1. Introduction

Martingale model belongs to the earliest models of financial asset prices. Its origin lies in the birth of probability theory and in the history of games of chance. It follows the principle of a fair game, i.e. the game which is neither in your favor nor your opponent's. Martingale is stochastic process \( \{P_t\} \) which satisfies the condition:

\[
E[P_t \mid P_{t-1}, P_{t-2}, \ldots] = P_t, \tag{1}
\]

equivalently, it is possible to write

\[
E[P_t P_{t-1} \mid P_{t-1}, P_{t-2}, \ldots] = 0. \tag{2}
\]

If \( P_t \) is the asset's price at time \( t \), the martingale hypothesis means that tomorrow's price is expected to be equal to today's price under the condition of the entire history development of the asset's price. The forecasting meaning follows: the martingale hypothesis implies that the "best" forecast (from the point of view of mean square error) of tomorrow's price is simply today's price.

One of the most important aspects of the martingale hypothesis is that nonoverlapping price changes are uncorrelated at all leads and lags [Davidson (1994), p. 230]. This means that there is not any systematic movement in price changes which would make effective the linear forecasting rule.

Historically, the martingale is closely related to hypothesis of efficient market which means that the information contained in past asset's prices is completely reflected in the current price. In efficient market it is not possible to profit by trading on the information contained in the asset's price history. Despite the fact that the modern financial economics considers the necessity of some trade-off between risk and expected returns, martingale is still powerful tool and has important applications in theory of asset prices. Another aspect of martingale is that it is basis for the development of a closely related model which is called random walk.
2. Random Walk

2.1 Random Walk 1 (RW1)

Process of random walk can be expressed in the following form:

\[ P_t = c + P_{t-1} + \varepsilon_t, \varepsilon_t \sim \text{IID}(0,\sigma^2), \]  

(3)

where \( c \) is the expected price change and IID(0,\( \sigma^2 \)) denotes that increments \( \varepsilon_t \) are independently and identically distributed with mean 0 and variance \( \sigma^2 \). This form of random walk is also fair game but independence of increments implies not only that they are uncorrelated, but also their any nonlinear functions are uncorrelated. It can be called Random Walk 1 or RW1.

As the random walk at time \( t \) and at some initial value \( P_0 \) at time 0 can be expressed in the form

\[ P_t = P_0 + ct + \sum_{i=0}^{t-1} \varepsilon_{t-i}, \]  

(4)

c conditional mean has form

\[ \mathbb{E}[P_t | P_0] = P_0 + ct \]  

(5)

and conditional variance is

\[ \text{D}[P_t | P_0] = \sigma^2 t. \]  

(6)

It follows that random walk is nonstationary and conditional mean and variance are both linear functions of time variable.

2.2 Random Walk 2 (RW2)

From practical analysis of financial asset's prices over longtime span (till hundred years) it is clear that the assumption of identically distributed increments is frequently not plausible. Therefore it is possible to relax the assumption of IID and introduce process with independent but not identically distributed (INID) increments \( \varepsilon_t \). It can be called Random Walk 2 or RW2. It is clear that RW1 is a special case of RW2. In contrast with RW1 process RW2 allows for unconditional heteroscedasticity in increments \( \varepsilon_t \).

2.3 Random Walk 3 (RW3)

More general version of random walk hypothesis can be obtained by replacing the independence assumption in process RW2 by assumption that increments \( \varepsilon_t \) are uncorrelated. This form of random walk is weakest, it can be called Random Walk 3 or RW3. RW1 and RW2 are special cases of RW3. Interesting example of process RW3 is process for which \( \text{Cov}[\varepsilon_t, \varepsilon_{t-k}] = 0 \) for all \( k \neq 0 \), but \( \text{Cov}[\varepsilon_t^2, \varepsilon_{t-k}^2] \neq 0 \). This form of process allows for conditional heteroscedasticity (for example model ARCH, GARCH).

