INVESTMENT FINANCE V ECF5103

Definition

Rate of return $P_1 = P_0(1+r)$ Discrete
(one period)

i.e., $r = \frac{P1}{Po} - 1$
↑
Price relative

$P1 = Po_e{}^r$ continuous

i.e., $r = \log_e\left(\frac{P_1}{P_0}\right)$

Adjustments for changes in the basis of quotation.

Dividends

- Add back cash dividend on date share is quoted XD
- Ignores - XD date is before cash is paid to shareholders
- Income tax on dividends

Rights Issues

Adjustment ignores

- Income tax issues
- Time value of delayed cash subscription
- Option value of delayed subscription

Discrete v. Continuous

Time	**Price**		**Discrete ROR**	**Cont. ROR**
t = o	Pt = 1		*	*
1	2		+100%	+69.3%
2	1		- 50%	-69.3%
		TOTAL	+ 50%	

Discrete ROR

- Biased (upwards) Estimate of Total ROR
- Is greater than continuous ROR
- Bias depends on variance of prices in sub periods
- Why biased? because

$$\frac{\mathrm{P1-Po}}{\mathrm{Po}}+\frac{\mathrm{P2-P1}}{\mathrm{P1}}\neq\frac{\mathrm{P2-Po}}{\mathrm{Po}}$$

Continuous ROR

- Unbiased estimate because total ROR

$$Ln\frac{(P1)}{(Po)}+Ln\frac{(P2)}{P1}=\ln\left(\frac{P1}{Po}\,x\,\frac{P2}{P1}\right)$$

$$=Ln\frac{(P2)}{Po}$$

Changes in Basis of Quotation

- Dividends Cum Div (with Div)

  Ex Div (without Div)

- Capitalisation Changes
  - Rights Issues
  - Bonus Share Issues
  - Share Splits
  - Returns of Capital, etc.

- Need P1 on a <u>Comparable Basis</u> with Po!

$$P1^* = Po\ (1+r)$$

Where $P1^* = P1 + ADJ$

- Conventional adjustments (for changes in basis of share price quotation):

ADJ = CASH Div. AMT - (dividends)

$ADJ = \frac{A}{B}(XR\text{-}CASH)$ - (Rights Bonus Issues and Share Splits and Consolidations

- Derivation of adjustment $ADJ = \frac{A}{B}(XR - CASH)$

A:B at Cash = issue terms

(EG: 2:5 at $1.25)

XR = Traded value of shares EX RIGHTS

CR = What shares would have traded for were they CUM RIGHTS

Investment Outlay	=	Portfolio Worth
B(CR) + A(CASH)	=	(A+B) XR
ADJ	=	CR-XR
now CR	=	(A+B)XR-A(CASH)B
i.e., ADJ	=	as above -XR
	=	(A+B)XR-A(CASH)-B(XR)/B
and ADJ	=	$\frac{A}{B}$(XR-CASH)

Example

Swan Tele 2-5 Rights Issue Dec 2008 XR Price = \$1.20

Cash = .85

Assume last Nov was \$1.30

Q. what was ROR Nov 30 to Dec 31?

$$\text{ROR} = \frac{\text{P1} + \text{ADJ}}{\text{Po}} - 1$$

ADJ = 2/5 (1.20-.85) = .14

$$\text{i.e., ROR} = \frac{1.20+.14}{1.30} - 1$$

= 3.08% (Discrete)

Bonus issue)

) NO CASH PAYABLE

Share split)

FAMA - THE BEHAVIOUR OF STOCK MARKET RETURNS

THE MEAN

$$E(\tilde{x}) = \sum_x \tilde{x} P(x) \qquad (1)$$

Discrete

Random variable

The mean of a continuous random variable

$$E(\tilde{x}) = \int_x x P(x) dx \qquad (2)$$

Standard deviation of a discrete random variable

$$\text{Variance} = \sigma^2(\tilde{x}) = E[\tilde{x} - E(\tilde{x})]^2 = \sum_x [\tilde{x} - E(\tilde{x})]^2 P(x) \qquad (3)$$

The variance of a continuous random variable

$$\sigma^2(\tilde{x}) = E\left([\tilde{x} - E(\tilde{x})]^2\right) = \int_x [\tilde{x} - E(\tilde{x})]^2 P(\tilde{x})dx \qquad (4)$$

Standard deviation

$$\sigma(\tilde{x}) = \sqrt{\sigma^2(\tilde{x})} \qquad (5)$$

Characterisation of Normal Distributions by their Means and Standard Deviations

For any normally distributed random variable, the probability that any drawing is within one standard deviation of the mean, i.e., in the interval

$$E \le (\tilde{x}) - \sigma(x) \le x \le E(\tilde{x}) + \sigma(\tilde{x})$$

is .6826

The possibility that the random drawing is in the interval

$$E(\tilde{x}) - 2\sigma(\tilde{x}) \le x \le E(\tilde{x}) + 2\sigma(\tilde{x})$$

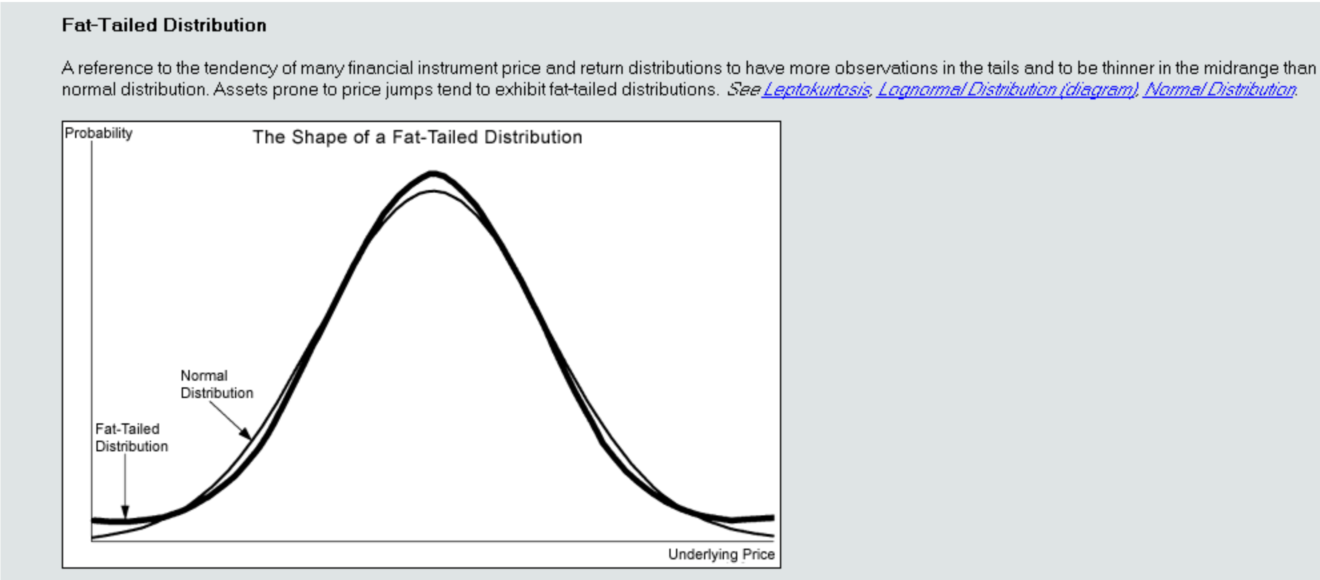
is .9550

Equivalently, for any normally distributed random variable ~, the transformed variable

$$\tilde{r} = \frac{\tilde{x} - E(\tilde{x})}{\sigma(\tilde{x})}$$

has the unit normal distribution with mean equal to 0 and standard deviation equal to 1.

