# The Variance-Free Characterization of Heteroscedastic Normal Variables with an Application in Financial Econometrics

Max V. Moldovan

Melbourne Business School - The University of Melbourne 200 Leicester Street, Carlton, Victoria, Australia, VIC 3053 Ph: (+613) 9349 8115. Fax: (+613) 9349 8133. Email: m.moldovan@mbs.edu

Nicholas A. Nechval

Department of Mathematical Statistics - University of Latvia Raina Blvd 19, Riga, Latvia, LV 1050 Ph: (+371) 7 144 050. Fax: (+371) 7 034 702. Email: nechval@junik.lv

#### Abstract

In the presented study it is shown how heteroscedastic normal variables with unknown variance can be characterized by a symmetric beta distribution of the first kind with a known parameter. The presented variance-free characterization technique is illustrated with testing for normality the empirically observed financial return time series. We further suggest one of the possible extensions of the presented method that can be used for statistical learning with applications in real-time and time-critical systems.

**Keywords**: Hypothesis testing; Beta distribution family; Financial econometrics; Realized volatility; End-of-sample instability.

#### 1 Introduction

Suppose the following assumingly zero-mean heteroscedastic serially uncorrelated normal random variable can be observed:

$$x_{1,1},\ldots,x_{1,k},\ldots,x_{T,1},\ldots,x_{T,k},$$
 (1)

where  $x_{t,i} \sim N\left(0, \frac{\sigma_t^2}{k}\right)$  and the variance  $\sigma_t^2$  is unknown. Since the variable  $x_{t,i}$ , (t = 1, ..., T, i = 1, ..., k) is serially uncorrelated,  $r_t = \sum_{i=1}^k x_{t,i}$  is distributed as  $r_t \sim N(0, \sigma_t^2)$ . It is required to test if  $x_{t,i}$  is indeed normally distributed, which

International Journal of Computing Anticipatory Systems, Volume 19, 2006 Edited by D. M. Dubois, CHAOS, Liège, Belgium, ISSN 1373-5411 ISBN 2-930396-05-9 is not trivial since the variance  $\sigma_t^2$  is unknown and unequal within different subperiods  $t=1,\ldots,T$ . In the following study a new test for heteroscedastic normality is presented that requires none or very little knowledge about the variance of an underlying variable. The test is illustrated on the example of testing for normality the financial return time series, though we believe that the results of the study can be of a practical value for a wide range of theoretical and applied disciplines.

The study is structured as follows. In Section 2 a non-asymptotic relationship between a normal distribution and beta distribution of the first kind is presented. In Section 3 the indicated relationship is applied in a variance-free test for heteroscedastic normality, which is illustrated with an application to the empirically observed financial return time series. Section 4 concludes with the summary of the main results and suggests the direction for further research.

# 2 A Non-Asymptotic Relation Between Normal and Beta Distributions

Having the observed sequence of variables similar to (1), it is required to test if  $x_{t,i}$ , (t = 1, ..., T, i = 1, ..., k) is a serially uncorrelated zero-mean normal variable. Recognising that the variance parameter is different for different t, a traditional way to start with would be the examination of standardized variables given by

$$z_t = \frac{r_t}{\sigma_t},\tag{2}$$

where  $r_t = \sum_{i=1}^k x_{t,i}$  and  $\sigma_t$  is the underlying standard deviation of  $r_t$  within each subperiod t. If the distribution of  $x_{t,i}$  given by (1) is indeed normal, then the resulted series  $z_t$ , (t = 1, ..., T) will be standard normally distributed N(0, 1). However, the standardization given by (2) is difficult to implement since the variance is unknown and therefore only variance estimates can be used for standardization. The following estimator, which is known in finance as a realized volatility estimator,

$$\sigma_{t,k}^2 = \sum_{i=1}^k x_{t,i}^2 \tag{3}$$

is unbiased  $(E(\sigma_{t,k}^2) = \sigma_t^2)$ , consistent  $(\sigma_{t,k}^2 \to \sigma_t^2)$  as  $k \to \infty$  and highly efficient variance estimator that does not require information outside any individual subperiod  $(t = 1, \ldots, T)$  (see Andersen and Bolerslev, 1998; Andersen et al., 2001). Note that the subscript k in  $\sigma_{t,k}^2$  indicates that it is an estimate of the true parameter  $\sigma_t^2$  computed by summing k squared variables as shown by (3). With the use of a realized volatility estimator, the standardization (2) takes the following form:

$$z_{t} = \frac{\sum_{i=1}^{k} x_{t,i}}{\sqrt{\sum_{i=1}^{k} x_{t,i}^{2}}} = \frac{r_{t}}{\sigma_{t,k}}.$$
(4)

