r;
  FORMAT Date Date9.;
RUN;
PROC SGPLOT DATA=st.return;
  SERIES X=Date Y=r;
RUN;
Testing for Autocorrelation

The following statements perform the Durbin-Watson test for autocorrelation in the returns for orders 1 through 3. The DWPROB option prints the marginal significance levels (p-values) for the Durbin-Watson statistics.

```
PROC AUTOREG DATA=st.return;
   TITLE2 "AUTOREG AR Test";
   MODEL r = / METHOD=ML DW=3 DWPROB;
RUN;
```
Durbin-Watson test

- If $e_t$ is the residual associated with the observation at time $t$, then the test statistic is

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

where $T$ is the number of observations.

- Since

$$d = \frac{\sum_{t=1}^{T} e_t^2 + \sum_{t=1}^{T} e_{t-1}^2 - 2 \sum_{t=1}^{T} e_t e_{t-1}}{\sum_{t=1}^{T} e_t^2}$$

$$\approx 2 - 2 \frac{\sum_{t=1}^{T} e_t e_{t-1}}{\sum_{t=1}^{T} e_t^2} \approx 2 - 2r = 2(1 - r)$$

where $r$ is the sample autocorrelation of the residuals.
Durbin-Watson test

- Since $d \approx 2(1 - r)$
  - Positive serial correlation $0 < r < 1 \implies 0 < d < 2$
  - Negative serial correlation $-1 < r < 0 \implies 2 < d < 4$

- To test for **positive autocorrelation** at significance $\alpha$:
  - If $d < d_{L, \alpha}$, the error terms are positively autocorrelated
  - If $d > d_{U, \alpha}$, there is **no** statistical evidence

- To test for **negative autocorrelation** at significance $\alpha$:
  - If $(4 - d) < d_{L, \alpha}$, the error terms are negatively autocorrelated
  - If $(4 - d) > d_{U, \alpha}$, there is **no** statistical evidence
    $(d_{L, \alpha}$ and $d_{U, \alpha}$ are lower and upper critical values)
Testing for Autocorrelation

<table>
<thead>
<tr>
<th>Order</th>
<th>DW</th>
<th>Pr &lt; DW</th>
<th>Pr &gt; DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.1419</td>
<td>0.9941</td>
<td>0.0059</td>
</tr>
<tr>
<td>2</td>
<td>1.9685</td>
<td>0.2979</td>
<td>0.7021</td>
</tr>
<tr>
<td>3</td>
<td>2.0859</td>
<td>0.9429</td>
<td>0.0571</td>
</tr>
</tbody>
</table>

Note: Pr<DW is the p-value for testing positive autocorrelation, and Pr>DW is the p-value for testing negative autocorrelation.

- Autocorrelation correction is needed.
- Generalized Durbin-Watson tests should not be used to decide on the autoregressive order.
Stepwise Autoregression

- Once you determine that autocorrelation correction is needed, you must select the order of the autoregressive error model to use. One way to select the order of the autoregressive error model is Stepwise Autoregression.

- The following statements show the stepwise feature, using an initial order of 5:

```plaintext
PROC AUTOREG DATA=st.return;
  TITLE2 "AUTOREG (fit P) for log returns";
  MODEL r = / METHOD=ML
              NLAG=5 BACKSTEP;
RUN;
```
### Stepwise Autoregression

#### Estimates of Autocorrelations

<table>
<thead>
<tr>
<th>Lag</th>
<th>Covariance</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.000430</td>
<td>1.000000</td>
</tr>
<tr>
<td>1</td>
<td>-0.00003</td>
<td>-0.065364</td>
</tr>
<tr>
<td>2</td>
<td>8.91E-6</td>
<td>0.020711</td>
</tr>
<tr>
<td>3</td>
<td>-0.00002</td>
<td>-0.038408</td>
</tr>
<tr>
<td>4</td>
<td>-0.000010</td>
<td>-0.016767</td>
</tr>
</tbody>
</table>

#### Backward Elimination of Autoregressive Terms

| Lag | Estimate  | t Value | Pr > |t| |
|-----|-----------|---------|-------|---|
| 2   | -0.014197 | -0.50   | 0.6158|   |
| 5   | -0.014863 | -0.53   | 0.5985|   |
| 3   | 0.034743  | 1.23    | 0.2181|   |
| 4   | -0.038273 | -1.36   | 0.1743|   |

#### Estimates of Autoregressive Parameters

<table>
<thead>
<tr>
<th>Lag</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.065364</td>
<td>0.028145</td>
<td>2.32</td>
</tr>
</tbody>
</table>
Testing for Heteroscedasticity

- One of the key assumptions of the ordinary regression model is that the errors have the same variance throughout the sample. This is also called the **homoscedasticity** model. If the error variance is not constant, the data are said to be Heteroscedastic.

