2.4.3 Variance Ratios

An important property of all three random walk hypotheses is that the variance of random walk increments must be a linear function of the time interval.[8] For example, under RW1 for log prices where continuously compounded returns $r_t \equiv \log P_t - \log P_{t-1}$ are IID, the variance of $r_t + r_{t-1}$ must be twice the variance of r_t. Therefore, the plausibility of the random walk model may be checked by comparing the variance of $r_t + r_{t-1}$ to twice the variance of r_t.[9] Of course, in practice these will not be numerically identical even if RW1 were true, but their ratio should be statistically indistinguishable from one. Therefore, to construct a statistical test of the random walk hypothesis using variance ratios, we require their sampling distribution under the random walk null hypothesis.

Population Properties of Variance Ratios

Before deriving such sampling distributions, we develop some intuition for the population values of the variance ratio statistic under various scenarios. Consider again the ratio of the variance of a two-period continuously compounded return $r_t(2) \equiv r_t + r_{t-1}$ to twice the variance of a one-period return r_t, and for the moment let us assume nothing about the time series of returns other than stationarity. Then this variance ratio, which we write as VR(2), reduces to:

$$\begin{aligned} \mathrm{VR}(2) &= \frac{\mathrm{Var}[r_t(2)]}{2\,\mathrm{Var}[r_t]} = \frac{\mathrm{Var}[r_t + r_{t-1}]}{2\,\mathrm{Var}[r_t]} \\ &= \frac{2\,\mathrm{Var}[r_t] + 2\,\mathrm{Cov}[r_t, r_{t-1}]}{2\,\mathrm{Var}[r_t]} \\ \mathrm{VR}(2) &= 1 + \rho(1), \end{aligned} \tag{2.4.18}$$

where $\rho(1)$ is the first-order autocorrelation coefficient of returns $\{r_t\}$. For any stationary time series, the population value of the variance ratio statistic VR(2) is simply one plus the first-order autocorrelation coefficient. In particular, under RW1 all the autocorrelations are zero, hence VR(2)=1 in this case, as expected.

In the presence of positive first-order autocorrelation, VR(2) will exceed one. If returns are positively autocorrelated, the variance of the sum of two

[8]This linearity property is more difficult to state in the case of RW2 and RW3 because the variances of increments may vary through time. However, even in these cases the variance of the sum must equal the sum of the variances, and this is the linearity property which the variance ratio test exploits. We shall construct tests of all three hypotheses below.

[9]Many studies have exploited this property of the random walk hypothesis in devising empirical tests of predictability; recent examples include Campbell and Mankiw (1987), Cochrane (1988), Faust (1992), Lo and MacKinlay (1988), Poterba and Summers (1988), Richardson (1993), and Richardson and Stock (1989).

one-period returns will be larger than the sum of the one-period return's variances; hence variances will grow faster than linearly. Alternatively, in the presence of negative first-order autocorrelation, the variance of the sum of two one-period returns will be smaller than the sum of the one-period return's variances; hence variances will grow slower than linearly.

For comparisons beyond one- and two-period returns, higher-order autocorrelations come into play. In particular, a similar calculation shows that the general q-period variance ratio statistic VR(q) satisfies the relation:

$$\mathrm{VR}(q) \equiv \frac{\mathrm{Var}[r_t(q)]}{q \cdot \mathrm{Var}[r_t]} = 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\rho(k), \tag{2.4.19}$$

where $r_t(k) \equiv r_t + r_{t-1} + \cdots + r_{t-k+1}$ and $\rho(k)$ is the kth order autocorrelation coefficient of $\{r_t\}$. This shows that VR(q) is a particular *linear* combination of the first $k-1$ autocorrelation coefficients of $\{r_t\}$, with linearly declining weights.

Under RW1, (2.4.19) shows that for all q, VR(q)=1 since in this case $\rho(k)$=0 for all $k \geq 1$. Moreover, even under RW2 and RW3, VR(q) must still equal one as long as the variances of r_t are finite and the "average variance" $\sum_{t=1}^{T} \mathrm{Var}[r_t]/T$ converges to a finite positive number. But (2.4.19) is even more informative for alternatives to the random walk because it relates the behavior of VR(q) to the autocorrelation coefficients of $\{r_t\}$ under such alternatives. For example, under an AR(1) alternative, $r_t = \phi r_{t-1} + \epsilon_t$, (2.4.19) implies that

$$\begin{aligned} \mathrm{VR}(q) &= 1 + 2\sum_{k=1}^{q-1}\left(1 - \frac{k}{q}\right)\phi^k \\ &= 1 + \frac{2}{1-\phi}\left[\phi - \frac{\phi^q}{q} - \frac{\phi - \phi^q}{q(1-\phi)}\right]. \end{aligned}$$

Relations such as this are critical for constructing alternative hypotheses for which the variance ratio test has high and low power, and we shall return to this issue below.

