A Runs Test for Stock-Market Prices with an Unobserved Trend

Nils Herger

Working Paper 24.01

A Runs Test for Stock-Market Prices with an Unobserved Trend

Nils Herger*

January, 2024

Abstract

To analyze whether stock-market prices follow a random walk, the algebraic sign of their returns has been compared with a coin toss, which is a prominent example for a Bernoulli trial with equiprobable outcomes. Like coin tosses, signed returns lend themselves for a simple runs test for randomness. However, they typically comprise an unobserved trend, and therefore represent Bernoulli trials whose theoretical outcome probability is not easily known. Fortunately, the Von Neumann algorithm can transform Bernoulli trials with unknown outcome probabilities into equiprobable outcomes. Thus, a runs test on correspondingly transformed returns can handle an unobserved stock-market trend.

JEL classification: C14, C22, G11, G14
Keywords: coin tossing; stock-market prices; random walk; runs test; Von Neumann algorithm

1 Introduction

In the popular imagination, gambling in casinos and investing in stock markets have much in common. On closer inspection, this analogy runs somewhat deeper than a mere cynical view on modern capitalism, insofar as the risk and luck associated with games of chance, including coin tossing, have indeed served as illustrative examples when mathematicians, such as Jakob Bernoulli (1655-1705), began to develop the very probability theory on which the modern analysis of stock-market prices is based.[1] Against this background, it is perhaps not surprising that early empirical work on the hypothesis that these prices follow a random walk, and can therefore not be predicted from their own past, compared their behavior with that of coin tosses, which are a textbook example for a Bernoulli trial with equiprobable outcomes (see Cowles and Jones, 1937, p.281; Osborne, 1959, pp.55ff.; Cowles, 1960, p.909). Another case in point is the runs test, which employs the number of uninterrupted sequences within a discrete variable to infer whether its realizations appear in a purely random manner. Obviously, this test can be applied to repeated coin tosses to ask whether their outcomes are fair in the sense that heads and tails have occurred, a priori, with equal probability (Siegel and Castellan, 1988, pp.58-59). That runs test for randomness have also been used to analyze the ordering of positively and negatively signed stock-market returns, and notably featured in the path-breaking contribution of Fama (1965a, pp.74-81) on the random-walk hypothesis, bears further testimony to the analogy mentioned at the outset.

In contrast to coin tosses, the stochastic behavior of stock-market prices continues to be the subject of a fierce debate with random and non-random walks down Wall Street both being propagated by popular books on this topic (Malkiel, 2011; Lo and McKinley, 1999). By and

*Study Center Gerzensee, E-mail: nils.herger@szgerzensee.ch.

[1]See Bernstein (1998) for a historical account of the overlaps between the analysis of games of chance, the development of probability theory, and modern theoretical and empirical research on financial markets.

large, this debate stems from the delicacies of statistical testing for randomness with prices that are not necessarily driven by an unchangeable, and more or less self-evident, stochastic process (Rapach and Zhou, 2013). With specific regard to the runs test for randomness, an outstanding feature is that stock markets are usually driven by an upward trend, which is not directly observable (see e.g. Campbell *et al.*, 1997, pp.33ff.). Then again, that the standard framework for the corresponding prices is a random walk *with* an unobserved drift term is not innocuous. For example, although Fama (1965a), Praetz (1969), Dryden (1970), and Jennergren and Korsvold (1974) all found that observed sequences comprising signed stock-market returns tend to produce significantly fewer runs than expected from Bernoulli trials with equiprobable outcomes, this finding leaves ample room for interpretation. In particular, as long as the properties and effects of the underlying trend are partially unknown, it remains unclear whether a shortage of runs is a reflection of non-randomness, or merely an artefact of a trend (or drift) giving rise to clustered return increases. By way of contrast, observing significantly fewer runs between heads and tails than expected from an equiprobable Bernoulli trial would be highly suspicious in a game of supposedly unbiased (or fair) coin tosses, for which the stochastic process is no source of controversy.

This paper endeavors to contribute to the literature on the runs test for randomness in stock-market prices by taking into account that they are almost always driven by unobserved trends. Based on a discussion of their effect on the derivation and interpretation of the runs test, the main contribution lies in suggesting a way to deal with unobserved trends by pretreating sequences comprising signed stock-market returns with an algorithm proposed by Von Neumann (1951). Originally, this algorithm was developed to transform coin tosses that might suffer from an unknown bias, meaning that there is an incalculable suspicion that heads or tails occur more often, into Bernoulli trials with equiprobable outcomes. This transformation is done by tossing the coin twice and retaining only the first observation when different outcomes arise. In principle, the Von Neumann unfair coin algorithm can be adapted to a sequence of stock-market returns with an unobserved trend, whose effect on the probability of observing a positive return is hard, or even impossible, to determine. By doing so, this paper aims at developing a more conclusive version of the runs test to determine whether signed stock-market returns appear in a completely random manner.

The paper is organised as follows. Setting the stage, the next section defines the conditions under which stock-market prices follow a random walk. This discussion starts with the size of returns, but mainly contemplates their algebraic sign, which provides the basis for the runs tests for randomness. Section 3 discusses this test and derives the corresponding test statistics when positive and non-positive stock market returns are not necessarily equiprobable outcomes. Section 4 explains why the effect of unobserved stock-market trends on the probability of observing a positively signed return is at most vaguely known, and why they undermine the interpretation of a standard runs test. Section 5 proposes a way to deal with this problem by adapting the Von Neumann algorithm to the case of stock-market prices with an unobserved trend. Section 6 contains an example applying the Von Neumann algorithm to more than 13'000 work-daily observations of the Dow Jones Industrial Average (DJIA) between 1970 and 2023. The final section concludes.

2 Random walk in stock-market returns

2.1 Conventional random walk

Consider a stock-market price process $\{P_t\}$ observed at time periods denoted by subscripts $t = 0, 1, 2, \ldots, T$. The logarithmic values of these prices, i.e. $p_t = \ln(P_t)$, are conventionally modelled through a recursive equation given by

$$p_t = \delta + p_{t-1} + \epsilon_t \qquad \text{with} \quad t = 0, 1, 2, \ldots, T, \tag{1}$$

whereby δ denotes an unobserved price trend, also called drift, and ϵ_t captures random disturbances with an expected value of 0 (see, among many others, Campbell *et al.*, 1997, pp.33ff; Lo and McKenzie, 1999, pp.19ff.). Because logarithmic differences represent percentage changes, the stock-market return r_t between $t-1$ and t is given by $r_t = p_t - p_{t-1}$. Note that r_0 cannot be calculated from the data. However, across the remaining observations, (1) lends itself to a rearrangement into

$$r_t = p_t - p_{t-1} = \delta + \epsilon_t \qquad \text{with} \quad t = 1, 2, \ldots, T, \tag{2}$$

whereby δ reflects the average return, and ϵ_t is typically thought to be drawn from some stochastic distribution. According to Fama (1965a, p.35), within this environment, the random-walk hypothesis is encapsulated in the following, interrelated conditions:

Definition 1: The stock-market returns r_t of (2) follow a random walk across $t = 1, 2, \ldots, T$, when the underlying price changes $p_t - p_{t-1}$ are (serially) independent, and conform to some probability distribution.