Unit Normal Distribution



Sample mean

$$\overline{X} = \sum_{i=1}^{T} \frac{xi}{T} \qquad (6)$$

The sample variance is

$$\sigma^2(x) = \sum_{i=1}^{T} (xi - x)^2 \, / \, (T-1) \qquad (7)$$

Sample variance = $s(x) = \sqrt{s^2(x)}$ (8)

Testing for Normality: The Studentised Range

A useful statistic for judging whether the distribution that generated a sample is normal is the studentised sample.

This is

$$SR = \frac{Max(xi) - Min(xi)}{S(x)} \qquad (9)$$

The SR is the maximum minus the minimum, measured in units of single standard deviation.

See Fama's evidence for US stocks.

TABLE 1.2

Frequency Distributions for Daily Returns on Dow-Jones Industrials

		INTERVALS															
		$\bar{R} - .5s(R) \leq R \leq \bar{R} + .5s(R)$		$\bar{R} - 1.0s(R) \leq R < \bar{R} - .5s(R)$ and $\bar{R} + .5s(R) < R \leq \bar{R} + 1.0s(R)$		$\bar{R} - 1.5s(R) \leq R < \bar{R} - 1.0s(R)$ and $\bar{R} + 1.0s(R) < R \leq \bar{R} + 1.5s(R)$		$\bar{R} - 2.0s(R) \leq R < \bar{R} - 1.5s(R)$ and $\bar{R} + 1.5s(R) < R \leq \bar{R} + 2.0s(R)$		$R < \bar{R} - 2s(R)$ and $R > \bar{R} + 2s(R)$		$R < \bar{R} - 3s(R)$ and $R > \bar{R} + 3s(R)$		$R < \bar{R} - 4s(R)$ and $R > \bar{R} + 4s(R)$		$R < \bar{R} - 5s(R)$ and $R > \bar{R} + 5s(R)$	
	T (1)	Expected no. (2)	Actual no. (3)	Expected no. (4)	Actual no. (5)	Expected no. (6)	Actual no. (7)	Expected no. (8)	Actual no. (9)	Expected no. (10)	Actual no. (11)	Expected no. (12)	Actual no. (13)	Expected no. (14)	Actual no. (15)	Expected no. (16)	Actual no. (17)
Allied Chemical	1,223	468.5	562	366.5	349	224.8	163	107.7	94	55.5	55	3.3	16	.08	4	.0007	2
Alcoa	1,190	455.8	521	356.6	343	218.7	172	104.8	85	54.1	69	3.2	7	.07	0	.0007	0
American Can	1,219	466.9	602	365.1	336	224.1	157	107.4	62	55.5	62	3.3	19	.08	6	.0007	3
AT & T	1,219	466.9	710	365.1	285	224.1	131	107.4	42	55.5	51	3.3	17	.08	9	.0007	6
American Tobacco	1,283	491.4	692	384.4	311	235.8	138	113.0	73	58.4	69	3.5	20	.08	7	.0008	4
Anaconda	1,193	456.9	513	357.4	331	219.3	204	105.1	88	54.3	57	3.2	8	.08	1	.0007	0
Bethlehem Steel	1,200	459.6	575	359.5	307	220.6	180	105.7	76	54.6	62	3.2	15	.08	4	.0007	1
Chrysler	1,692	648.0	736	506.9	493	311.0	259	149.1	117	77.0	87	4.6	16	.11	4	.0010	1
Du Pont	1,243	476.1	539	372.4	363	228.5	195	109.5	80	56.5	66	3.4	8	.08	3	.0007	1
Eastman Kodak	1,238	474.2	546	370.9	379	227.5	162	109.1	85	56.3	66	3.3	13	.08	2	.0007	2
General Electric	1,693	648.4	784	507.2	479	311.2	222	149.2	111	77.0	97	4.6	22	.11	5	.0010	1
General Foods	1,408	539.3	632	421.8	423	258.8	194	124.0	84	64.1	75	3.8	22	.09	3	.0008	1
General Motors	1,446	553.8	682	433.2	396	265.8	203	127.4	103	65.8	62	3.9	13	.09	6	.0009	3
Goodyear	1,162	445.0	539	348.1	331	213.6	164	102.4	71	52.9	57	3.1	10	.07	4	.0007	2
International Harvester	1,200	459.6	529	359.5	365	220.6	182	105.7	61	54.6	63	3.2	15	.08	4	.0007	1
International Nickel	1,243	476.1	587	372.4	362	228.5	149	109.5	72	56.5	73	3.4	16	.08	6	.0007	0
International Paper	1,447	554.2	643	433.5	442	266.0	180	127.5	100	65.8	82	3.9	19	.09	5	.0009	0
Johns Manville	1,205	461.5	526	361.0	363	221.5	163	106.2	91	54.8	62	3.2	11	.08	3	.0007	1
Owens Illinois	1,237	473.7	591	370.6	323	227.4	188	109.0	69	56.3	66	3.3	20	.08	3	.0007	1
Procter & Gamble	1,447	554.2	726	433.5	389	266.0	171	127.5	71	65.8	90	3.9	20	.09	6	.0009	2
Sears	1,236	473.4	666	370.3	305	227.2	144	108.9	58	56.2	63	3.3	21	.08	8	.0007	5
Standard Oil (California)	1,693	648.4	776	507.2	468	311.2	233	149.2	121	77.0	95	4.6	14	.11	5	.0010	1
Standard Oil (New Jersey)	1,156	442.8	582	346.3	314	212.6	139	101.8	70	52.5	51	3.1	12	.07	3	.0007	2
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	127.4	85	65.8	86	3.9	18	.09	4	.0009	0

The SR is the maximum minutes the minimum, measured in units of a single standrd deviation. See Fama's evidence for US stocks.

TABLE 1.3

Extreme Values and Studentized Ranges for Daily Returns on the Dow-Jones Industrials

	(1) SMALLEST RETURN	(2) LARGEST RETURN	(3) STUDENTIZED RANGE (*SR*)	(4) *T*
Allied Chemical	-.0718	.0838	10.83	1,223
Alcoa	-.0531	.0619	7.33	1,190
American Can	-.0623	.0675	11.30	1,219
AT & T	-.1038	.0989	20.07	1,219
American Tobacco	-.0800	.0724	12.62	1,283
Anaconda	-.0573	.0600	7.87	1,193
Bethlehem Steel	-.0725	.0619	10.32	1,200
Chrysler	-.0805	.1008	10.51	1,692
Du Pont	-.0599	.0515	10.79	1,243
Eastman Kodak	-.0443	.0779	9.23	1,238
General Electric	-.0647	.0565	9.59	1,693
General Foods	-.0468	.0625	9.00	1,408
General Motors	-.0976	.0829	14.31	1,446
Goodyear	-.0946	.1743	16.79	1,162
International Harvester	-.0870	.0687	11.17	1,200
International Nickel	-.0592	.0567	9.36	1,243
International Paper	-.0507	.0533	8.67	1,447
Johns Manville	-.0687	.1193	11.96	1,205
Owens Illinois	-.0637	.0606	10.08	1,237
Procter & Gamble	-.0635	.0656	11.06	1,447
Sears	-.1073	.0606	14.48	1,236
Standard Oil (California)	-.0633	.0674	9.85	1,693
Standard Oil (New Jersey)	-.1032	.1007	18.29	1,156
Swift & Co.	-.0675	.0628	9.18	1,446
Texaco	-.0593	.0548	8.84	1,159
Union Carbide	-.0456	.0394	8.17	1,118
United Aircraft	-.1523	.0849	13.81	1,200
U.S. Steel	-.0539	.0555	8.06	1,200
Westinghouse	-.0804	.0863	11.22	1,448
Woolworth	-.0674	.0896	13.63	1,445
Averages	-.0727	.0746	11.28	1,310

Source: Adapted from Eugene F. Fama, "The Behavior of Stock Market Prices," *Journal of Business* 38 (January 1965): 51.