Some authors (e.g. Andersen et al., 2000; Bollen and Inder, 2002; Djupsjöbaska, 2002; Kayahan et al., 2002, among others) used the standardization given by (4) in the empirical studies expecting a resulted variable  $z_t$ , (t = 1, ..., T) to be standard normally distributed. However, the authors did not recognise that for finite k, the variable  $z_t$  is no longer expected to be standard normally distributed due to the positive correlation between the absolute of the numerator  $|r_t|$  and the denominator  $\sigma_{t,k}$  in (4). Intuitively, such correlation will result in shifting the random variable  $z_t$  given by (4) towards zero, inducing platykurtosis. Furthermore, for finite k, maximum and minimum of  $z_t$  are also finite. For demonstrating this, suppose the following function is given:

$$z_k = \frac{\sum_{i=1}^k x_i}{\sqrt{\sum_{i=1}^k x_i^2}}, \quad (x_i \in (-\infty, +\infty), \ k \in 2, 3, \dots, \infty).$$
 (5)

It is required to find the minimum and maximum of the given function. In order to do this, the partial derivatives of the function (5) with respect to each  $x_i$ , (i = 1, ..., k) should be taken and jointly set to zero:

$$f(x_j)' = \frac{1}{\sqrt{\sum_{i=1}^k x_i^2}} - \frac{x_j \cdot \sum_{i=1}^k x_i}{\sqrt{(\sum_{i=1}^k x_i^2)^3}}, \qquad (j = 1, 2, \dots, k),$$
 (6)

$$\begin{cases} f(x_1)' = 0 \\ \dots \\ f(x_k)' = 0 \end{cases}$$
 (7)

The system (7) has a multiple solution given by  $x_1 = \ldots = x_k$ ,  $(x_i \in (-\infty, +\infty))$ . Therefore the function (5) has  $\min(z_k) = -\sqrt{k}$  for  $x_i < 0$  and  $\max(z_k) = \sqrt{k}$  for  $x_i > 0$ :

$$\min_{x_i < 0} / \max_{x_i > 0} (z_k) = \frac{\sum_{i=1}^k x_i}{\sqrt{\sum_{i=1}^k x_i^2}} = \frac{k \cdot x_1}{\sqrt{k \cdot x_1^2}} = \frac{\pm k}{\sqrt{k}} = \pm \sqrt{k},$$

$$(x_1 = \dots = x_k).$$
(8)

As a result,  $z_k \in (-\sqrt{k}, \sqrt{k})$ , which differs from  $(-\infty, +\infty)$  of a normal variable. The above arguments lead to the conclusion that for finite values of k > 1,  $z_t$  given by (4) follows a specific distribution which is not normal. Thompson (1935) and Pearson and Chandra Sekar (1936) studied a variable given by

$$z_{t,i} = \frac{x_{t,i} \cdot \sqrt{k}}{\sqrt{\sum_{i=1}^{k} x_{t,i}^2}} = \frac{x_{t,i} \cdot \sqrt{k}}{\sigma_{t,k}},\tag{9}$$

and concluded that it is distributed with a Pearson type II distribution, a distribution that belongs to the beta family (see Devroye, 1986). However, the authors did not explore the distribution of  $z_t$  given by (4) more likely because they worked

