Testing for random walk hypothesis with or without measurement error

Tin-Yu Jonathan Hui

A thesis submitted in partial fulfilment of the requirement for the degree of Master of Science of Imperial College London and Diploma of Imperial College London

September 2012

ABSTRACT

In biological study it is common to take observations over time and thus there is strong need in developing tools to analyse time series data. Variations of dynamics over time are often modelled for estimation and forecasting. The random walk process in particular is one of the most commonly used models and serves as the null hypothesis in many theories. As a result, this paper adopts the Q-statistic and Variance-Ratio (VR) Test to test the random walk hypothesis and assesses their testing power against AR(1) model. It is shown that both the Q-statistic and VR Test are valid tests in testing against AR(1) and VR has a higher testing power. Testing power decreases when ϕ approaches unity as the AR(1) process becomes less distinguishable from random walks. In population genetics allele frequencies fluctuate along generations, known as genetic drift. If there exist external forces such as selection, the drift will then have a trend and drive the allele to extinction or fixation. This motivates the second test for a trend in random walk. Both Likelihood Ratio Test and modified Variance-Ratio can test for trends within random walk models. It is shown that the power of the two tests increases with the trend-to-standard deviation $\frac{\mu}{\sigma}$ ratio. In general the Likelihood Ratio Test provides a better power in testing for a trend. In a more realistic scenario, measurement error is assumed in the observations and some newer tests are developed. In this paper the $\sqrt{T}\hat{\rho}(2)$ statistic and 3-point Variance-Ratio are introduced. The overall testing power of $\sqrt{T}\hat{\rho}(2)$ statistic is higher in both high and low signal-to-noise ratio.

CONTENTS

ZERO: OVERVIEW

In many biological experiments data is recorded across a time interval as biologists want to study the dynamics of a situation, in particular, the change in dynamics over time. As a result, the modelling and analysis of time series data becomes important to biological study.

The random walk serves the null model in many biological hypothesises. In the Neutral Theory of molecular evolution, the change in allele frequency can be by chance alone without any predetermined way and this process is known as genetic drift (Hartl & Clark, 1997). Wright-Fisher model (Fisher, 1930; Wright, 1931) states that the transition probability follows binomial distribution and the proportions can be approximated by diffusion process, which is a continuous random walk process. In morphological evolution, many variations fall on continuous scale, and random walk models are used to model quantitative traits. Neutral Theory of Ecology (Hubbell, 2001) assumes that biodiversity arises at random and the equivalence between individuals. This result in species within a community follows a random walk in composition. In mathematical ecology, random walks can be used to model population dynamics and animals or cells movement (Codling, Plank & Benhamou, 2008). Furthermore, population viability analysis and extinction risk can be assessed under the assumption of diffusion process (Lande & Orzack, 1988). Apart from biological applications, random walk plays an important role in physical sciences and economics, for instance, in econometrics, geometric Brownian motion is used to model share price movements and lead to the development of derivatives pricing model (Black & Scholes 1973). Many alternative models are raised, such as the niche theory in ecology, density-dependence model (Ricker, 1954) in population dynamics and mean-reverting models.

The random walk model may not be the most suitable model but many results are derived from random walks. Its mathematical simplicity and elegance make it popular and random walks are often the key assumption of the theories above. Hence, before applying these theories to the real data, it is interesting to ask whether the data follows a random walk. If the data satisfies the random walk hypothesis, the results and predictions from the theories are applicable to the real situation. It would otherwise violate the assumption and the conclusions drawn by these theories would be unrealistic.

Some tests have been developed for testing the random walk hypothesis. Cowles-Jones statistic (Cowles & Jones, 1937) and run test are some of the earliest non-parametric tests for random walks. Some tests are based on autocorrelation function such as the Q-statistic (Box & Pierce, 1970). Fuller (1976) provides asymptotic distributions for autocorrelation coefficients and allows the development of random walk testing in later ages. More recently, Lo and MacKinlay (1988) develop the Variance-Ratio test to test for various forms of random walks.

This paper aims to (1) evaluate the performance of existing tests for random walk hypothesis, in terms of statistical power of the tests against a specific alternative hypothesis, (2) develop and evaluate tests to discriminate trended random walk and non-trended random walk, and finally (3) discuss problems faced by the tests due to the existence of measurement errors and provide alternative testing methods.

The organization of this paper is as follows: Section 1 gives formal definition to time series and random walk process. Some important characters about random walk are derived and these facts are useful in building testing methods for later sections. Section 2 introduces two testing methods for random walk hypothesis, namely the Q-statistic and Variance-Ratio Test, and their null distribution and testing power are compared and analysed. Section 3 is a continuum of the previous section. If random walk pattern is once confirmed, then the next question will be whether a trend exists in such a random walk. This section provides testing methods for trends in random walk and evaluation of the performance of the tests. Section 4 discusses the case when measurement error is added into random walk hypothesis. Traditional tests for random walk are ineffective under this condition and thus new testing methods are introduced based on some features about random walk with measurement error, such as autocorrelation and variance structure. Section 5 is to provide a method to test for a trend in random walk with measurement error. Finally, section 6 recapitulates and discusses the limitation of this study.

ONE: DEFINITIONS

1.1 Introduction

Time series data is often viewed as a collection of values, $\{X_t: t = 0,1,\cdots,n\}$ in which the subscript t is the time at which the datum X_t is observed. In more common terminology, time series is a sequence of observations on a subject at different time points. Traditional statistical tools that are based on independent samples are not ideal because of the correlated structure of time series data. As a result a branch of statistical methods are built under this scenario to analyse data across time. Many statistical models have been used to fit in time series data, or ultimately, to predict the future. In particular, random walk model is one of these types.

This study begins with a formal definition to random walk process. According to Campbell, Lo & MacKinlay (1997), there are more than one definition for random walk, depending on the nature of increments, and the dependence that exists between increments in different distinct time intervals.

1.2 Random walk 1: Independent and identically distributed increments

The simplest version of the random walk hypothesis is the independent and identically distributed (IID) increments. It assumes that all increments are independently drawn from the same distribution with the same mean and variance. The simplest form of the dynamics is the following:

$$X_t = X_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim IID\ (0, \sigma^2) \qquad \text{[equation 1.1]}$$

and the increment is defined as:

$$\begin{aligned} r_t &= X_t - X_{t-1} \\ &= \varepsilon_t,\ \ \varepsilon_t \sim IID\ (0, \sigma^2) \qquad \text{[equation 1.2]} \end{aligned}$$

where $\{X_t\}$ is the process, $\{\varepsilon_t\}$ is distributed with mean 0 and variance σ^2, and $\{r_t\}$ is the increment sequence. The assumption of IID increments is often too strong and theoretical, but it provides good insight about the behaviour of random walk in general. The clear formulation and definition is useful for simulation purposes as well.

The most common distributional assumption of the increments $\{\varepsilon_t\}$ is normality. The process is given by the following equation:

$$X_t = X_{t-1} + \varepsilon_t,\ \ \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) \qquad \text{[equation 1.3]}$$

It is equivalent to the discrete version of Brownian motion, sampled at equal-spaced intervals. Simulation of 10 paths of random walk of this type can be seen in figure 1.1.

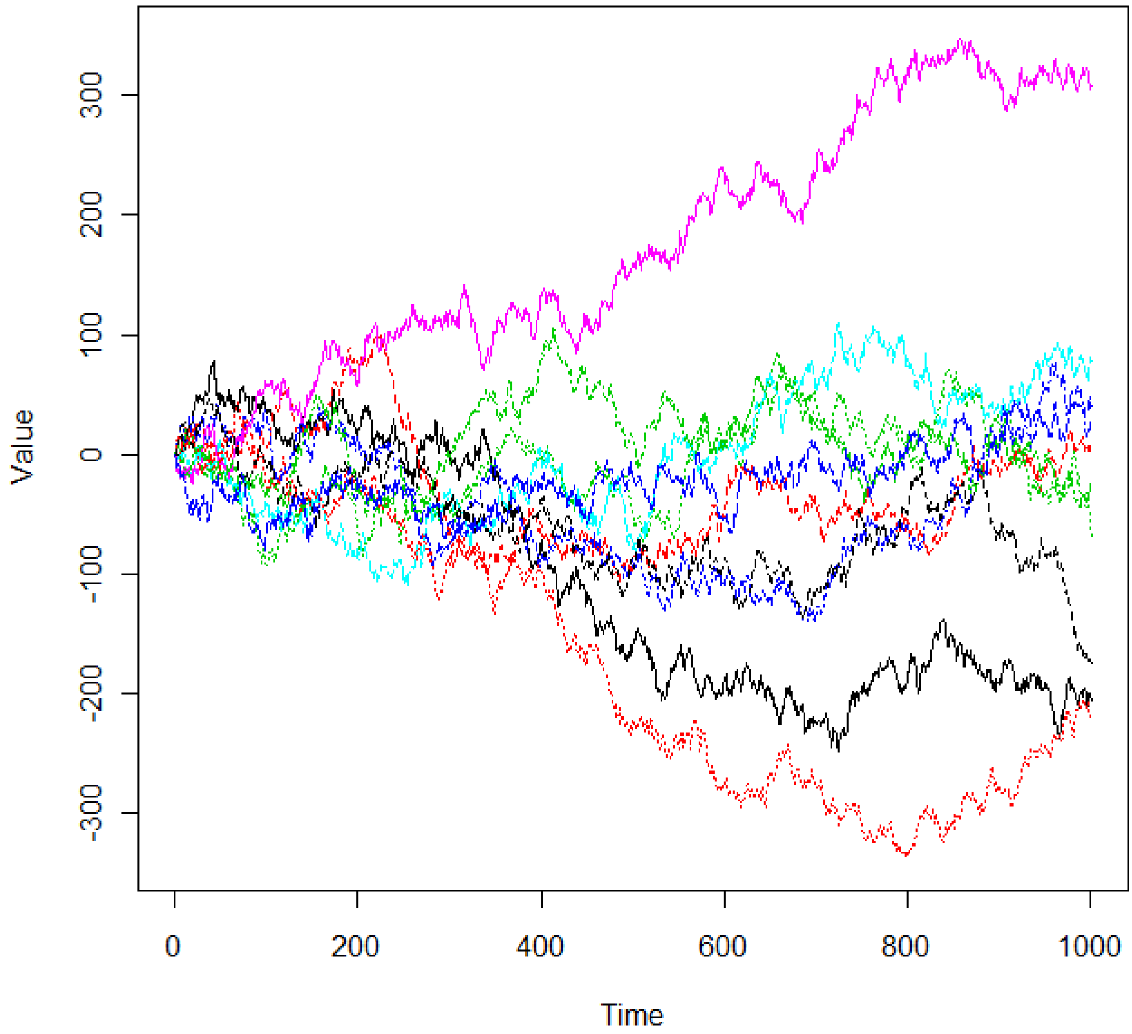

Figure 1.1. 10 simulated random walk trajectories of length 1000 with normally distributed increments.

1.3 Random walk 2: Independent increments

The second type of random walk is independent increments. It assumes all increments are independent but can be drawn from different distributions. It is a more general case than independent and identically distributed increments as it allows unconditional heteroskedasticity in the increments. In other words, time-variation fluctuation is allowed in any of the form as long as the increments are independent. Independent is a strong assumption that not only disjoint increments are uncorrelated, but it also implies any of the non-linear functions of increments are uncorrelated:

$$Cov(f(r_h), g(r_k)) = 0, for\ any\ f, g\ and\ disjoint\ h, k \qquad \text{[equation 1.4]}$$

1.4 Random walk 3: Uncorrelated increments

Further relaxing the assumption of independence yields the third definition of random walk. This is an even more general version of random walk hypothesis which only requires uncorrelated increments. In this case, for every pair of distinct increments, $Cov(r_h, r_k) = 0$, but where the functions of these increments may not be 0. For instance, $Cov(r_h^2, r_k^2) \neq 0$. This is the weakest form of random walk hypothesis among the three definitions.

Despite the differences in their definition, all three types of random walk share some of the common properties. In particular, the conditional mean and variance of random walk are:

$$E[X_t|X_0] = X_0 + \mu t \qquad \text{[equation 1.5]}$$

$$Var[X_t|X_0] = \sigma^2 t \qquad \text{[equation 1.6]}$$

Conditional on the initial value X_0, the conditional mean and variance are both linear with time. This implies random walk processes are non-stationary because of unbounded and increasing variance.