Comment 1: When $\delta \neq 0$, (2) represents a random walk with drift.

It is well known that the random-walk property has profound implications for the behavior of stock-market prices and returns. In particular, when the conditions of Definition 1 are fulfilled, these prices and returns follow no regular pattern and, hence, cannot be predicted from their own past in a meaningful way (Fama, 1965b, p.56). Considering this far-reaching implication, it is perhaps not surprising that various statistical tests for random walks have been applied to a vast number of stock-market prices and indices. A key challenge for this research has arisen from the fact that the stochastic distribution underlying the returns of (2) is neither obvious, nor necessarily unchangeable across time. Rather, the behavior of stock-market prices and returns is characterized by substantial model instability and, hence, uncertainty (Rapach and Zhou, 2013). In this regard, the effects of extreme value distributions (see e.g. Jansen and De Vries, 1992), and questions on the appropriate modelling of time-varying return volatility loom particularly large (see e.g. Andersen *et al.*, 2006). When random-walk tests are susceptible to this model uncertainty, their results can remain inconclusive.[2]

2.2 Random walk in signed stock-market returns

Focusing on the algebraic sign, rather than the value, of stock-market returns provides a way to avoid some of the model uncertainty underlying r_t. To represent this algebraic sign, a binomial indicator variable I_t^r is constructed, which adopts a value of 1 in case r_t is positive and 0 otherwise, that is

$$I_t^r = \begin{cases} 1 & \text{if} \quad r_{t+1} = p_{t+1} - p_t > 0 \quad \text{with probability} \quad \pi \\ 0 & \text{if} \quad r_{t+1} = p_{1+t} - p_t \leq 0 \quad \text{wih probability} \quad 1 - \pi. \end{cases} \tag{3}$$

Signed stock-market returns formed the basis for some of the earliest tests on the random-walk hypothesis (Campbell *et al.* 1997, pp.35ff.). Probably, the simplicity to calculate their results was a major consideration when Cowles and Jones (1937) or Fama (1965a) introduced these tests. However, since computational power has become easily available, it is rather the above-mentioned model uncertainty as regards the stochastic process of stock-market returns that provides the key advantage for focusing on binomial outcomes encapsulated in

[2] A case in point are mean-reversion tests, where (2) is rearranged into a regression equation given by

$$r_t = \delta + \phi(r_{t-1} - \delta) + \epsilon_t.$$

Within this regression equation, returns deviate from the random-walk property when their current value r_t is correlated with pervious return deviations from the drift δ, i.e. when $\phi \neq 0$. Fama and French (1988) as well as Poterba and Summers (1988) report cases, where the parametrization of these mean-reversion tests does affect their result.

I_t^r. Indeed, the assumptions for testing the random-walk hypothesis by means of (3) are essentially restricted to the probability of observing a positive return.

Statistically, I_t^r represents a Bernoulli trial with two outcomes, whose probability is denoted by, respectively, π and $1-\pi$. Furthermore, across time periods $t = 1, 2, \ldots, T$, the resulting Bernoulli trials give rise to a sequence of random variables $I_t \in (0, 1)$ displaying the chronology of signed stock-market returns. For an illustration of this type of sequence, contemplate the positive and non-positive returns calculated from 15 observations of the Dow Jones Industrial Index (DJIA) at the beginning of the year 1970. Owing to nontrading days on weekends or public holidays, the dates of these recordings are irregular. Also, as mentioned above, r_0 and, in turn, I_0^r cannot be calculated from the data. However, the chronology according to (3) over the next 14 values of the DJIA is given by

Year 1970:	2.Jan	5.Jan	6.Jan	7.Jan	8.Jan	9.Jan	12.Jan	13.Jan	14.Jan	15.Jan	16.Jan	19.Jan	20.Jan	21.Jan	22.Jan	
$\mathbf{I_t^r}$:	NA	1	0	0	1	0	0	0	0	0	0	0	1	1	1	$\rightarrow$
DJIA:	809.2	811.3	803.7	801.8	802.1	798.1	790.5	788.0	787.2	785.0	782.6	776.1	777.9	782.3	786.1	...

To test whether this sequence reflects a purely random ordering, contemplate the conditional probabilities of the possible transitions between pairs of consecutive return indicators I_t^r and I_{t+1}^r. In particular, $\pi^{1|1}$ defines the probability of observing a positive return in $t+1$ given a positive return in t. The counterpart of observing a negative return conditional on a positive return occurs with probability $\pi^{0|1} = 1 - \pi^{1|1}$. In a similar vein, $\pi^{0|0}$ pertains to the case of successive negative returns and $\pi_t^{1|0} = 1 - \pi^{0|0}$ refers to the conditional probability with the superscript indicating that a positive return will occur given a negative return. Commonly, conditional probabilities are displayed through the Markov-chain transition matrix, which is here given by

$$\begin{array}{ccc} & & I_{t+1}^r \\ & & \begin{array}{cc} 1 & \quad 0 \end{array} \\ I_t^r & \begin{array}{c} 1 \\ 0 \end{array} & \begin{pmatrix} \pi^{1|1} & \pi^{0|1} \\ \pi^{1|0} & \pi^{0|0} \end{pmatrix}. \end{array} \tag{4}$$

The transition matrix (4) represents the short-term dynamics of I_t^r in terms of describing how the algebraic sign of stock-market returns depends on that of past observations. In principle, the stochastic transition between signed returns can either follow some regular pattern, including clusters or recurrent cycles, or reflect a completely random arrangement. Recall from Definition 1 that randomness in stock markets typically requires that current prices, and the sequence of returns derived from them, are (serially) independent of past prices and returns (see also Campbell *et al.*, 1997, pp.28ff.; Lo and McKinley, pp.3ff.). In signed returns, according to the standard statistical definition, this kind of independence occurs when conditional probabilities coincide with their corresponding marginal, or unconditional, probability.[3] More specifically, with a marginal probability of π for observing a positive stock-market return, statistical independence across the four possible contingencies of the transition between I_t^r and I_{t+1}^r occurs when

$$\begin{aligned} \pi^{1|1} &= \pi \qquad & \pi^{0|1} &= 1 - \pi \\ \pi^{1|0} &= \pi \qquad & \pi^{0|0} &= 1 - \pi. \end{aligned} \tag{5}$$

Substituting (5) into (4) yields a simplified 2×2 transition matrix for the Markov chain given by

[3] Here, independence refers only to adjacent returns. In addition, it would be possible to look at independence across returns separated by more than one period. The empirical example of Sec. 6 will deal with this issue in a pragmatic way by contemplating stock-price data at the daily, weekly, and monthly frequency.

$$
\begin{array}{ccc}
 & & I^r_{t+1} \\
 & & \begin{array}{cc} 1 & 0 \end{array} \\
I^r_t & \begin{array}{c} 1 \\ 0 \end{array} & \begin{pmatrix} \pi & 1-\pi \\ \pi & 1-\pi \end{pmatrix}.
\end{array}
\tag{6}
$$

In (6), complete randomness manifests itself in a probability of signed stock-market returns as indicated by I^r_{t+1} that does not depend on the realization of I^r_t.