with a single-period T=1 and examined a more general case with subtracting a mean estimate from each  $x_{1,i},\ (i=1,\ldots,k)$ , which leads to  $\sum_{i=1}^k (x_{1,i}-\bar{x}_1)=0$ ,  $(\bar{x}_1=k^{-1}\sum_{i=1}^k x_{1,i})$ . Taking into account one additional degree of freedom obtained by not subtracting a mean  $\bar{x}_t$ , we state that the distributions of  $z_t$  given by (4) and  $z_{t,i}$  given by (9) are identical, though this result does not directly follow from the analytical proofs given in Thompson (1935) and Pearson and Chandra Sekar (1936). To justify this result we first can present any single variable  $\tilde{x}_j=x_{t,i},\ (\tilde{x}_j\sim N(0,\tilde{\sigma}_j^2),\tilde{\sigma}_j^2=\frac{\sigma_t^2}{k})$  given by (1) as a sum of k independently identically normally distributed variables  $x_j=\sum_{h=1}^k y_{j,h},\ (y_{j,h}\sim N(0,\frac{\tilde{\sigma}_j^2}{k}))$ . Performing the same analytical manipulations with  $x_j=\sum_{h=1}^k y_{j,h}$  as we did with the initial variable  $r_t=\sum_{i=1}^k x_{t,i},\ (r_t\sim N(0,\sigma_t^2))$  clearly will lead to the conclusion that  $\tilde{z}_j=\frac{\tilde{x}_j}{\sqrt{\sum_{h=1}^k y_{j,h}^2}}=\frac{\tilde{x}_j}{\tilde{\sigma}_{j,k}}$  is distributed with the same distribution as  $z_t$  given by (4).

Since the variance of  $y_{j,h}$  is  $\tilde{\sigma}_j^2 = \frac{\sigma_t^2}{k}$  by construction we have  $\tilde{z}_j = \frac{\tilde{x}_j}{\left(\frac{\sigma_{t,k}}{\sqrt{k}}\right)} = \frac{\tilde{x}_j \cdot \sqrt{k}}{\sigma_{t,k}}$ ,

which is equivalent to (9). On this ground, it can be concluded that the distributions of  $z_t$  given by (4) and  $z_{t,i}$  given by (9) follow the same Pearson type II probability distribution as proven by Thompson (1935) and Pearson and Chandra Sekar (1936).

Further, knowing the upper and lower limits of the variable  $z_t$  given by (4), it is straightforward to perform the following normalization:

$$\dot{z}_t = \frac{z_t + \sqrt{k}}{2 \cdot \sqrt{k}}, \quad (t = 1, \dots, T, \ k = 2, 3, \dots, \infty). \tag{10}$$

Noticing that  $z_t \in (-\sqrt{k}, \sqrt{k})$ , it is easy to show that  $\dot{z}_t \in (0, 1)$ . Moreover, it has been found that  $\dot{z}_t$  given by (10) is distributed with a symmetric beta distribution of the first kind with a known parameter:

$$\dot{z}_t \sim \beta_1 \left( \frac{k-1}{2} \right), \quad (t = 1, \dots, T, \ k = 2, 3, \dots, \infty),$$
(11)

which is a well studied theoretical distribution (see Jambunathan, 1954, among others). This result is confirmed by the following equality:

$$\frac{\sqrt{k-1} \cdot (\dot{z}_t - 0.5)}{\sqrt{\dot{z}_t - \dot{z}_t^2}} = \frac{z_t \cdot \sqrt{k-1}}{\sqrt{k-z_t^2}}.$$
 (12)

The left-hand side of (12) is a well known transform of symmetric beta of the first kind distributed variables to Student t distributed variables (see Devroye, 1986). The right-hand side of (12), after taking into account one additional degree of freedom, is given by Thompson (1935) as a transform of variables  $z_t$  given by (4) to Student t distributed variables. In fact, both sides of (12) give a variable distributed as Student t with (k-1) degrees of freedom, which is consistent with Thompson (1935) and

Pearson and Chandra Sekar (1936) and enough for drawing the conclusion that  $\dot{z}_t$  given by (10) is distributed with a symmetric beta distribution of the first kind.

To summarise the section, we have demonstrated how normally distributed variables given by (1) can be transformed to symmetric beta of the first kind distributed variables by the strict mathematical manipulations given by (4) and (10). The indicated relation is non-asymptotic, in contrast to the asymptotic convergence of a symmetric beta of the first kind towards normal when a parameter of beta goes to infinity, and this enables the use of the presented relation for real-world finite-sample problems.