Distributions of monthly returns closer to normal than distribution of daily returns.

- Concludes we can perhaps use the normal distribution as a working hypothesis.

A Model of the Behaviour of Stock Prices

Any variable whose value changes over time in an uncertain way is said to follow a stochastic process.

These processes can be classified as discrete time - where changes are only taking place as fixed points in time, or continuous time - where changes can take place at any time.

We will look at a continuous variable, continuous time stochastic process for stock options. This will help subsequently in understanding the pricing of options and the derivative securities.

The Markov Property

A Markov process is a particular type of stochastic process where only the present state of the process is relevant for predicting the future. The past history of the process and the way in which the present has emerged from the past are irrelevant.

If the price of BHP follows a Markov process and the price is $5.00, the past history of price movement of BHP is irrelevant in predicting the future as this corresponds to the weak form of market efficiency. If not, it should be possible to predict the future using technical analysis.

Discrete and Continuous Models

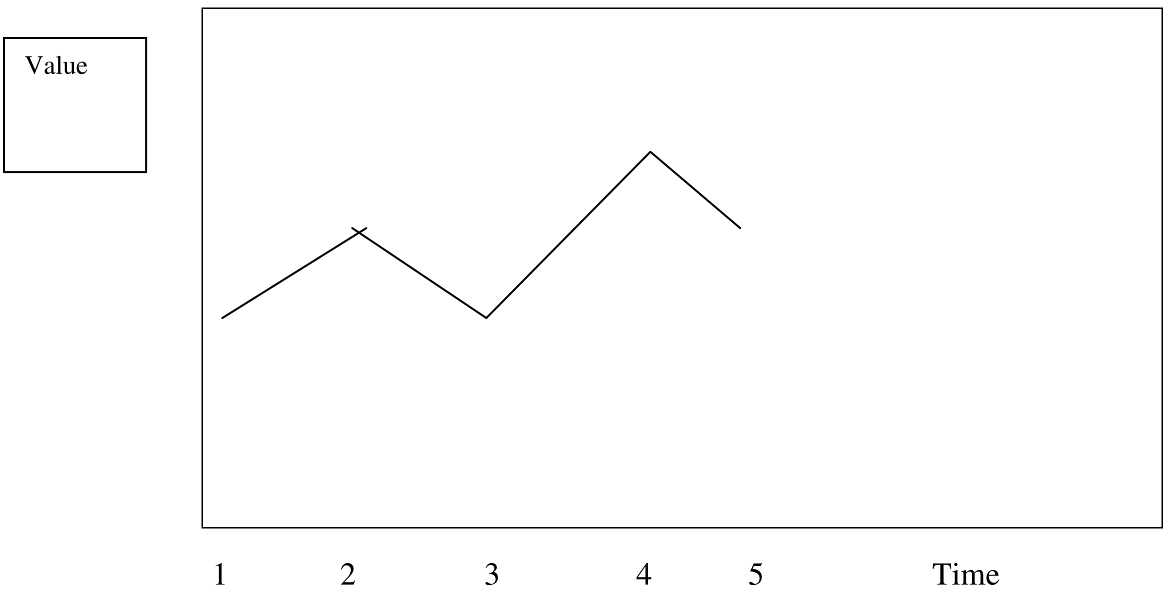
We begin by considering a discrete time random walk description.

$W(t+1) = W(t) + e(t+1)$; $w(o) = W$ e~i.i.d N(0,1)

The variable t represents time and is measured in discrete integer increments from -∞ to +∞. For convenience, we take time 0 as the present. The random variable e(t) is serially chosen from a normal (Gaussion) distribution with mean zero and unit variance. The draws through time are independent of each other and identically distributed (i.i.d).

W(t) is the level of the cumulant of e(t), it is called a random walk because it appears that W takes random steps up and down through time. Early stockmarket theorists used the random walk to describe the level of stock prices.

Diagram 1 - Discrete time random walk - one observation per period



We would specify the time interval for various periods.

Suppose we have a period $\frac{1}{n}$ for an arbitrary integer n>1.

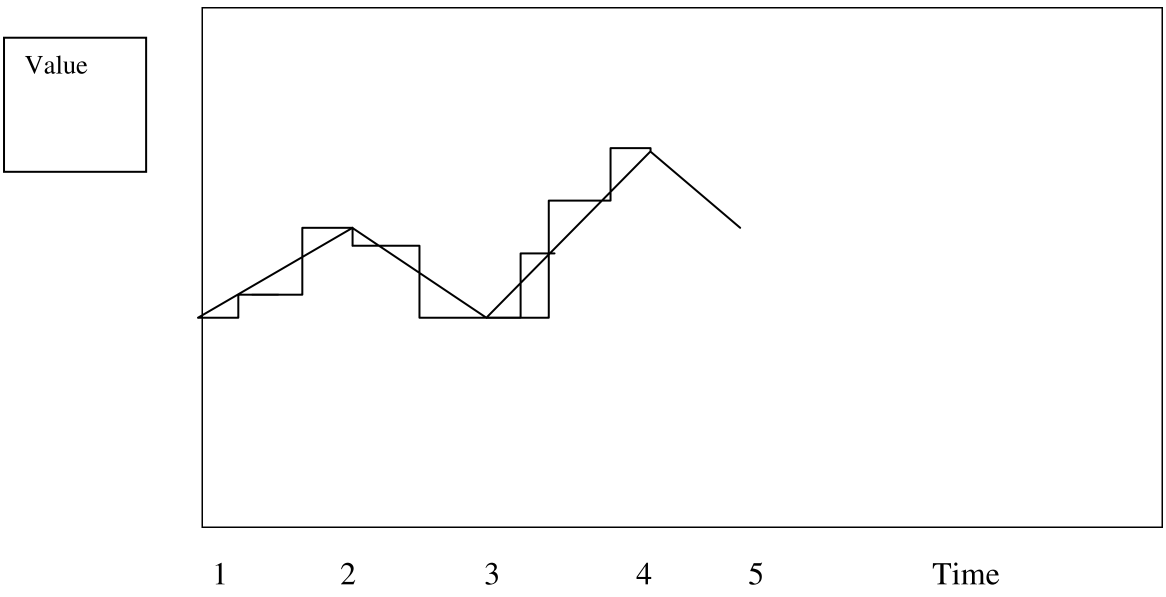
Suppose we were to describe the ~ process which has the same characteristics as a random walk but is observed more frequently:

$W(t+\Delta)=W(t)+e(t+\Delta)j \quad W(0)=W0,$~ i.i.:d $N(0,\Delta)$

This newly defined process has the same expected drift and variance over n periods as the first process does over one period.