1.5 Section summary

There are three definitions of random walk hypothesis. IID increment is of the strictest sense, while uncorrelated increment is the most relaxed form. This paper adopts the simulation-based approach to evaluate the testing methods and therefore IID normally distributed increments are usually assumed under H_0.

TWO: TESTING THE RANDOM WALK HYPOTHESIS

2.1 Introduction

From the three definitions of random walk, it can be seen that even for the weakest form of random walk, the uncorrelated increments, they imply that all autocorrelations of the increments are zero. This is one of the important features of random walk model and several tests had been derived based this consequence. In particular, this section focuses on the Q-statistic and Variance-Ratio (VR) Test.

Sample h-lag autocovariance of increments can be calculated using the formula below:

$$\hat{\gamma}(h) = \frac{1}{T}\sum_{t=1}^{T-h}(r_{t+h} - \bar{r})(r_r - \bar{r}),\ 0 \leq h < T \qquad \text{[equation 2.1]}$$

Where $r_t = X_t - X_{t-1}$ is the increment and $\bar{r}$ is the sample mean of all the increments. Sample autocorrelation coefficients can be obtained directly:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)} \qquad \text{[equation 2.2]}$$

2.2 Q-statistic

Box and Pierce (1970) develop an intuitive quantity Q using sample autocorrelations. The Q-statistic is the sum of squared autocorrelations up to a desired lag:

$$Q = T\sum_{k=1}^{q}\hat{\rho}^2(k) \qquad \text{[equation 2.3]}$$

As the sample autocorrelation coefficients are asymptotically independent and normally distributed, under H_0, the Q-statistic is the sum of independent squared normal random variables and thus asymptotically distributed as chi-square with q degrees of freedom. Ljung and Box (1978) modify the Q-statistic for finite-sample correction:

$$Q' = T(T+2)\sum_{k=1}^{q}\frac{\hat{\rho}^2(k)}{T-k} \qquad \text{[equation 2.4]}$$

It can be seen that in the Q'-statistic the sample autocorrelations are not equally weighted for small sample-size correction. The Q'-statistic is also chi-squared distributed with q degrees of freedom.

2.3 Variance-Ratio Test

Another property of random walk hypothesis is that the variance of increments is directly proportional to the length of time interval, as stated in equation 1.6. For example, if weekly increments are considered, their variance should be 7 times the daily increments. The plot of variances against different lengths of time interval from 10 simulated random walks is shown in figure 2.1. The Variance-Ratio Test makes use of this property by comparing the variances of the increments from different lengths of time interval. Ideally under H_0, the random walk hypothesis with IID normal errors, variance from q-period increment should be q times larger than that of one-period increment. Lo & MacKinlay (1988) suggest the Variance-Ratio test statistic, denoted as $\overline{VR}(q)$, is the ratio between two variances from different time intervals:

$$\overline{VR}(q) = \frac{\bar{\sigma}_c^2}{\bar{\sigma}_a^2} \qquad \text{[equation 2.5]}$$

with other estimates:

$$\hat{\mu} = \frac{1}{nq}\left(X_{nq} - X_0\right) \qquad \text{[equation 2.6]}$$

$$\bar{\sigma}_a^2 = \frac{1}{nq-1}\sum_{k=1}^{nq}(X_k - X_{k-1} - \hat{\mu})^2 \qquad \text{[equation 2.7]}$$

$$\bar{\sigma}_c^2 = \frac{1}{m}\sum_{k=q}^{nq}\left(X_k - X_{k-q} - q\hat{\mu}\right)^2 \qquad \text{[equation 2.8]}$$

$$m = q(nq - q + 1)\left(1 - \frac{q}{nq}\right) \qquad \text{[equation 2.9]}$$

$\bar{\sigma}_a^2$ and $\bar{\sigma}_c^2$ are the 1 and q-period unbiased variance estimators for overlapping increments. $nq + 1$ is the total length of the sequence. For large sample, the distribution of the test statistic is asymptotically normal:

$$\sqrt{nq}(\overline{VR}(q) - 1) \overset{a}{\sim} N\left(0, \frac{2(2q-1)(q-1)}{3q}\right) \qquad \text{[equation 2.10]}$$

For convenience, the test statistic is scaled standard normal:

$$\sqrt{nq}(\overline{VR}(q) - 1)\left(\frac{2(2q-1)(q-1)}{3q}\right)^{-1/2} \overset{a}{\sim} N(0,1) \qquad \text{[equation 2.11]}$$

The relationship between Variance-Ratio Test and autocorrelation coefficients may not be explicit, but through some mathematical derivations it can be proved that the $VR(q)$ statistic is the weighted sum of the first (q-1) autocorrelation coefficients with declining weights:

$$VR(q) = 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho(k)$$ [equation 2.12]

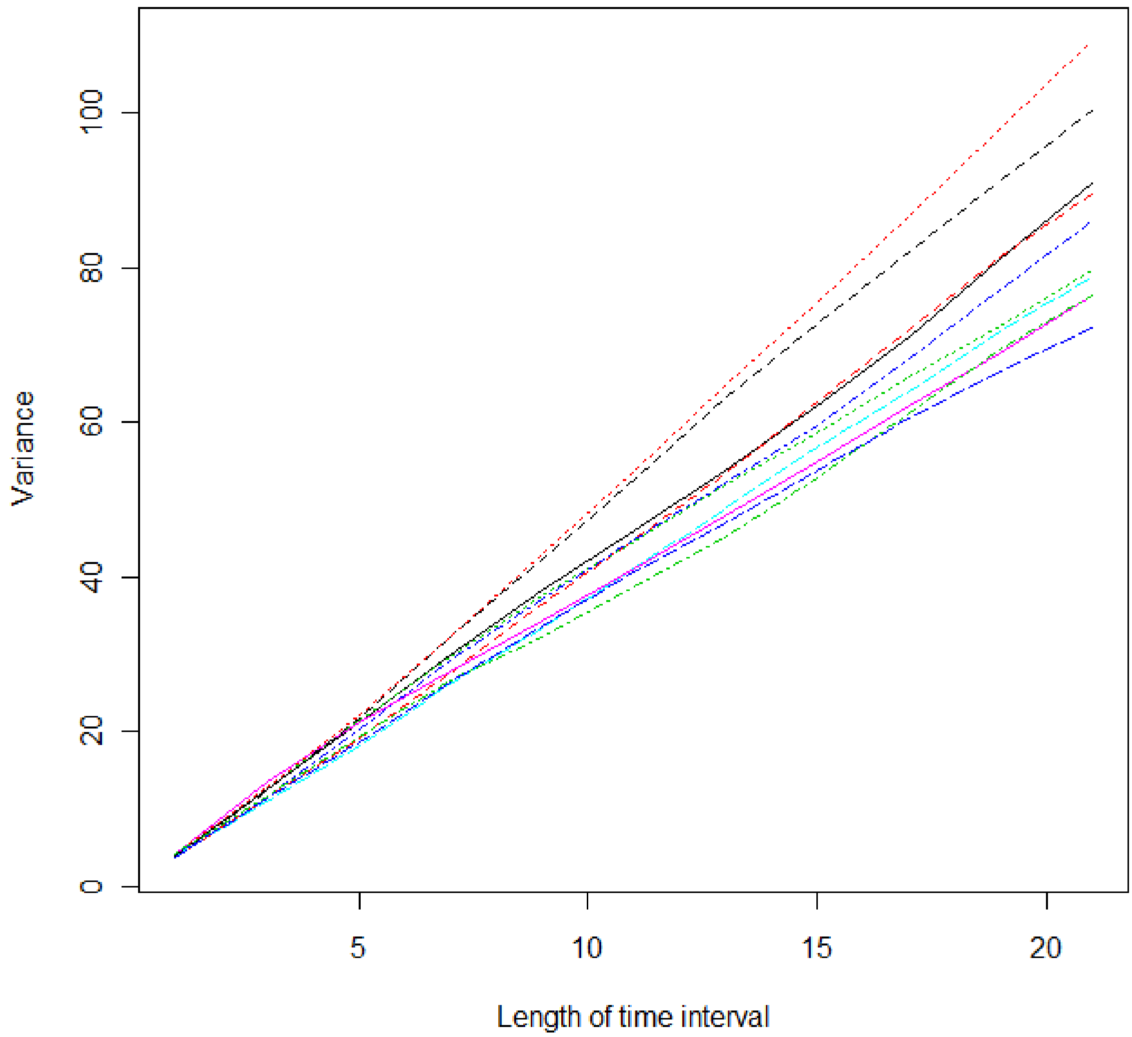

Figure 2.1. The plot of variance of increment against time interval from 10 simulated random walks. A linear relationship is expected.

2.4 Empirical distributions of test statistics under H_0

Although it has been shown that both test statistics (asymptotically) follow some theoretical distributions, it is more desirable to use empirical distributions under H_0 through simulation to enhance accuracy in the case of finite sample size. The null distributions of the two test statistics are generated from 20000 simulated random walks with length 100. Both test statistics are then computed with $q = 10$ for Q'-statistic, while $q = 11$ for Variance-Ratio Test. The q in Variance-Ratio Test is one bigger than that in the Q'-statistic because from equation 2.12 VR test statistic is the weighted sum of first (q-1) autocorrelations. As a result, both statistics are now composed of the first 10 sample autocorrelation coefficients. Therefore a fair comparison can be made using the same amount of information.

The empirical distributions of the Q'-statistic and VR statistic under H_0 are shown in figure 2.2. From the same distributions, empirical critical values of the two tests at 5% significance level (α) are shown in table 2.1. It is noteworthy that VR statistic is a two-sided test and its 2.5- and 97.5-percentile are reported.

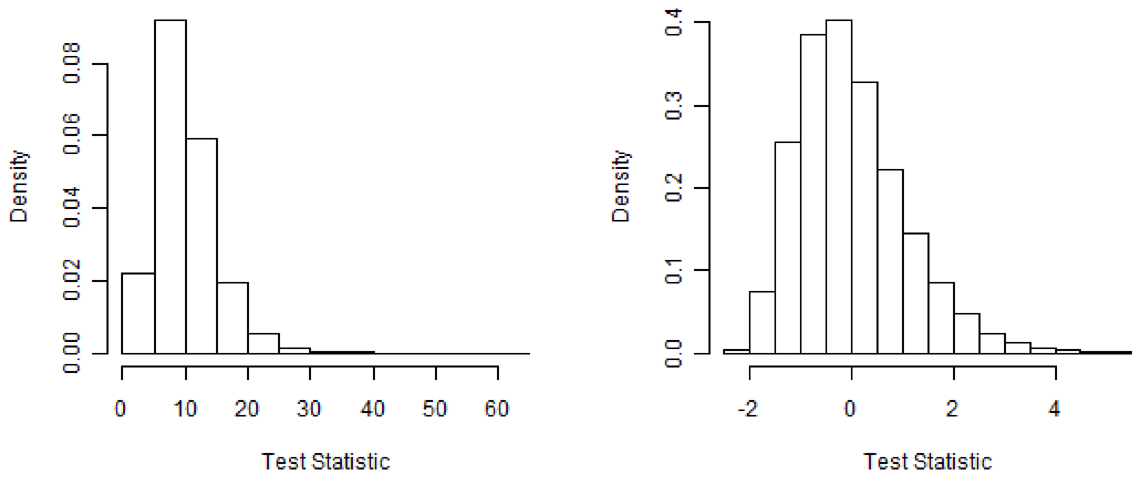

Figure 2.2. Empirical null distributions of Q'-statistic with q=10 (left), and Variance-Ratio statistic with q=11 (right).

Table 2.1. Empirical critical values at 5% significance level

Test statistic	Critical value(s) at $\alpha = 5\%$
Q'-statistic ($q = 10$)	19.044
Variance-Ratio ($q = 11$)	(-1.593, 2.480)

Hence rejection rule of H_0 can be seen easily. For using Q'-statistic, if the test statistic exceeds 19.044 then H_0 is rejected. Similarly, if the Variance-Ratio test statistic is smaller than -1.593, or greater than 2.480, the random walk hypothesis is rejected at 5% significance level. These rejection rules shown in table 2.1 will be used in power analysis below.

2.5 Power analysis

To further understand the Q'-statistic and Variance-Ratio Test, their testing power is examined. The power of a test is defined as the probability of rejecting H_0 at a certain significance level α given H_1 is true. In this study, an empirical method is adopted and power can be viewed as the fraction of trials which the test is able to reject the null hypothesis given the data is knowingly drawn under H_1.

