NONPARAMETRIC STATISTICAL TESTS FOR THE RANDOM WALK IN STOCK PRICES

Patricia Cerrito, Dennis Olson, and
Krzysztof Ostaszewski

ABSTRACT

A recently developed nonparametric test for independence in time series data asserts that two random variables, such as current and past stock returns, are independent if and only if their joint probability density equals the product of their marginal densities. We introduce a variant of the test statistic that is not degenerate under the null hypothesis and also use it to test whether random variables are identically distributed. The nonparametric tests show that recent United States' and Canadian stock index returns are neither independently nor identically distributed. The random walk is rejected for both daily and monthly data.

Advances in Quantitative Analysis of Finance and Accounting, Volume 6, pages 27–36.

ISBN: 0-7623-0327-1

I. INTRODUCTION

Robinson (1991) recently proposed an entropy-based nonparametric test for independence in time series data based upon the Kullback-Leibler information criterion. Two random variables, such as current and past stock returns, are independent if and only if their joint probability density equals the product of their marginal densities. The joint and marginal densities can be estimated by nonparametric kernel density estimation and with this information it is also possible to test whether two variables are identically distributed. Specifically, two random variables are identically distributed if and only if the marginal densities are equal. Rejection of either of the two null hypotheses of independence or identical distribution constitutes a rejection of the random walk hypothesis. This entropy-based test is more likely to detect dependence than other techniques, yet it will not reject the random walk hypothesis in instances where the data really reject a specific parametric form or some assumption about the returns distribution.

A growing body of literature suggests that stock returns do not follow a random walk, except perhaps in the short-run when noise appears to mask predictable patterns. For example, Lo and MacKinlay (1988) found positive serial correlation in weekly stock returns and rejected the random walk using a variance ratio test. Fama and French (1988) found negative serial correlation in stock returns over three to five year time horizons using regression-based tests for dependence. Both of these tests assume linearity and normality, but less restrictive models like McQueen and Thorley's (1991) modified runs test using a Markov chain and Durlauf's (1991) test for nonlinearity in the spectral distribution function also indicate stock return predictability. Kim, Nelson, and Startz (1991) further generalize random walk tests by randomizing the data before applying variance ratio and regression tests that do not assume normality. They show that the previous evidence for mean reversion has been overstated. However, they find that returns in recent years display mean aversion or persistence, which might be interpreted as evidence against the random walk. While randomization tests using variance ratios have considerable power and could be used to test the random walk hypothesis, Robinson's (1991) test appears to be more general, computationally feasible, and a more natural test for the random walk hypothesis. Another possibility is to test for deterministic chaos or predictability in time series data using a BDS statistic developed by Brock, Dechert, and Scheinkman. It is a test for independence based upon the correlation dimension. Brock, Hsieh, and LeBaron (1991) note that the BDS statistic is not asymptotically consistent, meaning that it may not detect all deviations away from an independent identical distribution. They suggest that Robinson's test may turn out to be preferable to the BDS statistic, but point out that the two tests may complement each other.

Robinson (1991) concludes from his entropy-based test that daily exchange rate movements are essentially random. Other work using kernel density estimation by Sentana and Wadhwani (1991) indicates some predictability for Japanese stock returns, while Diebold and Nason (1990) fail to reject the random walk for various weekly spot exchange rate series. Prescott and Stengos (1990) find that nonparametric density estimates of weekly gold returns have within-sample predictive power, but no out-of-sample forecasting ability. More recently, Frennberg and Hansson (1993) used variance ratios to show that monthly and annual Swedish stock returns did not follow a random walk over the years 1919 to 1990. Also using variance ratio tests, MacDonald and Power (1993) showed that the random walk could not be rejected for weekly returns for many United Kingdom stocks during the 1980s. The literature to date suggests that financial time series are generally random at high frequencies, such as for daily stock returns. Lower frequency data, such as monthly and annual stock returns, show greater predictability.

The purpose of this paper is to develop a general entropy-based nonparametric test for the random walk hypothesis. Robinson's (1991) proposed test for independence is improved by introducing a variant of the test statistic that is not degenerate under the null hypothesis. Also, nonparametric density estimators are used to test for an identical distribution—the second and often neglected part of the random walk hypothesis. The test is illustrated for United States and Canadian stock returns using both high frequency daily data and low frequency monthly data.