Suppose we examine a process that is the same as used in diagram 1 but is observed 4 times as frequently.

Diagram 2 - Discrete time random walk - 4 observations per period



Now consider the process as $\Delta \to dt$

$W(t+dt) = W(t)+e(t+dt)$

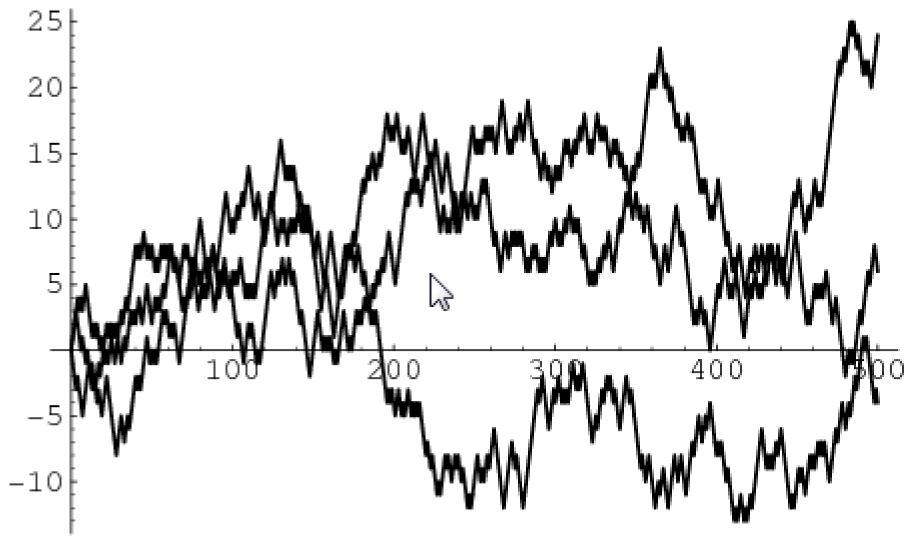
$W(0)=WO, e \sim i.i.d\ N(O,d)$

and define $dW(t) \equiv W(t+dt)-W(t)$

We define dt as the smallest possible real number such $dt^{\alpha} = 0$ whenever $\alpha>1$ (heuristically).

Either of these processes dw(t) or e(t+dt) is referred to as white noise. Figure (3) shows a discrete time random walk approaching the continuous limit.

Figure 3 - Discrete Time Random Walk - approaching the continuous limit: three different examples of random walks



Recall that dw may be thought of as a normally distributed random variable with mean zero and variance dt. The following six properties follow by construction.

1. E[dw(t)] = 0
2. E[dw(t)dt] = E[dW(t)]dt = 0
3. E[dw(t)²] = dt

Property 1 follows by construction, the mean of this normally distributed variable is zero.

Property 2 uses the property that the expectation of the product of a random variable (dw) and a constant (dt) equals the constant times the expected value of the random variable.

Property 3 uses the property for any distribution with zero mean that the expected value of the squared random variable is the same as the variance.

4. $Var\left[dw(t)^2\right] = E\left[dw(t)^4\right] - E^2\left[dw(t)^2\right] = 3dt^2 - dt^2 = 0$

5. $E\left[\left(dw(t)dt\right)^2\right] = e\left[dw(t)^2\right]dt^2 = 0$

6. $V\left[dw(t)dt\right] = E\left[\left(dw(t)dt\right)\right] = E\left[\left(dw(t)dt\right)^2\right] - E^2\left[dw(t)dt\right] = 0$

Property 4 follows from the knowledge of the fourth central moment of the standard normal distribution $(u_4 = 3)$, and the heuristic definition of dt that provides that $dt^2 = 0$.

Property 5 follows immediately from Properties 2 and 3.

Property 6 follows from Properties 2 and 5.

These properties are important because they demonstrate that the variance of a function of a random variable vanishes in Properties 4 and 6. Also the expectation operator is redundant if the variance of a random variable is zero. Therefore, we have

$E[f(dw)]=f(dw)$ if $Var[f(dw)] = 0$

These properties give rise to three multiplication rules:

Rule 1 $dw(t)^2 = dt$

Rule 2 $dw(t)dt = 0$

Rule 3 $dt^2 = 0$

The standard Wiener process has many properties, some of which we provided below.

1. W(t) is continuous in t.
2. W(t) is nowhere differentiable.
3. W(t) is a process of unbounded variation.
4. W(t) is a process of bounded quadratic variation.
5. The conditional distribution of W(u) given W(t), for u>t, is normal with mean W(t) and

   variance (U-t).
6. The variance of a forecast W(u) increases indefinitely as $u \rightarrow \infty$.

Property 1 holds because dw, although it is a random variable, is of infinitesimal magnitude. W is not differentiable (Property 2) since the left and right differentials are not the same; they are independent random variables. Property 3 states (without proof) that the continuous random walk path has infinite length. However, Property 4 states that the sum of squared changes in W is finite, and does in fact equal t.

Properties 5 and 6 discuss the distribution of W(u) given w(t) for u>t. Recall that W is an integral (a sum) of random variables dw. The sum of normal distributed random variables is also normal: the mean of the sum is the sum of the means, and the variance of the sum equals the sum of the variances if the correlations are all zero. This is the same as property 5. Property 6 simply mentions the property that the variance of an ever-expanding sum of normally distributed independent random variable will grow indefinitely.

The standard Wiener process is inappropriate for much financial modelling. However, we can write quite general continuous stochastic processes as functions of standard Wiener processes. For example, consider once again a discrete random walk with generalised drift and heteroscedasticity (i.e., changing variance) that depend on both X(t) and t:

$$X(t+1) = X(t) + \alpha(X(t),t)+\sigma(x(t),t)e(t+1)$$

$$X(o) = Xo, e \sim i.i.d.\ N(O,1)$$

If we choose a sub interval of length Δ that mimics the behaviour of this process, we can write:

$$X(t+\Delta)=X(t)+\alpha(X(t),t)\Delta+\sigma(X(t),t)e(t+\Delta)$$

$$X(0)=Xo, e \sim i.i.d\ No,\Delta)$$

As we let $\Delta \rightarrow dt$, we see that

$$dX(t) = \alpha(X(t),t)dt+\sigma(x(t)t)dw(t);X(o)=Xo$$

which is the description of a generalised univariate Wiener process.

From this point we drop t as an argument of the X and W processes, the time dependence will be understood.

$$dX = \alpha(X,t)dt+\sigma(X,t)dw;\ X(o) = Xo$$

How can we interpret the statement

$$dx = \alpha dt + \sigma dw$$

Suppose for the moment that α and σ are constant. The term dW is a normally distributed random variable, with mean zero and variance dt: the statement says that d is also a random variable, a linear function of a normal random variable, which is itself normally distributed. The random variable dx has mean "αdt" and variance "$\sigma^2 dt$".