To evaluate the power of the two tests, a specific alternative hypothesis of the first order Autoregressive model (i.e. AR(1) model) is used. The AR(1) model is expressed in the following form:

$$X_t = \phi X_{t-1} + \varepsilon_t \qquad \text{[equation 2.13]}$$

ϕ is the autoregressive coefficient ranging from -1 to 1 for stationarity. Clearly when ϕ reaches one it becomes a random walk model same as the one shown in equation 1.3. The choice of AR(1) model as the alternative hypothesis because it is the discretized version of Ornstein-Uhlenbeck (OU) process. It has mean-reverting property and declining autocorrelations which behaves differently from random walks. Its autocorrelation structure is:

$$\rho(h) = \phi^h, h \geq 0 \qquad \text{[equation 2.14]}$$

In this simulation, the alternative hypothesis is AR(1) process with ϕ set to be 0.7, 0.8 and 0.9. 20000 simulated sequences with length 100 are generated for each value of ϕ, and power can be viewed as the proportion of successfully rejecting H_0 at the 5% significance level, using the empirical critical values shown in table 2.1. Results from the above simulation are listed in table 2.2.

Table 2.2. The empirical power of the Q'-statistic (q=10) and Variance-Ratio Test (q=11) against AR(1) model with ϕ equals 0.7, 0.8 and 0.9, based on 20000 replications.

		Q' statistic (q=10)	VR (q=11)
ϕ	0.7	0.2095	0.85675
	0.8	0.12790	0.48475
	0.9	0.07025	0.12950

Table 2.2 shows that the power of both the Variance-Ratio Test and the Q'-statistic decreases when ϕ approaches 1. This is because when ϕ gets closer to unity, the time series behaves more like a random walk, and the separation between H_0 and H_1 becomes less distinct. It can be seen that a random walk is a special case when ϕ is exactly 1. As a result it is more difficult to discriminate the two processes and thus a decrease in testing power is expected.

By comparing the two tests, Variance-Ratio Test has a larger power than the Q'-statistic for all values of ϕ. It has been shown that both test statistics make use of the same information from the data but with different formulations. From equation 2.3 the Q-statistic is the sum of first q squared autocorrelations with equal weights. Q'-statistic, which is modified from the Q-statistic, is also composed of the first q squared autocorrelations with slightly unequal weights. It is observed from equation 2.4 that the weights are increasing with the lag. Higher-order autocorrelations tend to be insignificant as long-term effect is smaller. Instead, the Variance-Ratio Test takes into account the fading higher-other autocorrelations, and puts more weight on the closer autocorrelations. The Variance-Ratio makes better use of the existing information in constructing the test statistic and therefore it deserves a higher power. Furthermore the Q'-statistic is a portmanteau statistic in nature which is not specific enough. In conclusion, Variance-Ratio Test is a better option than the Q'-statistic in testing random walk hypothesis.

2.6 Section summary

Both Q-statistic (and Q'-statistic) and Variance-Ratio Test have random walk as the null hypothesis. Both tests are formulated from sample autocorrelations from the increment sequence. The results from power analysis shown in table 2.2 suggest that the testing power against AR(1) process decreases when ϕ approaches unity, and Variance-Ratio Test has higher power than Q'-statistic in general.

THREE: TESTING FOR TREND IN RANDOM WALK

3.1 Introduction

After testing for random walk hypothesis as shown in section 2, it is interesting to ask whether a trend (or drift) exists in the time series. For trended random walk there exists a direction on the increments, while the expectation of increments is zero for non-trended random walk. For instance, population biologists worry about whether a trend is driving a population towards fixation or extinction. Figure 3.1 shows some trajectories of random walk under different values of μ. The form of trended random walk is:

$$X_t = \mu + X_{t-1} + \varepsilon_t,\ \ \varepsilon_t \sim IID\ (0, \sigma^2) \qquad \text{[equation 3.1]}$$

and the increment becomes:

$$\begin{aligned} r_t &= X_t - X_{t-1} \\ &= \mu + \varepsilon_t,\ \ \varepsilon_t \sim IID\ (0, \sigma^2) \qquad \text{[equation 3.2]} \end{aligned}$$

The term μ is the trend and shows the general direction of increments. Two testing methods are provided to test for μ in this section: Likelihood Ratio Test and modified Variance-Ratio Test.

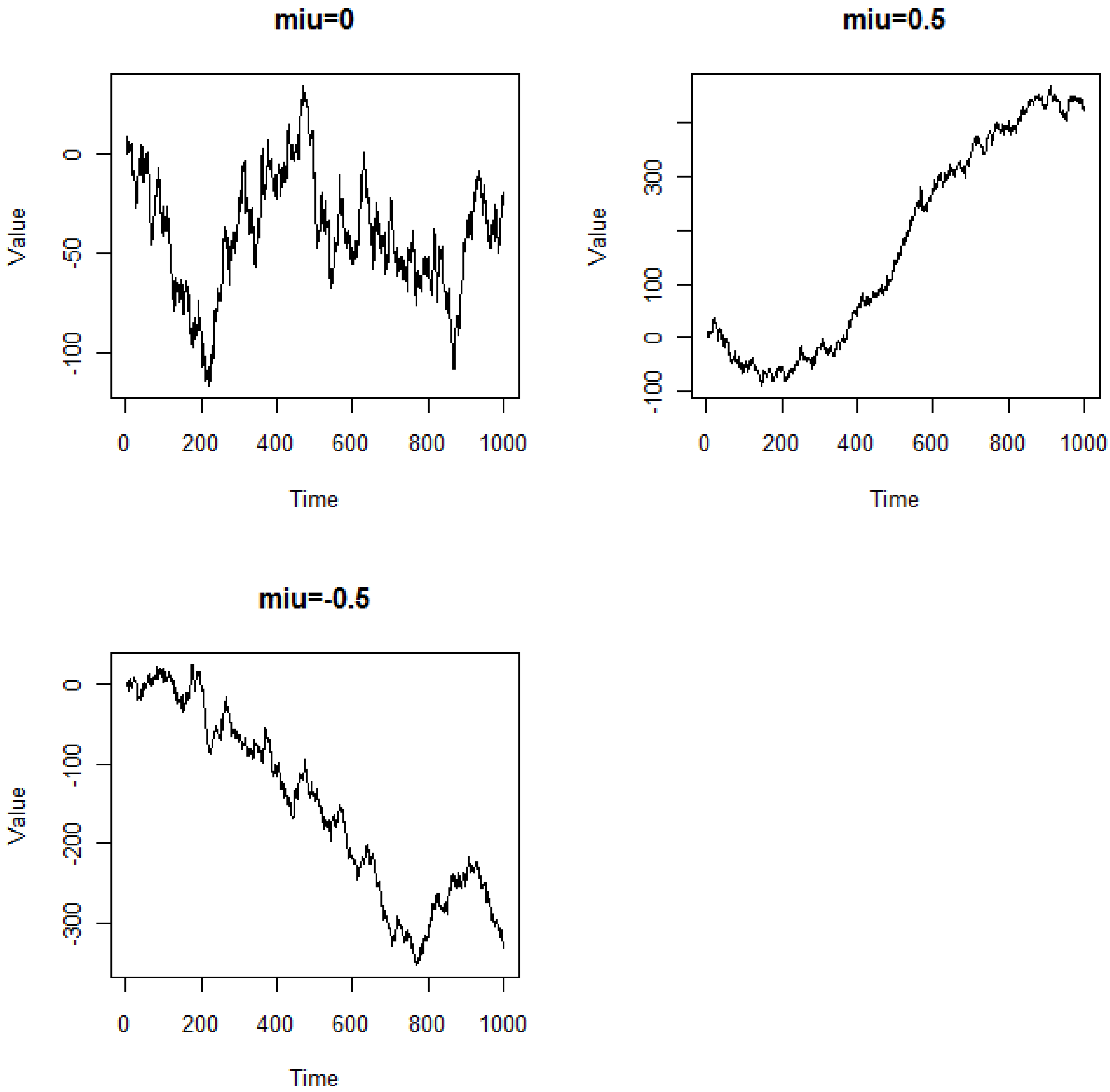

Figure 3.1. Plots of trajectory with different drift μ. Top left: $\mu = 0$. Top right: $\mu = 0.5$. Bottom left: $\mu = -0.5$.

<u>3.2 Likelihood Ratio Test</u>

If IID increments are assumed and the distribution of error terms is known, the Likelihood Ratio Test (LRT) provides a parametric test for μ. Particularly, for the case of discretized Brownian motion, in which the increments are normally distributed, the null and alternative hypothesis are:

$$\mathrm{H}_0: X_t = X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$$

$$\mathrm{H}_1: X_t = \mu + X_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma^2) \qquad \text{[equation 3.3]}$$

Thus the Likelihood Ratio test statistic is:

$$LRT = -2\, ln\big(L(H_0)\big) + 2\, ln\big(L(H_1)\big) \quad \text{[equation 3.4]}$$

$L(H_0)$ and $L(H_1)$ are the maximized likelihoods under H_0 and H_1 respectively. The Likelihood Ratio Test compares the ratio of the two likelihood values (or the difference of the log-likelihoods). If μ is not significant than the likelihoods under both cases should be roughly the same. Theoretically the test statistic is distributed as chi-square with one degree of freedom owning to the difference in the number of parameters between H_0 and H_1.

It is also noticed that under the assumption of random walk with IID increment, LRT is equivalent to t-test for $H_0: \mu = 0$ for sufficient large sample size. The equivalent t-statistic is:

$$t - statistic = \frac{\sum_1^T r_t}{\sigma\sqrt{T}} \quad \text{[equation 3.5]}$$

The t-statistic is the sample mean of all increment terms divided by its standard error. This t-statistic converges to standard normal when sample size T is large.

3.3 Modified Variance-Ratio Test

Another approach adopts the method in Variance-Ratio (VR) Test with μ set to zero. In traditional VR test, trend is allowed and μ is estimated from observed data as shown in equation 2.6. Instead if μ is pre-defined to zero, equation 2.6 is replaced by equation 3.6:

$$\hat{\mu} = 0 \quad \text{[equation 3.6]}$$

and variances are estimated with $\hat{\mu} = 0$ using equation 2.7 and 2.8, then all trended random walks will have shifted variance estimates, and result in a distorted test statistic. As a result, a test is developed by comparing the observed test statistic to the critical values under H_0.

3.4 Empirical distributions of test statistics under H_0

Empirical distributions are used for more accurate comparison. To obtain the empirical distributions of LRT statistic under H_0, 20000 simulated random walks of length 100 without trend are sampled. Then for each of the randomly drawn time series, the maximized likelihoods are calculated under both H_0 and H_1 and thus the null distributions and critical values of the test statistic can be drawn. Similarly, the null distribution and percentiles of modified VR statistic can be drawn using the same set of simulated sequence, with $q = 11$. Significance level is 5% for all comparisons. The empirical distributions of the two statistics under H_0 are shown in figure 3.2. The empirical critical values at 5% significance level are also displayed in table 3.1.

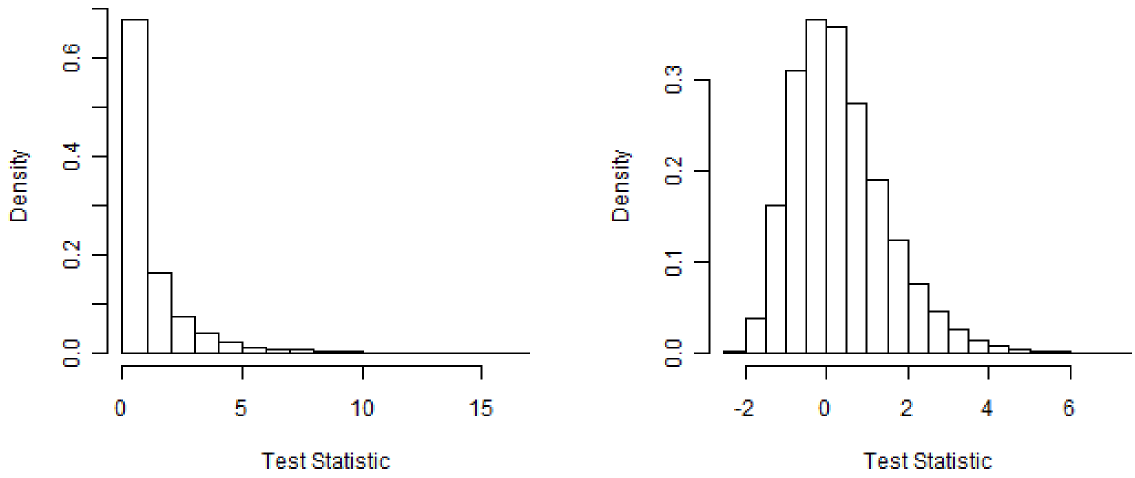

Figure 3.2. Empirical null distribution of LRT statistic (left), and modified Variance-Ratio statistic with q=11 (right).