II. NONPARAMETRIC DENSITY ESTIMATION

Let $x_1, x_2, x_3, \ldots$ be the sequence of annual, monthly, or daily returns for the index considered. Returns are calculated by logarithmic differencing the levels of the indexes (I_t) from one period to the next, that is $x_t = In(I_t) - In(I_{t-1})$. Assume that the returns series forms a discrete, stationary, strong mixing, time Markov process and that each x_t is a continuous random variable with common marginal density $f_1(x)$. Denote $f(x, y) = f(x_t, x_{t+1})$ as the common bivariate density for (x_t, x_{t+1}). For forecasting purposes, the transitional function is

$$p(x, y) = \frac{f(x, y)}{f_1(x)} \tag{1}$$

We want to find a robust estimate for $p(x, y)$, which is a conditional density function representing the best forecast of tomorrow's return given the information available today. Due to the well-known phenomena of serial correlation and time varying volatilities, we do not expect the data to be independent over time. Hence, we use nonparametric density estimation, as described in Silverman (1986) and extended to an arbitrary strong mixing process by Ahmad (1982), to estimate f_1 and $f(x, y)$. For any arbitrary density function $h(x)$, its kernel density estimate is

$$\hat{h}(x) = \frac{1}{na_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{a_n}\right), \tag{2}$$

where $X_1, \ldots, X_n$ is a sample of n data points from the strong mixing process considered, and x is the point at which the kernel density function is evaluated. Also, a_n is a constant (or a sequence of constants in the more general case) representing bandwidth and satisfying the property $a_n \to 0$ as $n \to \infty$. It is also called window width, or the smoothing parameter since it controls the smoothness of the kernel density by setting the size of the neighborhood about x over which local averaging of observations occurs. The kernel function $K(x)$ is a probability density function with finite mean and variance satisfying the condition $\int_{-\infty}^{+\infty} K(x)dx = 1$. Several functions satisfy these criteria and they represent alternative weighting schemes for local averaging within the bandwidth about the point x. The choice of kernel functions is somewhat arbitrary, but we prefer the optimal or Epanechnikov kernel. Samiaddin and El-Sayyad (1990) show that it is the only admissible kernel, meaning that no other kernel estimator is more efficient. Also, since calculation of the optimal kernel does not require exponentiation, there is a considerable savings in computational time relative to using the more popular Gaussian kernel.

The bivariate density, $h(x, y)$, is estimated by

$$\hat{h}(x, y) = \frac{1}{na_n^2} \sum_{j=1}^{n} K\left(\frac{x - X_j}{a_n}, \frac{y - Y_j}{a_n}\right), \tag{3}$$

where $K(x, y)$ is a known bivariate kernel density function.

III. TESTING FOR INDEPENDENT IDENTICAL DISTRIBUTIONS

Parametric tests for dependence such as the chi-square goodness of fit test and tests for correlation involving the F statistic are sensitive to deviations away from normality. Since financial time series data are generally characterized by nonnormality and nonlinearity, we wish to devise a test for dependence that holds for any returns distribution without reference to an equilibrium pricing model. This is accomplished using Robinson's (1991) test for independence. It states that two random variables x and y are independent if and only if their joint bivariate probability density function $f(x, y)$ equals the product of the marginal densities $g_1(x)$ and $g_2(y)$ for all x and y. The nonparametric hypothesis test is

$$H_0\colon f(x, y) = g_1(x)g_2(y), \tag{4}$$

versus

$$H_1\colon f(x, y) \neq g_1(x)g_2(y). \tag{5}$$

Note that there is an implicit assumption of stationarity that must be met for this test to be valid. If the data are not stationary it is not possible to estimate either the joint or marginal densities. Also, the test does not provide information about the form of dependence. If the null hypothesis is rejected we know that the data follow a Markov chain rather than a random walk. Further testing is necessary to discover the order of the Markov chain.

There are two difficulties in testing the independence hypothesis. First, data used to estimate the joint density function are generally the same data that are used to estimate the marginal densities. This problem can be overcome by assuming homogeneity or by using two different sets of data. We use different data points to estimate the joint and marginal densities and emphasize this fact by denoting the marginal densities by $g_1(x)$ and $g_2(y)$ rather than the more natural notation of $f_1(x)$ and $f_2(y)$. The second problem is that the most natural test statistic

$$\delta = \iint [f(x, y) - g_1(x)g_2(y)]^2 dx\, dy \tag{6}$$

is degenerate under the null hypothesis. Ahmad and Cerrito (1993) proposed an estimator $\hat{\delta}$ which eliminates the problem of degeneracy. The estimator can be defined by taking two random samples of (x_t, x_{t+1}) points. We denote m as the number of data points used to estimate the joint density. The data are represented by the values $(x_1, y_1), \ldots, (x_m, y_m)$. The n data points used to estimate the marginal densities are given by $(u_1, w_1), \ldots, (u_n, w_n)$. The estimate $\hat{f}_m$ of the joint density is given by

$$\hat{f}_m(x, y) = \frac{1}{ma_m^2} \sum_{i=1}^{m} K\left(\frac{x - X_i}{a_m}, \frac{y - Y_i}{a_m}\right), \tag{7}$$

where K is the known bivariate kernel density and the marginal densities are estimated by

$$\hat{g}_1(u) = \frac{1}{nb_n} \sum_{j=1}^{n} K_1\left(\frac{u - U_j}{b_n}\right), \tag{8}$$

$$\hat{g}_2(u) = \frac{1}{nb_n} \sum_{j=1}^{n} K_1\left(\frac{w - W_j}{b_n}\right), \tag{9}$$

In this framework, $K = K_1^2$ and an estimate of δ can be defined by

$$\hat{\delta} = \frac{1}{ma_m^2} \sum_{i=1}^{m} \sum_{j=1}^{m} K\left(\frac{X_i - X_j}{a_m}, \frac{Y_i - Y_j}{a_m}\right)$$

$$+\frac{1}{n^2 b_n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} K\left(\frac{U_i - U_j}{b_n}, \frac{W_i - W_j}{b_n}\right)$$

$$-\frac{1}{m a_m^2 m \sum_{j=1}^{n} D_{j,n}(\gamma)} \sum_{i=1}^{m} \sum_{j=1}^{n} D_{j,n}(\gamma)\, K\left(\frac{X_i - U_j}{a_m}, \frac{Y_i - W_j}{a_m}\right)$$

$$-\frac{1}{n b_n^2 \sum_{i=1}^{m} C_{i,m}(\gamma)} \sum_{i=1}^{m} \sum_{j=1}^{n} C_{i,m}(\gamma)\, K\left(\frac{U_i - X_j}{b_n}, \frac{W_i - Y_j}{b_n}\right), \tag{10}$$

where the bandwidth parameters $\{a_m\}$ and $\{b_n\}$ are real numbers satisfying the condition that $a_m^2 \to 0$ as $m a_m^2 \to \infty$ and $b_n^2 \to 0$ as $n b_n^2 \to \infty$. Also, $C_{i,m}(\gamma)$ and $D_{j,n}(\gamma)$ are real numbers that vary with sample size. They are included by the researcher to prevent degeneracy under the null hypothesis and defined as

$$\frac{m \sum_{i=1}^{m} C_{i,m}(\gamma)}{\left(\sum_{i=1}^{m} C_{i,m}(\gamma)\right)^2} \to C^2(\gamma) > 1 \text{ as } m \to \infty, \text{ and} \tag{11}$$

$$\frac{n \sum_{i=1}^{n} D_{i,n}(\gamma)}{\left(\sum_{j=1}^{n} D_{j,n}(\gamma)\right)^2} \to D^2(\gamma) > 1 \text{ as } n \to \infty. \tag{12}$$

Ahmad and Cerrito (1993) show that $\hat{\delta}$ is a consistent estimator, such that the value of $(\hat{\delta} - \delta)\sqrt{\frac{mn}{m+n}}$ is asymptotically normal with a mean of zero and a variance (σ^2) of

$$\sigma^2 = 4\left\{\int f^3 - \left(\int f^2\right)^2\right\} + 4\left\{\int (g_1 g_2)^3 - \left[\int (g_1 g_2)^2\right]^2\right.$$

$$+ [C^2(\gamma) + D^2(\gamma)]\left[\left\{\int f^2 g_1 g_2 - \left(\int f g_1 g_2\right)^2\right\} + \left\{\int (g_1 g_2)^2 f\right.\right.$$

$$- \int (fg_1g_2)^2\}\Big] - 6\left\{\int f^2 g_1g_2 - \int f^2 \int fg_1g_2 + \int fg_1g_2 - \int fg_1g_2 \int (g_1g_2)^2\right\}, \tag{13}$$

where f is compact notation for $f(x, y)$ and g_1 and g_2 denote $g_1(x)$ and $g_2(y)$. Under the null hypothesis, the variance reduces to

$$\sigma_0^2 = 2\,[C^2(\gamma) + D^2(\gamma)]\,[\int f^3 - (\int f^2)^2]. \tag{14}$$

In addition to estimating $\hat{\delta}$, an estimate of the variance ($\hat{\sigma}^2$) must be obtained. Under H_0, we can use a combined sample $\mathbf{z}_1, \ldots, \mathbf{z}_N$ of $(x_1, y_1), \ldots, (x_m, y_m)$ and $(u_1, w_1), \ldots, (u_n, w_n)$, where $N = m + n$. Then,

$$\hat{\sigma}^2 = 8\left\{\int \left[\hat{h}_N(\mathbf{z})\right]^2 dH_N(\mathbf{z}) - \left[\int \hat{h}_N(\mathbf{z}) dH_N(\mathbf{z})\right]^2\right\}, \text{ where} \tag{15}$$

$$\hat{h}_N(\mathbf{z}) = \frac{1}{N_{CN}} \sum_{i=1}^{N} \left(\frac{\mathbf{z} - Z_i}{C_N}\right), \tag{16}$$

given that $H_N(\mathbf{z})$ is the empirical distribution function calculated numerically and C_N is the bandwidth parameter for the combined sample. Since $\hat{\sigma}^2$ converges in probability to σ^2, the null hypothesis can be rejected for sufficiently large values for m and n if

$$\delta^* = \frac{\sqrt{\dfrac{mn}{m+n}}}{\hat{\sigma}} > z_{\alpha/2} \tag{17}$$

for the two-sided alternative H_1: $f(x, y) \neq g(x)g(y)$.

The random walk hypothesis asserts that the distributions of x and y, or returns on Day 1 and Day 2, are identical. An entropy-based test for identical distribution is based upon the equality of the marginal densities. The null hypothesis is

$$H_0\colon g_1(x) = g_2(y). \tag{18}$$

The alternative hypothesis that the distributions are not identical is

$$H_1\colon g_1(x) \neq g_2(y). \tag{19}$$

The values of a_m, b_n, and C_N must be estimated. As discussed by Silverman (1986), a rough estimate of bandwidth is sufficient. We use the following so called "naive" estimators

$$a_m = 1.06\, S_m / m^{1/5} \tag{20}$$

$$b_n = 1.06\, S_n / n^{1/5} \tag{21}$$

$$C_N = 1.06\, S_N / N^{1/5}, \tag{22}$$

where S_m, S_n, and S_N are the standard deviations of the respective samples. The nonparametric test for independence is not robust when bandwidth is underestimated, but it performs well when bandwidth is overestimated. Since the naive estimator is easy to calculate and tends to slightly overestimate bandwidth, it is a good choice for testing the null hypothesis of independence. Depending upon the correlation factor of the distribution this test has power for hypothesis testing for large sample sizes and generally performs well for sample sizes larger than 50 data points. This test should be useful for high frequency data and for frequencies as low as monthly returns.

Simulations were run to test the statistics $\hat{\delta}$ and δ^*. For example, taking a sample of $n = 50$ and $\rho = .4$ from a bivariate normal distribution, a sample correlation of $\hat{\rho} = .397$ was obtained. The estimator $\hat{\delta} = 2.18$ with a p-value of .03 rejects H_0. Other arbitrary values of ρ were selected and similar results were obtained for other simulated data sets. This suggests that the nonparametric test using $\hat{\delta}$ compares favorably with alternative parametric tests. Since more information is used, $\hat{\delta}$ has higher power tests for independence. By running simulations, it was discovered that $\hat{\delta}$ is robust, even for very poor estimates of bandwidth. For the function

$C_{i,m}(\gamma) = \{ {}^{1-\gamma \ i \ even}_{1+\gamma \ i \ odd} \}$ and $m = 50$, $\hat{\delta}$ and therefore δ^* was optimized at a value of $\gamma = 20$. This value was used for all data sets for both $C_{i,m}(\gamma)$ and $D_{i,n}(\gamma)$ for sample sizes of 50 to 100. Prior to nonparametric density estimation, z-scores were used to normalize the data.

IV. EMPIRICAL RESULTS

We examine monthly United States stock returns from the Dow-Jones Industrial Average for 1982–1992 and dividend-inclusive monthly returns from the Center for Research in Security Prices (CRSP) value-weighted index from 1964–1989. Daily United States returns are from the Dow Jones Industrials for 1982–1992 and Canadian daily returns data are from the Toronto Stock Exchange index for 1982 to 1987. Beginning with the daily returns series for the Dow-Jones Industrials, we randomly selected 75 observations for $g_1(x)g_2(y)$ and 75 observations for $f(x, y)$. The selected observations are not sequential. Assuming that $\hat{\delta} - \delta$ is asymptotically normal, the test for equality of the two marginal distributions rejects the null hypothesis that the returns distributions are identical (i.e., that $g_1(x) = g_2(y)$) at the 5% confidence level. Similar results were obtained using other samples of up to 400 observations and for other data sets.

Graphing the joint density for the returns distribution calculated under the assumption of independence and the same distribution estimated assuming dependence showed that the distributions were shaped differently, which is indicative of possible dependence in day to day returns. More formally, independence is rejected based upon the nonparametric estimate of $\delta^* = 1.96$, which is significant at 5% confidence level. Hence, Dow-Jones daily returns are not independent, nor is the returns distribution identical over time.

The same tests are performed for daily Canadian returns data to see whether results are unique to the United States. Using the same sized random samples of 75 observations, we obtain a value of $\delta^* = 2.0$, which rejects independence at the 5% confidence level. Equality of the marginal distributions also is rejected at the 5% confidence level. Although previous tests have rejected the random walk for low frequency monthly or annual data, the nonparametric test shows that even high frequency noisy daily returns do not follow a random walk.

For monthly dividend-inclusive CRSP value-weighted index returns, the test statistic is $\delta^* = 12.89$, which is significant at the 1% level. Also, equality of the marginal distributions for Day 1 and Day 2 is rejected at the 5% confidence level. Similar results were obtained using monthly data from the Dow Jones index. As expected, less noisy monthly data provide even stronger rejections of independence and the random walk.

V. CONCLUSIONS

We have modified and extended Robinson's (1991) entropy-based test for independence to provide a complete nonparametric test for the random walk in time series data. Nonparametric density estimation has been used to identify both independence and whether the distribution of stock returns is identical over time. Using the most general test known, we have shown that daily and monthly stock returns do not follow a random walk. Instead, the stock returns examined follow a Markov process and further testing is necessary to determine the order of this process. As expected, the rejections of the random walk are more decisive for low frequency monthly returns data, but our results may be unique in that we also reject the random walk using for noisy, high frequency, daily returns data.

REFERENCES

Ahmad, I. A. 1982. "Nonparametric Estimation of the Location and Scale Parameters Based on Density Estimation." *Annals of the Institute of Statistical Mathematics* 34: 39–55.

Ahmad, I. A., and P. Cerrito. 1993. "Goodness of Fit Tests Based on the L_2 Norm of Multivariate Probability Density Functions." *Nonparametric Statistics* 2: 169–181.

Brock, W. A., D. A. Hsieh, and B. LeBaron. 1991. *Nonlinear Dynamics, Chaos, and Instability.* Cambridge: MIT Press.

Diebold, F. X., and J. A. Nason. 1990. "Nonparametric Exchange Rate Predictions." *Journal of International Economics* 28: 315–332.

Durlauf, S. N. 1991. "Spectral Based Testing of the Martingale Hypothesis." *Journal of Econometrics* 50: 355–376.

Fama, E. F., and K. R. French. 1988. "Permanent and Temporary Components of Stock Returns." *Journal of Political Economy* 96: 246–273.

Frennberg, P., and B. Hansson. 1993. "Testing the Random Walk Hypothesis on Swedish Stock Prices: 1919–1990." *Journal of Banking and Finance* 17: 175–191.

Kim, M. J., C. R. Nelson, and R. Startz. 1991. "Mean Reversion in Stock Price: A Reappraisal of the Evidence." *Review of Economic Studies* 58: 515–528.

Lo, A. W., and A. C. MacKinlay. 1988. "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test." *Review of Financial Studies* 1: 41–66.

MacDonald, R., and D. Power. 1993. "Persistence in UK Share Returns: Some Evidence from Disaggregated Data." *Applied Financial Economics* 3: 27–38.

McQueen, G., and S. Thorley. 1991. "Are Stock Returns Predictable? A Test Using Markov Chains." *Journal of Finance* 46: 239–263.

Prescott, D. M., and T. Stengos. 1990. "Testing for Forecastable Nonlinear Dependence in Weekly Gold Rates of Return." Working paper, Department of Economics, University of Guelph.

Robinson, P. M. 1991. "Consistent Nonparametric Entropy-Based Testing." *Review of Economic Studies* 58: 437–453.

Samiaddin, M., and G. El-Sayyad. 1990. "On Nonparametric Kernel Density Estimates." *Biometrika* 77:4: 865–874.

Sentana, E., and S. Wadhwani. 1991. "Semi-parametric Estimation and the Predictability of Stock Market Returns: Some Lessons from Japan." *Review of Economic Studies* 58: 547–563.

Silverman, B. 1986. *Density Estimation for Statistics and Data Analysis* London: Chapman and Hall.