- The following statements use the ARCHTEST= option to test for heteroscedasticity:

```plaintext
PROC AUTOREG DATA=st.return;
   TITLE2 "AUTOREG arch Test";
   MODEL r = / METHOD=ML ARCHTEST;
RUN;
```
Testing for Heteroscedasticity

- Portmanteau Q Test

For nonlinear time series models, the portmanteau test statistic based on squared residuals is used to test for independence of the series

\[
Q(q) = T(T + 2) \sum_{i=1}^{q} \frac{r(i; \hat{e}_t^2)}{N - i}
\]

where

\[
r(i; \hat{e}_t^2) = \frac{\sum_{t=i+1}^{T} (\hat{e}_t^2 - \hat{\sigma}^2)(\hat{e}_{t-i}^2 - \hat{\sigma}^2)}{\sum_{t=1}^{T} (\hat{e}_t^2 - \hat{\sigma}^2)^2}
\]

\[
\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{e}_t^2
\]
Testing for Heteroscedasticity

- Lagrange Multiplier Test for ARCH Disturbances

Engle (1982) proposed a Lagrange multiplier test for ARCH disturbances. Engle’s Lagrange multiplier test for the qth order ARCH process is written

$$LM(q) = \frac{TW'Z(Z'Z)^{-1}Z'W}{W'W}$$

where

$$W = \left(\frac{\hat{\epsilon}_1^2}{\hat{\sigma}^2}, \ldots, \frac{\hat{\epsilon}_T^2}{\hat{\sigma}^2}\right)'$$

$$Z = \begin{bmatrix}
1 & \hat{\epsilon}_0^2 & \ldots & \hat{\epsilon}_{-q+1}^2 \\
\vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
1 & \hat{\epsilon}_{N-1}^2 & \ldots & \hat{\epsilon}_{N-q}^2
\end{bmatrix}$$

The presample values \((\hat{\epsilon}_0^2, \ldots, \hat{\epsilon}_{-q+1}^2)\) have been set to 0
## Testing for Heteroscedasticity

The $p$-values for the test statistics strongly indicate heteroscedasticity.

<table>
<thead>
<tr>
<th>Order</th>
<th>Q</th>
<th>Pr &gt; Q</th>
<th>LM</th>
<th>Pr &gt; LM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.3583</td>
<td>0.0669</td>
<td>3.3218</td>
<td>0.0684</td>
</tr>
<tr>
<td>2</td>
<td>20.6045</td>
<td>&lt;.0001</td>
<td>19.7499</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3</td>
<td>30.4729</td>
<td>&lt;.0001</td>
<td>27.3240</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4</td>
<td>32.0168</td>
<td>&lt;.0001</td>
<td>27.5946</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>5</td>
<td>40.3500</td>
<td>&lt;.0001</td>
<td>32.2671</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>6</td>
<td>45.0011</td>
<td>&lt;.0001</td>
<td>34.6383</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>7</td>
<td>53.2330</td>
<td>&lt;.0001</td>
<td>38.9149</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>8</td>
<td>74.7145</td>
<td>&lt;.0001</td>
<td>52.6820</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>
Fitting AR(p)-GARCH(P,Q)