The difficulty lies in changing levels of α and σ. Changes may depend on the level of X, the passage of time, or both. The accumulation of these normal random variables can yield distributions of future values that follow many distributions: for example:

ARITHMETIC BROWNIAN MOTION

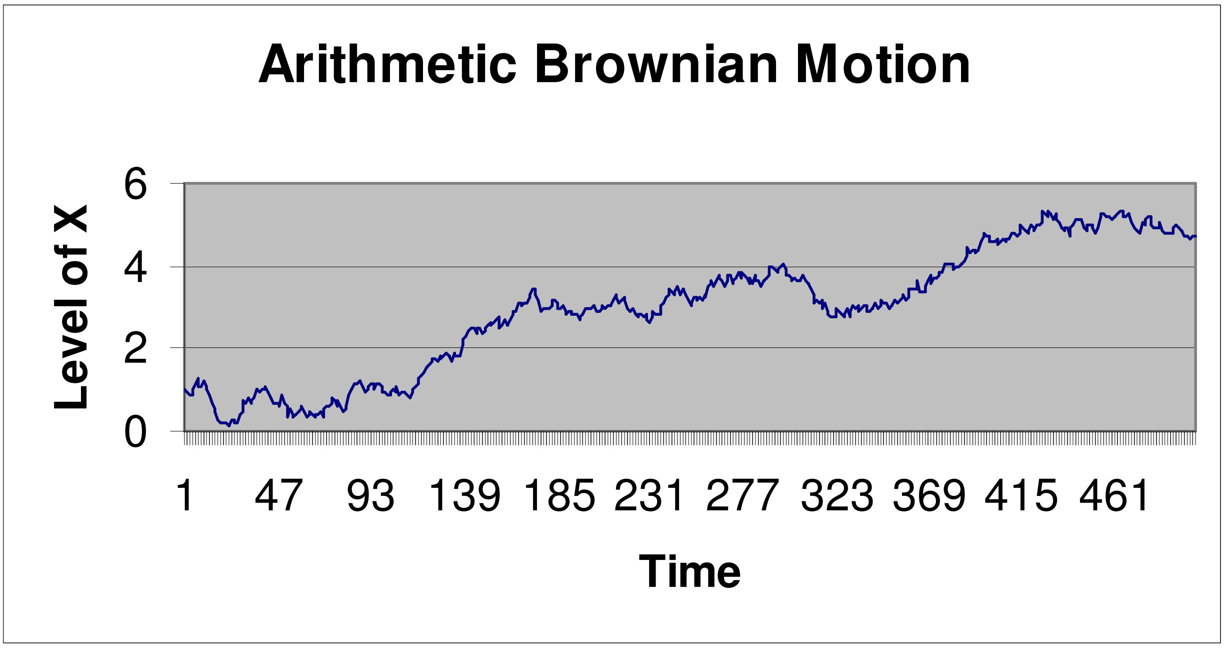
$$dx = \alpha dt+\sigma dw$$

Let $\alpha(X,t) = \alpha$ and $\sigma(X,t) = \sigma$, two constants then the process X is s aid to follow arithmetic Brownian motion with drift α and volatility σ. The process is an appropriate specification for economic variables that grow at a linear rate and exhibit increasing uncertainty. The process X has the following properties (among others):

1. X may be positive or negative
2. If u>t, then Xu is a future value of the process relative to time t. The distribution of Xu, given Xt is normal with mean $Xt+\alpha(u-t)$ and a standard deviation of $\sigma\sqrt{u-t}$.
3. The variance of a forecast Xu tends to infinity as u does (given t, Xt).

The three properties indicate that arithmetic Brownian motion is appropriate for variables that can become positive or negative, have normally distributed forecast errors, and have forecast variance that increases linearly in time.

The diagram below demonstrates a sample arithmetic Brownian motion path with positive drift (α>0).



GEOMETRIC BROWNIAN MOTION

$dx = \alpha X dt + \sigma X dw$

Let $\alpha(X,t) = \alpha X$ and $\sigma(X,t) = \sigma X$. The process X is then said to follow geometric Brownian motion with drift α and volatility σ.

The process is appropriate for economic variables that grow exponentially at an average r ate of α and have volatility proportional to the level of the variable. The process also exhibits increasing forecast uncertainty.

The process X has the following properties (among others)

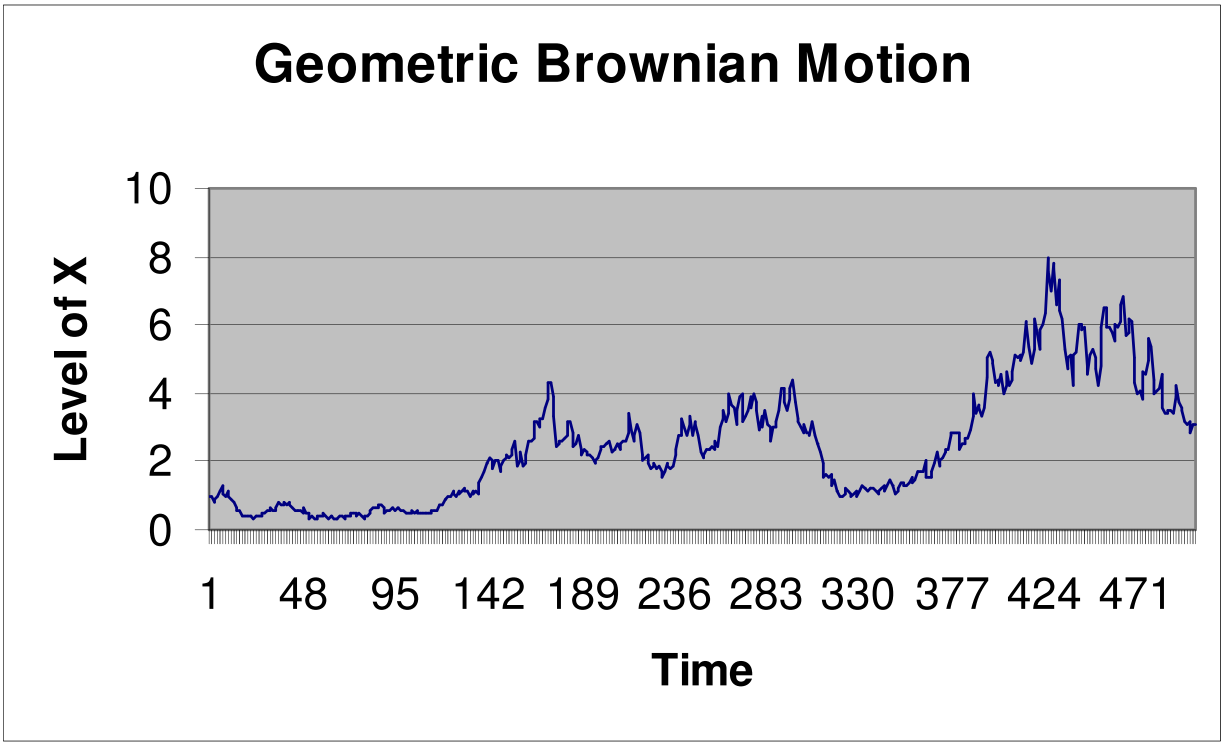
1. If X starts at a positive value, it will remain positive.

2. X has an absorbing barrier at 0: Thus if X hits zero (a zero probability result) then X will remain at zero.

3. The conditional distribution of Xu, given Xt is log normal. The conditional mean of ln(Xu) for u>t is ln(Xt)+α(u-t)-$\frac{1}{2}\sigma^2(u-t)$ and the conditional standard deviation of ln(Xu) is $\sigma\sqrt{[u-t]}$. ln(Xu) is normally distributed. The conditional expected value of Xu is Xt exp[α(u-t)].