Sampling Distribution of $\widehat{\mathrm{VD}}(q)$ and $\widehat{\mathrm{VR}}(q)$ under RW1

To construct a statistical test for RW1 we follow the exposition of Lo and MacKinlay (1988) and begin by stating the null hypothesis H_0 under which the sampling distribution of the test statistics will be derived.[10] Let p_t denote the log price process and $r_t \equiv p_t - p_{t-1}$ continuously compounded returns.

[10]For alternative expositions see Campbell and Mankiw (1987), Cochrane (1988), Faust (1992), Poterba and Summers (1988), Richardson (1993), and Richardson and Stock (1989).

Then the null hypothesis we consider in this section is[11]

$$H_0: \quad r_t = \mu + \epsilon_t, \qquad \epsilon_t \text{ IID } \mathcal{N}(0, \sigma^2).$$

Let our data consist of $2n+1$ observations of log prices $\{p_0, p_1, \ldots, p_{2n}\}$, and consider the following estimators for μ and σ^2:

$$\hat{\mu} \equiv \frac{1}{2n}\sum_{k=1}^{2n}(p_k - p_{k-1}) = \frac{1}{2n}(p_{2n} - p_0) \tag{2.4.20}$$

$$\hat{\sigma}_a^2 \equiv \frac{1}{2n}\sum_{k=1}^{2n}(p_k - p_{k-1} - \hat{\mu})^2 \tag{2.4.21}$$

$$\hat{\sigma}_b^2 \equiv \frac{1}{2n}\sum_{k=1}^{n}(p_{2k} - p_{2k-2} - 2\hat{\mu})^2. \tag{2.4.22}$$

Equations (2.4.20) and (2.4.21) are the usual sample mean and variance estimators. They are also the *maximum-likelihood* estimators of μ and σ^2 (see Section 9.3.2 in Chapter 9). The second estimator $\hat{\sigma}_b^2$ of σ^2 makes use of the random walk nature of p_t: Under RW1 the mean and variance of increments are linear in the increment interval, hence the σ^2 can be estimated by one-half the sample variance of the increments of even-numbered observations $\{p_0, p_2, p_4, \ldots, p_{2n}\}$.

Under standard asymptotic theory, all three estimators are strongly consistent: Holding all other parameters constant, as the total number of observations $2n$ increases without bound the estimators converge almost surely to their population values. In addition, it is well known that $\hat{\sigma}_a^2$ and $\hat{\sigma}_b^2$ possess the following normal limiting distributions (see, for example, Stuart and Ord [1987]):

$$\sqrt{2n}\,(\hat{\sigma}_a^2 - \sigma^2) \overset{a}{\sim} \mathcal{N}(0, 2\sigma^4) \tag{2.4.23}$$

$$\sqrt{2n}\,(\hat{\sigma}_b^2 - \sigma^2) \overset{a}{\sim} \mathcal{N}(0, 4\sigma^4). \tag{2.4.24}$$

However, we seek the limiting distribution of the *ratio* of the variances. Although it may readily be shown that the ratio is also asymptotically normal with unit mean under RW1, the variance of the limiting distribution is not apparent since the two variance estimators are clearly *not* asymptotically uncorrelated.

But since the estimator $\hat{\sigma}_a^2$ is asymptotically efficient under the null hypothesis RW1, we may use Hausman's (1978) insight that the asymptotic

[11]We assume normality only for expositional convenience—the results in this section apply much more generally to log price processes with IID increments that possess finite fourth moments.

variance of the difference of a consistent estimator and an asymptotically efficient estimator is simply the difference of the asymptotic variances.[12] If we define the variance difference estimator as $\widehat{\mathrm{VD}}(2) \equiv \hat{\sigma}_b^2 - \hat{\sigma}_a^2$, then (2.4.23), (2.4.24), and Hausman's result implies:

$$\sqrt{2n}\,\widehat{\mathrm{VD}}(2) \overset{a}{\sim} \mathcal{N}(0, 2\sigma^4). \tag{2.4.25}$$

The null hypothesis H can then be tested using (2.4.25) and any consistent estimator $2\hat{\sigma}^4$ of $2\sigma^4$ (for example, $2(\hat{\sigma}_a^2)^2$): Construct the standardized statistic $\widehat{\mathrm{VD}}(2)/\sqrt{2\hat{\sigma}^4}$ which has a limiting standard normal distribution under RW1, and reject the null hypothesis at the 5% level if it lies outside the interval $[-1.96, 1.96]$.