Table 3.1. Empirical critical values at 5% significance level

Test statistic	Critical value(s) at $\alpha = 5\%$
Likelihood Ratio	3.929
Modified Variance-Ratio	(-1.431, 2.995)

A simple rejection rule for H_0 for both tests can be directly obtained from table 3.1 above.

3.5 Power analysis

The performance of the two tests is evaluated by their power. To find out the power against some specific H_1, the alternative hypothesis is a trended random walk with μ

set to be 0.05, 0.5 and 5 respectively and for each μ 20000 simulated sequences with length 100 are generated. It is also suggested that the magnitude of process error σ will affect the power; therefore the whole simulation is repeated three times with σ equals 0.1, 1, and 10. μ and σ are set in these values because several $\frac{\mu}{\sigma}$ ratios can be compared. The power of both tests is then calculated by the fraction of trails rejecting H_0 using the empirical critical values listed in table 3.1.

The empirical power of the two tests under different μ and σ combinations are listed in table 3.2.

Table 3.2. The empirical power analysis of Likelihood Ratio Test (LRT) and Variance-Ratios (VR).

		LRT			VR (q=11)		
		μ			μ		
		0.05	0.5	5	0.05	0.5	5
σ	0.1	0.99830	1.00000	1.00000	0.93060	1.00000	1.00000
	1	0.07690	0.99890	1.00000	0.05775	0.93140	1.00000
	10	0.04910	0.07550	0.99880	0.05050	0.05610	0.93770

It can be seen from Table 3.2 that the power of both tests is higher when μ is larger. This result is expected, because the separation between null and alternative hypothesis is larger, and thus it is easier to distinguish their difference. From the same table it appears that power decreases with increasing σ. Large σ indicates there are large fluctuations in the sequence, and a small trend becomes undetectable. The confidence interval for μ is wider and a more obvious trend is needed to reject H_0. By combination these two observations, it can be seen that the $\frac{\mu}{\sigma}$ ratio is the parameter determining the power of the test. When $\frac{\mu}{\sigma}$ ratio is 0.5, the power of the test is 99.8% for LRT, and about 93% for modified VR. The power of both tests approaches one when $\frac{\mu}{\sigma}$ ratio increases to 5 and above. For smaller $\frac{\mu}{\sigma}$ ratio, such as 0.05, the power of the two tests decreases dramatically. When the $\frac{\mu}{\sigma} = 0.005$, the two tests have virtually no testing power as power of about 5% is just random occurrence. It is also observed that under the same $\frac{\mu}{\sigma}$ ratio, power of the test remains constant.

By comparing the two tests, Likelihood Ratio Test has a larger power than Variance-Ratio in most of the cases. It is because the primary objective of Variance-Ratio Test is to test for random walk hypothesis, rather than trends in random walk. Although the test statistic is responsive to trend after setting $\hat{\mu} = 0$, it is not its main goal and thus the power against trended random walk is low. On the contrary, the Likelihood Ratio Test is specifically designed to test the existence of μ. Through this direct testing on μ, it is equivalent to do a t-test for mean on all the increments. Thus Likelihood Ratio Test should be more powerful than Variance-Ratio Test.

3.6 Section summary

A trend gives a direction to a random walk. Both the Likelihood Ratio and modified Variance-Ratio Test can be used to test for the existence of trend. It is shown that when the actual trend in a random walk is more significant, the tests can detect the trend more easily and thus the power increases. However, large variation within the sequence hinders the trend and makes it less observable. Therefore the $\frac{\mu}{\sigma}$ ratio is the primary parameter determining testing power. By comparing the two tests, The likelihood Ratio test is a more specific and powerful test to test for μ under the random walk hypothesis.

FOUR: RANDOM WALK WITH MEASUREMENT ERROR AND TESTING

4.1 Introduction

In many real world applications the underlying process cannot be observed directly. Especially in biological studies, data, often recorded by humans, may unavoidably contain errors other than controlled factors. The true value is unknown and can only be observed through experiments. These realizations contain both the original process signal as well as the measurement error. As a result, the whole process becomes a composite of two stages, the state process and the observation process, and is represented by the diagram below (figure 4.1):

$$
\begin{array}{ccccccc}
X_t & \rightarrow & X_{t+1} & \rightarrow & X_{t+2} & \rightarrow & \cdots \\
\downarrow & & \downarrow & & \downarrow & & \\
Y_t & & Y_{t+1} & & Y_{t+2} & &
\end{array}
$$

Figure 4.1. State-space model representation.

In this representation $\{X_t\}$ is the underlying state (unobserved, latent) process and $\{Y_t\}$ is the observed process. This is a class of state-space model and, in particular, under the random walk hypothesis with measurement error, as well as normality of the errors, the whole process is expressed by two equations:

$$
\begin{cases} X_t = X_{t-1} + w_t \ , \ w_t \sim IID(0, \sigma_w^2) \\ Y_t = X_t + v_t \quad , \ v_t \sim IID(0, \sigma_v^2) \end{cases} \qquad \text{[equation 4.1]}
$$

Direct inference on the true process $\{X_t\}$ is not possible; one can only infer the true underlying process through the observed process $\{Y_t\}$.

Ten sample trajectories of random walk with measurement error are simulated and shown in figure 4.2. Apparently they look "random" but they do not satisfy the definitions of random walk stated in the previous section. The autocorrelation structure of the observed process is different from that of a pure random walk process without measurement error. Consequently, traditional tests mentioned in section 2 will reject the random walk hypothesis even when the underlying process is truly a random walk. As a result, the existing tests cannot distinguish random walks with the existence of measurement error from truly non-random walks, and the use of these tests becomes ineffective and unreliable. The following example illustrates how these tests behave when data contains measurement errors.

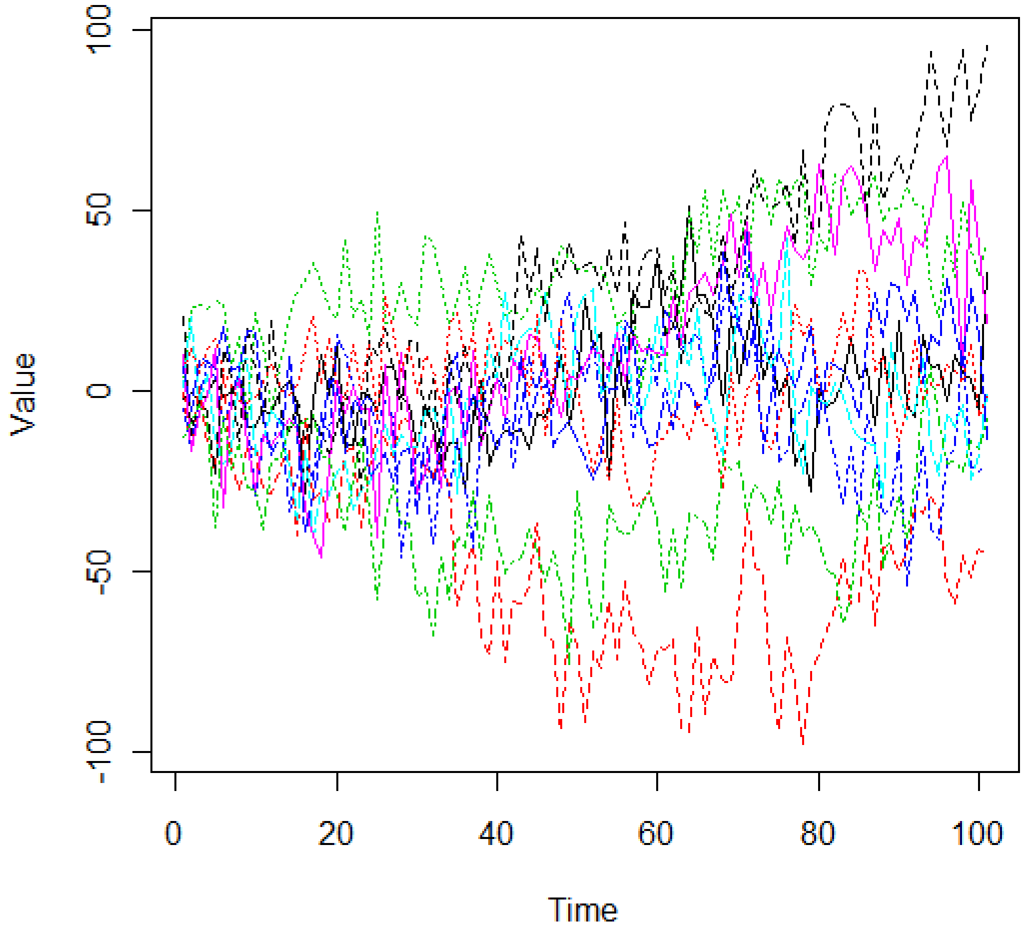

Figure 4.2. Trajectories of 10 simulated random walks with measurement errors, assuming normal errors.

<u>4.2 Behaviour of existing tests under random walk with measurement error</u>

To begin with, one of the trajectories in figure 4.2 is chosen as an example. Values of the selected sequence are listed below in table 4.1:

Table 4.1. List of values of the sample trajectory.

[01]	-1.81655	-0.30133	-2.76778	-0.7029	3.380842	-1.66207
[07]	-0.16852	5.039456	-9.67842	-9.57742	-5.18699	-5.71473
[13]	0.585821	3.139541	-3.59503	-27.8596	-2.25952	10.23557
[19]	-0.86211	12.80279	-15.253	-15.0609	-0.31863	-1.0598
[25]	-5.14825	7.003242	7.094107	-1.67617	-3.8548	-15.6257
[31]	-3.80385	-9.49214	-21.633	-14.4179	-14.4163	-27.0201
[37]	-6.51582	3.125526	-20.2519	-13.9449	-10.7426	-11.5044
[43]	-10.4998	-15.7534	-6.10289	-6.84369	0.18829	10.23492
[49]	-9.16149	2.638747	26.03019	11.99228	16.29775	-23.1206
[55]	10.85808	-3.72252	27.82792	23.65392	23.43384	36.60498
[61]	15.5756	17.81813	27.85755	51.20211	26.62963	26.91942
[67]	19.38787	-3.59969	20.76472	2.166269	15.5617	27.16592
[73]	3.412395	10.90992	-0.6915	4.651	-20.1487	-14.6836
[79]	-27.6904	0.550463	-4.16697	-2.3711	3.396161	14.51513
[85]	3.760468	6.084593	-9.49301	11.85654	-3.16182	19.64315
[91]	-3.67911	-6.93572	15.94753	6.366132	7.206813	-2.35697
[97]	10.1617	5.252345	3.72218	-5.53919	33.133	

The sample autocorrelation of increments from this selected example is plotted in figure 4.3 and it can be seen that the sample autocorrelation of lag one is not zero. The dotted line on figure 4.3 shows the 5% critical values for testing $H_0: \rho = 0$. Clearly it violates the property of random walk that all autocorrelation coefficients are zero.