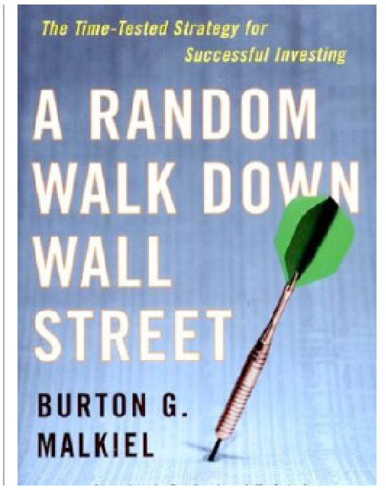
4. The variance of a forecast of Xu tends to infinity as u does.

Geometric Brownian Motion



Geometric Brownian motion(GBM) is often used to model security values, since the proportional changes in security price are independent and identically normally distributed. It can also be used to model anything that is positive and increases (on average) at a constant exponential rate.

Some key players in this area of finance.



Return rates as a random walk

Bachelier's thesis 1900
Black, Scholes, Merton 1973
Cox, Ross, Rubenstein 1979

Properties of Univariate Financial Series

1. Are financial series stationary?

We need some model to formally address the issue.

Traditionally, the response has been to assume that the financial series y_t, t = 1 ....T is a covariance stationary process, i.e., a process for which the auto-covariances $\gamma_j = E(y_t, y_{t-j})$ exist.

Generally, one also wishes to impose some memory conditions upon the rate at which γ_j dies out with j. At a minimum, one will need that $\lim_{j \to \infty} \gamma_j = 0$

A series cannot be independently distributed if any of γ_j, j = 1 ....∞ are non-zero.

The auto-covariance can be estimated and converted into autocorrelations
$\hat{\rho}_j = \hat{\gamma}_0^{-1}\tilde{\gamma}_j$, where upon tests are performed upon whether ρ_1, ρ_2 ....are zero.

Most econometric programs provide this information along with suitable standard errors for

$$\hat{\rho}_j = \left(\sum_{t=j+1}^{T} y^2_{t-j}\right)^{-1} \sum_{t=j+1}^{T} y_t y_{t-j},$$

the standard errors depend on the exact hypothesis being tested, but if it is that ρ_j j = I, ...∞ are zero, they will be $T^{-1/2}$

In the following arguments, we will generally treat random variables as having zero expectation.

There is a great deal of dependence in stock prices and very little in returns.

A response to this emerged in the theory of efficient markets. Here it is argued that the change in stock prices should be independent over time.

If it were not true, opportunities for profit would emerge and these should have been competed away.

This leads to a particular type of independence, viz. the "unit root" or "random walk" model.

$$y_t = y_{t-1} + et \qquad (1)$$

and then the question has a dual nature: does yt have the structure of (1) and is Δyt independent?

This is a time series view of the hypothesis "Accounting/Finance" would take an 'event study' viewpoint, studying the reaction of yt to various events to assess whether all the information is incorporated into prices. To do the latter requires detailed individual stock behaviour. Here, new "events" are subsumed into et.

Two different issues need to be addressed. The first has been investigated in a number of ways.

writing (1) as

$$y_t = \rho y_{t-1} + \rho t \qquad (2)$$

we might test H0: $\rho = 1$ vs H1: $\rho < 1$ using yt and y_{t-1}.

This leads to tests for a unit root of the Phillips-Peron, Dickey-Fuller type which feature in many econometric programs.

The evidence for a unit root in stock prices and no unit root in stock returns is very strong.

1.2 Are financial series independently distributed over time?

The evidence suggest that there is a unit root in stock prices and exchange rates, that there is not one in earnings, and that the evidence is mixed for exchange rates, but we have not yet addressed the issue of dependence in $\Delta yt = et$ which is an integral part of the efficient markets' hypothesis.

There are a number of ways in which this issue has been addressed.

(a) Computation of the auto-correlation function of Δyt followed by tests that serial correlation coefficients are zero. Generally, these re found to be zero, except if data has been measured such that there is an overlapping component.

(b) A different viewpoint is to be had by thinking of the impact of news et. With no dependence, the short-run and long-run impact of news is the same, i.e., $dyt / det = 1 \quad \partial y_{t+d} / \partial et = 1$

However, if the Δyt process has dependence, e.g. is a MA(1)$\varepsilon_t + \alpha\varepsilon_{t-1}$, then (1) can be reformulated as $\partial yt / \partial \varepsilon t = 1 \quad \partial y_{t+\infty} / \partial \varepsilon t = (1+\alpha)$

In general, for Δyt being an MA(q), $\varepsilon t + \alpha, \varepsilon_{t-1} \varepsilon t - 1 + \ldots\ldots\ldots + \alpha q \varepsilon t - q$, $\partial_{yt+\infty} / \partial et = (1 + d_1 + d_2 + \ldots. + \alpha q)$, leading to the idea of testing for dependence by testing Ho: $\alpha_1 + \alpha_2 + \alpha q = 0$. Testing if the sum of α's is zero is likely to be more powerful than testing if the individual α's are zero, because only a scalar is being tested and it is more likely to be precisely estimated than any of its components.

Fama and French (1988) "Permanent and Temporary Components of Stock Prices", *Journal of Political Economy*, 96, pp246-273, work with a model of permanent and temporary components in a series and concentrate upon the question of whether there is a temporary component in returns or not.

Fama and French (1988) form y_{t+k} - yt and yt - y_{t-k}, i.e., kth forward and backward being differences of the series yt, and to regress the former or the latter. With yt as the log of stock prices r_{t+k} = y_{t+k} = yt is the continuously compounded kth period return and it will be the sum of the one-period returns.

$r_{t+k} = y_{t+k} - y_{t+k-1}, r_{t+k-1} = y_{t+k-1} - y_{t+k-2}$, etc.

In large samples, the numerator of the regression coefficient will tend to $E\left[\left(\sum_{j=1}^{k} r_{t+j}\right)\left(\sum_{j=0}^{k-1} r_{t-j}\right)\right]$ and Fama and French are therefore testing if this is

zero. To appreciate what is being tested set k= 1, 2, 3, giving

$$E(r_{t+1}r_t) = \gamma_1 (k=1)$$

$$E[(r_{t+1} + r_{t+2})(r_t + r_{t-1} + r_{t-2})] = \gamma_1 + 2\gamma_2 + \gamma_3 (k=2)$$

$$E[(r_{t+1} + r_{t+2} + r_{t+3})(r_t + r_{t-1} + r_{t-2})] = \gamma_1 + 2\gamma_2 + 3\gamma_3 + 2\gamma_4 + \gamma_3 (k=3)$$

Thus, one is essentially testing if a weighted average of the auto-covariances is zero, rather than whether the auto-covariances themselves are. The reason we test a weighted average is because we are implicitly testing if the scalar var $\left(y_t^*\right)$ is zero. We get this model because we have a particular model of deviations from efficiency in mind.