## 3 A Variance-Free Test for Heteroscedastic Normality

In the following section we demonstrate how the property given by (11) can be used in a variance-free test for heteroscedastic normality. We illustrate the test with an application to the empirically observed financial return time series, which according to the Efficient Market Hypothesis (EMH) by Fama (1970) and Mixture of Distributions Hypothesis (MDH) by Clark (1973) are expected to be zero-mean serially uncorrelated normally distributed. Considering the series  $x_{t,i}$  given by (1) and  $r_t = \sum_{i=1}^k x_{t,i}$  as intradaily and daily continuously compounded financial returns respectively, the test can be stated in the form of the following hypothesis:

$$H_0: r_t \sim N(0, \sigma_t^2), \ \left(r_t = \sum_{i=1}^k x_{t,i}, \ x_{t,i} \sim N\left(0, \frac{\sigma_t^2}{k}\right). \right.$$

$$Cov(x_i, x_j) = 0, \ i \neq j)$$
(13)

is tested against

$$H_A: r_t \text{ is not } N(0, \sigma_t^2), \ \left(r_t = \sum_{i=1}^k x_{t,i}, \ x_{t,i} \sim N\left(0, \frac{\sigma_t^2}{k}\right), \right.$$

$$Cov(x_i, x_j) = 0, \ i \neq j).$$
(14)

 $H_0$  should not be rejected if the variable  $\dot{z}_t$  given by (10) is distributed with a symmetric beta distribution of the first kind given by (11). It should be noted that this test is simple, as opposed to composite, because we do not treat  $\sigma_{t,k}^2$  given by (3) as a variance estimate when calculate a characteristic variable  $\dot{z}_t$  in (10). Thus, the suggested test is variance-free. The obvious weakness of the test is the required equality of variances of  $x_{t,i}$ ,  $(i=1,\ldots,k)$  within each day t. However, the empirical evidence suggest that the variance of intradaily financial returns, in general, has the same pattern within each day, though it is not necessary equal for different days (see Andersen and Bolerslev, 1997; Areal and Taylor, 2002). Relying on this empirical observation, we use the weighting procedure suggested by Areal and Taylor (2002) for maximizing the efficiency of the estimator given by (3):

$$\sigma_{t,k}^2 = \sum_{i=1}^k w_i x_{t,i}^2,\tag{15}$$

where  $w_i = \frac{1}{k \cdot \lambda_i}$ , following Taylor and Xu (1997)  $\hat{\lambda_i} = \frac{\sum_{t=1}^T x_{t,i}^2}{\sum_{t=1}^T \sum_{j=1}^k x_{t,j}^2}$  and  $\sum_{i=1}^k \lambda_i w_i = 1$ . Now a test for parameters of  $x_i$  and  $x_i$ 

1. Now a test for normality of  $r_t$  takes the form of the following hypothesis:

$$H_0: r_t \sim N(0, \sigma_t^2), \ \left(r_t = \sum_{i=1}^k x_{t,i}, \ x_{t,i} \sim N\left(0, \frac{w_i \sigma_t^2}{k}\right), \right.$$

$$Cov(x_i, x_j) = 0, \ i \neq j)$$
(16)

is tested against

$$H_A: r_t \text{ is not } N(0, \sigma_t^2), \ \left(r_t = \sum_{i=1}^k x_{t,i}, \ x_{t,i} \sim N\left(0, \frac{w_i \sigma_t^2}{k}\right), \right.$$

$$Cov(x_i, x_j) = 0, \ i \neq j).$$
(17)

As before,  $H_0$  should not be rejected if the variable given by

$$\dot{z}_t = \frac{z_t + \sqrt{k}}{2 \cdot \sqrt{k}}, \qquad \left(z_t = \frac{\sum_{i=1}^k w_i x_{t,i}}{\sqrt{\sum_{i=1}^k w_i x_{t,i}^2}}, \ t = 1, \dots, T, \ k = 2, 3, \dots, \infty\right)$$
(18)