3. Variance Ratios

3.1 Definition of Variance Ratios

Net return on the asset between dates \( t - 1 \) and \( t \) is defined as

\[ R_t = \frac{P_t}{P_{t-1}} - 1, \]  

(7)
**gross return** is defined as

\[ R_t + 1 = \frac{P_t}{P_{t-1}}. \]  

(8)

Continuously compounded return \( r_t \) is defined as the natural logarithm of gross return:

\[ r_t = \ln(R_t + 1) = \ln \frac{P_t}{P_{t-1}} = p_t - p_{t-1}, \]

(9)

where \( p_t \equiv \ln P_t \). Continuously compounded multi-period return is the sum of continuously compounded single-period returns

\[ r_t(k) = \ln[(R_t+1)(R_{t+1}+1) \ldots (R_{t+k+1}+1)] = r_t + r_{t+1} + \ldots + r_{t+k+1}. \]

(10)

The random walk increments are linear function of time variable, this is very important property of random walk. In RW2 and RW3 it is more difficult to state this property than in RW1 as the variances of individual increments \( r_t \) can change through time. But still, the variance of the sum have to equal the sum of the variances. For logarithmic version of process RW1, i.e.

\[ p_t = c + p_{t-1} + \varepsilon_t, \varepsilon_t \sim \text{IID}(0, \sigma^2), \]

(11)

the variance of \( r_t(k) \) have to be \( k\sigma^2 \). Therefore, the acceptability of the random walk model can be checked by comparing the variance of \( r_t(k) \) to \( k \) times the variance of \( r_t \). Lo and MacKinlay (1988, 1989) exploited this idea for development of test for random walk.

Consider the ratio of variance of two-period continuously compounded return \( r_t(2) = r_t + r_{t+1} \) to twice the variance of a one period return \( r_t \). Assume that returns create stationary process. Variance ratio has form

\[ VR(2) = \frac{D[r_t(2)]}{2D[r_t]} = \frac{D[r_t + r_{t+1}]}{2D[r_t]} = \frac{2D[r_t] + 2Cov[r_t, r_{t+1}]}{2D[r_t]}, \]

(12)

it follows that

\[ VR(2) = 1 + \rho(1), \]

(13)

where \( \rho(1) \) is the autocorrelation coefficient of the first-order of process \( \{r_t\} \). For process RW1 all the autocorrelations are zero, so \( VR(2) = 1 \). When there is positive first order autocorrelation, \( VR(2) > 1 \), it means that variances will grow faster than linearly. When there is negative first order autocorrelation, \( VR(2) < 1 \), variances will grow slower than linearly.

Variance ratio for \( q \) periods is

\[ VR(q) = \frac{D[r_t(q)]}{qD[r_t]} = 1 + 2\sum_{k=1}^{q-1}(1 - \frac{k}{q})\rho(k), \]

(14)

where \( \rho(k) \) is the autocorrelation coefficient of the \( k \)th order of process \( \{r_t\} \). Under RW1 all autocorrelation coefficients are zero, so \( VR(q) = 1 \). Even under RW2 and RW3, \( VR(q) = 1 \) as long as the variances of \( r_t \) are finite.

### 3.2 Sampling Distribution of Variance Ratios under RW1

Under conditions (9) and (11) the null hypothesis of RW1 is

\[ H_0: r_t = c + \varepsilon_t, \varepsilon_t \sim \text{IID}(0, \sigma^2). \]

(15)

The normality assumption is considered for convenience, the Lo's and MacKinley's results apply more generally to process with IID increments with finite forth moments.
Assume stationary time series of $2n + 1$ log prices $p_0, p_1, \ldots, p_{2n}$. The estimators of $c$ and $\sigma^2$ have forms

$$\hat{c} = \frac{1}{2n} \sum_{k=1}^{2n} (p_k - p_{k-1}) = \frac{1}{2n} (p_{2n} - p_0),$$