Volatility

Volatility refers to the movement in an asset price from one point in time to the next. Since volatility reflects the uncertainty in expected asset returns, it is central to most investment decisions as a proxy for risk. The measurement and determination of volatility has become relatively more important with increased financial innovation, i.e., the development of trading instruments designed to neutralise specific forms of risk. Unfortunately, volatility is neither uniquely defined, nor is it likely to be constant over time. Some of the work on this topic is summarised in the following slides.

If the current price of an equity is P_t and the price next period is P_{t+1}, the simplest definition of volatility is

$$\sigma_{t+1}^2 = \left(P_{t+1} - P_t\right)^2 \qquad (1)$$

where σ_{t+1}^2 denotes the volatility over the period $t, t+1$.

Yet there is no unique definition of volatility. The most general definition is that volatility is the expected movement in price from its anticipated value in the next period.

It is often estimated by

$$\sigma_{t+1}^2 = \left(P_{t+1} - E_t P_{t+1}\right)^2 \qquad (2)$$

where $E_t P_{t+1}$ is the anticipated or expected price of the asset at t+1, based on information at t. Clearly, any definition of volatility depends on the definition of expected price in the next period. In the so-called naive case, where we replace the expected price by its current price, that is $E_t P_{t+1} = P_t$, we obtain the definition 1.

Why is volatility important?

Volatility reflects the uncertainty in the expected value of an asset. Hence it becomes one measure of the risk premium associated with an asset. The risk premium can be used in many ways, inter alia in the valuation of assets and in the pricing of options conditional on the asset. For these purposes, the volatility which is important is the volatility of returns, where returns are defined logarithmically by

$$R_{t+1} = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) \qquad (3)$$

where D_{t+1} is the dividend in the period t to t+1. We define the volatility of returns analogously to the volatility of prices, substituting R_t for P_t in(1) and (2).

One valuation method for equities is the dividend discount model. In this method, the valuation is determined by the expected future dividend stream discounted by a risk-adjusted discount rate. The usual method of risk adjustment is obtained using the capital asset pricing model. However, it is also possible to use more direct measures of risk adjustment from volatility models, so that the risk-adjusted discounted rate is then

$$\rho = rf + g(\sigma) \qquad (4)$$

where ρ is the risk-adjusted discount rate, rf is the risk-free rate and $g(\sigma)$ is a function of the average volatility of the asset. The function g depends on the volatility model used.

Options and other derivative assets are also priced on the basis of the average volatility of returns on the underlying stock.

What causes volatility?

Volatility is caused by the reaction of traders to market signals. These include:

1. **Information**
   (a) Political, e.g., election announcements, leadership changes.
   (b) Economics, e.g., balance of payments announcements, interest rate changes.
   (c) External market, e.g., the effect of Wall Street, and of the futures market on the Australian Stock Exchange.
   (d) Internal market, e.g., earnings announcements, resource discourses.

2. **Past Price Changes (leverage effects)**

   It is usually assumed that the volatility of an equity is an increasing function of financial leverage. If the price of equity falls, the consequent leverage effects typically induce greater volatility.

3. **Expectations of Volatility**

   For most financial variables, the current value of the variable is often an important determinant of the next period's value, simply because expectations are partially self-fulfilling. This is no less true for volatility. In periods of high uncertainty, we tend to expect this uncertainty to persist so that volatility will depend on previous volatility.

4. **Noise Trading**

   Many traders may trade on noise (unsystematic or random signals) as if it were information. This imparts a common level of volatility to financial markets.

How to model volatility?

The principal reason for modelling and not just estimating volatility is to forecast it for use in option pricing models and in the determination of risk premia.

Clearly, any model of volatility proceeds in two stages: we must first specify the expected returns process E_tR_{t+1} and then specify the process governing volatility. The expected returns process may be a constant $\left(E_t R_{t+1} = C\right)$ as in Black-Scholes pricing, or of some other form, such as a martingale process $\left(E_t R_{t+1} = R_t\right)$. A point estimate of volatility of returns in the period t,t+1 is then given by $\hat{u}^2_{t+1}$ where

$$\hat{u}_{t+1} = R_{t+1} - \hat{E}_t R_{t+1} \tag{5}$$

the estimated deviation of returns from their estimated anticipated values.

There are then essentially two methods for modelling volatility:

1. The easiest method is to consider the variable $\hat{u}^2_{t+1}$ and to describe its evolution. This is called the <u>stylised facts</u> model of volatility. We can for example run regressions such as:

a) $\hat{u}^2{}_t = a_o + a_1, \hat{u}^2{}_{t-1} + \ldots\ldots\ldots\ldots + a_r \hat{u}^2{}_{t-r} + V_t$ (6a)

b) $\hat{u}^2{}_t = a_o + a_1, UNBP_t + a_r UNINF_t + \ldots\ldots\ldots\ldots + V_t$ (6b)

c) $\hat{u}^2{}_t = a_o + a_1 \mid P_{t-1} - E_{t-2} P_{t-1} \mid , a_1 < 0 + V_t$ (6c)

where V_t is some positive error term.

The first of these regressions is the volatility persistence model, that is the volatility (or its estimate) depends on volatility up to r periods before.

The second of these regressions is an information model, asserting that volatility depends on unanticipated changes in the balance of payments (UNBP); and unanticipated changes in inflation (UNINF), and other macroeconomic information. The third regression is a leverage effect regression - so if the price falls further than we expect, volatility increases.

These models can be combined into a more general model of volatility which includes volatility persistence, information effects and leverage effects.

Furthermore, when all the coefficients except for a_o are zero, each model becomes the standard historical estimate of volatility used in Black-Scholes, that is:

$$\hat{u}_t^2 = a_o + v_t \tag{7}$$

2. A second method is to model the expected return and the volatility simultaneously. The advantage here is that the point estimate of volatility in (5) is replaced by a population value of volatility which is estimated period by period. This includes the recent set of models referred to as the ARCH and GARCH models. Briefly, these models can be written as:

$$R_{t+1} = E_t\left(R_{t+1}\right) + u_t \tag{8a}$$

$$\sigma_{t+1}^2 = f(It) \tag{8b}$$

where I_t is the information set at time t, which includes past returns, past values of the innovations u_t^2, and past values of volatilities σ_t^2.

The important point in ARCH-GARCH models is that volatility and not just a point estimate of volatility is being modelled - furthermore, (8a) and (8b) are estimated jointly.

Within ARCH-GARCH models, there are various ways to represent (8b). The simplest is the linear ARCH model, which is specified as:

$$\sigma_{t+1}^2 = \alpha_o + \alpha, u_t^2 + \ldots\ldots\ldots + \alpha q u_{t+1}^2 - q \tag{9}$$

In this case, volatility is prescribed to be a function of past squared innovations in the expected returns equation. This will capture the effect that in periods of high uncertainty (when u_t is large), volatility will also be high. Obviously, more elaborate models of volatility are possible. One generalisation is to the linear GARCH model, given by:

$$\sigma_{t+1}^2 = \alpha_o + \sum_{i=1}^{a} \alpha_i u_{t+1-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t+1-j}^2 \tag{10}$$

where the volatility depends on both past squared innovations and past volatilities. This captures the expected volatility model referred to above. The process in (9) is called an ARCH (q) volatility model, and the process in (10) a GARCH (p,q) volatility model. There are several important points associated with these models.

1. Estimation

These models are readily estimated using maximum likelihood techniques. For example, they are standard routines in many recent packages, including TSP version 4.2.