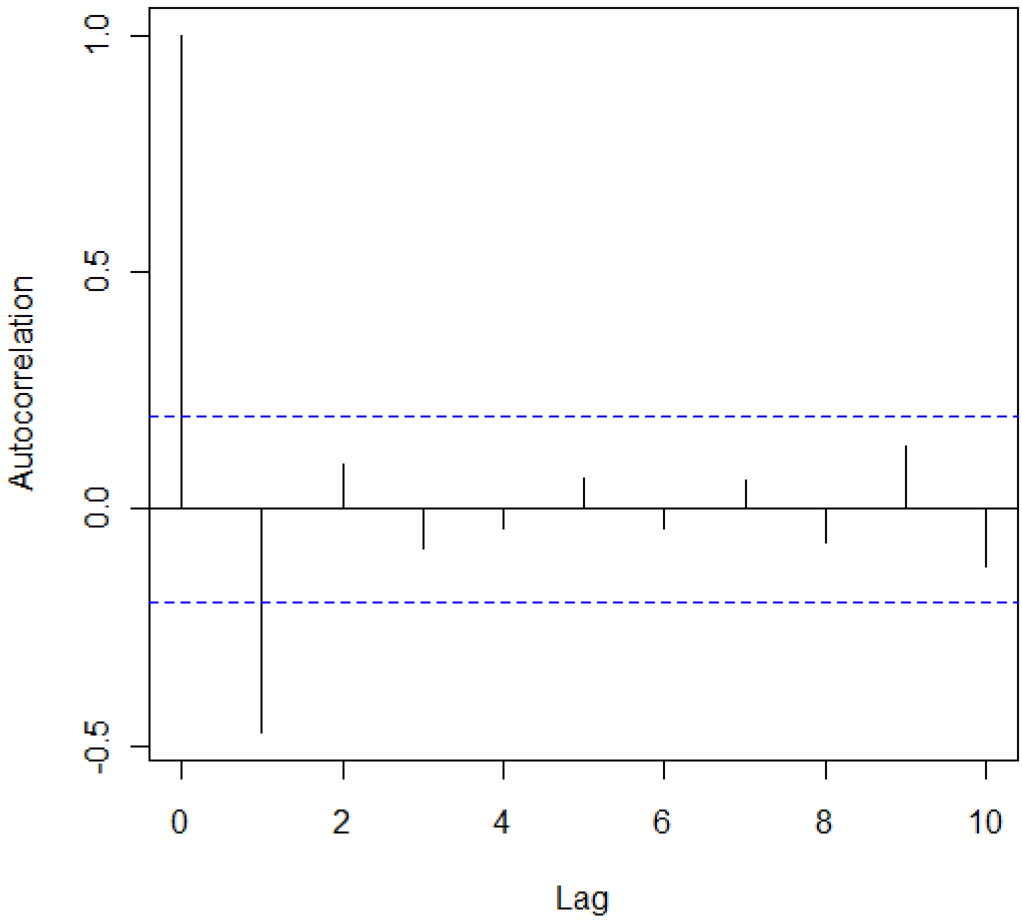

Figure 4.3. Autocorrelation function plot up to lag 10 of the increments from the selected sequence.

The Q'-statistic is then applied to the data to test the random walk hypothesis. By equation 2.4, the test statistic of this example is 29.827 when $q = 10$. Under H_0 the empirical critical value is 19.044 for 5% significance level (Table 2.1). The data has a test statistic larger than the critical value mainly because of the non-zero value of $\hat{\rho}(1)$. The Q'-statistic rejects the null hypothesis despite the fact that the underlying process is a random walk. Similar result can be found for Variance-Ratio Test. The *VR* statistic using formulae 2.5-2.9 is -2.267 with $q = 11$. Using the empirical critical values as shown in table 2.1 before, the lower bound is -1.593 and the null hypothesis is rejected at the 5% level.

Similar conclusions can be drawn for other trajectories. In short, the usual testing methods, in particular the Q'-statistic and Variance-Ratio Test, are no longer applicable in recognizing random walk with measurement error and development of alternative tests under such conditions is thus required.

4.3 Autocorrelation and variance structure

If the process errors and measurement errors are normally distributed, serially and mutually independent, the whole process can be represented in a state-space model with two equations:

$$\begin{cases} X_t = X_{t-1} + w_t \ , \ w_t \sim IID(0, \sigma_w^2) \\ Y_t = X_t + v_t \quad\ , \ v_t \sim IID(0, \sigma_v^2) \end{cases} \qquad \text{[equation 4.2]}$$

w_t and v_t are the process error and measurement error respectively. Their magnitude is determined by two parameters, σ_w^2 and σ_v^2. The first order differencing (or the increment) of the observed process, denoted as ΔY_t, is:

$$\Delta Y_t = Y_t - Y_{t-1}$$

$$\Delta Y_t = (X_t + v_t) - (X_{t-1} + v_{t-1})$$

$$\Delta Y_t = (X_t - X_{t-1}) + v_t - v_{t-1}$$

$$\Delta Y_t = w_t + v_t - v_{t-1} \qquad \text{[equation 4.3]}$$

To understand the characteristics of this model, the autocovariance function is considered:

$$\gamma_0 = Cov(\Delta Y_t, \Delta Y_t) = \sigma_w^2 + \sigma_v^2 + \sigma_v^2$$

$$\gamma_0 = \sigma_w^2 + 2\sigma_v^2 \qquad \text{[equation 4.4]}$$

$$\gamma_1 = Cov(\Delta Y_t, \Delta Y_{t-1})$$

$$\gamma_1 = Cov(w_t + v_t - v_{t-1}, w_{t-1} + v_{t-1} - v_{t-2}) = Cov(-v_{t-1}, v_{t-1})$$

$$\gamma_1 = -\sigma_v^2 \qquad \text{[equation 4.5]}$$

$$\gamma_2 = Cov(\Delta Y_t, \Delta Y_{t-2})$$

$$\gamma_2 = Cov(w_t + v_t - v_{t-1}, w_{t-2} + v_{t-2} - v_{t-3})$$

$$\gamma_2 = 0$$

$$\gamma_k = 0 \ for\ k \geq 2 \qquad \text{[equation 4.6]}$$

It is clear that autocovariance is zero for $k \geq 2$ because there is no overlapping terms between ΔY_t and ΔY_{t+k} anymore. The autocorrelation function can also be derived easily:

$$\rho(0) = \frac{\gamma_0}{\gamma_0} = 1 \qquad \text{[equation 4.7]}$$

$$\rho(1) = \frac{\gamma_1}{\gamma_0} = \frac{-\sigma_v^2}{\sigma_w^2 + 2\sigma_v^2} \qquad \text{[equation 4.8]}$$

$$\rho(k) = 0 \; for \; k \geq 2 \qquad \text{[equation 4.9]}$$

To further understand the behaviour of random walk with measurement error, the graph of variance against time interval of increments is plotted and shown in figure 4.4. Compare to the similar plot (figure 2.1) without measurement error, it is noted that the straight line relationship still holds, except that the graph does not pass through the origin. Owning to the existence of measurement error, the whole curve is shifted upwards. The vertical intercept reflects the magnitude of measurement error σ_v^2, while the slope of the line is the process error σ_w^2.

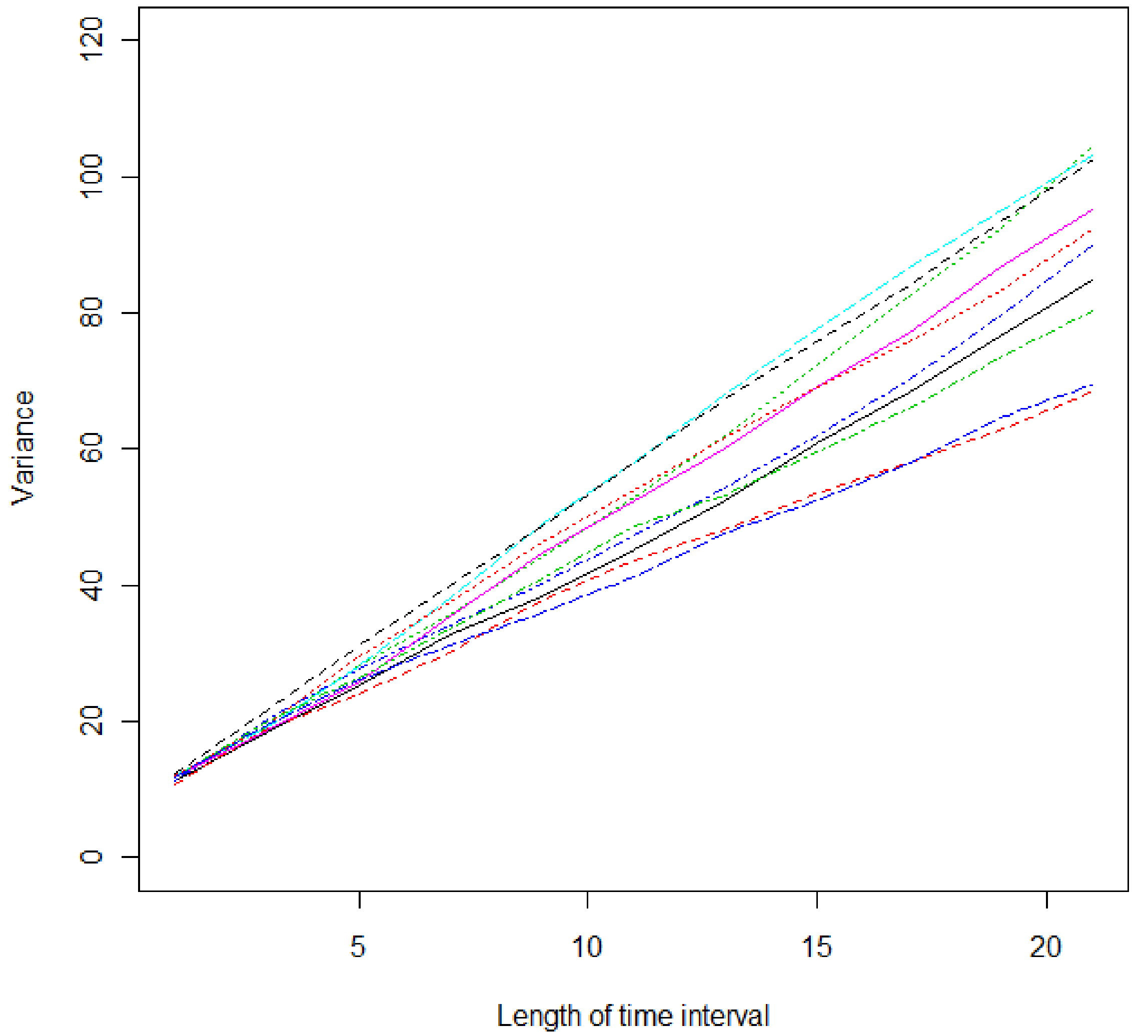

Figure 4.4. The plot of variances of increments against time interval from 10 simulated random walk with measurement error. A straight line relationship is obtained with non-zero vertical intercept.

Section 4.5 and 4.6 make use of these properties of random walk with measurement error to develop some valid testing methods.

4.4 Signal to noise ratio

The observed process is expressed in a signal plus noise equation. The signal is unobservable with process variance σ_w^2 and the noise comes from the measurement error with variance σ_v^2. The signal-to-noise ratio is defined as:

$$SNR = \frac{\sigma_w^2}{\sigma_v^2} \qquad \text{[equation 4.10]}$$

The SNR is an important factor in state-space models and characterizes the feature of the models. If the SNR is large, more information is contained about the underlying process from the observation, but when SNR is too small the observation reduces to a trivial constant mean white noise model. It follows that testing methods (or test statistics) behave differently under various SNR. Hence, some tests perform better under certain range of SNR.

