# HANDOUT 03 (REVISED!)

# TIME SERIES ANALYSIS - I

Outline of time series lecture:



## I. Stationarity

A key concept of time series processes is stationarity. A time series is (*covariance*) stationary when it has the following 3 characteristics:

1) exhibits ***mean reversion*** in that it fluctuates around a constant long-run mean;
2) has a finite variance that is time-invariant;
3) has a theoretical correlogram that diminishes as the lag length increases.

In its simplest terms a time series $Y_t$ is said to be ***stationary*** if:

(a) $E(Y_t)$=constant for all t;
(b) $Var(Y_t)$=constant for all t; and
(c) $Cov(Y_t, Y_{t+k})$=constant for all t and all $k \neq 0$

or if its mean, variance and covariance remain constant over time.

In the literature, a *covariance stationary* process is also referred to as a *weakly stationary*, *second-order stationary* or *wide-sense stationary* process. Note that a *strongly stationary* process need not have a finite mean and/or variance.

In this course, we will only consider covariance stationary series so that there is no ambiguity in using the terms stationary and covariance stationary interchangeably.

The $E(Y_t)$, $Var(Y_t)$ and $Cov(Y_t, Y_{t+k})$ would remain the same whether the observations for the time series were, for example, from 1975 to 1985 or from 1985 to 1995.

Stationarity is important because, if the series is non-stationary, all the typical results of the classical regression analysis are not valid.

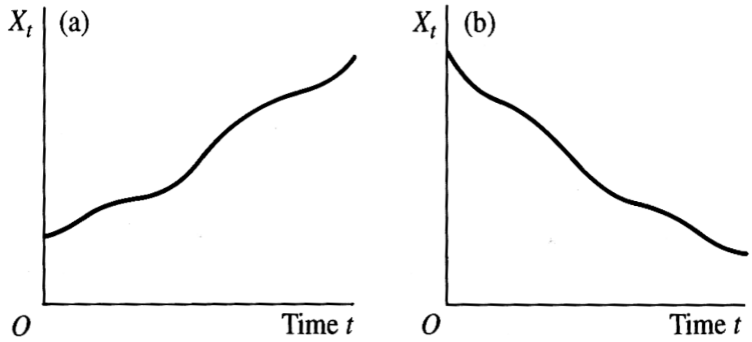
Regressions with non-stationary series may have no meaning and therefore called "*spurious*".

Shocks to a stationary time series are necessarily temporary; over time, the effects of the shock will dissipate and the series will revert to its long-run mean level.

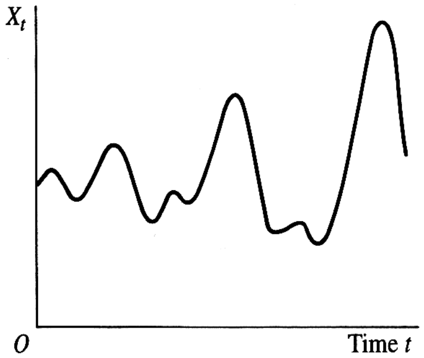
As such, long-term forecasts of a stationary series will converge to the unconditional mean of the series.

A time series is non-stationary if it fails to satisfy any part of the above conditions (a) to (c).

For example, the time series trending consistently upwards or downwards (such as those illustrated below) are almost certain not to satisfy $E(Y_t)$=constant for all t condition since their mean values appear to change over time.




In addition, however, the time series illustrated below is also likely to be non-stationary since, although its mean may be constant, its variance appears to be increasing over time.




In this lecture, we consider models of non-stationary time series, i.e., series $\{Y_t\}$ whose first and second moments (means and covariances) are functions of time. These involve all series with a trend. Trends, which can be either *deterministic* (like a time trend) or *stochastic*, will obviously produce non-stationarities.

An example of nonstationary stochastic process is where the mean of the process is itself a specific function of time: *deterministic trend* situation. This can be described as:

$Y_t = Y_0 + a_0 t + u_t$ $\qquad$ $a_0 \neq 0$ and $u_t$ is a white noise.

It is called a deterministic trend because a fixed value $a_0$ is added for each time $t$.

Another example of nonstationary stochastic process is where a series may be drift slowly upwards or downwards purely as a result of the effects of stochastic (or random) shocks: *stochastic trend* situation.