(16)

$$\hat{\sigma}_a^2 = \frac{1}{2n} \sum_{k=1}^{2n} (p_k - p_{k-1} - \hat{c})^2,$$

(17)

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

(18)

Estimators (16) and (17) are maximum-likelihood estimators of $c$ and $\sigma^2$. Estimator (18) exploits the random walk property, i.e. the fact that the mean and variance of increments are linear in the increment interval. The above mentioned estimators are consistent and

$$\sqrt{2n}(\hat{\sigma}_a^2 - \sigma^2) \sim_{a} N(0,2\sigma^4),$$

(19)

$$\sqrt{2n}(\hat{\sigma}_b^2 - \sigma^2) \sim_{a} N(0,4\sigma^4),$$

(20)

where "$\sim_{a}$" means asymptotic distribution. Lo and MacKinley (1988) proved that

$$\sqrt{2n}(VR(2) - 1) \sim_{a} N(0,2),$$

(21)

where

$$VR(2) = \frac{\hat{\sigma}_b^2}{\hat{\sigma}_a^2}$$

(22)

is the two-period variance ratio statistic. The null hypothesis can be tested by the standardized statistic $\sqrt{2n}(\hat{VR}(2) - 1)/\sqrt{2}$ which has asymptotically standard normal distribution. It means that if the value of standardized statistic lies outside interval [-1.96, 1.96], RW1 can be rejected at the 5% significance level.

The generalization of ratio statistic is straightforward. Let time series consists $nq + 1$ observations $p_0, p_1, \ldots, p_{nq}$, where $q$ is any integer greater than one, than we have the following estimators

$$\hat{c} = \frac{1}{nq} \sum_{k=1}^{nq} (p_k - p_{k-1}) = \frac{1}{nq} (p_{nq} - p_0),$$

(23)

$$\hat{\sigma}_a^2 = \frac{1}{nq} \sum_{k=1}^{nq} (p_k - p_{k-1} - \hat{c})^2,$$

(24)

$$\hat{\sigma}_b^2 (q) = \frac{1}{nq} \sum_{k=1}^{n} (p_{qk} - p_{qk-q} - q\hat{c})^2.$$ 

(25)

Than $q$-period variance ratio statistic has form

$$\hat{VR}(q) = \frac{\hat{\sigma}_b^2 (q)}{\hat{\sigma}_a^2}$$

(26)

and it can be proved that

$$\sqrt{nq}(\hat{VR}(q) - 1) \sim_{a} N(0,2(q-1)).$$

(27)

Instead of estimator $\hat{\sigma}_b^2 (q)$ it is possible to apply alternative estimator
\[ \hat{\sigma}^2_c(q) = \frac{1}{nq} \sum_{k=q}^{nq} (p_k - p_{k-q} - \hat{c})^2 \]  

which use overlapping \( q \)-period returns. It contains \( nq - q + 1 \) terms instead of \( n \) terms of estimator (25). Variance ratio test based on this estimator is more powerful.

The another refinement involves correcting the bias in the variance estimators \( \hat{\sigma}^2_a \) and \( \hat{\sigma}^2_c(q) \). The unbiased estimators have forms

\[ \bar{\sigma}^2_a = \frac{1}{nq-1} \sum_{k=1}^{nq} (p_k - p_{k-1} - \hat{c})^2, \]

\[ \bar{\sigma}^2_c(q) = \frac{1}{m} \sum_{k=q}^{nq} (p_k - p_{k-q} - q\hat{c})^2, \]

where

\[ m = q(nq - q + 1) \left(1 - \frac{q}{nq}\right). \]

\( q \)-period variance ratio statistic has form

\[ \sqrt{nq} (V R(q) - 1) \sim_d N \left(0, \frac{2(2q-1)(q-1)}{3q}\right). \]

Statistics (27) and (33) can be simply standardized to have asymptotically standard normal distribution.