4.5 $\sqrt{T}\hat{\rho}(2)$ statistic

Referring to equation 4.9, it is known that $\rho(2)$ is zero for a random walk with measurement error. A test can be constructed directly from this consequence by testing H_0: $\rho(2) = 0$ through sample autocorrelation coefficient. The sample autocorrelation coefficient of lag 2, denoted by $\hat{\rho}(2)$, can be calculated from the data by equation 2.1-2.2 in section 2. This can be a useful test statistic and the next step is to find out the distribution of $\hat{\rho}(2)$ under H_0. Fuller (1976) states that under some typical conditions sample correlations are normally distributed, with mean the actual $\rho(2)$, which is zero in this case. Mathematically speaking:

$$\sqrt{T}\big(\hat{\rho}(2) - \rho(2)\big) \sim N(0\,, g^2) \qquad \text{[equation 4.11]}$$

T is the number of increments. g^2 is the variance of $\sqrt{T}\hat{\rho}(2)$ and Fuller also gives the formula for calculating g^2:

$$g^2 = \sum_{\ell=-\infty}^{\infty} \begin{bmatrix} \rho(\ell)\rho(\ell) + \rho(\ell+3)\rho(\ell-3) - 2\rho(3)\rho(\ell)\rho(\ell-3) \\ -2\rho(3)\rho(\ell)\rho(\ell-3) + 2\rho(3)\rho(3)\rho^2(\ell) \end{bmatrix}$$

[equation 4.12]

The calculation of g^2 involves the sum of other autocorrelation coefficients. Nevertheless, by using the properties in equations 4.7-4.9 the calculation can be much simplified because all autocorrelations are zero except $\rho(-1)$, $\rho(0)$ and $\rho(1)$, hence:

$$g^2 = \sum_{\ell=-\infty}^{\infty}[\rho(\ell)\rho(\ell) + \rho(\ell+3)\rho(\ell-3)]$$

$$g^2 = \sum_{\ell=-\infty}^{\infty}\rho^2(\ell) + 0$$

$$g^2 = \sum_{\ell=-1}^{1}\rho^2(\ell)$$

$$g^2 = \rho^2(-1) + \rho^2(0) + \rho^2(1)$$

$$g^2 = 1 + 2\rho^2(1) \qquad \text{[equation 4.13]}$$

The results from equations 4.11 and 4.13 show that $\sqrt{T}\hat{\rho}(2)$ is normally distributed with mean 0 and variance $1+2\rho^2(1)$. Hence the rejection rule for H_0 is:

$$\left|\sqrt{T}\hat{\rho}(2)\right| > z_{1-\alpha/2}\sqrt{1+2\rho^2(1)} \qquad \text{[equation 4.14]}$$

where $z_{1-\alpha/2}$ is the $(1-\frac{\alpha}{2})$-percentile of standard normal, $N(0,1)$. For example, the significance level α is usually 5% and thus $z_{1-\alpha/2}$ is 1.96. However, $\rho(1)$ is not a fixed quantity and varies between data. Hence, an upper and lower bound of $\rho(1)$ under H_0, as illustrated by the two cases below, are used to generalize the test. Recall the formula of $\rho(1)$ (from equation 4.8):

$$\rho(1) = \frac{-\sigma_v^2}{\sigma_w^2+2\sigma_v^2}$$

Case 1: When σ_w^2 is much greater than σ_v^2 (i.e. SNR ratio is high), the process converges to a pure random walk (because of extremely small measurement error) and thus $\rho(1)$ becomes 0:

$$\sigma_w \gg \sigma_v$$

$$\rho(1) = \frac{-\sigma_v^2}{\sigma_w^2+2\sigma_v^2} = \frac{-\frac{\sigma_v^2}{\sigma_w^2}}{\frac{\sigma_w^2}{\sigma_w^2}+2\frac{\sigma_v^2}{\sigma_w^2}} \to \frac{0}{1+0} = 0 \qquad \text{[equation 4.15]}$$

Case 2: Consider the reverse case when σ_v^2 is much bigger than σ_w^2:

$$\sigma_v \gg \sigma_w$$

$$\rho(1) = \frac{-\sigma_v^2}{\sigma_w^2+2\sigma_v^2} = \frac{-\frac{\sigma_v^2}{\sigma_v^2}}{\frac{\sigma_w^2}{\sigma_v^2}+2\frac{\sigma_v^2}{\sigma_v^2}} \to \frac{-1}{0+2} = -\frac{1}{2} \qquad \text{[equation 4.16]}$$

$\rho(1)$ converges to $-\frac{1}{2}$ but it cannot be smaller than that under random walk hypothesis with measurement error.

By combining the two cases, $-\frac{1}{2}$ and 0 are the lower and upper bound of $\rho(1)$ under H_0 and therefore the boundaries of g^2 are:

$$-\frac{1}{2} < \rho(1) < 0 \Rightarrow 0 < \rho^2(1) < \frac{1}{4}$$

$$\Rightarrow 1 < 1+2\rho^2(1) < \frac{3}{2} \qquad \text{[equation 4.17]}$$

Equation 4.17 shows that g^2 can only range between 1 and $\frac{3}{2}$ under H_0. Figure 4.5 plots the relationship between the theoretical variance of $\sqrt{T}\hat{\rho}(2)$ statistic and a range of SNR. Furthermore the theoretical critical values of the $\sqrt{T}\hat{\rho}(2)$ statistic using equation 4.14 for different SNR values at 95% confidence level are shown in figure 4.6:

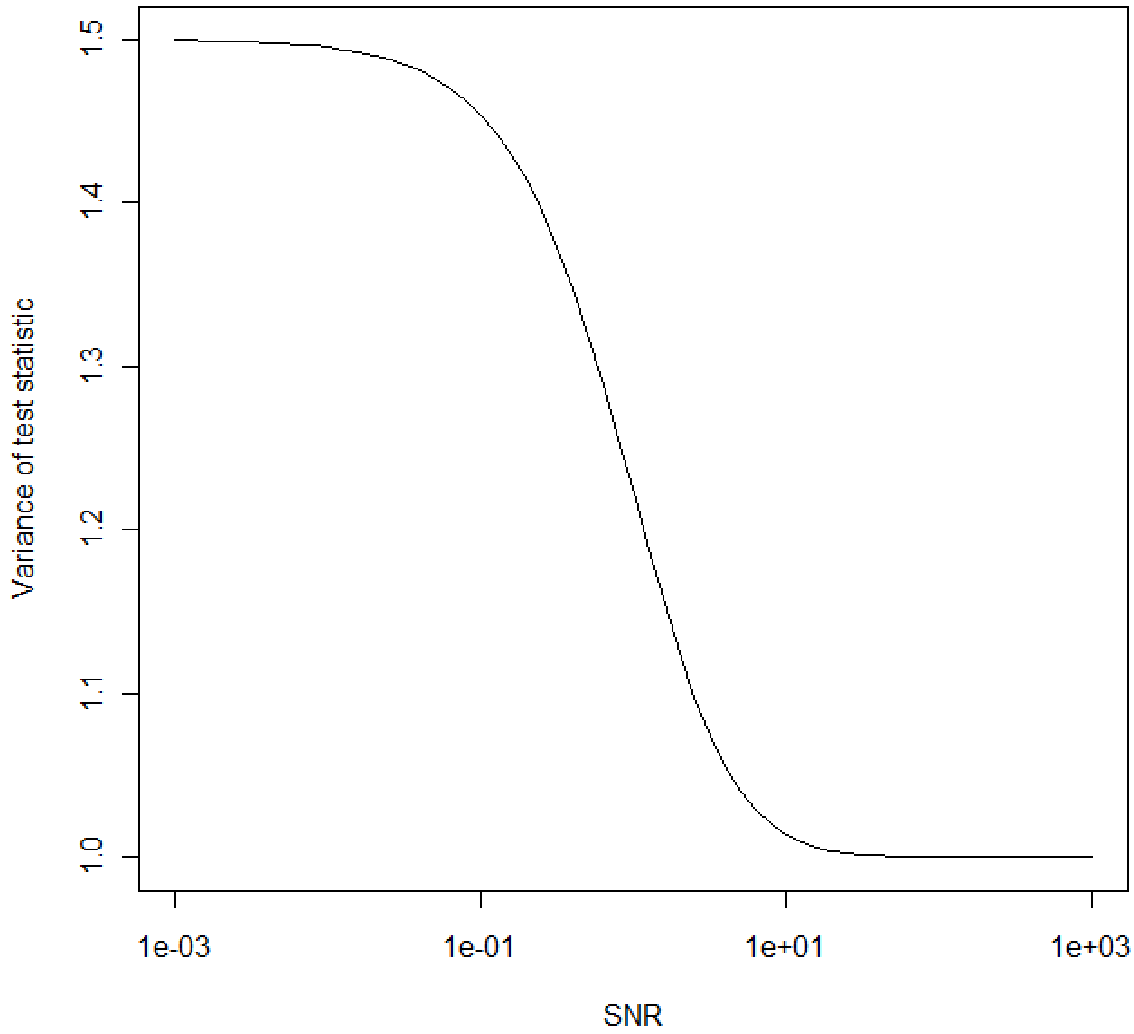

Figure 4.5. Plot of variance of $\sqrt{T}\hat{\rho}(2)$ statistic with different σ_w^2/σ_v^2 (SNR).

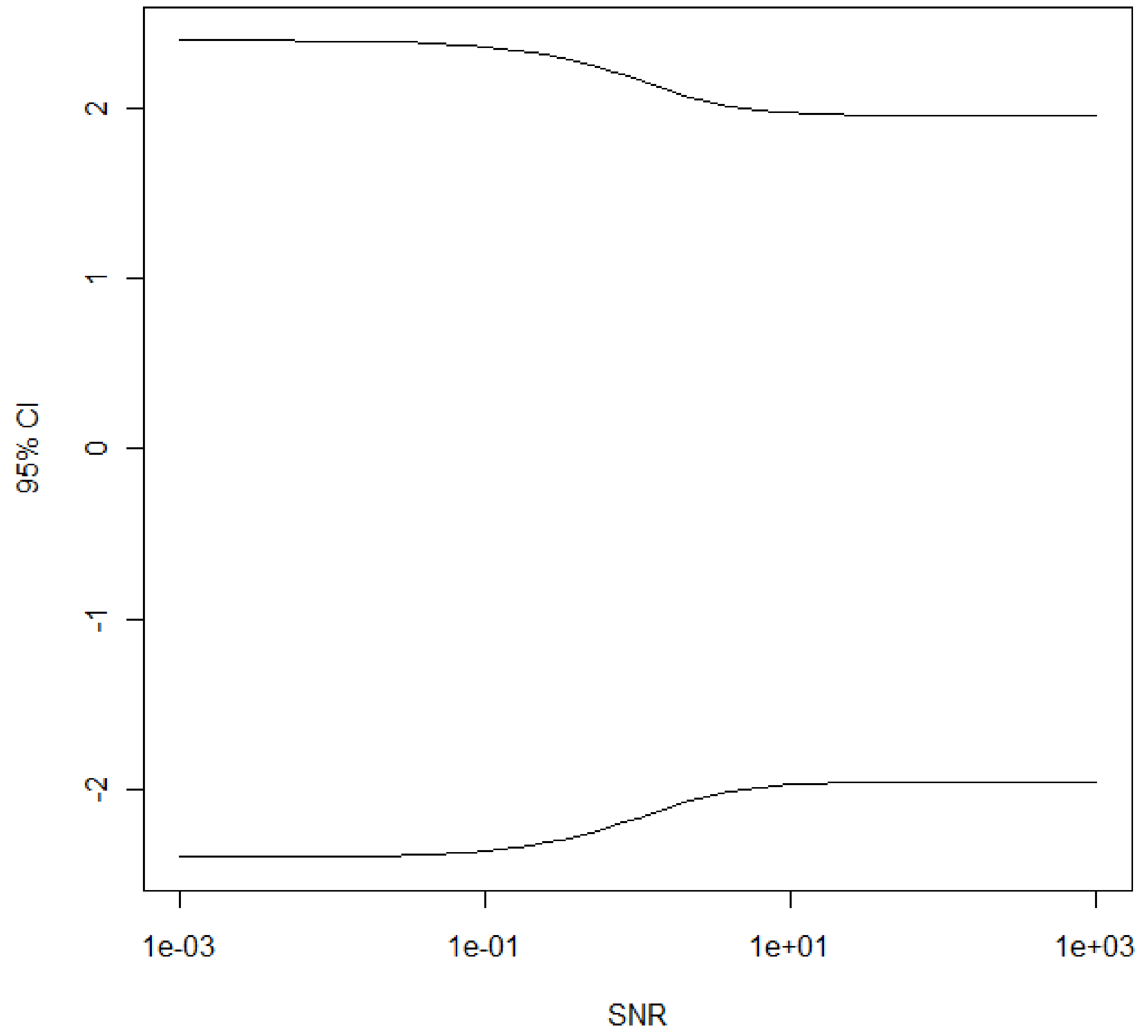

Figure 4.6. Relationship between theoretical critical values of $\sqrt{T}\hat{\rho}(2)$ statistic and σ_w^2/σ_v^2 (SNR) at 95% confidence level.

It is noticed that from figure 4.5, that the variance of the test statistic increases with decreasing SNR with the upper bound $\frac{3}{2}$. If there is prior knowledge on the range of σ_w^2/σ_v^2, for instance, a researcher is confident that in one particular experiment the measurement error will not exceed a certain fraction of the process error, a proper g^2 can be chosen from the above relationship shown in figure 4.5, and hence a theoretical rejection region can be determined by equation 4.14 and figure 4.6. Instead, the use of $g^2 = \frac{3}{2}$ is the most conservative choice in order not to falsely reject the null hypothesis. It is because when g^2 is set to be $\frac{3}{2}$, the type I error of the test is guaranteed to be less than the chosen level α under all circumstances, no matter what the real SNR the data has. As a result, the most conservative rejection rule is:

$$\left|\sqrt{T}\hat{\rho}(2)\right| > z_{1-\alpha/2}\sqrt{1.5} \qquad \text{[equation 4.18]}$$

where $z_{1-\alpha/2}$ is the $(1-\frac{\alpha}{2})$-percentile from $N(0,1)$. In particular, the most conservative 5% critical values are -2.400 and 2.400.