As alluded to at the outset, coin tosses provide a textbook example for Bernoulli trials with statistical independence. Then again, the analogy between coin tossing and signed stock-market returns is far from being perfect. For example, constant returns, which have no practical similarity in coin tossing, are a conceivable stock-market outcome alongside positive and negative returns. In (3), constant and negative returns have somewhat arbitrarily been grouped together into a non-positive category with $I^r_{t+1} = 0$.[4] This asymmetry can imply that positive and non-positive returns are not necessarily equiprobable outcomes, i.e. $\pi \neq \frac{1}{2}$ (Campbell *et al.*, 1997, pp.35-36). As discussed in more detail in Section 4, the presence of an unobserved trend, as emphasized in Comment 1, only adds to this problem when testing the random-walk hypothesis through I^r_t. To frame this discussion, the following definition and comment stipulate the interrelated conditions of the random-walk hypothesis when contemplating signed stock-market returns.

Definition 2: The signed stock-market returns r_t, as indicated by I^r_t of (3), follow a random walk across $t = 1, 2, \dots, T$, when the transition probabilities between I^r_t and I^r_{t+1} are serially independent, and, hence, each time period t has the same a priori probability π of witnessing a positive stock-market return.

Comment 2: For stock-market prices, random walks can be consistent with $\pi \neq \frac{1}{2}$.

3 Conventional runs test for stock-market returns

The runs test for randomness was introduced by Fama (1965a) to the analysis of stock markets, and subsequently also employed by Praetz (1969), Dryden (1970), and Jennergren and Korsvold (1974) to infer whether signed returns are serially independent and, hence, fulfill a key property of the random-walk hypothesis of Definition 2.[5] In statistics, a run refers to an uninterrupted sequence of any length, across which data share the same characteristic, reflect similar events, or represent like objects (see Bradley, 1968, p.251; Siegel and Castellan, 1988, p.58). Typical applications involve binomial variables, such as coin tosses where runs summarize successions of heads and tails. In the analysis of stock markets, the focus lies on signed returns, with a run comprising unbroken sequences of increasing or decreasing prices. The question is whether the corresponding ordering encapsulates some regular pattern, or is completely random. As an illustration, contemplate the runs that arise within the above-mentioned 14 observations of the DJIA.

$$
\begin{array}{ll}
\underline{\textbf{Year 1970}}: & \textbf{5.Jan} \quad \textbf{6.Jan} \quad \textbf{7.Jan} \longrightarrow \\
\mathbf{I^r_t}: & \underbrace{1}_{\text{run 1}} \ \underbrace{0 \quad 0}_{\text{run 2}} \ \underbrace{1}_{\text{run 3}} \underbrace{0\ 0\ 0\ 0\ 0\ 0\ 0}_{\text{run 4}} \underbrace{1\ 1\ 1}_{\text{run 5}} \ \underbrace{\dots}_{\text{etc.}} \\
\textbf{Runs:} &
\end{array}
$$

When testing for randomness in stock-market returns, the focus lies usually on the number of runs, denoted by N_{runs}, observed across $t = 1, 2, \dots T$ (see e.g. Campbell *et al.*, 1997,

[4] Fama (1965a) as well as Jennergren and Korsvold (1974) consider a separate category for zero returns (or constant prices). Usually, zero-valued returns occur very rarely in stock-market data.

[5] In statistics, an early comprehensive derivation of the distribution of runs can be found in Mood (1940). Runs have been used for testing various scenarios, including whether two samples have been randomly drawn from the same population (Wald and Wolfowitz, 1940), whether the probability of success—e.g. having no breakdown in a production process—remains constant across time (Mosteller, 1941), or whether runs up (e.g. sequences of increasing observations) and runs down (e.g. sequences of decreasing observations) reflect a purely random ordering (Moore and Wallis, 1943). For textbook discussions of the runs test, see Bradley (1968, Ch.11 and Ch.12), Siegel and Castellan (1988, Ch.4.5.), or Gibbons and Chakraborti (2003, Ch.3.).

pp.39-41).[6] In principle, deviations from the benchmark of complete randomness—implying a lack of independence—can manifest themselves in too few runs, which could signal that price increases are clustered due to momentum effects or trends, but also in too many runs, which could result from some cyclical pattern. Obviously, the implementation of the runs test warrants a clearly defined benchmark of randomness. In concurrence with Definition 2, an ordering of signed stock-market returns shall here be considered as random when I_t^r of (3) has, a priori, the same probability π of recording a value of 1 across all time periods $t = 1, 2, \ldots, T$ (see also Bradley, 1968, pp.250, 277).

Crucial components of the runs test are the theoretical mean and standard deviation of N_{runs}. When making the corresponding derivation based on the statistical theory of Bernoulli trials, textbooks have often had examples with equiprobable outcomes in mind (see e.g. Siegel and Castellan, 1988, pp.58-60). To analyze the behavior of stock-market prices, this scenario is clearly inadequate (see Comment 2). For obtaining a more general derivation of the theoretical mean and standard deviation of N_{runs}, it will be helpful to know the probability of observing a reversal, which refers to cases of sign switches of stock-market returns between t and $t{+}1$. To compute these reversals, as indicated by R_t^r, recall from Definition 2 that under independence and complete randomness, positive returns occur with a constant probability π. Hence, the probability that they occur in succession at t and $t+1$ is given by π^2. In a similar vein, the probability of successive negative returns is given by $(1-\pi)^2$. Consequently, the probability of observing a sign reversal equals $1 - \pi^2 - (1-\pi)^2 = 2\pi(1-\pi)$.[7] Taken together, the binomial variable describing a sign reversal is characterized by

$$R_t^r = \begin{cases} 0 & \text{with probability} \quad \pi^2 + (1-\pi)^2 \\ 1 & \text{wih probability} \quad 2\pi(1-\pi). \end{cases} \tag{7}$$

Intuitively, the theoretical mean of the number of runs under complete randomness, as denoted by $\mu(N_{\text{runs}})$, reflects that they are by definition delimited by the sign reversals indicated by R_t^r. Consequently, $\mu(N_{\text{runs}})$ simply comprises the sum of the initial run that automatically appears at $t = 1$ and the expected number of sign reversals, which occur independently and with constant probability $2\pi(1-\pi)$ across the $T-1$ remaining observations. The formal result, as derived in Appendix A, is given by

$$\mu(N_{\text{runs}}) = 1 + (T-1)2\pi(1-\pi). \tag{8}$$

A more complicated derivation, which is again relegated to Appendix A, yields the corresponding standard deviation, which is for large values of T approximately given by

$$\sigma(N_{\text{runs}}) \approx 2\sqrt{(T-1)\pi(1-\pi)[1-3\pi(1-\pi)]}. \tag{9}$$

For large samples, the distribution of the runs test converges to a normal distribution with mean $\mu(N_{\text{runs}})$ and standard deviation $\sigma(N_{\text{runs}})$ (see e.g. Siegel and Castellan, 1988, p.62; Gibbons and Chakraborti, 2003, pp.83-84). Hence, the null-hypothesis that the number of runs does not deviate from a completely random ordering can be tested via a conventional z-test for mean differences. Formally, the test statistic $z(N_{\text{runs}})$ is a function of the difference between the observed and expected number of runs, e.g. $N_{\text{runs}} - \mu(N_{\text{runs}})$, standardized by σ_{runs}.

Runs test for randomness (with $0 < \pi < 1$):

$$\begin{aligned} z(N_{\text{runs}}) \quad &= \quad \frac{N_{\text{runs}} - \mu(N_{\text{runs}})}{\sigma_{\text{runs}}} \overset{a}{\sim} N(0,1) \\ &\text{with} \quad \mu(N_{\text{runs}}) = 1 + (T-1)2\pi(1-\pi) \\ &\text{and} \quad \sigma(N_{\text{runs}}) \approx 2\sqrt{(T-1)\pi(1-\pi)[1-3\pi(1-\pi)]} \end{aligned} \tag{10}$$

[6] In principle, runs tests can also be based on the maximal length of runs.

[7] As a predecessor of the runs test, Cowles and Jones (1937) developed a statistical test for randomness in stock markets based on the ratio between the number of these types of sequences, and sign reversals. For a discussion, see Campbell *et al.* (1996, pp.34-38).

Significant mean deviations according $z(N_{\text{runs}})$ imply a rejection of the null-hypothesis of complete randomness in signed stock-market returns. Further to the discussion above, the corresponding test is typically two-sided, because too many as well as too few runs are conceivable alternatives to the hypothesis of completely random stock-market returns.

4 Complications with the unobserved stock-market trend

Thanks to the focus on signed returns, the runs test is less susceptible to potential disturbances from outliers, structural breaks, parameter instability, or other thorny features of stock-market price processes. Still, the theoretically expected mean of the number of runs $\mu(N_{\text{runs}})$ and the corresponding standard deviation $\sigma(N_{\text{runs}})$ depend on the underlying probability π of observing a positively signed return. Therefore, by design, the runs test leaves room for interpretation when this outcome probability is not exactly known. By the same token, it is hard to distinguish between the effect of a given value of π and certain non-random patterns, such as momentum effects, on the number of runs (Bradley, 1968, p.280; Siegel and Castellan, 1988, p.59). In the classical coin tossing example, these issues hardly matter, as a fifty-fifty chance of observing heads and tails can theoretically be presupposed. By setting $\pi = \frac{1}{2}$, the calculation of the theoretical mean of (8), the corresponding standard deviation of (9), and ultimately the test statistic of the number of runs under randomness (10) is indeed straightforward. Conversely, as emphasized above, in this regard the analogy between coin tossing and stock-market prices is wanting. In particular, further to Comment 2, there is no reason why positive and non-positive stock-market returns should, a-priori, be equiprobable outcomes.

When testing for random walks in stock-market returns, the complications associated with the probability π are further aggravated by the fact that the underlying trend, as reflected by the drift term δ of (2), is not directly observable. However, unless the value of δ as well as its mapping into π are known, it remains unclear what conclusions can be drawn from the test statistic of (10). For example, the observation of relatively few runs of positive stock-market returns could either reflect non-random momentum effects, or a random walk with a sufficiently strong upward trend. In principle, various trends, as comprised in δ, can give rise to a broad range of probabilities π and, in turn, different numbers of runs expected under complete randomness (Campbell *et al.*, 1997, p.40). Moreover, even if stock-market trends were observable, uncovering their mapping into the outcome probability π would be anything but straightforward. For example, although it sounds plausible that upward trends, i.e. $\delta > 0$, should make positively signed stock-market returns relatively more likely, i.e. $\pi > \frac{1}{2}$, this relationship cannot be taken for granted. The reason is that trends could also result from larger average increments of upward than of downward price changes. Conceptually, the connections between δ and π, as described by the function $\pi = Pr(r_t > 0) = f(\delta)$, depend on the model postulated for the price process of (1). Consequently, in all attempts to establish $\pi = Pr(r_t > 0) = f(\delta)$, questions about model uncertainty could reenter through the back door. Hence, if the key advantages of the runs test are to be retained, a different approach is warranted.

5 A runs test for stock-market returns with equiprobable outcomes

Provided that the unobserved trend has an unclear effect on the probability of observing a positive return, and therefore undermines the interpretation of the runs test, removing this effect would facilitate the analysis of stock-market prices. Fortunately, Von Neumann (1951) has developed an algorithm for the similar task of transferring coin tosses, which are suspected to suffer from an unobserved bias (i.e. $\pi \neq \frac{1}{2}$), into equiprobable outcomes (i.e.

$\tilde{\pi} = \frac{1}{2}$).[8] More specifically, consider a potentially unfair coin with an unknown probability π of observing heads, which is an outcome labelled by H, and an equally unknown probability $1-\pi$ of observing tails, which is an outcome labelled by T. Despite the suspected bias hidden in π, the Von Neumann algorithm can reproduce a fair (or fifty-fifty) outcome through the following three-step procedure.

Von Neumann algorithm (for coin tosses):

1. Toss the potentially biased coin twice.
2. In case the same outcomes (i.e., HH or TT) arise, return to step 1.
3. In case mixed outcomes (i.e., HT or TH) arise, retain the observation of the first coin toss.

To understand why under randomness and independence $\tilde{\pi}(H) = \tilde{\pi}(T) = \frac{1}{2}$ necessarily results from the algorithm, recall from the discussion above that the probability of observing either HH or TT in step 2 is given by $\pi^2 + (1-\pi)^2 = 1 - 2\pi(1-\pi)$.[9] Crucially, in step 3, the probabilities of observing either HT or TH are the same and given by $\pi(1-\pi)$. Across the three steps, $\tilde{\pi}$ summarizes the joint probabilities of either restarting over at step 2, i.e. $(1-2\pi(1-\pi))\tilde{\pi}$, or arriving at step 3, i.e. $\pi(1-\pi)$. Hence, $\tilde{\pi} = (1-2\pi(1-\pi))\tilde{\pi} + \pi(1-\pi)$. Solving this expression for $\tilde{\pi}$ yields $\tilde{\pi} = \frac{1}{2}$ for the mixed outcomes (i.e., HT or TH). The crucial ingredient for arriving at this result is statistical independence, which implies that heads (and tails) have, a priori, the same probability of being observed across coin tosses. Without this property, the Von Neumannn trick would not work (see e.g. Samuelson, 1968, p.1526).

Consider now the analogy between dealing with potentially biased coin tosses and stock-market prices driven by an unobserved trend, whose effect on the probability π for a positive return is at most partially known. Then again, it would be helpful when these signed stock-market returns could be converted into a sequence, within which positive and non-positive returns are equiprobable. To this end, the following version of the Von Neumann algorithm can be applied to the signed stock-market returns encapsulated in the binomial variable I_t^r of (3).

Von Neumann algorithm (for signed stock-market returns):

1. Arrange the sequence of stock-market return indicators I_t^r observed across $t = 1, 2, \ldots, T$ into non-overlapping pairs of consecutive observations.
2. Drop pairs with sequences of positive or negative returns (e.g. pairs between t and $t+1$ with $I_t^r = I_{t+1}^r$).
3. Retain the first observation of pairs, within which a reversal occurs (e.g. pairs between t and $t+1$ with $I_t^r \neq I_{t+1}^r$).

Under the assumption of randomness and independence, this algorithm should give rise to a transformed sequence with $\tilde{T}$ observations of positively or non-positively signed stock-market returns that occur with equal probability (i.e. $\tilde{\pi} = \frac{1}{2}$). Given this probability, the calculation of the theoretical mean $\mu(\tilde{N}_{\text{runs}})$ and standard deviation $\sigma(\tilde{N}_{\text{runs}})$ of this transformed sequence from (8) and (9) is straightforward. As a result, the following, simplified test statistic for the runs test for randomness arises.

[8] Subsequently, Hoeffding and Simons (1970) analyzed the The Von Neumann algorithm in greater detail.

[9] This probability results directly from the indicator variable of observing a reversal R_t^r according to (7).

Runs test for randomness (with $\tilde{\pi} = \frac{1}{2}$):

$$
\begin{aligned}
z(\tilde{N}_{\text{runs}}) \quad &= \quad \frac{\tilde{N}_{\text{runs}} - \mu(\tilde{N}^{\text{runs}})}{\sigma(\tilde{N}^{\text{runs}})} \overset{a}{\sim} N(0,1) \\
&\text{with} \quad \mu(\tilde{N}_{\text{runs}}) = 1 + (\tilde{T} - 1)/2 \\
&\text{and} \quad \sigma(\tilde{N}_{\text{runs}}) = \frac{\sqrt{\tilde{T} - 1}}{2}
\end{aligned}
$$

As an illustration for a sequence of binomial variables transformed by the Von Neumann algorithm, contemplate again the stock-market returns resulting from the observations of the DJIA at the beginning of the year 1970. The above-mentioned three-step procedure to obtain equiprobable outcomes under randomness and independence yields

$$
\begin{array}{ll}
\underline{\text{Year 1970}}: & \text{5.Jan 6.Jan} \longrightarrow \\
\mathbf{I^r_t}: & \underbrace{\underbrace{1 \quad 0}_{1}}_{\text{run } \tilde{1}} \quad \underbrace{\underbrace{0\,1}_{0}\ \underbrace{0\,0}_{\text{drop}}\ \underbrace{0\,0}_{\text{drop}}\ \underbrace{0\,0}_{\text{drop}}\ \underbrace{0\,1}_{0}\ \underbrace{1\,1}_{\text{drop}}}_{\text{run } \tilde{2}} \quad \underbrace{\underbrace{\ldots}_{\text{etc.}}}_{\text{run } \ldots} \\
\textbf{Runs:} &
\end{array}
$$

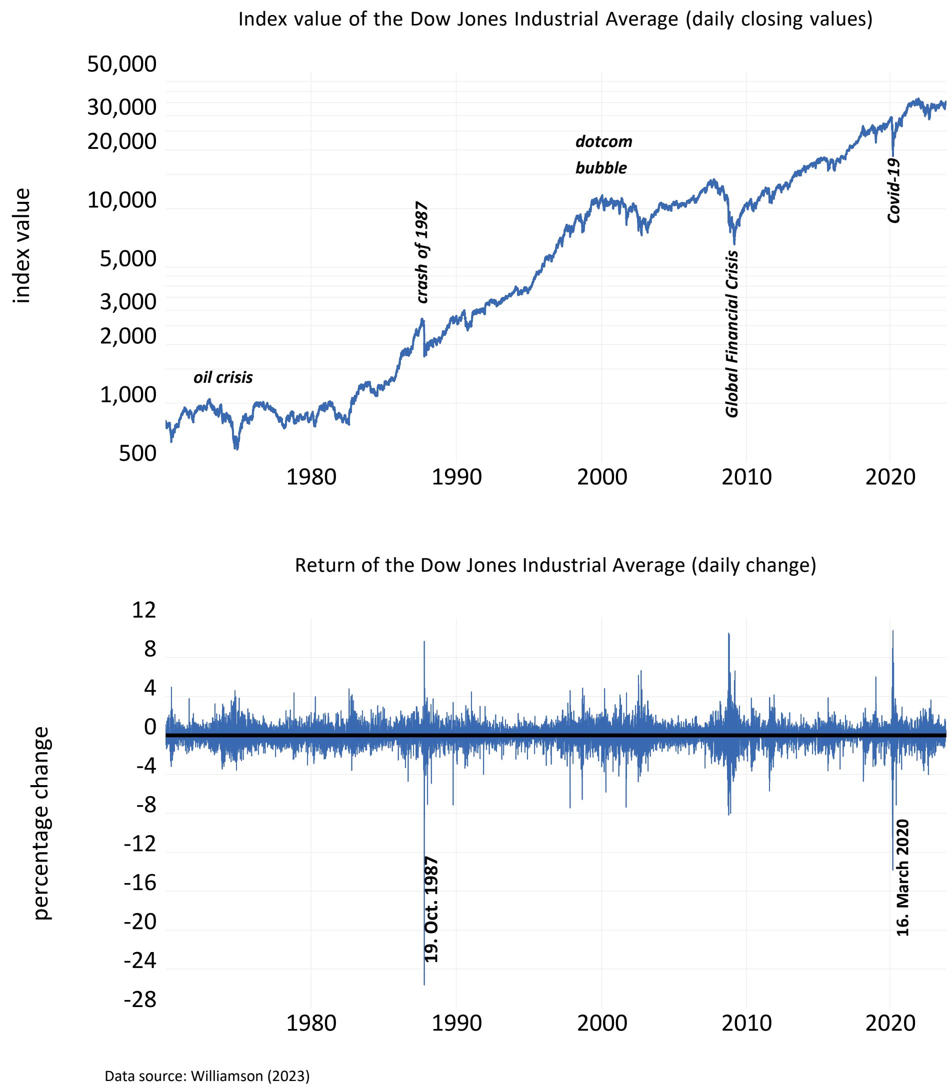
6 Example: Dow Jones Industrial Average (1970-2023)

To illustrate the runs test for randomness in the analysis of stock-market returns with an example, this section contemplates the Dow Jones Industrial Average (DJIA) for the period beyond the beginning of 1970. In particular, on a logarithmic scale, the top panel of Figure 1 depicts more than 13'0000 daily closing values of the DJIA up until the end of the year 2023. The resulting line graph underscores that stock-market prices p_t tend to follow an upward trend, which is constantly affected by erratic shocks and recurrent crashes that sometimes degenerate into outright crises. Within the present sample, the economic turbulences following the oil-price shock of 1973, the crash of 1987, the dotcom bubble around the turning of the millennium, the instability after the Global Financial Crisis of 2008, and the economic disruptions amid the Covid-19 pandemic represent outstanding boom-and-bust episodes on the stock market.

The bottom panel of Figure 1 depicts the corresponding daily returns as calculated from (2). Reflecting the just-mentioned erratic development of the DJIA, these returns follow no simple stochastic process. In particular, various episodes of economic, financial, and political crises are reflected in complex patterns of return volatility. Moreover, crashes have manifested themselves in outliers, such as the negative daily return of -13.8 per cent on 16. March 2020, and -25.6 per cent on 19. October 1987.

By means of (2), the returns of the bottom panel of Figure 1 can be coded into positive and non-positive signs. The resulting realisation of the indicator variable I_t^r provides, in turn, the basis for the runs test for randomness of (10). Table 1 summarizes the corresponding results. In particular, the top panel focuses on the daily closing values between 1970 and 2023 with a total of 13,569 observations, among which 6,589 runs have been recorded. As discussed in Sec. 4, whether or not this number reflects a scenario with too many or too few runs with respect to the benchmark of complete randomness depends on the probability π of observing an increase of the DJIA. In this regard, column (1) contemplates the case of equiprobability, i.e. $\pi = 0.5$, under which the theoretically expected number of runs $\mu(N_{\text{runs}})$, as calculated from (8), is 6,785. The corresponding standard deviation $\sigma(N_{\text{runs}})$, as calculated from (9), is 58.24. A runs test based on the standardized mean difference of (10) yields a z-statistic of -3.37, which is statistically different from 0 at any conventionally used level of rejection. This result would suggest that the DJIA contains too few runs between 1970 and 2023, which coincides with the findings reported in Fama (1965a,

Figure 1: Dow Jones Industrial Average (1970-2023)



p.76) or Jennergren and Korsvold (1974, p.179). Nevertheless, it remains unclear whether this finding provides evidence against the hypothesis of randomly arranged returns, as the upward trending behavior of the DJIA is hardly compatible with equiprobable positive and negative returns. These doubts are reinforced by the fact that higher values of π affect the just-mentioned statistical conclusion as reported in columns (2) to (10), across which the theoretically expected number of runs declines considerably. In particular, when π lies above 0.55, the runs test is barely significant or becomes even non-significant. Under these scenarios, the observed number of runs within the daily returns of the DJIA of Figure 1 would no longer be significantly lower than expected under pure randomness. In case $\pi = 0.59$, the observed number of runs even exceeds the theoretical benchmark of 6,565 of column (10). Although, with a z-value of 0.40, the corresponding deviation is non-significant.

The main conclusion that can be drawn from the runs tests of columns (1) to (10) of the top panel of Table 1 depends, obviously, on the value of π. In principle, the probability of observing an increase of the DJIA can be estimated from the data. Within the current sample, 7,085 out of 13,569 observations reported an increase implying an estimated probability of 7,085/13,569 $\approx$ 0.522. Despite this estimate, the following question about the interpretation of the results remains: Would a rejection of the hypothesis of randomness at a given value of π merely imply that the DJIA follows a random walk with drift, or reflect some genuine non-random pattern? To avoid this question, column (11) reports the result of a runs test, where the DJIA has been transformed through the Von Neumann algorithm of Sec. 5. Recall from discussion above that under randomness and independence, the corresponding three-step procedure necessarily yields equiprobable outcomes, which lend themselves for a runs test that is no longer susceptible to discussions about the outcome probability. Of course, a downside of reconstructing a sequence of the DJIA with $\tilde{\pi} = \frac{1}{2}$ is that around three quarters of the observations are dropped within the process. However, owing to the large number of closing values of the DJIA at the daily frequency, a sample with 3,267 observations, among which 1,656 runs are observed, remains. For the current example, the corresponding runs test yields a z-value of 0.77 and, hence, does not reject the hypothesis that signed returns of the DJIA represent a purely random ordering.

In the empirical analysis of stock-market prices, daily data can suffer from distortions due to nontrading on weekends or public holidays (Lo and McKinley, 1999, pp.26-27). Furthermore, the focus on sequences with daily closing prices cannot reveal non-random patterns that potentially arise over somewhat longer time horizons. To address these issues, the middle of Table 1 focuses on a sample of the DJIA with end-of-the-week observations (e.g. Fridays), which encompasses 2,712 observations and 1,287 runs between 1970 and 2023. In essence, the results of the runs tests with weekly data coincide with those of daily data. In particular, in columns (1) to (10), there tend to be fewer observed runs when compared with the benchmark of pure randomness. However, the significance of these differences depends on the value of π. Furthermore, when avoiding questions as regards this probability by means of the Von Neumann algorithm, there are 331 runs among 649 observations. According to the z-statistics of column (11), in this case, the hypothesis of pure randomness cannot be rejected.[10] A potential drawback of the runs test for randomness is illustrated by the bottom panel, where the data have been summarized into end-of-month observations. Within this sample, there are only 646 observations and 323 runs left. The resulting lack of statistical power could explain why no statistically significant results arise for all values of π between 0.5 and 0.59. The Von Neumann transformation of column (11) aggravates the situation in the sense of leaving only 166 observations and 86 runs.

[10] Similar results arise when contemplating weekly data based on mid-week, i.e. Wednesday, observations.

Table 1: Results of runs tests on the Dow Jones Industrial Average (1970-2023)

		Conventional runs test with different probabilities for a positive return										V. Neumann Alg.	
Prob. of pos. return	π	0.50	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59	$\tilde{\pi}$ (fix)	0.50
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)		(11)
		Daily changes of closing price											
Nr. of observations	T						13,569					$\tilde{T}$	3,267
Obs. nr. of runs	N_{runs}						6,589					$\tilde{N}_{\text{runs}}$	1,656
Theoretical nr. of runs	$\mu(N_{\text{runs}})$	6,785	6,782	6,774	6,761	6,742	6,717	6,687	6,589	6,611	6,565	$\mu(\tilde{N}_{\text{runs}})$	1,634
Std. of runs	$\sigma(N_{\text{runs}})$	58.24	58.26	58.33	58.45	58.61	58.81	59.06	59.34	59.66	60.01	$\sigma(\tilde{N}_{\text{runs}})$	28.57
Runs test	$z(N_{\text{runs}})$	-3.37***	-3.32***	-3.17***	-2.94***	-2.60***	-2.18**	-1.66	-1.06	-0.27	0.40	$z(\tilde{N}_{\text{runs}})$	0.77
		Weekly changes between end-of-week observations											
Nr. of observations	T						2,712					$\tilde{T}$	649
Obs. nr. of runs	N_{runs}						1,287					$\tilde{N}_{\text{runs}}$	331
Theoretical nr. of runs	$\mu(N_{\text{runs}})$	1,357	1,356	1,354	1,352	1,348	1,343	1,330	1,330	1,322	1,313	$\mu(\tilde{N}_{\text{runs}})$	325
Std. of runs	$\sigma(N_{\text{runs}})$	26.02	26.04	26.08	26.13	26.61	26.29	26.40	26.52	26.67	26.82	$\sigma(\tilde{N}_{\text{runs}})$	12.73
Runs test	$z(N_{\text{runs}})$	-2.68***	-2.65***	-2.58***	-2.47***	-2.32**	-2.13**	-1.89*	-1.62	-1.30	-0.95	$z(\tilde{N}_{\text{runs}})$	0.47
		Monthly changes between end-of-month observations											
Nr. of observations	T						646					$\tilde{T}$	166
Obs. nr. of runs	N_{runs}						323					$\tilde{N}_{\text{runs}}$	86
Theoretical nr. of runs	$\mu(N_{\text{runs}})$	324	323	323	322	321	320	319	317	315	313	$\mu(\tilde{N}_{\text{runs}})$	83.5
Std. of runs	$\sigma(N_{\text{runs}})$	12.70	12.70	12.72	12.74	12.78	12.82	12.88	12.94	13.01	13.08	$\sigma(\tilde{N}_{\text{runs}})$	6.42
Runs test	$z(N_{\text{runs}})$	-0.04	-0.03	0.01	0.07	0.12	0.21	0.32	0.45	0.57	0.76	$z(\tilde{N}_{\text{runs}})$	0.39

Notes: This table reports estimates of the runs test. The null-hypothesis is that the signed returns of the DJIA are randomly arranged. The theoretical number of runs $\mu(N_{\text{runs}})$ has been calculated by means of (8). The standard deviation $\sigma(N_{\text{runs}})$ has been calculated by means of (9). The runs test has been calculated by means of (10) and the resulting z-statistic follows a standard normal distribution.

7 Conclusion

To this day, a lively debate exists whether stock-market returns follow a random walk, and can therefore not be predicted from their own past. This debate reflects major challenges in testing for randomness in stock-market prices, whose underlying stochastic process is neither self-evident, nor unchangeable across time. To shed new light on this question, this paper has revisited the runs test that focuses on the algebraic sign of stock-market returns to test for randomness. A crucial advantage of using these signed returns is that they can be interpreted as Bernoulli trials similar to coin tosses, whose statistical distribution is well known. However, in contrast to heads and tails, positive and non-positive stock market returns typically comprise an unobserved trend and are, therefore, not necessarily equiprobable outcomes. Against this background, this paper has derived the statistical distribution of the runs test under various probabilities of observing a positive return. Furthermore, to avoid questions as to whether this probability reflects a genuine trend or some form of non-randomness, this paper has turned to an algorithm developed by John von Neumann to convert coin tosses that might suffer from an unobserved bias into equiprobable outcomes. When adapting this algorithm to the example of the Dow Jones Industrial Average (DJIA) since 1970, the hypothesis that the corresponding returns are randomly arranged cannot be rejected.

8 References

Andersen, Torben G., Bollerslev, Tim., Christoffersen, Peter F., and Diebold, Francis X. (2006) "Volatility and Correlation Forecasting." In *Handbook of Economic Forecasting, Volume 1*, edited by Graham Elliot, Clive W.J. Granger, and Allan Timmermann, 777-878. Elsevier.

Bernstein, Peter L. (1998) *Against the Gods—The Remarkable Story of Risk*. John Wiley & Sons.

Bradley, James V. (1968) *Distribution-Free Statistical Tests*. Prentice Hall.

Campbell, John Y., Lo, Andrew W., and McKinley, A. Craig. (1997) *The Econometrics of Financial Markets*. Princeton University Press.

Cowles, Alfred. (1960) "A Revision of Previous Conclusion Regarding Stock Price Behavior.", *Econometrica*, 28, 909-615.

Cowles, Alfred, and Jones, Herbert E. (1937) "Some a Posteriori Probabilities in Stock Market Action." *Econometrica*, 5, 280-294.

Dryden, Myles M. (1970) "A statistical study of U.K. share prices." *Scottish Journal of Political Economy*, 17, 369-389.

Fama, Eugene F. (1965a) "The Behavior of Stock-Market Prices." *The Journal of Business*, 105, 38-34.

Fama, Eugene F. (1965b) "Random Walks in Stock Market Prices." *Financial Analysts Journal*, 5, 55-59.

Fama, Eugene F., and French, Kenneth R. (1988) "Permanent and Temporary Components of Stock Prices." *Journal of Political Economy*, 96, 246-273.

Gibbons, Jean D., and Chakraborti, Subhabrata. (2003) *Nonparametric Statistical Inference*. Marcel Dekker (4^{th} ed.).

Hoeffding, Wassily., and Simons, Gordon. (1970) "Unbiased Coin Tossing with Biased Coin." *The Annals of Mathematical Statistics*, 41, 341-352.

Jansen Dennis W., and De Vries, Casper G. (1991) "On the frequency of large stock returns: putting booms and busts into perspective." *Review of Economics and Statistics* 73,18-24.

Jennergren, L. Peter., and Korsvold, Paul E. (1974) "Price Formation in the Norwegian and Swedish Stock Markets: Some Random Walk Tests." *The Swedish Journal of Economics*, 76, 171-185.

Lo, Andrew W., and MacKinley, A. Craig. (1999) *A Non-Random Walk Down Wall Street*. Princeton University Press.

Malkiel, Burton G. (2011) *A Random Walk Down Wall Street*. W.W. Norton & Company.

Mood, Alexander M. (1940) "The distribution theory of runs." *The Annals of Mathematical Statistics*, 11, 367-392.

Mosteller, Fredrick. (1941) "Note on an application of runs to quality control charts." *Annals of Mathematical Statistics*, 12, 228-232.

Osborne, M.F.M. (1959) "Brownian Motion in the Stock Market", *Operations Research*, 7, 145-173.

Poterba, James M., and Summers, Lawrence. (1988) "Mean Reversion in Stock Returns: Evidence and Implications." *Journal of Financial Economics*, 22, 27-59.

Praetz, Peter D. (1969) "Australian share prices and the random walk hypothesis." *Australian Journal of Statistics*, 11, 123-139.

Rapach, David, and Zhou, Guofu. (2013) "Forecasting Stock Returns." In *Handbook of Economic Forecasting, Volume 2A*, edited by Graham Elliot, and Allan Timmermann, 328-383. Elsevier.

Samuelson, Paul. (1968) "Constructing An Unbiased Random Sequence." *Journal of the American Statistical Association*, 63, 1526-1527.

Siegel, Sidney., and Castellan, N. John. (1988) *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill (2^{nd} ed.).

Von Neumann, John. (1951) "Various Techniques Used in Connection with Random Digits." *Monte Carlo Method, Applied Mathematical Series*, 12, 36-38.

Wald, Abraham., and Wolfowitz, Jacob. (1940) "On a test whether two samples are from the same population." *The Annals of Mathematical Statistics*, 11, 147-162.

Wallis, W. Allen., and Moore, Geoffrey H. (1941) "A Significance Test for Time Series Analysis." *Journal of the American Statistical Association*, 36, 401-409.

Williamson, Samuel H. (2023) "Daily Closing Values of the DJA in the United States, 1885 to Present", MeasuringWorth.

A Deriving the mean and standard deviation of N_{runs}

This appendix provides a concise derivation of the mean, variance, and standard deviation of the number of runs, i.e. N_{runs}, as a function of the (constant) probability $\pi \in [0,1]$ of observing $I_t^r = 1$. The derivations draw on Gibbons and Chakraborti (2003, pp.82-83). See also Campbell *et al.* (1997, pp.39-41).

The mean number of runs, i.e. $\mu(N_{\text{runs}})$ of (9), involves the initial run at $t = 1$, and the expected number of sign-reversals R_t^r across the remaining observations at $t = 2, 3, \ldots, T$, that is

$$\mu(N_{\text{runs}}) = 1 + E\Big[\sum_{t=2}^{T} R_t^r\Big] = 1 + \sum_{t=2}^{T} E[R_t^r]. \tag{11}$$

Using (7), under statistical independence, $E[R_t^r]$ is given by

$$E[R_t^r] = 2\pi(1-\pi)\cdot 1 + (1 - 2\pi(1-\pi))\cdot 0 = 2\pi(1-\pi).$$

Across $t = 2, 3, \ldots, T$, sign-reversals R_t^r represent independent Bernoulli trials. Hence,

$$\sum_{t=2}^{T} E[R_t^r] = (T-1)2\pi(1-\pi). \tag{12}$$

Inserting (12) into (11) yields the mean[11] of (8), that is

$$\mu(N_{\text{runs}}) = 1 + (T-1)2\pi(1-\pi).$$

Under independence, the corresponding variance $\sigma^2(N_{\text{runs}})$ can also be broken into the (zero-valued) contribution of the initial run at $t = 1$, and the contributions of the overlapping Bernoulli random variables R_t^r across $t = 2, 3, \ldots, T$, that is

$$\begin{aligned}
\sigma^2(N_{\text{runs}}) &= \underbrace{Var[1]}_{=0} + Var\Big[\sum_{t=2}^{T} R_t^r\Big] = (T-1)Var[R_t^r] + \sum_{\substack{s=2\\ s\neq t}}^{T}\sum_{\substack{t=2\\ t\neq s}}^{T} Cov[R_s^r, R_t^r] \\
&= (T-1)E[(R_t^r)^2] - (T-1)^2(E[R_t^r])^2 + \sum_{\substack{s=2\\ s\neq t}}^{T}\sum_{\substack{t=2\\ t\neq s}}^{T} Cov[R_s^r, R_t^r].
\end{aligned} \tag{13}$$

Using (7), the first terms on the right-hand side of (13) are given by

$$(T-1)E[(R_t^r)^2] - (T-1)^2(E[R_t^r])^2 = (T-1)2\pi(1-\pi)\underbrace{(1^2)}_{=1} - (T-1)^2(2\pi(1-\pi))^2. \tag{14}$$

The covariance term $\sum\sum Cov[R_s^r, R_t^r]$ of (13) captures joint sign reversals between R_s^r and R_t^r. They can either occur within triplets, or between two (disjoint) pairs observations. These cases have different probabilities:

[11] Instead of contemplating the probabilities π and $1-\pi$, the mean of N_{runs} is often reported through the frequency of the outcomes (see e.g. Wald and Wolfowitz, 1940, p.151). When m refers to the number of positive stock-market returns counted across $T-1$ observations, probability and frequency are connected via $\pi = m/(T-1)$. In a similar vein, for n negative returns, we have $(1-\pi) = n/(T-1)$. Inserting these expressions into (12) yields the formula reported in e.g. Bradley (1968, p.262) or Siegel and Castellan (1988, p.62), that is

$$\mu(N_{\text{runs}}) = 1 + \frac{2mn}{T-1}.$$

i. Joint sign reversals within triplets can either occur with probability $\pi(1-\pi)\pi$, or with probability $(1-\pi)\pi(1-\pi)$. Hence, across the $2(T-2)$ possible triplets, the corresponding probability is given by

$$2(T-2)[\pi(1-\pi)\pi + (1-\pi)\pi(1-\pi)] = (T-2)2\pi(1-\pi).$$

ii. The disjoint pairs involve 4 combinations between two sign reversals, which all have either probability $\pi(1-\pi)$ or equivalently $(1-\pi)\pi$. Hence, across the remaining $(T-1)(T-2) - 2(T-1) = (T-2)(T-3)$ joint sign reversals, the disjoint pairs' probability is given by

$$4(T-2)(T-3)\pi^2(1-\pi)^2.$$

Taken together, we have that

$$\sum_{\substack{s=2\\ s\neq t}}^{T}\sum_{\substack{t=2\\ t\neq s}}^{T} Cov[R_s^r, R_t^r] = (T-2)2\pi(1-\pi) + 4(T-2)(T-3)\pi^2(1-\pi)^2. \tag{15}$$

Substituting (14) and (15) back into (13) yields

$$\sigma^2(N_{\text{runs}}) = (T-1)2\pi(1-\pi) - (T-1)^2(2\pi(1-\pi))^2 + (T-2)2\pi(1-\pi) + 4(T-2)(T-3)\pi^2(1-\pi)^2.$$

Rearranging and simplifying yields

$$\begin{aligned}\sigma^2(N_{\text{runs}}) &= 4(T-1)\pi(1-\pi)\Big[\frac{1}{2} + \frac{1}{2}\frac{T-2}{T-1} - \Big((T-1) - \frac{(T-2)(T-3)}{T-1}\Big)\pi(1-\pi)\Big] \\ &= 4(T-1)\pi(1-\pi)\Big[\frac{1}{2} + \frac{1}{2}\frac{T-2}{T-1} - 3\frac{T-2}{T-1}\pi(1-\pi)\Big].\end{aligned}$$

For large values of T, we have approximately that

$$\sigma^2(N_{\text{runs}}) \approx 4(T-1)\pi(1-\pi)[1 - 3\pi(1-\pi)].$$

Taking the square root yields $\sigma(N_{\text{runs}})$ of (9).