Stochastic Methods Lab, Jacobs University Bremen, Fall 2022 — Session 24, Nov. 30, 2022

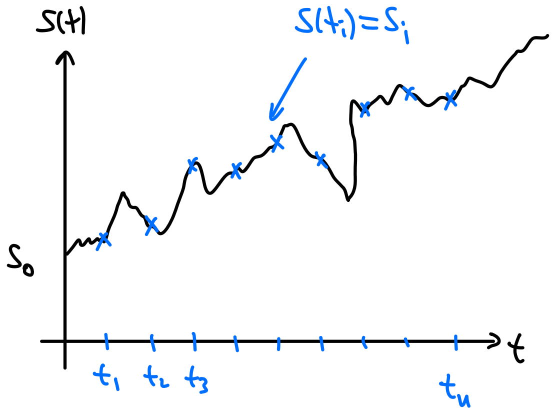
Prof. Dr. Soeren Petrat

---

## 5. Parameter Estimates for Time Series

Our stock price model is geometric Brownian motion (GBM):

$$dS = \mu S\,dt + \sigma S\,dW \quad \text{, with solution} \quad S(t) = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W(t)}$$




Time Series: we sample $S(t)$ at times $t_1, \dots, t_n$, which gives us $S(t_i) = S_i$.

Now, let us consider the yields or log-returns, i.e., the $r_i$ s.t. $S(t_i) = S(t_{i-1}) e^{r_i}$

$$\Rightarrow r_i := \ln \frac{S_i}{S_{i-1}} = \ln S_i - \ln S_{i-1}$$

For GBM, we find $r_i = \ln S_0 e^{(\mu - \frac{\sigma^2}{2})t_i + \sigma W(t_i)} - \ln S_0 e^{(\mu - \frac{\sigma^2}{2})t_{i-1} + \sigma W(t_{i-1})}$

$$= \ln S_0 + (\mu - \tfrac{\sigma^2}{2})t_i + \sigma W(t_i) - \left[\ln S_0 + (\mu - \tfrac{\sigma^2}{2})t_{i-1} + \sigma W(t_{i-1})\right]$$

$$= (\mu - \tfrac{\sigma^2}{2})\underbrace{(t_i - t_{i-1})}_{\Delta t_i} + \sigma\big(\underbrace{W(t_i) - W(t_{i-1})}_{\Delta W_i}\big)$$

$\Rightarrow$ According to our model/assumption, the $r_i$'s are $\underbrace{\text{normally}}$ and independently distributed

(because this is how we defined $\Delta W_i$)

but with expectation not necessarily 0 and variance not necessarily 1

Let us choose $\Delta t_i = \Delta t$. Then our theoretical prediction from our model is:

- expectation: $\mathbb{E}(r_i) = \mathbb{E}\left( (\mu - \frac{\sigma^2}{2})\Delta t + \sigma \Delta W_i \right) = (\mu - \frac{\sigma^2}{2})\Delta t + \sigma \underbrace{\mathbb{E}(\Delta W_i)}_{=0}$

$$= (\mu - \frac{\sigma^2}{2})\Delta t$$

- variance: $Var(r_i) = Var\left( (\mu - \frac{\sigma^2}{2})\Delta t + \sigma \Delta W_i \right) = \underbrace{Var((\mu - \frac{\sigma^2}{2})\Delta t)}_{=0} + Var(\sigma \Delta W_i)$

$$= \sigma^2 \underbrace{Var(\Delta W_i)}_{\Delta t} = \sigma^2 \Delta t \qquad (std = \sigma\sqrt{\Delta t})$$

$$\Rightarrow \quad r_i \sim \mathcal{N}\left( (\mu - \frac{\sigma^2}{2})\Delta t, \sigma\sqrt{\Delta t} \right)$$

normal  expectation  standard deviation

Now we want to match the theoretical prediction to our data. We get

- sample mean/average: $\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i$
- sample variance: $\sigma_r^2 = \frac{1}{\underbrace{(n-1)}}\sum_{i=1}^{n} (\bar{r} - r_i)^2$

These are still random variables, bc. of the many different ways one can sample a given stochastic process.

↳ "unbiased sample variance", which comes out when we average over all possible ways of sampling $S(t)$.

(But $\frac{1}{n-1} - \frac{1}{n} \sim O(\frac{1}{n^2})$, so difference to regular variance of the one given sample is small.)