The asymptotic distribution of the two-period variance ratio statistic $\widehat{\mathrm{VR}}(2) = \hat{\sigma}_b^2/\hat{\sigma}_a^2$ now follows directly from (2.4.25) using a first-order Taylor approximation or the delta method (see Section A.4 of the Appendix):[13]

$$\widehat{\mathrm{VR}}(2) \equiv \frac{\hat{\sigma}_b^2}{\hat{\sigma}_a^2}, \qquad \sqrt{2n}\,(\widehat{\mathrm{VR}}(2) - 1) \overset{a}{\sim} \mathcal{N}(0, 2). \tag{2.4.26}$$

The null hypothesis H_0 can be tested by computing the standardized statistic $\sqrt{2n}(\widehat{\mathrm{VR}}(2)-1)/\sqrt{2}$ which is asymptotically standard normal—if it lies outside the interval $[-1.96, 1.96]$, RW1 may be rejected at the 5% level of significance.

Although the variance ratio is often preferred to the variance difference because the ratio is scale-free, observe that if $2(\hat{\sigma}_a^2)^2$ is used to estimate $2\sigma^4$, then the standard significance test of VD=0 for the difference will yield the same inferences as the corresponding test of VR−1=0 for the ratio since:

$$\frac{\sqrt{2n}\widehat{\mathrm{VD}}(2)}{\sqrt{2\hat{\sigma}_a^4}} = \frac{\sqrt{2n}(\hat{\sigma}_b^2 - \hat{\sigma}_a^2)}{\sqrt{2}\,\hat{\sigma}_a^2} = \frac{\sqrt{2n}(\widehat{\mathrm{VR}}(2) - 1)}{\sqrt{2}} \sim \mathcal{N}(0, 1). \tag{2.4.27}$$

Therefore, in this simple context the two test statistics are equivalent. However, there are other reasons that make the variance ratio more appealing

[12]Briefly, Hausman (1978) exploits the fact that any asymptotically efficient estimator of a parameter θ, say $\hat{\theta}_e$, must possess the property that it is asymptotically uncorrelated with the difference $\hat{\theta}_a - \hat{\theta}_e$, where $\hat{\theta}_a$ is any other estimator of θ. If not, then there exists a linear combination of $\hat{\theta}_e$ and $\hat{\theta}_a - \hat{\theta}_e$ that is more efficient than $\hat{\theta}_e$, contradicting the assumed efficiency of $\hat{\theta}_e$. The result follows directly, then, since:

$$\begin{aligned} \mathrm{aVar}[\hat{\theta}_a] = \mathrm{aVar}[\hat{\theta}_e + \hat{\theta}_a - \hat{\theta}_e] &= \mathrm{aVar}[\hat{\theta}_e] + \mathrm{aVar}[\hat{\theta}_a - \hat{\theta}_e] \\ \Rightarrow \quad \mathrm{aVar}[\hat{\theta}_a - \hat{\theta}_e] &= \mathrm{aVar}[\hat{\theta}_a] - \mathrm{aVar}[\hat{\theta}_e]. \end{aligned}$$

where $\mathrm{aVar}[\cdot]$ denotes the asymptotic variance operator.

[13]In particular, apply the delta method to $f(\hat{\theta}_1, \hat{\theta}_2) \equiv \hat{\theta}_1/\hat{\theta}_2$ where $\hat{\theta}_1 \equiv \hat{\sigma}_b^2 - \hat{\sigma}_a^2$, $\hat{\theta}_2 \equiv \hat{\sigma}_a^2$, and observe that $\hat{\sigma}_b^2 - \hat{\sigma}_a^2$ and $\hat{\sigma}_a^2$ are asymptotically uncorrelated because $\hat{\sigma}_a^2$ is an efficient estimator.