A process involving stochastic trend can be written as:

$$Y_t = Y_0 + \sum_{t=1}^{T} u_t$$

Here $\sum_{t=1}^{T} u_t$ is the sum of past stochastic terms. Hence $\sum_{t=1}^{T} u_t$ is often called the *stochastic trend*.

This term arises because a stochastic component $u_t$ is added for each time $t$, and because it causes the time series to trend in unpredictable directions. If the variable $Y_t$ is subjected to a sequence of positive shocks ($u_t$>0) followed by a sequence of negative shocks ($u_t$<0), it will have the appearance of *wandering upward*, then *downward*.

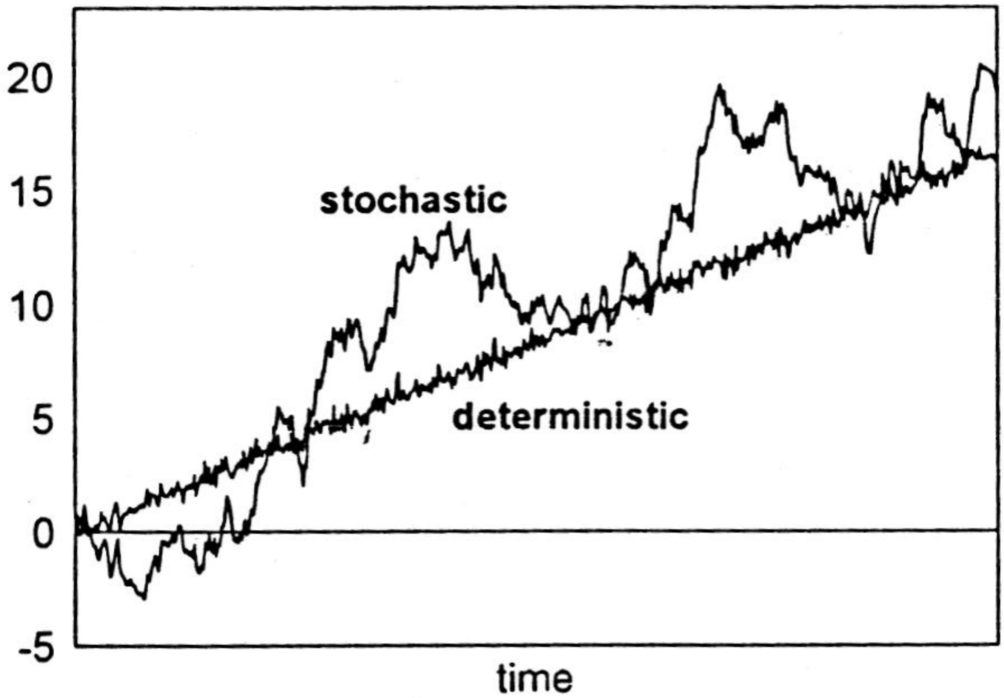



Figure X Stochastic and deterministic trend (Cheremza)

(Stock and Watson, p.588)

A *deterministic trend* is a *nonrandom function* of time. For example, a deterministic trend might be linear in time: if inflation had a deterministic linear trend so that it increased by 1.5 percentage point per quarter, this trend could be written as 1.5*t*, where *t* is measured in quarters. In contrast, a *stochastic trend* is *random* and varies over time. For example a stochastic trend in inflation might exhibit a prolonged period of increase followed by a prolonged period of decrease (Figure 14.1, p.561)

Like many econometricians, we think it is more appropriate to model economic time series a having stochastic trends rather than deterministic trends.

Economics is a complicated stuff. It is hard to reconcile the predictability implied by a deterministic trend with the complications and surprises faced year after year by workers, businesses and governments. For example, although US inflation rose through the 1970s, it was neither destined to rise forever nor destined to fall again. Rather, the slow rise of inflation is now understood to have occurred because of bad luck and monetary policy mistakes, and its taming was in large part a consequence of tough decisions made by the Board of Governors of the Federal Reserve. Similarly, the $/£ exchange rate trended down from 1972 to 1985 and subsequently drifted up, but these movements too were the consequences of complex economic forces; because these forces change unpredictably, these trends are usually thought of as having a large unpredictable, or random component.

Below we present a simplest model of a stochastic trend: the pure *random walk* model.

### *A. The Random Walk Model*

The simplest model of a variable with a stochastic trend is the random walk.