4.6: 3-point Variance-Ratio Test

This test makes use of the straight line relationship between the variance and the time interval of increments, as shown in figure 4.4 earlier. Under the random walk model with measurement error, if variances are sampled with at several lengths of time interval, these variances should form a straight line on the variance-time interval plane. Therefore if the sample variances substantially deviate from a straight line, the null hypothesis of random walk with measurement error is likely to be rejected.

A straight line can always be drawn by joining any two points on the plane. As a result, using two time intervals, as the traditional Variance-Ratio Test uses, is not sufficient to test for straight line. Hence, a third point is needed.

Similar to the procedure in Variance-Ratio Test, the 3-point Variance-Ratio considers the variance estimates of 1-period increments and q-period increments, denoted by $\bar{\sigma}_1^2$ and $\bar{\sigma}_q^2$ respectively (Top left, Figure 4.7). A line is drawn by joining the two variances (Top right, Figure 4.7), and the mid-point of the line is $\frac{\bar{\sigma}_1^2+\bar{\sigma}_q^2}{2}$ at length of increment $\frac{q+1}{2}$ (Bottom left, Figure 4.7). After that the true estimate of variance, called $\bar{\sigma}_{mid}^2$ with interval length $\frac{q+1}{2}$ is obtained (assume q is an odd number to ensure $\frac{q+1}{2}$ is a whole number).

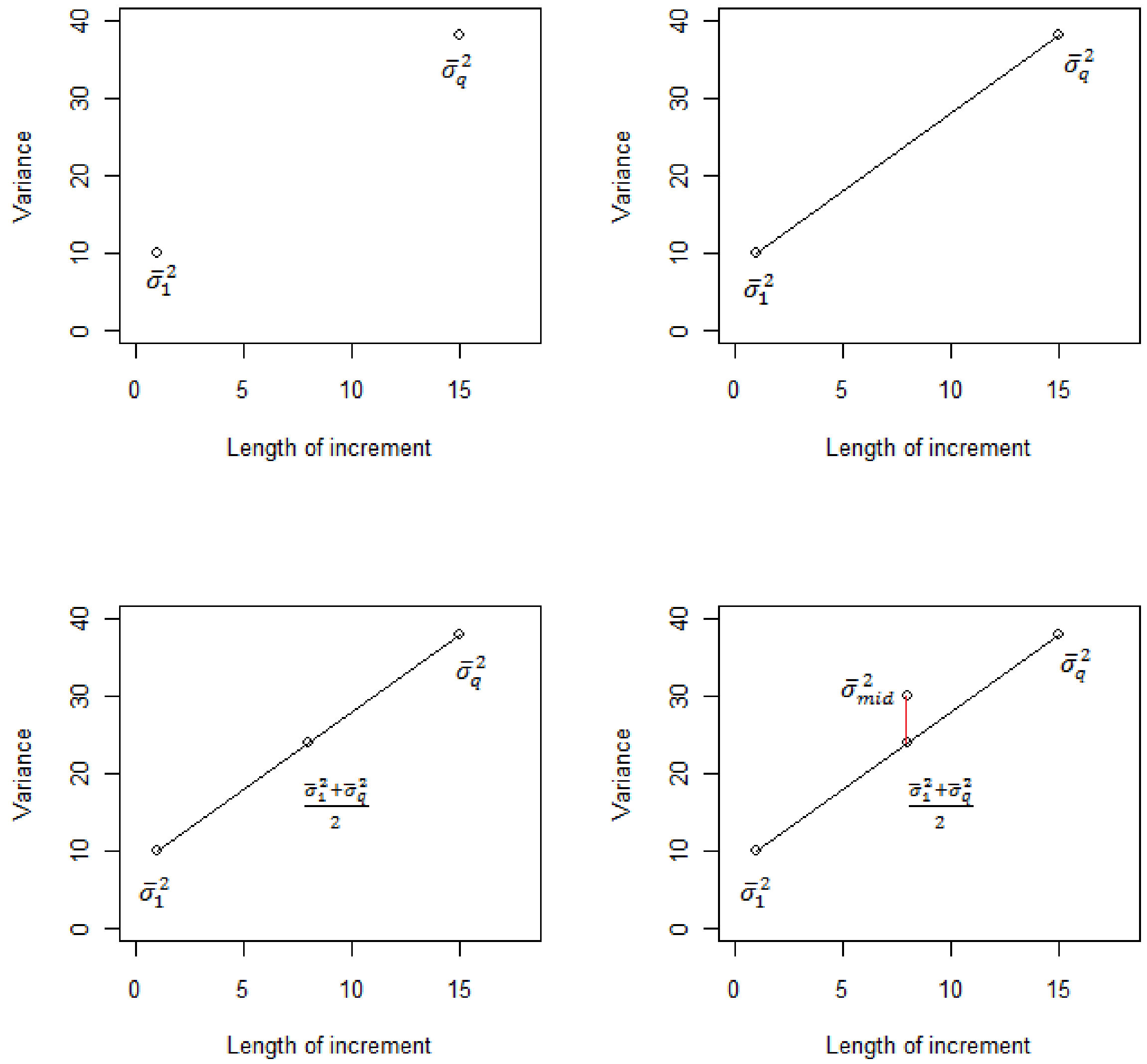

Figure 4.7. Diagram showing the construction of 3-point Variance Ratio test statistic.

Under H_0, the three variance estimates, $\bar{\sigma}_1^2$, $\bar{\sigma}_{mid}^2$, and $\bar{\sigma}_q^2$, should form a straight line. It implies that $\bar{\sigma}_{mid}^2$ should be very close to the mid-point of the linear interpolation of $\bar{\sigma}_1^2$ and $\bar{\sigma}_q^2$ (Bottom right, Figure 4.7). The distance of $\bar{\sigma}_{mid}^2$ from the mid-point, $\frac{\bar{\sigma}_1^2+\bar{\sigma}_q^2}{2}$, can be viewed as the deviation from the straight line hypothesis. This paper suggests the following test statistic:

$$Test\ statistic = ln(\frac{2\bar{\sigma}_{mid}^2}{\bar{\sigma}_1^2+\bar{\sigma}_q^2}) \qquad \text{[equation 4.19]}$$

If the straight line assumption holds, this quantity should be close to zero. The value of q in this test is set to 15 for this study and hence the variances of 1, 8, and 15-period increment are considered.

4.7 Empirical percentiles of test statistics under H_0

It is noticed that for both tests the null distributions and percentiles vary with SNR. Furthermore, although the asymptotic distribution of $\sqrt{T}\hat{\rho}(2)$ statistic is known, the exact distribution of 3-point VR under H_0 is too complex and has yet to be found. As a result, empirical 5% critical values are found by simulation over a wide range of SNR from 10^{-3} to 10^{3}, with 100 the length of sequence. The plots of empirical 5% critical values against SNR, similar to the one shown in figure 4.6 but using simulated data, are generated for both tests. The two plots can be shown in figure 4.8 below.

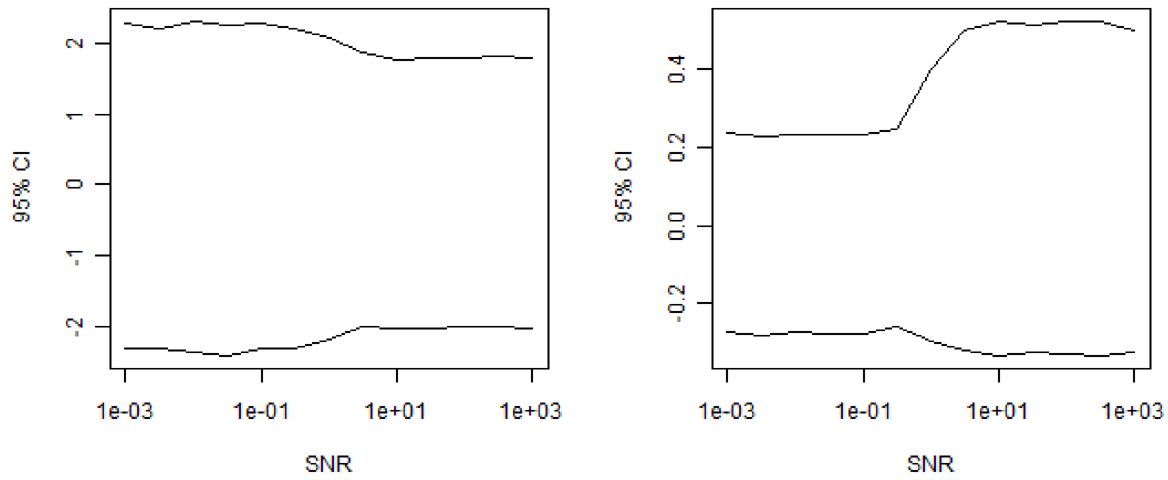

Figure 4.8. the relationship between critical values at 5% significance level and SNR, (left) $\sqrt{T}\hat{\rho}(2)$ statistic, and (right) 3-point Variance-Ratio.

In particular, the 5% empirical critical values of both test statistics at SNR = 0.01 and SNR = 100 are reported and displayed in table 4.2.

Table 4.2. Critical values at SNR=0.01 and SNR=100 at 5% significance level.

	Critical Values at $\alpha = 5\%$	
	SNR=0.01	SNR=100
$\sqrt{T}\hat{\rho}(2)$	(-2.317, 2.310)	(-2.009, 1.805)
3-point Variance-Ratio (q=15)	(-0.275, 0.234)	(-0.329, 0.522)

From the results in table 4.2 and figure 4.8, 5% rejection rules of H_0 can be obtained directly for both tests under the desired SNR.

4.8 Power analysis

After obtaining the empirical 5% critical values, the power of the tests can be analysed through simulation. An AR(1) with measurement error is chosen as the specific alternative and is expressed in the following state-space model:

$$\begin{cases} X_t = \phi X_{t-1} + w_t \ , \ w_t \sim IID(0, \sigma_w^2) \\ Y_t = X_t + v_t \qquad , \ v_t \sim IID(0, \sigma_v^2) \end{cases} \qquad \text{[equation 4.20]}$$

ϕ is fixed at 0.5 in this analysis, with SNR is chosen to be 0.01 and 100. For each value of SNR, 20000 samples of AR(1) with measurement error with length 100 are generated. By using the 5% critical values from table 4.2, power can be calculated as the fraction of trials being rejected using respective critical values under each SNR. Also, the use of the most conservative 5% critical values for $\sqrt{T}\hat{\rho}(2)$ statistic (equation 4.18) is included in the power analysis for comparison. The empirical power of the tests can be shown in table 4.3.

Table 4.3. Empirical power of $\sqrt{T}\hat{\rho}(2)$ and 3-point VR at 5% significance level with SNR = 0.01 and 100.

	SNR = 0.01	SNR = 100
$\sqrt{T}\hat{\rho}(2)$	0.0512	0.2482
3-point Variance-Ratio ($q = 15$)	0.0488	0.05135
$\sqrt{T}\hat{\rho}(2)$ with most conserve bound	0.0437	0.1317

Table 4.3 shows that power of the two tests at SNR = 100 is higher than that of SNR = 0.01. This result agrees with the knowledge about SNR, that the higher the

signal-to-noise ratio is, the more information is retained from the underlying process. The resulting test statistics, which is calculated from the observed process, can provide better judgement at higher SNR.

By comparing the two tests, $\sqrt{T}\hat{\rho}(2)$ statistic has a higher testing power than 3-point Variance-Ratio in both high and low SNR. At low SNR, the power of both tests is around 5%. However, at high SNR, $\sqrt{T}\hat{\rho}(2)$ has a much higher power of about 25%, while 3-point VR does not have a sharp increase in testing power.

In the $\sqrt{T}\hat{\rho}(2)$ statistic, the use of empirical critical values always gives a higher power. At low SNR, the empirical 5% boundary is close to the most conservative one and therefore the power of the test using both critical values is similar. With high SNR, the most conservative boundary is too wide compared to the empirical one and thus a significant drop in testing power is anticipated. However, the use of the most conservative critical values guarantees the type I error is no more than the desired level (5% in this case).

