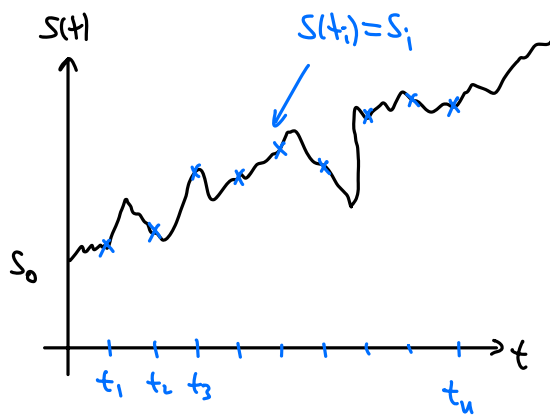


5. Parameter Estimates for Time Series

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

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



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. $S(t_i) = S(t_{i-1}) e^{r_i}$

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

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

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

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

$\Rightarrow$ According to our model/assumption, the r_i 's are normally and independently distributed
 (because this is how we defined ΔW_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: $\mathbb{E}(r_i) = \mathbb{E}\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma \Delta W_i\right) = \left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \underbrace{\sigma \mathbb{E}(\Delta W_i)}_{=0}$
 $= \left(\mu - \frac{\sigma^2}{2}\right)\Delta t$

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

$\Rightarrow r_i \sim \mathcal{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t, \sigma \sqrt{\Delta t}\right)$
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: $\sigma_r^2 = \frac{1}{(n-1)} \sum_{i=1}^n (\bar{r} - r_i)^2$

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 $\Sigma(t)$.

(But $\frac{1}{n-1} - \frac{1}{n} \sim \mathcal{O}(\frac{1}{n^2})$, 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 $\text{Var}(r_i) \approx \sigma_r^2$, i.e.,

we approximate

$$\begin{aligned} \bullet \sigma &= \sqrt{\frac{\text{Var}(r_i)}{\Delta t}} \quad \text{by} \quad \hat{\sigma} = \sqrt{\frac{\sigma_r^2}{\Delta t}} = \frac{\sigma_r}{\sqrt{\Delta t}} \\ \bullet \mu &= \frac{\mathbb{E}(r_i)}{\Delta t} + \frac{\sigma^2}{2} \quad \text{by} \quad \hat{\mu} = \frac{\bar{r}}{\Delta t} + \frac{\hat{\sigma}^2}{2} \end{aligned} \quad \left. \vphantom{\begin{aligned} \bullet \sigma &= \sqrt{\frac{\text{Var}(r_i)}{\Delta t}} \\ \bullet \mu &= \frac{\mathbb{E}(r_i)}{\Delta t} + \frac{\sigma^2}{2} \end{aligned}} \right\} \text{These are the parameters for GBM that we read off from our data.}$$

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

↳ one can show (we will see this numerically in the HW) that $\text{Var}(\hat{\sigma}) = \frac{\mathbb{E}(\hat{\sigma})^2}{2n}$, so

$\hat{\sigma}$ as a random var. is $\sim \text{const}$ for $n \rightarrow \infty$, and thus can reliably be estimated (by simply choosing n large enough)

↳ on the other hand, $\text{Var}(\hat{\mu})$ 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!

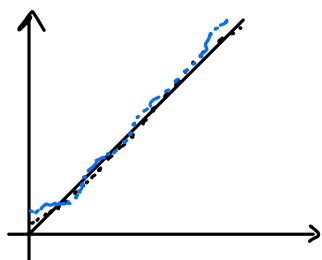
But we assumed that the r_i 's are normally and independently distributed, so this should be checked for our data!

Test assumption of normality: QQ plot (see HW 4 Problem 5)

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

• sort $\tilde{r}_i$

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



here, a visual inspection is more helpful than other indicators

Test assumption of independence:

We can look for correlations by looking at the covariance

$$\begin{aligned}\text{Cov}(X, Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\end{aligned}$$

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

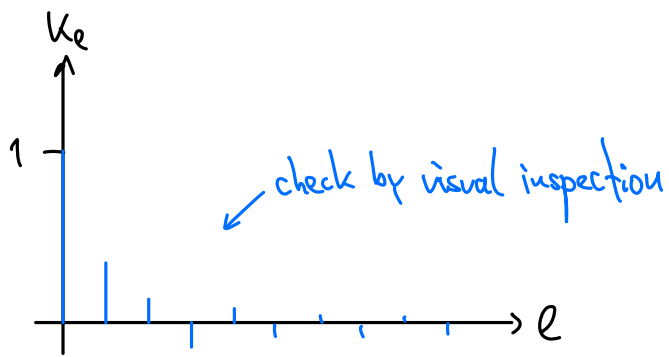
In the context of time series, we want to consider the autocorrelation fct. (ACF)

$$K_e = \frac{\text{Cov}(\vec{r}_i, \vec{r}_{i-e})}{\sqrt{\text{Var}(r_i)} \sqrt{\text{Var}(r_{i-e})}} \quad \text{with } e \text{ called the "lag"}$$

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., $\mathbb{E}(r_i r_{i-e}) = \frac{1}{n-e} \sum_{i=e+1}^n r_i r_{i-e}$.)

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

- So:
- $K_e = 1$ means perfect correlation
 - $K_e = -1$ means perfect anticorrelation
 - $|K_e| \approx 0$, the r_i 's are very uncorrelated



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

$\Rightarrow$ tradeoff: choose Δt small enough so $\text{Var}(G_r)$ small enough, but not too small that the assumption of independence is violated.