- The following statements fit an AR(1)-GARCH model for the return r. The GARCH=(P=1,Q=1) option specifies the GARCH conditional variance model. The NLAG=1 option specifies the AR(1) error process.

```sas
PROC AUTOREG DATA=st.return;
   TITLE2 "AR=1 GARCH(1,1)";
   MODEL r = / METHOD=ML NLAG=1
       GARCH=(p=1,q=1);
   OUTPUT OUT=st.rout HT=variance
       P=yhat LCL=low95 UCL=high95;
RUN;
```
Fitting AR(1)-GARCH(1,1)

| Variable | DF | Estimate  | Standard Error | t Value | Approx Pr > |t| |
|----------|----|-----------|----------------|---------|-------------|---|
| Intercept| 1  | 0.001778  | 0.000476       | 3.74    | 0.0002      |
| AR1      | 1  | 0.0512    | 0.0330         | 1.55    | 0.1203      |
| ARCH0    | 1  | 7.8944E-6 | 1.6468E-6      | 4.79    | <.0001      |
| ARCH1    | 1  | 0.0804    | 0.009955       | 8.08    | <.0001      |
| GARCH1   | 1  | 0.9041    | 0.0110         | 81.91   | <.0001      |

\[ r_t = 0.001778 + v_t \]
\[ v_t = -0.0512v_{t-1} + \varepsilon_t, \]
\[ \varepsilon_t = \sigma_t Z_t, \]
\[ \sigma_t^2 = 7.89 \times 10^6 + 0.0804\varepsilon_{t-1}^2 + 0.9041\sigma_{t-1}^2 \]
Fitting GARCH(1,1)

Parameter Estimates

| Variable | DF | Estimate | Standard Error | t Value | Approx Pr > |t| |
|----------|----|----------|----------------|---------|--------------|-----------------------|
| Intercept| 1  | 0.001789 | 0.000500       | 3.58    | 0.0003       |
| ARCH0    | 1  | 8.2257E-6| 1.7139E-6      | 4.80    | <.0001       |
| ARCH1    | 1  | 0.0824   | 0.0103         | 8.04    | <.0001       |
| GARCH1   | 1  | 0.9014   | 0.0114         | 78.91   | <.0001       |

\[
r_t = 0.001789 + \varepsilon_t, \]

\[
\varepsilon_t = \sigma_t Z_t, \]

\[
\sigma_t^2 = 8.23 \times 10^6 + 0.0824 \varepsilon_{t-1}^2 + 0.9014 \sigma_{t-1}^2 \]
## Model Comparison

### AR(1)-GARCH(1,1)

<table>
<thead>
<tr>
<th></th>
<th>GARCH Estimates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SSE</td>
<td>0.535053</td>
<td>Observations</td>
<td>1258</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000425</td>
<td>Uncond Var</td>
<td>0.000508</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>3208.050</td>
<td>Total R-Square</td>
<td>0.0048</td>
</tr>
<tr>
<td>SBC</td>
<td>6380.415</td>
<td>AIC</td>
<td>6406.101</td>
</tr>
<tr>
<td>MAE</td>
<td>0.014387</td>
<td>AICC</td>
<td>6406.053</td>
</tr>
<tr>
<td>MAPE</td>
<td>115.6916</td>
<td>HQC</td>
<td>6396.448</td>
</tr>
<tr>
<td>Normality Test</td>
<td>1079.950</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### GARCH(1,1)

<table>
<thead>
<tr>
<th></th>
<th>GARCH Estimates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SSE</td>
<td>0.537626</td>
<td>Observation</td>
<td>1258</td>
</tr>
<tr>
<td>MSE</td>
<td>0.000427</td>
<td>Uncond Var</td>
<td>0.000508</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>3206.712</td>
<td>Total R-Square</td>
<td></td>
</tr>
<tr>
<td>SBC</td>
<td>-6384.876</td>
<td>AIC</td>
<td>6405.425</td>
</tr>
<tr>
<td>MAE</td>
<td>0.014407</td>
<td>AICC</td>
<td>6405.393</td>
</tr>
<tr>
<td>MAPE</td>
<td>113.7514</td>
<td>HQC</td>
<td>6397.702</td>
</tr>
<tr>
<td>Normality Test</td>
<td>1054.273</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Prediction
VaR (Value at Risk)
Summary

- Models for conditional variance (risk)
  - ARCH
  - EWMA
  - GARCH

- Numerical experiment
  - ARIMA+GARCH in R
  - AR+GARCH in SAS
  - VaR in SAS
This presentation was revised from my students’ presentation.

青出于蓝，而胜于蓝

"We've considered every potential risk except the risks of avoiding all risks."

HARDIN