4.9 Section summary

This section explains how existing tests fail with the presence of measurement error and provides two testing methods based on the autocorrelation and variance structure of random walk model with measurement error. It has been shown that the SNR is a key parameter in state-space model and thus plays an important role in power analysis. Both $\sqrt{T}\hat{\rho}(2)$ and 3-point Variance-Ratio can be used to test for random walk with measurement error and the former test provides a larger power on testing against an AR(1) model.

FIVE: TREND IN RANDOM WALK WITH MEASUREMENT ERROR

5.1 Introduction

Under the scenario with measurement error, it is sensible to ask whether a trend exists after accepting the random walk hypothesis, similar to the procedure in section 3. It should be noticed that increments are correlated with their preceding terms due to the existence of measurement error, as explained in equation 4.5. Consequently, the direct use of all the increment terms in the Likelihood Ratio Test, suggested in section 3.2, is not valid with the existence of measurement error. The likelihood function used in section 3.2 is based on independent increments, and thus the joint likelihood is the product of all independent likelihoods (or density). For correlated increments, conditional likelihoods should be considered and the construction of such a test becomes unavoidably cumbersome.

5.2 Likelihood Ratio test

Assuming independent and normal errors, the state-space form of a trended random walk with measurement error is:

$$\begin{cases} X_t = \mu + X_{t-1} + w_t \ , & w_t \sim IID(0, \sigma_w^2) \\ Y_t = X_t + v_t \ , & v_t \sim IID(0, \sigma_v^2) \end{cases} \qquad \text{[equation 5.1]}$$

where μ is the trend. The increment of the observed process, denoted by ΔY_t, is:

$$\begin{aligned} \Delta Y_t &= Y_t - Y_{t-1} \\ &= (X_t + v_t) - (X_{t-1} + v_{t-1}) \\ &= (X_t - X_{t-1}) + v_t - v_{t-1} \\ &= \mu + w_t + v_t - v_{t-1} \end{aligned} \qquad \text{[equation 5.2]}$$

Instead of using all the increments, every other increment (say, all the odd number terms) is selected. The covariance structure of every other increment is as follows:

$$\begin{aligned} & Cov(\Delta Y_t \, , \Delta Y_{t-2}) \\ &= Cov(\mu + w_t + v_t - v_{t-1} \, , \mu + w_{t-2} + v_{t-2} - v_{t-3}) \\ &= 0 \end{aligned} \qquad \text{[equation 5.3]}$$

Note that adding a constant μ has no effect to the covariance structure and therefore it has the same form as shown in equation 4.6. Zero correlation (and also covariance) usually does not imply independence, but under the assumption of H_0 stated at the top of this section, that $\{w_t\}$ and $\{v_t\}$ are serially independent, and mutually independent to each other, it can be deduced that the sequence of every second increment is independent. Also, from equation 5.2, ΔY_t is the sum of 3 independent and normally distributed variables. Thus ΔY_t is also normally distributed, with mean μ and variance $\sigma_w^2 + 2\sigma_v^2$:

$$\Delta Y_t \sim N(\mu\,, \sigma_w^2 + 2\sigma_v^2) \qquad \text{[equation 5.4]}$$

By combining the two consequences, independence and normality of every other increment, the same Likelihood Ratio Test in section 3.2 can be applied. Although it is not desirable to discard half of the information by considering only every other element, this guarantees independence and fulfils the assumptions in Likelihood Ratio Test.

As the test is the same Likelihood Ratio Test as the one shown in section 3.2, the null distribution of test statistic can be referred to figure 3.2.

5.3 Power analysis

It has been shown in section 3.2 that the Likelihood Ratio Test in this specific case is equivalent to a t-test with $H_0: \mu = 0$. As a result, the equivalent test statistic is:

$$t - statistic = \frac{\Delta \bar{Y}_t}{\sigma/\sqrt{T/2}} \qquad \text{[equation 5.5]}$$

where $\Delta \bar{Y}_t$ is the sample mean of every other increments, and $\sigma = \sqrt{\sigma_w^2 + 2\sigma_v^2}$. This test statistic follows standard normal when T is sufficiently large.

The power of the test follows the result from section 3.4 and depends on the ratio $\frac{\mu}{\sigma}$ as described. The variance of the test is the sum of process error and twice the measurement error. Hence a smaller trend can be detected with small variance, while a more significant trend is needed for large σ to reject H_0.

5.4 Section summary

It is shown that the Likelihood Ratio Test suggested in section 3.2 can be directly applied to random walk with measurement error when only every other increment is considered. It is proved by its correlation structure that most the consequences from section 3 about Likelihood Ratio Test still hold in this condition.

SIX: DISCUSSIONS AND LIMITATIONS OF THIS STUDY

This paper evaluates existing testing procedures for the random walk hypothesis and provides alternative methods with the presence of measurement error. However, there are some limitations about these testing methods that users have to be aware of, or take them as a general guidance to these tests.

First, many of the test statistics in this paper, such as Variance-Ratio Test in section 2.3 (and also the modified VR in section 3.3), Likelihood Ratio Test in section 3.2, and the $\sqrt{T}\hat{\rho}(2)$ statistic in section 4.5, increments are assumed to be IID in the null hypothesis. Some may even specify the distribution of the process or measurement error. As mentioned at the beginning of this paper, there are more than one definition of random walk, and IID or normality of increments is the strictest among the three cases. Test statistics are generated by simulation in which IID normal increments are assumed. Therefore, null distributions and rejection criteria are based on such a restricted condition. However in reality, data rarely follows IID assumption and the actual dependency between increments is unknown. The null distributions under the weaker forms of random walk, for instance, independent (but not identically distributed) increments and uncorrelated increments, are usually undetermined and can only be approximated with the IID case. A discrepancy in the null distributions between the stronger and weaker form of random walk leads to testing error and misjudgement. It is known that Variance-Ratio Test can be used to test and derive the null distribution for the weaker form of random walk by heteroskedasticity-consistent methods (White, 1980; White & Domowitz ,1984). Lo & MacKinlay (1988) can be referred for further details.

Under the presence of measurement error, a random walk can be represented in a state-space model. The persistency of such an extra source of randomness in the observed sequence hinders the effect of the underlying process, and thus testing methods for random walk fail. The Kalman Filter (Kalman, 1960) is an algorithm to produce an optimal estimate of the underlying state process by sequence of predictions and updating processes. This method assumes linear models while the Extended Kalman Filter (EKF) can handle non-linear models. These methods are useful but often considered as abstruse and too computational-intensive to be implemented for hypothesis testing. As this paper aims to provide some intuitive testing methods that relies on simple estimations and rejection rules and can be readily applied by every biologist, the $\sqrt{T}\hat{\rho}(2)$ statistic and 3-point Variance-Ratio in section 4 are thus introduced. However $\sqrt{T}\hat{\rho}(2)$ considers only one autocorrelation coefficient but not higher-order autocorrelations. There may be some time series with zero $\rho(2)$ but are not a random walk in nature. A more comprehensive test should be constructed in the future to capture higher-order autocorrelation. The exact or asymptotic null distribution of the 3-point Variance-Ratio Test is unknown and can only be generated empirically. Hence, it is difficult to understand and analyse the characteristics of the test rigorously.

Another limitation under measurement error is the estimation of signal-to-noise ratio (SNR). In the Kalman Filter environment, the SNR can be estimated simultaneously with a linear model, for instance, by the Newton-Raphson algorithm. However, a model is assumed before estimation and a more robust model-invariant estimate of SNR is needed. Furthermore, in the 3-point Variance-Ratio test critical values depend on SNR and a prior knowledge of SNR is required. Unlike $\sqrt{T}\hat{\rho}(2)$ statistic, in which the most conservative boundary can be used without estimating SNR, 3-point Variance-Ratio behaves very differently in high and low SNR and type I error cannot be well controlled. Ideally, a robust testing method should be developed that works well in a wide range of SNR.

Apart from SNR estimation, the tests for random walk with measurement error have a rather low power when the length of sequence (T) is about 100. It is expected that power will increase with the length sequence. By simulation, it can be shown that $\sqrt{T}\hat{\rho}(2)$ statistic can achieve a power of 96% when $T = 1000$ in testing against AR(1) model with $\phi = 0.5$. However, having over a thousand data points in a biological dataset is usually too optimistic. As a result, testing power in general remains low for small sample size.

Finally, many of the biological data, that would be interesting to tell for deviations from random walks, are multidimensional or compositional data. Especially in genetics data, allele frequencies are reported in proportional form which ranges from 0 to 1. Even after proper transformations (logit or probit transform), the increments may not be normally distributed. With three or more categories, the compositional data forms a multivariate random walk. There is not much context on testing random walk hypothesis for multivariate or compositional data and more investigation have to be done in the near future.

SEVEN: CONCLUSION

The Q-statistic (and also Q'-statistic) and Variance Ratio Test are introduced to test for random walk hypothesis in section 2. The power of two tests against AR(1) process is analysed and it is shown that overall the Variance-Ratio Test has higher testing power than the Q'-statistic.

This paper also shows that the Likelihood Ratio Test and modified Variance-Ratio are valid tests for trends in random walks under IID assumption. In particular, the Likelihood Ratio Test is a direct test on μ and equivalent to a t-test for the mean. Therefore it has a higher testing power. The ratio $\frac{\mu}{\sigma}$ plays an important role in determining the power of the tests; the testing power is higher when $\frac{\mu}{\sigma}$ is large. This result is particularly useful when testing genetic drift.

Because of the autocorrelation structure, the Q'-statistic and Variance Ratio test are no longer reliable in identifying random walk process with presence of measurement error. Consequently, the $\sqrt{T}\hat{\rho}(2)$ statistic and 3-point Variance-Ratio are introduced and served as alternatives to existing methods. The $\sqrt{T}\hat{\rho}(2)$ statistic asymptotically follows the normal distribution, while the distributional form of the 3-point Variance-Ratio can only be estimated through simulation. Furthermore, the signal-to-noise (SNR) ratio affects the power and critical values of the tests; higher SNR values retain more information about the underlying sequence thus has a higher testing power.

By taking every other element from the increments, Likelihood Ratio Test suggested in section 3.2 can be applied to test for trends in random walk with measurement error. Furthermore, under the assumption of IID increments, the Likelihood Ratio Test is equivalent to a t-test for mean with sufficient large sample size.

REFERENCE

Bod, P., Blitz, D., Franses, P.H. & Kluitman, R. (2002) An unbiased variance estimator for overlapping returns. *Applied Financial Economics*, 12, 155-158.

Bolker, B.M. (2008) *Ecological Models and Data in R*. Princeton University Press. pp.348-354.

Brunsdon, T.M. & Smith, T.M.F. (1998) The Time Series Analysis of Compositional Data. *Journal of Official Statistics*. 14(3), 237-253.

Campbell, J.Y., Lo, A.W. & MacKinlay, A.C. (1997) *The Econometrics of Financial Markets.* Princeton University Press. pp.27-82.

Casella, G. & Berger, R.L. (2001) *Statistical Inference*. 2nd edition. Duxbury Press. pp.312-323, pp.374-278.

Chan, N.H. (2002) *Time Series: Application to Finance*. 1st edition. Wiley-Interscience. pp.137-150.

Hartl, D.L. & Clark, A.G. (1997) *Principles of Population Genetics*. 3rd edition. Sinauer Associates. pp.267-313.

Lande, R., Engen, S., Saether, B.E. (2003) *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press. pp.29-34.

Shumway, R.H. & Stoffer, D.S. (2010) *Time Series Analysis and Its Applications: With R Examples*. 3rd edition. Springer, pp.86-143, pp.335-341.

Taylor, H.M. & Karlin, S. (1998*) An Introduction To Stochastic Modeling.* 3rd edition. Academic Press. pp.524-534.

Welch, G. & Bishop, G. (2006) An Introduction to Kalman Filter. Department of Computer Science, University of North Carolina at Chapel Hill.

ACKNOWLEDGEMENTS

Writing this paper is probably one of the most challenging tasks in my study and I could not have finished it on my own.

First I would like to thank my supervisor, Professor Austin Burt, for his patience and professional advice throughout the whole research project. This study would not have reached this stage without his guidance. It is truly a rewarding experience to work with him and his research group. I would also like to thank my fellow classmates in Quantitative Biology MSc class of 2012 for bringing us fruitful interactions and discussions. In particular, our course director, Dr Dan Reuman, has helped organize our course and injected new ideas to the class. I am grateful for many supports and assistance from the Silwood Park community.

Last, but not least, I would like to thank my family who are six thousand miles away in Hong Kong for their encouragement and understanding.