3.3 Sampling Distribution of Variance Ratios under RW3

Financial time series are characteristic by changing volatility over time. It is therefore practically interesting and useful to consider variance ratios test for RW3. To allow for more general forms of heteroskedasticity Lo and MacKinley (1988) developed a special approach of testing random walk hypothesis. This approach can be applied to much broader of log price processes than the IID or IIN increment process. Let \( r_t = c + \epsilon_t \), Lo and MacKinley defined the compound null hypothesis \( H_0^* \), which assumes that \( p_t \) possesses uncorrelated increments but allows for quite general forms of heteroscedasticity, including deterministic changes in variance and ARCH processes.

Under the condition \( H_0^* \), \( VR(q) \) still approaches one. There is the question what is the form of its asymptotic variance. Lo and MacKinley (1988) supposed that under very general conditions it is possible to write

\[ VR(q) = \hat{\rho}(k) + 2 \sum_{k=1}^{q-1} (1 - \frac{k}{q}) \hat{\rho}(k), \]

where "\( =_a \)” means asymptotically equals. They supposed that \( \hat{\rho}(k) \) are asymptotically uncorrelated. The asymptotic variance \( \theta(q) \) of \( VR(q) \) can be calculated as the weighted sum of \( \hat{\rho}_i \)'s, which are the asymptotic variances of autocorrelations \( \hat{\rho}(k) \). Lo and MacKinley (1988) showed that
(a) the statistic $\overline{VR}(q) - 1$ converge almost surely to zero for all $q$ as $n$ increases without bound.

(b) 

$$\hat{\delta}_k = \frac{nq \sum_{j=k+1}^{\infty} (p_j - p_{j-1} - \hat{c})^2 (p_{j-k} - p_{j-k-1} - \hat{c})^2}{\sum_{j=1}^{\infty} (p_j - p_{j-1} - \hat{c})^2}$$  \hspace{1cm} (35)

is a heteroscedasticity-consistent estimator of $\delta_k$.

(c) heteroscedasticity-consistent estimator of $\theta(q)$ has form

$$\hat{\theta}(q) = 4 \sum_{k=1}^{q-1} \left(1 - \frac{k}{q}\right)^2 \hat{\delta}_k.$$  \hspace{1cm} (36)

The following standardized test statistic for testing of the null hypothesis $H_0^*$ can be used

$$\psi(q) = \frac{\sqrt{nq(\overline{VR}(q) - 1)}}{\sqrt{\hat{\theta}(q)}} \sim \text{N}(0,1).$$  \hspace{1cm} (37)

4. The Random Walk Hypothesis for Czech Stock Market Prices

4.1 Daily Data

First, in our analysis we will use daily time series of Czech stock market price index PX50. The series is from September 7, 1993 to March 28, 2000. The behavior if its log transformation shows Figure 1. It is clear that time series is non-stationary and it is probable that it contains stochastic trend i.e. that it is generated by integrated process. The log differences of this time series are on Figure 2.

![Figure 1](image-url)
Now there is a question if the time series of log transformation of PX50 is random walk. Table 1 contains the values of the sample autocorrelation function for the log differences. Many of these values are relatively high which would indicate that daily time series PX50 is not random walk. From Figure 2 it is also clear that this time series is characteristic by changing volatility. It could be modeled for example by some form of ARCH or GARCH model. It is therefore useful to test RW3 hypothesis by heteroscedasticity robust variance ratios test.