For large $n$, we should be able to replace $\mathbb{E}(r_i) \approx \bar{r}$ and $\mathrm{Var}(r_i) \approx \sigma_r^2$, i.e.,

we approximate

- $\sigma = \sqrt{\dfrac{\mathrm{Var}(r_i)}{\Delta t}}$ by $\boxed{\hat{\sigma} = \sqrt{\dfrac{\sigma_r^2}{\Delta t}} = \dfrac{\sigma_r}{\sqrt{\Delta t}}}$
- $\mu = \dfrac{\mathbb{E}(r_i)}{\Delta t} + \dfrac{\sigma^2}{2}$ by $\boxed{\hat{\mu} = \dfrac{\bar{r}}{\Delta t} + \dfrac{\hat{\sigma}^2}{2}}$

These are the parameters for GBM that we read off from our data.

Comment: • as written above, $\bar{r}$ and $\sigma_r$ are still random variables

↳ one can show (we will see this numerically in the HW) that $\mathrm{Var}(\hat{\sigma}) = \dfrac{\mathbb{E}(\hat{\sigma})^2}{2n}$, so $\hat{\sigma}$ as a random var. is $\sim$ const for $n \to \infty$, and thus can reliably be estimated (by simply choosing $n$ large enough)

↳ on the other hand, $\mathrm{Var}(\hat{\mu})$ does not in general become smaller for large $n$, so it cannot be reliably estimated.

↳ Luckily, we do not need the parameter $\hat{\mu}$ for the option pricing!

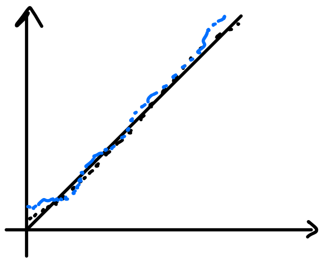
But we assumed that the $r_i$'s are normally and independently distributed, so this should be checked for our data!

<u>Test assumption of normality:</u> QQ plot (see HW4 Problem 5)

recall: • rescale $\tilde{r}_i = \dfrac{r_i - \bar{r}}{\sigma_r}$

• sort $\tilde{r}_i$

· then plot against sorted sample of $\underbrace{\text{standard}}_{\mathbb{E}=0,\ Var=1}$ normal distribution



here, a visual inspection is more helpful than other indicators

## Test assumption of independence:

We can look for correlations by looking at the covariance

$$Cov(X,Y) = \mathbb{E}\big((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\big)$$

$$= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

If $X$ and $Y$ are independent, then $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ and $Cov(X,Y) = 0$.

In the context of time series, we want to consider the **autocorrelation fct. (ACF)**

$$\boxed{K_\ell = \frac{Cov(\vec{r}_i, \vec{r}_{i-\ell})}{\underbrace{\sqrt{Var(r_i)}\sqrt{Var(r_{i-\ell})}}} \qquad \text{with } \ell \text{ called the "lag"}}$$

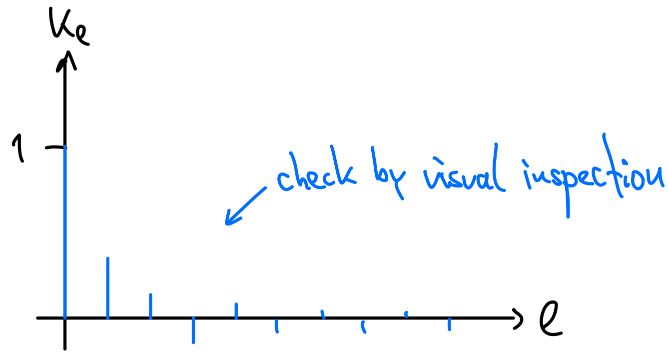
Note: Here, we look at the Cov and Var of $\vec{r} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$. (Equation is indep. of $i$.)
(E.g. $\mathbb{E}(r_i r_{i-\ell}) = \frac{1}{n-\ell} \sum_{i=\ell+1}^{n} r_i r_{i-\ell}$.)

Recall: $Cov(X,X) = Var(X)$, so $K_\ell$ is normalized in the sense that $K_0 = 1$ and
(since by Cauchy-Schwarz $\int fg \le \sqrt{\int |f|^2}\sqrt{\int g^2}$, i.e., $Cov(X,Y) \le \sqrt{Var(X)Var(Y)}$), $|K_\ell| \le 1$.

So: • $K_\ell = 1$ means perfect correlation

• $K_\ell = -1$ means perfect anticorrelation

• $|K_\ell| \approx 0$, the $r_i$'s are very uncorrelated




Note: For stock data, there might be "inertia effects", i.e., some autocorrelation if $\Delta t$ was chosen too small.

$\Rightarrow$ tradeoff: choose $\Delta t$ small enough so $\mathrm{Var}(\sigma_r)$ small enough, but not too small that the assumption of independence is violated.