A time series $Y_t$ is said to follow a random walk if the change in $Y_t$ is independent and identically distributed (iid):

$$Y_t = Y_{t-1} + u_t \qquad (1)$$

where $u_t$ is independent and identically distributed (iid). We will, however, use the term random walk more generally to refer to a time series that follows Equation (1) where $u_t$ has a conditional mean zero:

$$E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots) = 0$$

The basic idea of a random walk is that the value of the series tomorrow is its value today plus an unpredictable change. Because the path followed by $Y_t$ consists of random "*steps*" $u_t$, that path is a "*random walk*".

If $Y_t = Y_{t-1} + u_t$ with $E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots) = 0$, then taking the expectation of both sides of $Y_t = Y_{t-1} + u_t$ yields:

$$E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots) = Y_{t-1} + \underbrace{E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots)}_{0}$$

$$E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots) = Y_{t-1}$$

which means that the conditional mean of $Y_t$ based on data trough time *t*-1 is $Y_{t-1}$. In other words, if $Y_t$ follows a random walk, then the best forecast of tomorrow's value is its value today.

- The best known example: Stock Prices

Some series have an obvious upward tendency in which case the best forecast of the series must include an adjustment for the tendency of the series to increase. This adjustment leads to an extension of the random walk model to include a tendency to move or "*drift*" in one direction or the other. This extension is referred to as a random walk with drift:

$$Y_t = a_0 + Y_{t-1} + u_t \qquad (2)$$

where $E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots) = 0$ and $a_0$ is the "*drift*" in the random walk. If $a_0$>0, then $Y_t$ increases on average:

$$Y_t - Y_{t-1} = a_0 + u_t$$

$$\Delta Y_t = a_0 + u_t$$

$$E(\Delta Y_t) = a_0 + \underbrace{E(u_t)}_{0}$$

$$E(\Delta Y_t) = a_0$$ [$\Rightarrow$ Hence, $Y_t$ increases on average]

In the random walk with drift model, the best forecast of the series tomorrow is the value of the series today plus the drift $a_0$ since:

$$Y_t = a_0 + Y_{t-1} + u_t$$

and taking the expectation of both sides yields:

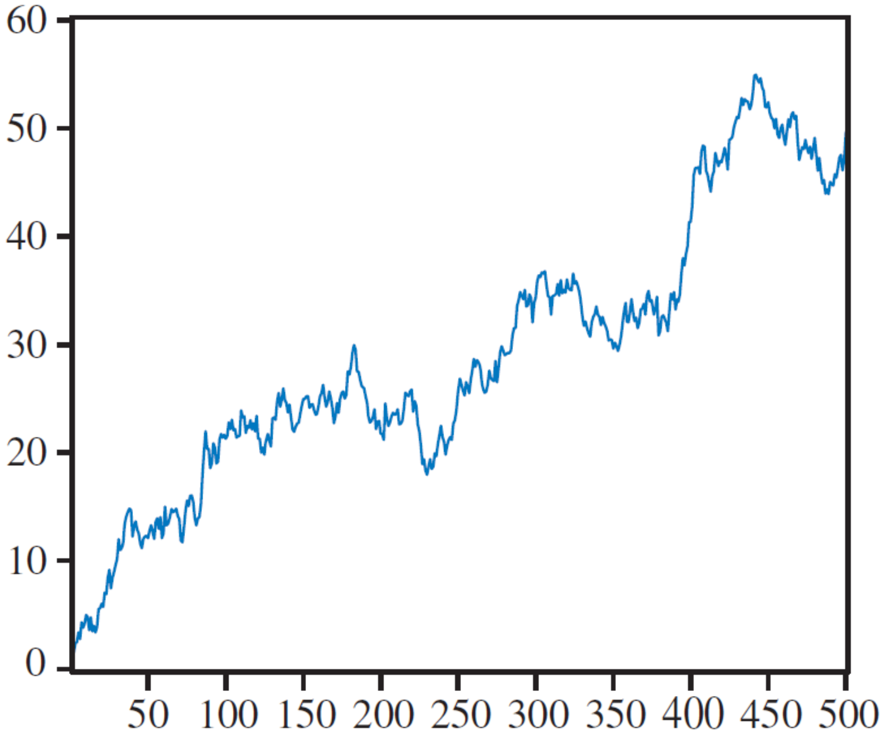
$$E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots) = a_0 + Y_{t-1} + \underbrace{E(u_t \mid Y_{t-1}, Y_{t-2}, \ldots)}_{0}$$

$$E(Y_t \mid Y_{t-1}, Y_{t-2}, \ldots) = a_0 + Y_{t-1}$$

**Example** An example of a time series that can be described by this model can be:

$$Y_t = 0.1 + Y_{t-1} + u_t$$

where its graph can be seen below (Judge and Hill, p.479):




Notice how the time series data appear to "*wandering*" as well as "*trending*" upward. In general, random walk with drift models show definite trends either upwards (when the drift is positive, $a_0$>0) or downward (when the drift is negative, $a_0$<0)

### *B. A Random Walk is Nonstationary*

*If $Y_t$ follows a random walk, then it is* *__not stationary__*: the variance of a random walk increases over time, so the distribution of $Y_t$ changes over time.

Consider $Y_t = Y_{t-1} + u_t$, we can write:

$$Y_1 = Y_0 + u_1$$
$$Y_2 = Y_1 + u_2 = (Y_0 + u_1) + u_2 = Y_0 + \sum_{t=1}^{2} u_t$$
.............................................................
$$Y_t = Y_{t-1} + u_t = Y_0 + \sum_{t=1}^{T} u_t$$

$$\Rightarrow Y_t = Y_0 + \sum_{t=1}^{T} u_t$$

Hence a random walk model contains an *initial value* (often set to zero because it is so far in the past that its contribution to $Y_t$ is negligible), $Y_0$, plus a component of *stochastic trend*: $\sum_{t=1}^{T} u_t$

Recalling that the $u_t$ are independent (since $u_t$ is iid), taking the expectation and the variance of $Y_t$ yields:

$$E(Y_t) = Y_0 + \underbrace{E[u_1 + u_2 + ... + u_T]}_{0}$$

$$\Rightarrow E(Y_t) = Y_0$$

$$Var(Y_t) = Var(u_1 + u_2 + ... + u_T)$$

since $u_t$ is iid we can write:

$$Var(Y_t) = \underbrace{Var(u_1)}_{\sigma_u^2} + \underbrace{Var(u_2)}_{\sigma_u^2} + ... + \underbrace{Var(u_T)}_{\sigma_u^2}$$

$$\Rightarrow Var(Y_t) = T.\sigma_u^2$$

On the other hand:

$$Y_{t-s} = Y_0 + \sum_{t=1}^{T} u_{t-s}$$

$$Var(Y_{t-s}) = Var(u_1 + u_2 + ... + u_{T-s})$$

since $u_t$ is iid we can write:

$$Var(Y_{t-s}) = \underbrace{Var(u_1)}_{\sigma_u^2} + \underbrace{Var(u_2)}_{\sigma_u^2} + ... + \underbrace{Var(u_{T-s})}_{\sigma_u^2}$$

$$\Rightarrow Var(Y_{t-s}) = (T-s).\sigma_u^2$$

In addition,

$$Cov(Y_t, Y_{t-s}) = E[Y_t - \underbrace{E(Y_t)}_{Y_0}]E[Y_{t-s} - \underbrace{E(Y_{t-s})}_{Y_0}].$$

$$Cov(Y_t, Y_{t-s}) = E[\underbrace{Y_t - Y_0}_{\sum_{t=1}^{T} u_t}]E[\underbrace{Y_{t-s} - Y_0}_{\sum_{t=1}^{T} u_{t-s}}]$$

$$Cov(Y_t, Y_{t-s}) = E[u_1 + u_2 + ... + u_T]E[u_1 + u_2 + ... + u_{T-s}]$$

$$Cov(Y_t, Y_{t-s}) = E[u_1^2 + u_2^2 + ... + u_{T-s}^2 + cross\ terms]$$

$$Cov(Y_t, Y_{t-s}) = E[u_1^2] + E[u_2^2] + ... + E[u_{T-s}^2] + \underbrace{E[cross\ terms]}_{=0 \text{ since no AC in } u_t}$$

$$Cov(Y_t, Y_{t-s}) = \underset{\sigma_u^2}{E[u_1^2]} + \underset{\sigma_u^2}{E[u_2^2]} + ... + \underbrace{E[u_{T-s}^2]}_{\sigma_u^2}$$

$$\Rightarrow Cov(Y_t, Y_{t-s}) = (T-s).\sigma_u^2$$

The correlation coefficient (autocorrelation coefficient or autocorrelation function) of the series $Y_t$:

$$\rho_s = \frac{Cov(Y_t, Y_{t-s})}{\sqrt{Var(Y_t)Var(Y_{t-s})}}$$

$$\rho_s = \frac{(T-s).\cancel{\sigma_u^2}}{\sqrt{T.\cancel{\sigma_u^2}\,(T-s).\cancel{\sigma_u^2}}}$$

$$\rho_s = \frac{(T-s)}{T^{1/2}(T-s)^{1/2}}$$

$$\rho_s = \frac{(T-s)^{1/2}}{T^{1/2}}$$

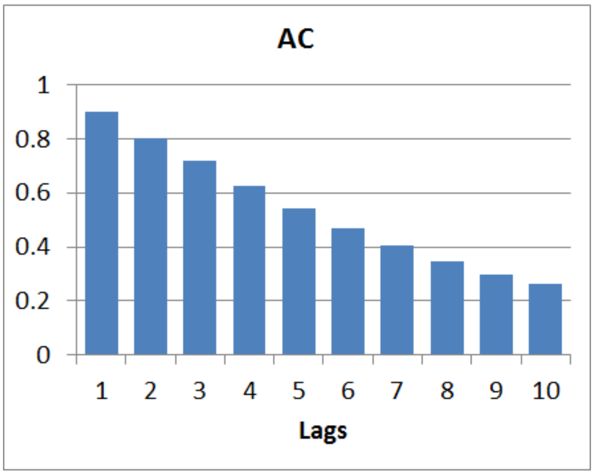
$$\Rightarrow \rho_s = \sqrt{\frac{T-s}{T}}$$

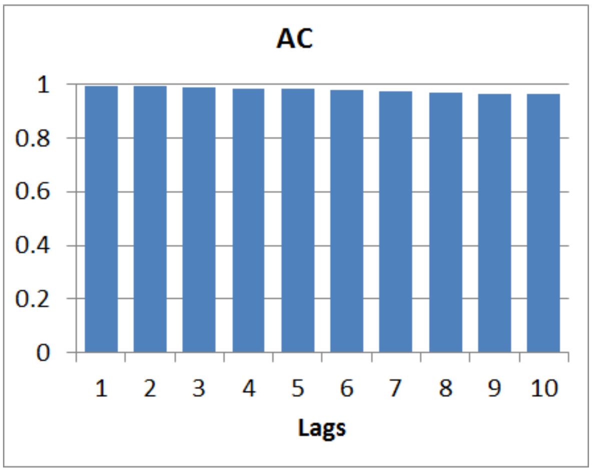
This result play an important role in the detection of nonstationary series. For small values of *s*, the ratio $\frac{T-s}{T}$ is approximately equal to unity. However, as *s* increases the values of $\rho_s$ will decline. Hence the sample *autocorrelation function (correlogram*[1]*)* for a *random walk* process will show *a slight tendency to fall*.

[1] As you know, if we plot the sample correlations $\hat{\rho}_s$ against s we obtain what is called a *correlogram* and hence

| **Correlogram for stationary series** | |
|---|---|
| ***s (lags)*** | ***AC*** |
| 1 | 0.900 |
| 2 | 0.803 |
| 3 | 0.718 |
| 4 | 0.629 |
| 5 | 0.545 |
| 6 | 0.470 |
| 7 | 0.408 |
| 8 | 0.348 |
| 9 | 0.299 |
| 10 | 0.266 |

| **Correlogram for nonstationary series (Pure random walk)** | |
|---|---|
| ***s (lags)*** | ***AC*** |
| 1 | 0.997 |
| 2 | 0.993 |
| 3 | 0.990 |
| 4 | 0.986 |
| 5 | 0.983 |
| 6 | 0.979 |
| 7 | 0.975 |
| 8 | 0.972 |
| 9 | 0.968 |
| 10 | 0.965 |




As seen, there is a dramatic difference between the correlogram for the stationary series and the nonstationary series.

For the stationary series the autocorrelations (correlation between $Y_t$ and $Y_{t-s}$) the column labeled AC in the table above, gradually die out, indicating that values further in the past are less correlated with the current value.

For the nonstationary series, the AC in table does not die out rapidly at all. The correlation between $Y_t$ and $Y_{t-10}$ is 0.965. Hence visual inspection of these functions can be a fixed indicator of nonstationarity.

### C. Random Walk with Drift

Consider $Y_t = a_0 + Y_{t-1} + u_t$, we can write:

$$Y_1 = a_0 + Y_0 + u_1$$
$$Y_2 = a_0 + Y_1 + u_2 = a_0 + (a_0 + Y_0 + u_1) + u_2 = 2a_0 + Y_0 + \sum_{t=1}^{2} u_t$$
.........................................................
$$Y_t = a_0 + Y_{t-1} + u_t = ta_0 + Y_0 + \sum_{t=1}^{T} u_t$$

$$\Rightarrow Y_t = Y_0 + a_0 t + \sum_{t=1}^{T} u_t$$

Hence the value of $Y$ at time $t$ is made up of an *initial value* ($Y_0$) and the *stochastic trend component* ($\sum_{t=1}^{T} u_t$) and now a *deterministic trend component* ($a_0 t$). It is called a deterministic trend since a fixed value $a_0$ is added for each time $t$.

*Hence the variable $Y_t$ wanders up and down (stochastic trend) as well as increases by a fixed amount at each time (deterministic trend).*

Recalling that the $u_t$ are independent (since $u_t$ is iid), taking the expectation and the variance of $Y_t$ yields:

$$Y_t = Y_0 + a_0 t + \sum_{t=1}^{T} u_t$$

$$E(Y_t) = Y_0 + a_0 t + E[u_1 + u_2 + ... + u_T]$$

and since $u_t$ is iid, $E[u_1 + u_2 + ... + u_T] = E[u_1] + E[u_2] + ... + E[u_T]$.

$$E(Y_t) = Y_0 + a_0 t + \underbrace{E[u_1]}_{0} + \underbrace{E[u_2]}_{0} + \ldots + \underbrace{E[u_T]}_{0}$$

$$\Rightarrow E(Y_t) = Y_0 + a_0 t$$

which is not constant, changes by $t$.

As for the variance, taking the variance of both sides of $Y_t = a_0 + Y_{t-1} + u_t$ yields:

$$Var(Y_t) = Var(u_1 + u_2 + \ldots + u_T)$$

since $u_t$ is iid we can write:

$$Var(Y_t) = \underbrace{Var(u_1)}_{\sigma_u^2} + \underbrace{Var(u_2)}_{\sigma_u^2} + \ldots + \underbrace{Var(u_T)}_{\sigma_u^2}$$

$$\Rightarrow Var(Y_t) = T.\sigma_u^2$$

Consequently, in the case of random walk with drift, both the constant mean and constant variance conditions for stationarity are violated.

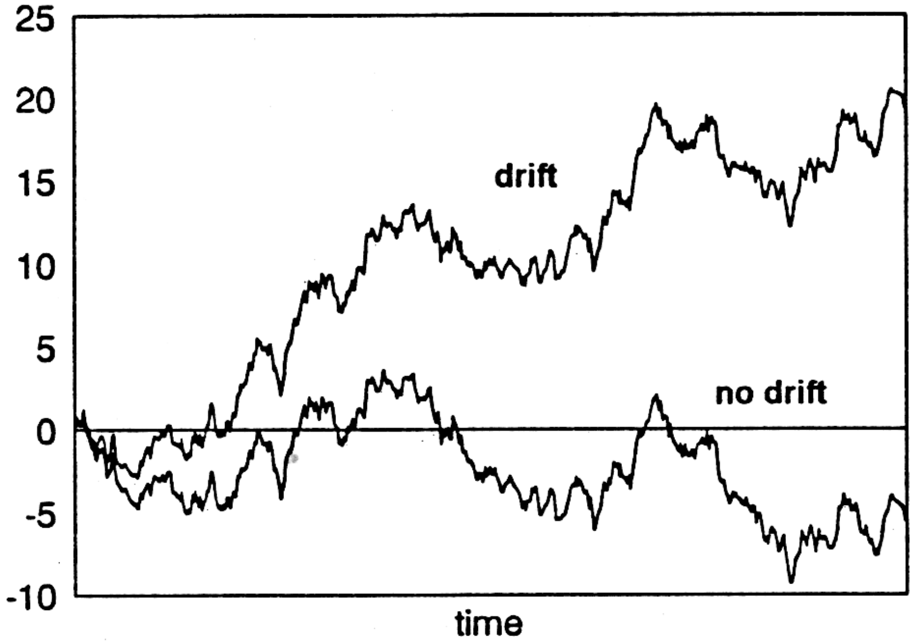



Figure Z. Stochastic trend with and without drift (Cheremza)

### *D. Random Walk with Drift and Trend*

We can extend the random walk model further by adding a time trend:

$$Y_t = a_0 + Y_{t-1} + a_2 t + u_t$$

$$Y_1 = a_0 + a_2 1 + Y_0 + u_1$$
$$Y_2 = a_0 + a_2 2 + Y_1 + u_2 = a_0 + a_2 2 + (a_0 + a_2 + Y_0 + u_1) + u_2 =$$
$$Y_2 = 2a_0 + 3a_2 + Y_0 + \sum_{t=1}^{2} u_t$$

.............................................................

$$Y_t = a_0 + a_2 t + Y_{t-1} + u_t = ta_0 + \left[\frac{t(t+1)}{2}\right] a_2 + Y_0 + \sum_{t=1}^{T} u_t$$

where we have used the formula for a sum of an arithmetic progression: $1 + 2 + 3 + ... + t = \dfrac{t(t+1)}{2}$

Hence:

$$\Rightarrow Y_t = Y_0 + a_0 t + \left[\frac{t(t+1)}{2}\right] a_2 + \sum_{t=1}^{T} u_t$$

Here note that this new term of $\left[\frac{t(t+1)}{2}\right] a_2$ has the effect of strengthening the trend behavior.

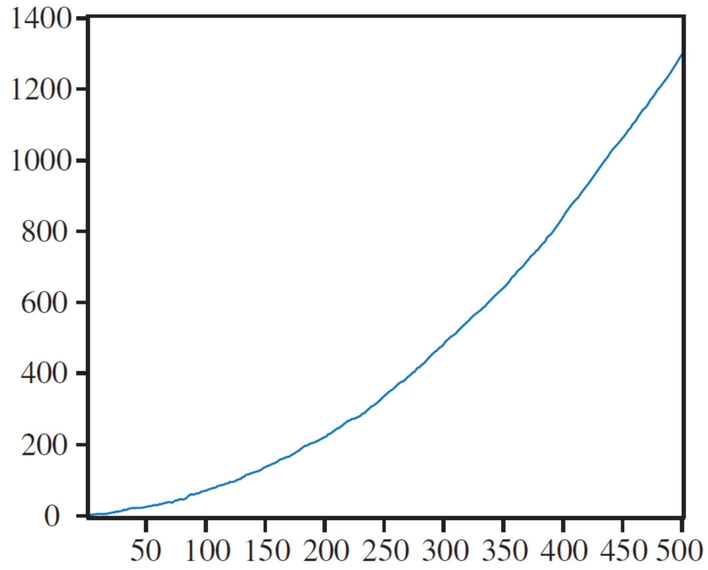
The mean and variance of a random walk with drift and trend are:

$$\Rightarrow E(Y_t) = a_0 t + Y_0 + \left[\frac{t(t+1)}{2}\right] a_2$$

$$\Rightarrow Var(Y_t) = T.\sigma_u^2$$

**Example** Consider $Y_t = 0.1 + Y_{t-1} + 0.01t + u_t$

The graph of this process is shown in the figure below. Note that the addition of a time trend variable *t* strengthens the trend behavior.



# References


- Cameron, Samuel (2005) *Econometrics*, McGraw Hill, Berkshire.
- Chauhan, S.P.S., (2009) *Microeconomics: An Advanced Treatise*, Eastern Economy Edition, New Delhi.
- Dougherty, Christopher (2007) *Introduction to Econometrics*, Oxford, New York.
- Erlat, Haluk (1997) *Introduction to Econometrics*, *Chapter 6: Autocorrelation*, Draft (corrected for misprints), Ankara.
- Ezekiel, Mordecai (1938) The Cobweb Theorem, *The Quarterly Journal of Economics*, Vol. 52, No. 2, pp. 255-280.
- Gujarati, D., and Porter (2011) *Basic Econometrics*, McGraw Hill, New York.
- Hill, R. C., Griffiths, W. E., and Judge, G. G., (2001) *Undergraduate Econometrics*, Second Edition, Wiley, New York.
- Kennedy, Peter (1998) *A Guide to Econometrics*, Fourth Edition, Blackwell, New york.
- Stock, J., and Watson, M. M., (2012) *Introduction to Econometrics*, Third Edition, Pearson, Boston.