Table 1

<table>
<thead>
<tr>
<th>$k$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}(k)$</td>
<td>0.4831</td>
<td>0.3132</td>
<td>0.2250</td>
<td>0.1235</td>
<td>0.0002</td>
<td>-0.0013</td>
<td>-0.0089</td>
<td>0.0010</td>
<td>0.0256</td>
<td>0.0996</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}(k)$</td>
<td>0.1025</td>
<td>0.1127</td>
<td>0.1371</td>
<td>0.1239</td>
<td>0.1379</td>
<td>0.1190</td>
<td>0.1092</td>
<td>0.1217</td>
<td>0.1152</td>
<td>0.0741</td>
</tr>
</tbody>
</table>

Table 2 contains the results of this test for $q = 2, 3, 4, 5, 10, 15, 20$. The first row contains the values of statistic $\overline{V}(q)$, the second row the values of testing criterion $\psi(q)$. These values are compared with 2.5% and 97.5% critical values of the standard normal distribution, i.e. the values -1.96, 1.96. As the all values of testing criterion $\psi(q)$ lie outside interval [-1.96, 1.96] the hypothesis of RW3 for daily time series of PX50 is rejected at the 5% significance level.

Table 2

<table>
<thead>
<tr>
<th>$q$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\overline{V}(q)$</td>
<td>1.4831</td>
<td>1.8529</td>
<td>2.1503</td>
<td>2.3782</td>
<td>2.8331</td>
<td>3.2155</td>
<td>3.7153</td>
</tr>
<tr>
<td>$\psi(q)$</td>
<td>4.7181</td>
<td>5.5868</td>
<td>6.0047</td>
<td>6.1431</td>
<td>5.3453</td>
<td>5.3428</td>
<td>5.9610</td>
</tr>
</tbody>
</table>
4.2 Monthly Data

Now, we will analyze monthly time series of Czech stock market price index PX50. The series is from September 1993 to March 2000. The behavior if its log transformation shows Figure 3. Also this time series is non-stationary and is generated by integrated process, its log differences are on Figure 4.

Table 3 contains the values of the sample autocorrelation function for the log differences. Generally, these values are smaller than in the case of daily data, but for lags 1, 2, 3 are still relatively high which would indicate that also monthly time series PX50 is not random walk. From Figure 4 it can be seen that also this time series is characteristic by changing volatility. Therefore, we will test RW3 hypothesis by heteroscedasticity robust variance ratios test.
Table 3

<table>
<thead>
<tr>
<th>$k$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}(k)$</td>
<td>0.4467</td>
<td>0.2286</td>
<td>0.0595</td>
<td>0.0072</td>
<td>-0.0824</td>
<td>-0.0617</td>
<td>-0.0333</td>
<td>-0.0374</td>
<td>-0.0287</td>
<td>0.0269</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$k$</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\rho}(k)$</td>
<td>-0.0678</td>
<td>-0.1495</td>
<td>-0.1031</td>
<td>-0.1655</td>
<td>-0.0538</td>
<td>-0.1313</td>
<td>-0.1155</td>
<td>-0.1157</td>
<td>0.0008</td>
<td>-0.0207</td>
</tr>
</tbody>
</table>

Table 4 contains the results of variance ratios test for $q = 2, 3, 4, 5, 10, 15, 20$. By comparing the values of $VR(q)$ and the values of testing criterion $\psi(q)$ with the results for daily data it is clear, that these values are considerably lower. But as the values of testing criterion $\psi(q)$ for $q = 2, 3, 4, 5, 10, 15$ lie outside interval $[-1.96, 1.96]$, we can conclude that even for monthly time series of PX50 the hypothesis RW3 is rejected at the 5% significance level. The values of testing criterion for $q = 2, 3, 4, 5, 10$ lie outside interval $[-2.58, 2.58]$, so it is possible to conclude that the hypothesis RW3 is rejected also at the 1% significance level.

Table 4

<table>
<thead>
<tr>
<th>$q$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>$VR(q)$</td>
<td>1.4467</td>
<td>1.7481</td>
<td>1.9285</td>
<td>2.0397</td>
<td>2.0895</td>
<td>1.9311</td>
<td>1.5810</td>
</tr>
<tr>
<td>$\psi(q)$</td>
<td>2.9086</td>
<td>3.4205</td>
<td>3.4797</td>
<td>3.3800</td>
<td>2.6889</td>
<td>2.0561</td>
<td>1.1129</td>
</tr>
</tbody>
</table>

References:


