Random walks, stocks, volatility and diversification

report

Sönke Beier

April 4, 2024

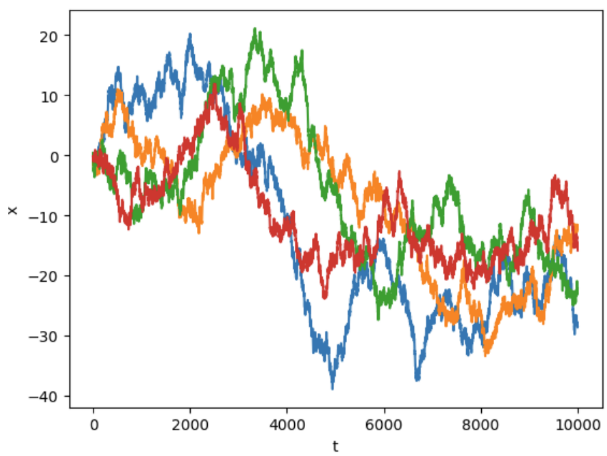
Contents

It is often simplistically assumed that the development of securities is like a random walk. In this project, we will learn what statistical differences we can recognize between both. The tasks of the report are adopted from [Set22].

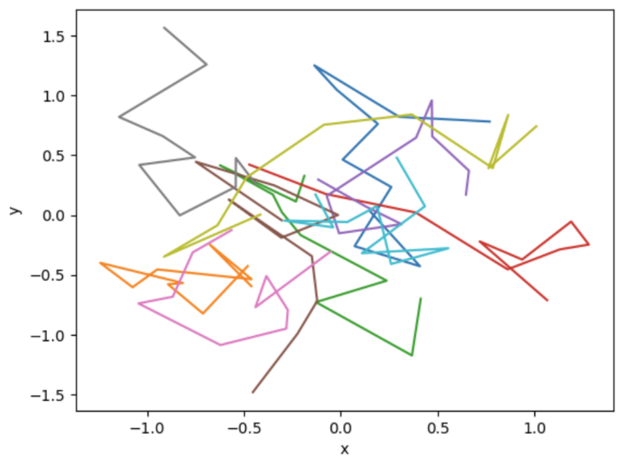
1 Computed Random walks

A random walk is a random process, that describes a path which consists of a series of random steps. In our case we have in 1d a stepsize from $-1/2$ to $1/2$. In more dimensions we will take this interval for each dimension. In figure 1 we see different examples of random walks, which also have a different number of steps N.

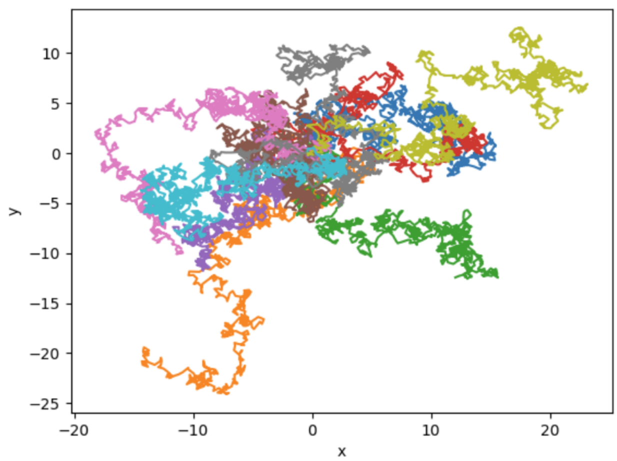
1.1 Trajectories of different random walks



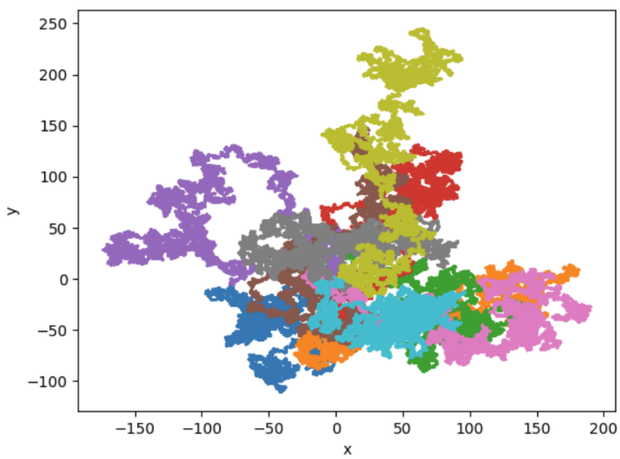
(a) Plot of 4 10,000-step one-dimensional random walks.



(b) Plot of 10 10-step two-dimensional random walks.



(c) Plot of 10 1000-step two-dimensional random walks.



(d) Plot of 10 100,000-step two-dimensional random walks.

Figure 1: Examples of random walks. Steps are uniformly distributed in the range $(-1/2, 1/2)$.

If we look at the net distance of random walks and multiply the number of steps by 100, we can see that the net distance increases by 10. This can be explained by the fact that the root mean square displacement

$$RMS := \sqrt{< distance^2 >} = \sqrt{N}$$

which is for explained at [Fow].

1.2 Endpoints of random walks of different lengths, the central limit theorem and the Gaussian distribution

In figure 2 we look at the endpoints of random walks with different timesteps. The square is created with 2d random walks with one step, as the x and y components can both take values between $-1/2$ and $1/2$. Diagonal steps are therefore longer. Since the central limit theorem applies to random walks, they have a Gaussian probability density for the number of steps $N \to \infty$. For many time steps, a rotational symmetry arises around the starting point.
The **gaussian probability density** is given by

$$\rho_G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{-x^2}{2\sigma^2} \tag{1}$$

With $\sigma = \sqrt{N}a$. a is here the root mean square step-size.

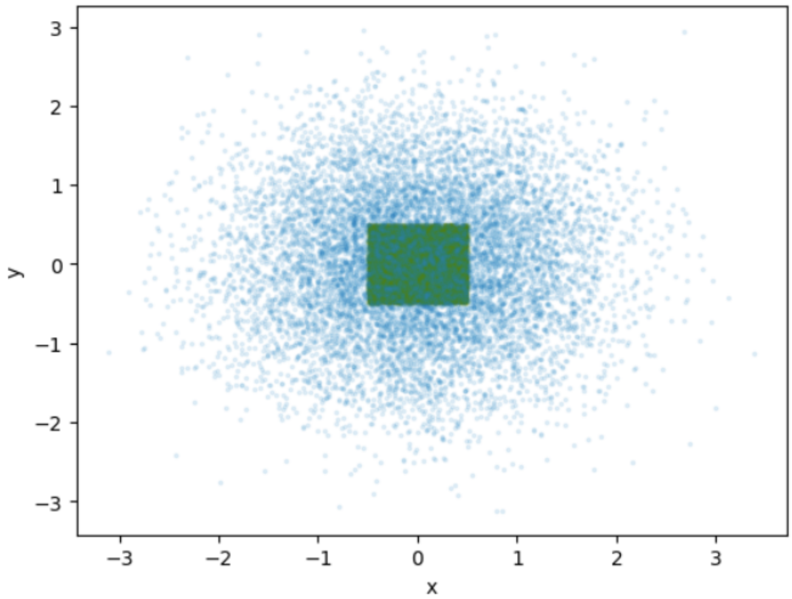
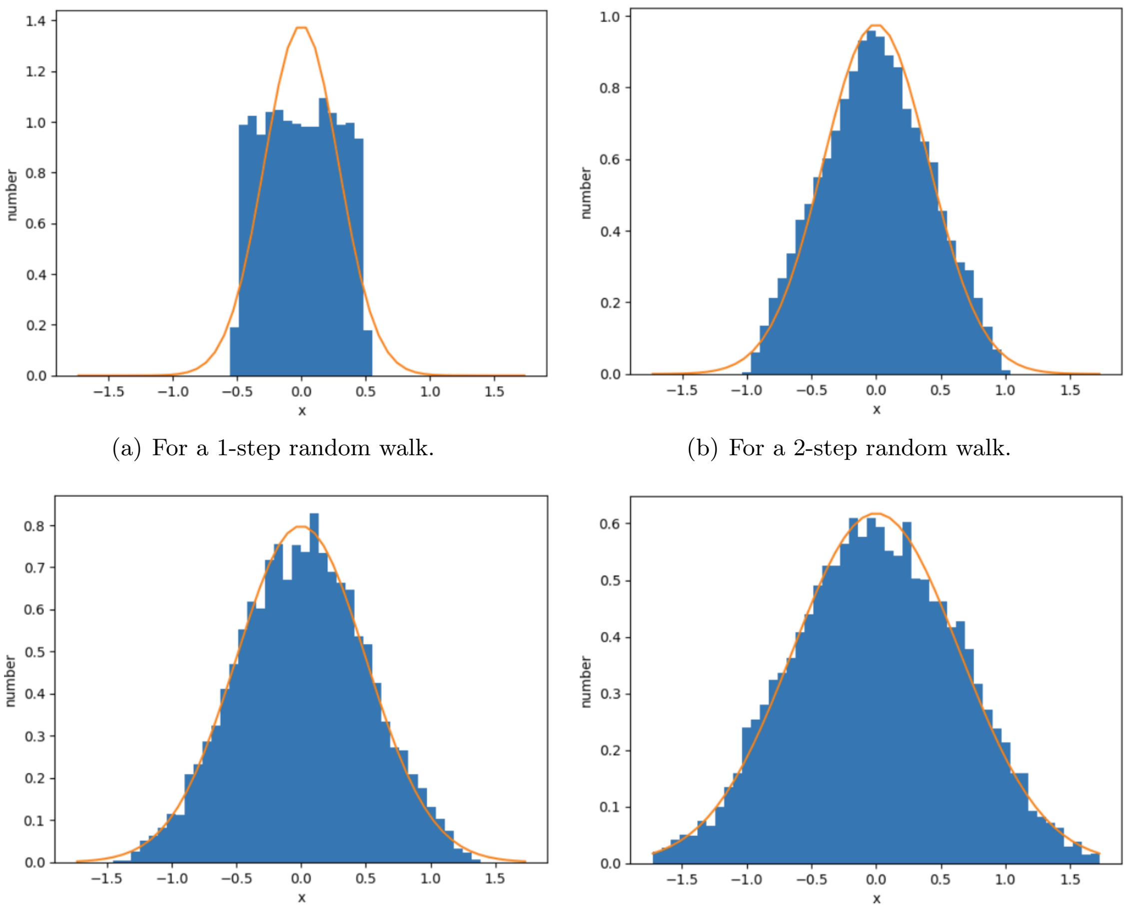


Figure 2: Scatter plot of the endpoints of 10,000 two-dimensional random walks with 1(green) and 10 (blue) steps. Steps are uniformly distributed in the range $(-1/2, 1/2)$

To see from which number of steps the Gaussian distribution is a good prediction for the endpoints, we created histograms in figure 3 for the x-values of the endpoints of random walks at different numbers of steps N. We can see that the Gaussian form is assumed from just a few timesteps. For one timesteps it shows the expected uniform distribution from $-1/2$ to $1/2$.



(a) For a 1-step random walk.

(b) For a 2-step random walk.

(c) For a 3-step random walk.

(d) For a 5-step random walk.

Figure 3: Histograms of 10,000 random walks each. Steps of the random walks are uniformly distributed in the range $(-1/2, 1/2)$. Additional the prediction for $N- > \infty$, a gaussian of the form of equation 1 was plottet.

For the prediction we had to calculate the **root mean square (RMS) step-size** a. In this case the RMS stepsize a for a random variable X with the probability distribution ρ with $\rho = c$ for $a < x < b$ and 0 otherwise can be calculated by:

$$\begin{aligned} < X^2 > &= \int_{-\infty}^{\infty} x^2 \rho dx = \int_a^b x^2 c dx \\ &= [1/3 c x^3]_a^b = 1/3 c b^3 - 1/3 c a^3 \\ &= \frac{b^3 - a^3}{3[b-a]} = 2/24 = 1/12 \\ \Rightarrow a &= \sqrt{< X^2 >} = 1/\sqrt{12} \end{aligned}$$

Here we used the property:

$$\begin{aligned} 1 &= \int \rho dx = \int_a^b c dx \\ &= [cx]_a^b = c[b-a] \\ \Rightarrow c &= \frac{1}{b-a} \end{aligned}$$

This result was also confirmed by a numerical calculation.

2 Stock prices

Now we want to compare the behavior of random walks with the behavior of financial data. It is said that the market trends are similar to those of a random walk. We will see what differences there are nevertheless.

2.1 Time evolution and external events

For this purpose, we use the data of the Standard and Poor's 500 index from 1982 onwards (see figure 4). It is a weighted average of the prices of 500 large companies in the US stock-market.

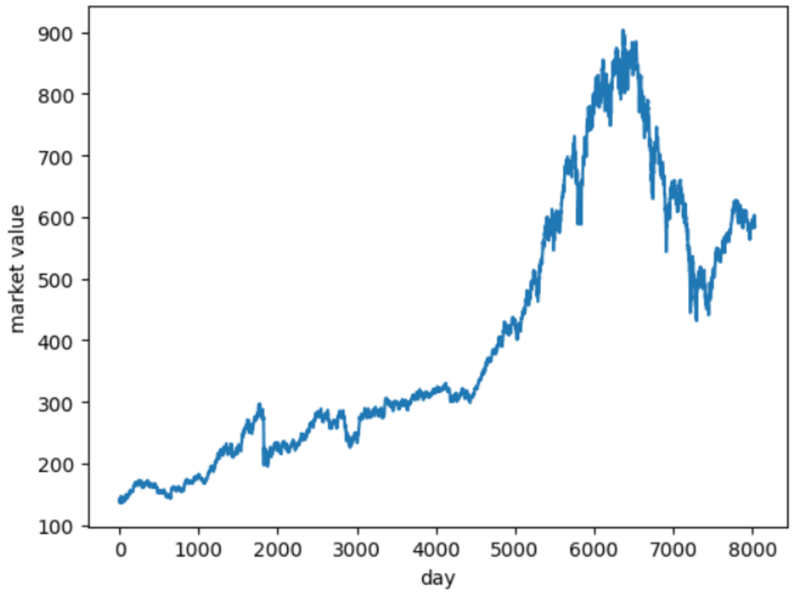
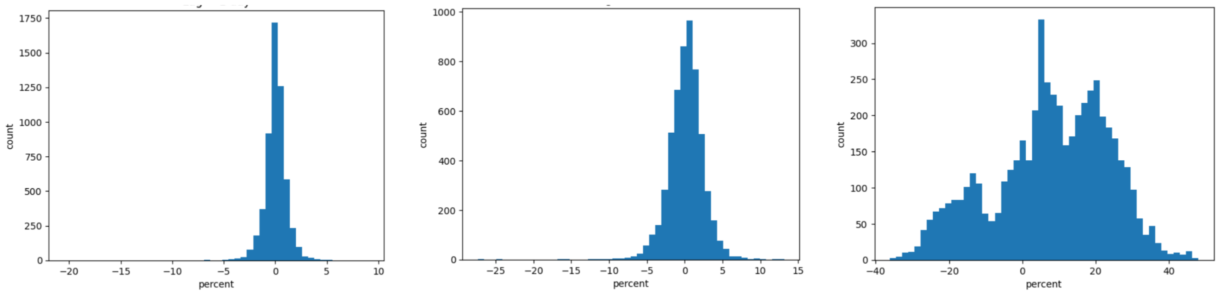


Figure 4: Plot of the Standard and Poor's 500 index (market value) versus time. The time starts in October 1982.

According to the given task, we look at the importance of external influences on the development of the market price. To do this, we look at the attack on the World Trade Center on September 11, 2001. It seems that the ongoing decrease after September 11 (around day 6903) was not triggered by the attack, as there was already a downward trend in the market before. Overall, the market is dominated by the dotcom bubble. Nevertheless, we can recognize various short events with strong changes over time.

2.2 Changes in different time scales

Now we have a closer look at the change in value at certain times. In figure 5 we can see that it would make sense to invest in the Standard and Poor index on an annual scale. For the other two time scales, the distribution is symmetrical around 0. The distribution on the annual scale is shifted towards positive returns. The graph for yearly changes looks different, because the development within a year also depends on larger external events. These do not follow the Central Limit Theorem, as there are not enough major events within a year. Larger events such as the dotcom bubble also influence several time steps in succession. This means that they are not really independent of each other (which is a prerequisite for the central limit theorem). This is more visible on larger time scales.As a result, the data here are not normally distributed. In addition, the stock value rises continuously. This increase is too small, to recognise it at the daily or weekly level. The smaller fluctuations dominate here.



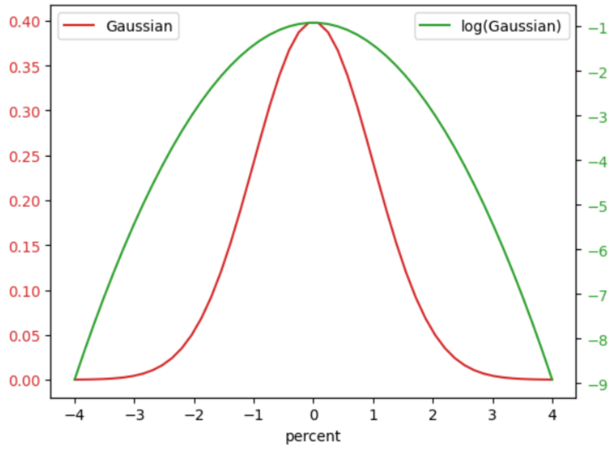
(a) Daily changes.

(b) Weekly changes.

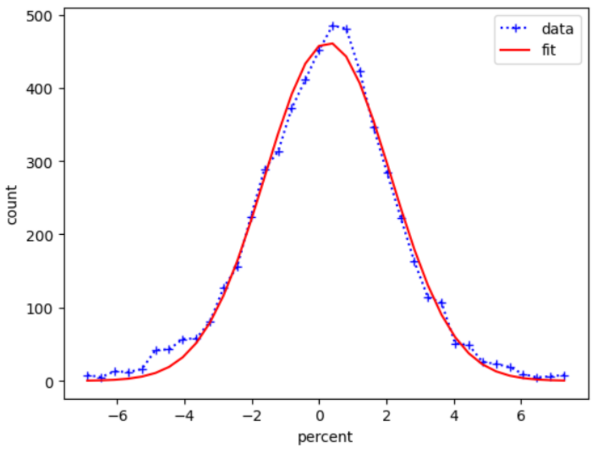
(c) Yearly changes.

Figure 5: Plot of per cent changes of the index over a specific time interval.

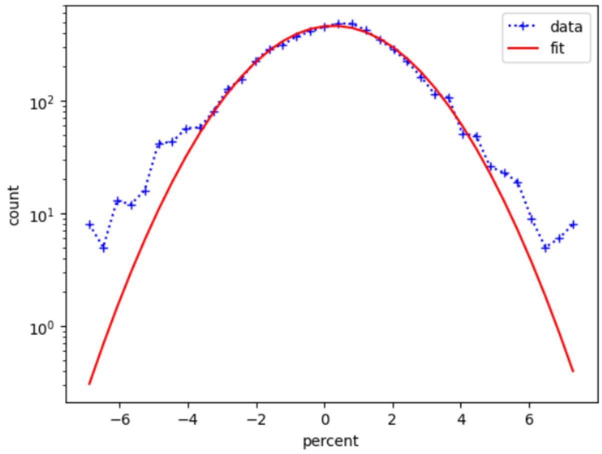
Now we compare the results once again with those that would be expected for a random walk. In figure 6 we have plotted the weekly changes against a Gaussian. For better visualization we use a log-lin plot here. In this plot, the Gaussian appears as an inverted parabola. If we look closer to (c) we can see that there are fat tails: there are more large percent changes then expected for a gaussian distribution.



(a) Gaussian and the logarithm of a gaussian, which looks like an inverted parabola.



(b) Lin-lin plot of the weekly percentage changes (blue) and the fittet gaussian (red).



(c) Log-lin plot of the weekly percentage changes (blue) and the fittet gaussian (red).

Figure 6: Comparison of the weekly percentage changes and a gaussian distribution.

2.3 Volatility of financial data

To assess the risk of securities, we can look at their volatility v_l, i.e. the standard deviation of the percentage return:

$$v_l = \sqrt{< (P_l(t) - \overline{P}_l)^2 >} \tag{2}$$

The volatility of the data is plotted in figure 2. A linear increase would be expected for a random walk. For our data, we see that the volatility increases more strongly for larger distances.

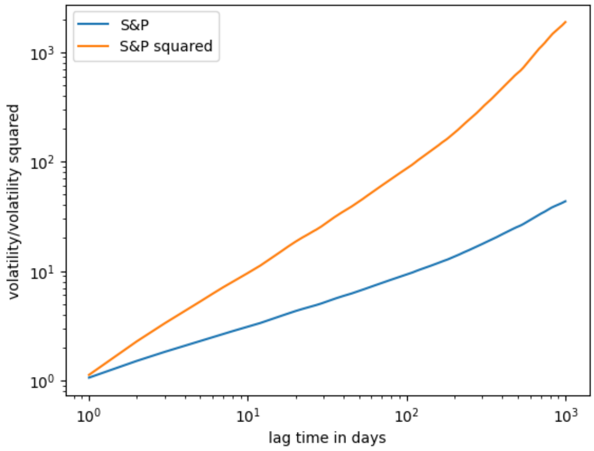


Figure 7: Volatility and volatility squared of a lag time from 1 to 100 days of the S&P data.

2.4 Strategy for a high return

For an index with 500 equally weighted uncorrelated stocks the return is

$$m_I = 1/500 \sum_{i=1}^{500} m_i = m$$

if the annual percentage returns of one stock $m_i = m$ for each individual stock i. The volatility of the index is

$$s_I = 1/500^2 \sum_{i=1}^{500} s_i = \frac{s}{500}$$

[1] if $s_i = s$ for each individual stock i.
So the mean return of the index and the individual stocks would be the same. The volatility is smaller for the index. So investing in the index is safer, because the variation of the trading price is lower.

3 Code

The code for the project can be found on Github at https://github.com/SoenBeier/random_walks_and_stocks

References

[Fow] Michael Fowler. *The One-Dimensional Random Walk*. https://galileo.phys.virginia.edu/classes/152.mf1i.spring02/RandomWalk.htm [Accessed: 02.01.2024].

[Set22] James Sethna. *Entropy, Order Parameters, and Complexity*. Clarendon Press, 2022. URL: https://sethna.lassp.cornell.edu/StatMech/.

[1]Follows from 9.1.8, 9.2.18 of https://www.uni-ulm.de/fileadmin/website_uni_ulm/mawi.inst.110/lehre/ws12/WR/Skript_9.pdf