follows a symmetric beta distribution of the first kind given by (11). Although the transformed test is no longer simple, since estimates of weights  $w_i$ , (i = 1, ..., k)are used for testing, it still does not require estimates of daily return variance  $\sigma_t^2$ and thus is variance-free. Further we give the empirical illustration of a test assuming that only three intradaily observations are available: open  $(x_{t,1})$ , middle-day  $(x_{t,2})$  and closing  $(x_{t,3})$  continuously compounded returns. The hypothesis given by (17) and (18) has been tested with two financial return time series. The return series have been computed from the foreign futures exchange rates between the Australian dollar and US dollar (AUD/USD) and between the Japanese yen and US dollar (JPY/USD). The data has been initially provided by Tick Data, Inc (www.tickdata.com) and covers the period from 02 January 1990 to 31 March 2000 (T=2586 trading days). The duration of each trading day t is 400 minutes. Trading is open Monday to Friday, from 7.20 a.m. to 2.00 p.m.  $x_{t,1}$  is the last return recorded before 8.10 a.m. and 8.00 a.m. for AUD/USD and JPY/USD series respectively. The middle-day return  $x_{t,2}$  is the last return recorded before 10.40 a.m. for both series. The closing return  $x_{t,3}$  is the last return recorded before 2.00 p.m. for both series within a day t. The days when at least one intradaily return  $x_{t,i}$ , (i = 1, 2, 3)could not be observed, have been excluded from the series. As a result, we have left with 2394 and 2436 daily observations  $(2394 \cdot 3 = 7182 \text{ and } 2436 \cdot 3 = 7308 \text{ intradaily}$ observations) for AUD/USD and JPY/USD return series respectively.

Before proceeding with the empirical test, we report some characteristics of size and power for the simple test given by (14) and (15) and its composite alternative given by (17) and (18) for k=3, though we should note that the power can be greater for values of k other than 3.

**Table 1**: The size and power of a simple test for normality given by (14) and (15) with k=3. The power is measured against Student t distributions with 2, 4 and 6 degrees of freedom and autocorrelated normal with autocorrelation coefficients (a.c.) equal to 0.05, 0.10 and 0.20. The number of Monte Carlo trials M=10000.

|                 | $N(0, \sigma_t^2)$ | $\overline{t(2)}$ | t(4)   | t(6)   | Autocorrelated Normal |        | rmal   |
|-----------------|--------------------|-------------------|--------|--------|-----------------------|--------|--------|
| T = 1000        |                    | i je              |        |        | a.c.=0.05             | 0.10   | 0.20   |
| $\alpha = 0.01$ | 0.0088             | 0.2242            | 0.0328 | 0.0177 | 0.0360                | 0.2322 | 0.9770 |
| 0.05            | 0.0465             | 0.5116            | 0.1229 | 0.0792 | 0.1452                | 0.5488 | 0.9990 |
| 0.10            | 0.0965             | 0.6983            | 0.2136 | 0.1445 | 0.2542                | 0.7255 | 1.0000 |
| T = 2500        |                    | 644               |        |        |                       | 14 LL  |        |
| $\alpha = 0.01$ | 0.0093             | 0.7560            | 0.0927 | 0.0377 | 0.1174                | 0.8010 | 1.0000 |
| 0.05            | 0.0488             | 0.9626            | 0.2671 | 0.1273 | 0.3548                | 0.9704 | 1.0000 |
| 0.10            | 0.0957             | 0.9938            | 0.4164 | 0.2222 | 0.5241                | 0.9933 | 1.0000 |
| T = 5000        |                    |                   |        |        |                       |        |        |
| $\alpha = 0.01$ | 0.0102             | 0.9990            | 0.2481 | 0.0749 | 1.0000                | 1.0000 | 1.0000 |
| 0.05            | 0.0508             | 1.0000            | 0.5513 | 0.2228 | 1.0000                | 1.0000 | 1.0000 |
| 0.10            | 0.0988             | 1.0000            | 0.7294 | 0.3628 | 1.0000                | 1.0000 | 1.0000 |
|                 |                    |                   |        |        |                       |        |        |

**Table 2**: The size and power of a composite test for normality given by (17) and (18) with k=3. The power is measured against Student t distributions with 2, 4 and 6 degrees of freedom and autocorrelated normal with autocorrelation coefficients (a.c.) equal to 0.05, 0.10 and 0.20. The number of Monte Carlo trials M=10000.

|                 | $N(0,\sigma_t^2)$ | t(2)   | t(4)   | t(6)      | Autocorre | lated No | rmal   |
|-----------------|-------------------|--------|--------|-----------|-----------|----------|--------|
| T = 1000        | )                 |        |        |           | a.c.=0.05 | 0.10     | 0.20   |
| $\alpha = 0.01$ | 0.0062            | 0.4271 | 0.0343 | 0.0100    | 0.0107    | 0.0295   | 0.2240 |
| 0.05            | 0.0398            | 0.5815 | 0.1081 | 0.0500    | 0.0579    | 0.1289   | 0.5610 |
| 0.10            | 0.0822            | 0.6773 | 0.1861 | 0.1032    | 0.1135    | 0.2346   | 0.7500 |
| T = 2500        | )                 |        |        | 1 - 1 - 1 |           |          |        |
| $\alpha = 0.01$ | 0.0090            | 0.6496 | 0.0591 | 0.0197    | 0.0218    | 0.1005   | 0.8404 |
| 0.05            | 0.0400            | 0.8094 | 0.1697 | 0.0741    | 0.0981    | 0.3256   | 0.9868 |
| 0.10            | 0.0853            | 0.8841 | 0.2739 | 0.1434    | 0.1845    | 0.5138   | 0.9989 |
| T = 5000        | )                 |        |        | 12_ 1     |           |          |        |
| $\alpha = 0.01$ | 0.0092            | 0.8544 | 0.1089 | 0.0277    | 0.0349    | 0.3262   | 0.9997 |
| 0.05            | 0.0437            | 0.9607 | 0.2752 | 0.1152    | 0.1489    | 0.7087   | 1.0000 |
| 0.10            | 0.0873            | 0.9867 | 0.4115 | 0.2078    | 0.2683    | 0.8711   | 1.0000 |

As can be seen from Table 1, the simple test has asymptotically correct size and keeps the substantial power even with a Student t(6) distribution, which is fairly close to a normal. The size of the composite variance-free test is slightly distorted, as can be seen from the column of Table 2 corresponding to  $N(0, \sigma_t^2)$ . The distortion of size is due to the use of estimates in the testing procedure. Further, the distorted size results on a lower power of the composite test with respect to its simple alternative as also demonstrated by Table 2. See Lloyd (2005) for the exposition of testing size and power issues.

For performing the actual empirical test, the characteristic variables given by (18) have been calculated and tested if they are distributed as  $\beta_1\left(\frac{k-1}{2}\right) \equiv \beta_1(1)$  with an application of the Kolmogorov-Smirnov (KS) test. The estimates of weights and *p*-values obtained from the KS test are given in the table below. According to

**Table 3**: The test for the symmetric  $\beta_1(1)$  distribution of the empirical characteristic series calculated as shown by (18). The failure to reject  $\beta_1(1)$  distribution of the characteristic series is equivalent to the failure to reject heteroscedastic serially uncorrelated normality of the initial return series.

|         | $w_1$  | $w_2$  | $w_3$  | <i>p</i> -value |
|---------|--------|--------|--------|-----------------|
| AUD/USD | 0.5148 | 1.5474 | 2.4318 | 0.0290          |
| JPY/USD | 0.5188 | 1.5582 | 2.3209 | 0.0649          |

Table 3,  $H_0$  given by (17) cannot be rejected at 99 and 95% confidence levels for AUD/USD and JPY/USD series respectively. These results provide the objective evidence in favour of the EMH and MDH since the test requires intradaily returns to be uncorrelated zero-mean normal variables.

## 4 Conclusion and Suggestion for Further Research

In the presented study it has been demonstrated that a random variable  $z_t$  given by (4) follows a specific platykurtic distribution with one parameter which is found to be identical to a Pearson type II distribution obtained in Thompson (1935) and Pearson and Chandra Sekar (1936). It has been proven by (6), (7) and (9) that the minimum and maximum of  $z_t$  given by (4) correspond to  $-\sqrt{k}$  and  $\sqrt{k}$  respectively. This result suggested the transformation given by (10) that leads to a new variable  $\dot{z}_t \in (0,1)$ . It has been confirmed by (12) that  $\dot{z}_t$  follows a symmetric beta distribution of the first kind with a known parameter  $\frac{(k-1)}{2}$ . This property has been used in a variance-free test for heteroscedastic normality. Two versions of the test, simple and composite, have been suggested. The composite version of the test given by (17) and (18) with k=3 has been applied to the empirically observed AUD/USD and JPY/USD returns on futures contracts. The null could not be rejected at 99 and

95% confidence levels for AUD/USD and JPY/USD series respectively, providing additional objective evidence in favour of the widely debated normality of financial returns.

We would like to note that the property given by (11) can be used in a group of alternative statistical procedures. For example, under the null, a heteroscedastic normal variable can first be characterized by a symmetric beta distribution with a known parameter, next with the use of a beta cumulative distribution function transformed to a probability integral distributed as U(0,1), and finally through a normal inverse function transformed to an invariant standard normal variable. Then the end-of-sample instability testing techniques suggested in Nechval (1988), Andrews (2003) and Moldovan (2003, pp. 40-50 and 69-71) can be applied for anticipation and detection of breaks in a random process. Such end-of-sample instability procedures can be used for statistical learning in real-time and time-critical systems.

#### References

- Andersen, T. and T. Bollerslev (1997) Intraday Periodicity and Volatility Persistence in Financial Markets, Journal of Empirical Finance 4, 115-158.
- Andersen, T. and T. Bollerslev (1998) Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts, International Economic Review 39, 885-905.
- Andersen, T., T. Bollerslev, F. Diebold and P. Labys (2000) Exchange Rates Returns Standardized by Realized Volatility are (Nearly) Gaussian, Multinational Finance Journal 4, 159-179.
- Andersen, T., T. Bollerslev, F. X. Diebold and P. Labys (2001), The Distribution of Realized Exchange Rate Volatility, Journal of the American Statistical Association, 96, 42-55.
- Andrews, D.W.K. (2003) End-of-Sample Instability Test, Econometrica 71, 1661-1694.
- Areal, N. and S. Taylor (2002) The Realized Volatility of FTSE-100 Futures Prices, The Journal of Futures Markets 22, 627-648.
- Bollen, B. and B. Inder (2002) Estimating Daily Volatility in Financial Markets Utilizing Intraday Data, Journal of Empirical Finance 9, 551-562.
- Clark, P. (1973) A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices, Econometrica 41, 135-156.
- Devroye, L. (1986) Non-Uniform Random Variate Generation, New York: Springer-Verlag.
- Djupsjöbaska, D. (2002) Properties of Realized Volatility and Correlations, Manuscript, http://ssrn.com/abstract=393469.
- Fama, E. (1970) Efficient Capital Markets: A Review of Theory and Empirical Work, Journal of Finance 25, 383-417.

- Jambunathan M. V. (1954) Some Properties of Beta and Gamma Distributions, The Annals of Mathematical Statistics, Vol. 25, No. 2, 401-405.
- Kayahan, B., T. Stengos and B. Saltoğlu (2002) Intra-Day Features of Realized Volatility: Evidence From an Emerging Market, International Journal of Business and Economics 1, 17-24.
- Lloyd, C. J. (2005) On Comparing the Accuracy of Competing Tests of the Same Hypothesis from Simulation Data, Journal of Statistical Planning and Inference 128, 497-508.
- Moldovan, M. V. (2003) Stochastic Modelling of Random Variables with an Application in Financial Risk Management, Manuscript, http://ssrn.com/abstract=484383.
- Nechval, N. A. (1988) A General Method for Constructing Automated Procedures for Testing Quickest Detection of a Change in Quality Control, Computers in Industry 10, 177-183.
- Pearson, E. S. and C. Chandra Sekar (1936) The Efficiency of Statistical Tools and a Criterion for the Rejection of Outlying Observation, Biometrica Vol. 28, No. 3/4, 308-320.
- Taylor, S. and X. Xu (1997) The Incremental Volatility Information in One Million Foreign Exchange Quotations, Journal of Empirical Finance 4, 317-340.
- Thompson, W. R. (1935) On the Criterion for the Rejection of Observations and the Distribution of the Ratio of Deviation to Sample Standard Deviation, The Annals of Mathematical Statistics Vol. 6, No. 4, 214-219.