2. Positivity

Volatility is a positive variable. This means that all the parameters α_i and β_j should be positive. In practice, this may be difficult to ensure except by using complicated estimation techniques. There are two easy ways:

1. If the lag length of the ARCH-GARCH process is kept short $\left(p_1 q_1 = 1 \text{ or } 2\right)$, the parameters usually remain positive.

2. We can reparametise the process. For the ARCH process, rewrite as:

$$\sigma_{t+1}^2 = \exp\left(\alpha_o + \alpha_1 u_t^2 + \ldots\ldots + \alpha q u_{t+1-q}^2\right) \tag{11}$$

where exp refers to the exponential function. In this case, the volatility is ensured to be positive and the process (11) is an example of a special class of volatility models called EGARCH.

3. Extensions

The ARCH-GARCH models can be expanded to include terms other than past innovations and past volatilities; for example, the leverage and information models of (6b) and (6c).

For example, the ARCH form of (6b) would be

$$\sigma_{t+1}^2 = \alpha_o + \alpha_1 u_t^2 + \alpha_2 \text{UNBP}_t + \alpha_3 \text{UNINF}_t \tag{12}$$

4. ARCH or GARCH

A GARCH (1,1) process is algebraically equivalent to an ARCH process with an infinite lag structure, so that long ARCH processes are usually better represented by a short GARCH process.

5. Frequency

ARCH and GARCH models are especially suitable for high frequency data, i.e., where the interval $(t_1 t+1)$ is one day or less. In particular, this includes data sampled at 15-minute intervals, at hourly intervals, and closing prices on successive days. ARCH and GARCH models are typically not suitable for models of monthly volatility, essentially because the volatility persistence tends to disappear in monthly data. Since for option pricing, the estimate of volatility required is that obtained as the time interval becomes small (the instantaneous volatility), ARCH-GARCH models are quite appropriate.

````````````````````

T Bollershev, R Y Chou, K F Kroner, "Arch modelling in Finance: A review of the theory and empirical evidence", *Journal of Econometrics* (1992), pp5-59.

Although volatility clustering has a long history as a relevant empirical regularity characterising high frequency speculative prices, it was not until recently that applied researchers in finance have explicitly recognised the importance of modelling time varying second order moments.

A key factor in most of these studies has been the Autoregressive Conditional Heteroskedasticity model (ARCH) introduced by Engle (1982). The paper provides an overview of developments in the formulation of ARCH models and a survey of some of the numerous empirical applications.

- Uncertainty plays a central role in finance.
- The uncertainty of speculative prices changes through time.
- One of the most prominent tools for modelling changing variances is ARCH.

ARCH

Following Engle (1982) we refer to all discrete time stochastic processes (Et) of the form

$$E_t = Z_t\sigma_t \qquad (1)$$

$$Z_t.\text{i.i.d.},\ E(Z_t) = 0,\ Var(Z_t) = 1 \qquad (2)$$

With $\sigma_t$ a time-varying, positive, and measurable function of the time t-1 information set, as an ARCH model. To begin, we consider $E_t$ as a univariate process.
````````````````````

By definition, E_t is serially uncorrelated with mean zero, but the conditional variance of E_t equals σ_t^2, which may be changing through time.

In most applications, E_t will correspond to the innovation in the mean for some stochastic process, say (yt) where

$$y_t = g(x_{t-1},b) + c_t \qquad (3)$$

and $g(x_{t-1},b)$ denotes a function of x_{t-1} and the parameter vector b, where x_{t-1} is in the time t-1 information set. To simplify the exposition, in most of the discussion below we shall assume that E_t itself is observable.

Let $f(Z_t)$ denote the density function for Z_t and let θ be the vector of all the unknown parameters in the model. By the prediction error decomposition, the log-likelihood function for the sample E_t, E_{t-1}, ...E, becomes apart from initial conditions

$$L(\theta) = \sum_{t=1}^{T}\left[\log f\left(E_t \sigma_t^{-1}\right) - \log \sigma_t\right] \qquad (4)$$

The second term in the summation is a Jacobian term arising from the transformation from Z_t to E_t. Note that (4) also defines the sample log-likelihood for $Y\tau$, $Y\tau$-1..................,Y_1, as given by (3). Given a parametric representation for Zt, maximum likelihood estimates for the parameters of interest can be computed from (4) by a number of different numerical optimisation techniques.

The above is very general and allows for a wide variety of models.

- The economic theory explaining the conditional variances is very limited.

The linear ARCH (q) model

As Engle (1982) suggests in his seminal paper, one possible parameterisation for σ_t^2 is to express σ_t^2 as a linear function of past squared values of the process.

$$\sigma_t^2 = w + \sum_{i=1}^{q} \alpha_i E_{t-i}^2 = w + \alpha(L) E_t^2 \qquad (5)$$

where w>0 and $\alpha_i \geq 0$, and L denotes the lag operator. This model is known as the linear ARCH (q) model. With financial data it captures the tendency for volatility clustering, i.e., for large (small) price changes to be followed by other large (small) price changes, but of unpredictable sign.

- In order to reduce the number of parameters and ensure a monotonic declining effect of more distant shocks, an ad hoc linearly declining lag structure was often imposed in many early applications of this model.

Maximum likelihood (ML) based inference procedures for the ARCH class of models under this distributional assumption are discussed in Engle (1982) and Pantula (1985).

The Linear GARCH (p.q) Model

An alternative and more flexible log structure is often provided by the Generalised ARCH or GARCH (p.q) model in Bollershev (1986).

$$\sigma_t^2 = w + \sum_{i=1}^{a} \alpha_i E_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 = w + \alpha(L) E_t^2 + \beta(L) \sigma_t^2 \qquad (7)$$

To ensure a well-defined process, all the parameters in the infinite order AR representation $\sigma_t^2 = \phi(L)E_t^2 = (1-\beta(L))^{-1}\alpha(L)E_t^2$ must be non-negative where it is assumed that the roots of the polynomial $\beta(\lambda)=1$ lie outside the unit circle.

For a GARCH (1.1) process, this amounts to ensuring that both α_1 and β_1 are non-negative. It follows also that E_t is covariance stationary if and only if $\alpha(1) + \beta(1)<1$.

Non-linear and non-parametric ARCH

In the GARCH (p.q) model (7), the variance depends on the magnitude and not on the sign of E_t. This is somewhat at odds with empirical work on the behaviour of stock prices which suggest that leverage effect may be present.

In the Exponential GARCH (p.q) or GARCH (p.q) model introduced by Nelson (1990), σ_t^2 is an asymmetric function of past E_ts as defined by (1) and (2) and

$$\log\sigma_t^2 = w + \sum_{i=1}^{q}\alpha_i\left(\phi Z_{t-i} + \gamma\left[\left|Z_{t-i}\right| - E\left|Z_{t-i}\right|\right]\right) + \sum_{i=1}^{p}\beta_i \log\sigma_{t-i}^2 \qquad (8)$$

Unlike the linear GARCH model in (7), there are no restrictions on the parameters α_1 and β_1 to ensure non-negativity of the conditional variances. Thus, the representation in (8) resembles an unrestricted ARMA (p.q) model for log σ_t^2. If $\alpha_i\phi < 0$, the variance tends to rise (fall) when E_{t-i} is negative (positive) in accordance with the empirical evidence for stock returns.