Financial Risk

Lecture 2

Returns

- We may be interested in what we may lose from time t_1 to time t_2
- Letting p_{t_1} and p_{t_2} denote the value of the portfolio at time t_1 and t_2, respectively, we define the *absolute return*

$$\bar{r}_{t_2} = p_{t_2} - p_{t_1}$$

- The *relative return*

$$\tilde{r}_{t_2} = \frac{p_{t_2} - p_{t_1}}{p_{t_1}}$$

- And the *log-return*

$$r_{t_2} = \ln\left(\frac{p_{t_2}}{p_{t_1}}\right)$$

Loss variable

- So if we are interested in what we may lose in dollars, euros or SEK we define the loss variable L as the negative absolute return

- If we are interested in what we may lose in terms of a percentage of the portfolio value we define the loss variable L as the negative relative return

- Note that (Taylor expansion) log-returns and relative returns will be close for small relative returns

Value at Risk (VaR)

- One of the most common notions in financial risk management is that of Value at Risk (VaR)

- VaR may be used to determine the amount of regulatory capital to set aside for different types of risks

- Letting L be loss variable for our portfolio we may define VaR as

$$VaR_p = inf\{x: P(L > x) \leq 1 - p\}$$

- So VaR gives us the smallest amount (or percentage) we may lose with a certain probability

Expected Shortfall (ES)

- Given that the VaR is exceeded, one may wonder how bad this can be

- We may quantify this in terms of the *expected shortfall*

$$ES_p = E\left[L \mid L > VaR_p\right]$$

Extremal events

- Typically we are interested in hedging big losses

- Therefore, we are interested in models of extremal events, i.e., models of big losses

- It turns out that ”standard” distributions, such as the normal distribution, are not sufficient

- Furthermore, the distribution of L is typically unknown…

Extreme Value Theory

- We know that we, eventhough the distribution of a (i.i.d.) sample may be unknown, may approximate the distribution of the sample mean by the normal distribution if the sample size is sufficiently large

- Is there a ”similar” scheme for the maximum or minimum of a (i.i.d.) sample of, say, log-returns $(r_1, \ldots, r_n)$?

$$M_n = max\{r_1, \ldots, r_n\} \text{ or } m_n = min\{r_1, \ldots, r_n\}$$

Extreme Values

- If we assume that the log returns are serially independent and have distribution function

$$F(x) = P(r_t \leq x)$$

it holds that

$$P(M_n \leq x) = P(r_1 \leq x, \dots, r_n \leq x) = \prod_{i=1}^{n} P(r_i \leq x)$$

$$= [F(x)]^n$$

Degeneration

- But what happens if we let the number of observations increase, i.e. let $n \to \infty$?

- Then $[F(x)]^n \to 0$ or $[F(x)]^n \to 1$ depending on if $x < u$ or $x \geq u$ where u is the upper end point of r_t (typically $u = \infty$ for log returns)

- So we need something more to get a non-trivial limit…

Appropriate sequences

- We need sequences $\{\alpha_n\}, \{\beta_n\}$ such that the distribution of

$$M_{n*} = \frac{M_n - \beta_n}{\alpha_n}$$

converges to a non-trivial limit

- We sometimes refer to $\{\alpha_n\}$ and $\{\beta_n\}$ and the scaling and location sequences, respectively

Limiting distributions

- It turns out that if the limit exists its distribution function will be (Generalized extreme value distribution, GEV)

$$F_*(x) = exp\left[-(1+\xi x)^{-1/\xi}\right]$$

for $x < -1/\xi$ if $\xi < 0$ and for $x > -1/\xi$ if $\xi > 0$

- The special case $\xi = 0$ gives

$$F_*(x) = exp[-exp(-x)]$$

for $-\infty < x < \infty$

Three types

- Type I, $\xi = 0$, the Gumbel distribution

$$F_*(x) = exp[-exp(-x)], -\infty < x < \infty$$

- Type II, $\xi > 0$, the Fréchet distribution

$$F_*(x) = \begin{cases} exp\left[-(1+\xi x)^{-1/\xi}\right], x > -1/\xi \\ 0 \qquad\qquad\qquad\qquad\quad , otherwise \end{cases}$$

- Type III, $\xi < 0$, the Weibull distribution

$$F_*(x) = \begin{cases} exp\left[-(1+\xi x)^{-1/\xi}\right], x < -1/\xi \\ 1 \qquad\qquad\qquad\qquad\quad , otherwise \end{cases}$$

In practice

- In real world problems we typically do not have to worry about what

$$F(x) = P(r_t \leq x)$$

looks like in order to fit extreme value distributions

- What we do have to worry about however is that returns typically not are independent or stationary

- In lecture 1 we proposed a way of imposing stationarity on the data at hand and a scheme for estimating and predicting VaR dynamically

Estimation for GEV

- For a given sample there is just one maximum or minimum
- Of course we cannot estimate parameters using just one observation
- One way of circumventing this problem is to divide the sample into non-overlapping blocks and then use the maximum from each block to estimate parameters

$$\{r_1, \dots, r_n | r_{n+1}, \dots, r_{2n} | \cdots | r_{(k-1)n+1}, \dots, r_{kn}\}$$

- It should be noted that r_t should be replaced by z_t when using the dynamic scheme proposed in lecture 1.

Estimation

- The method described above is referred to as the ”Block maxima method”.
- For sufficiently large blocks the block maxima should follow the GEV distribution
- The block maxima may be considered a sample from the GEV distribution

Estimation

- If we denote the block maxima $\{M_{1,n}, \dots, M_{k,n}\}$, the pdf needed for the ML estimation is (exercise)

$$f(M_{i,n}) = \frac{1}{\alpha_n}\left[1 + \frac{\xi_n(M_{i,n} - \beta_n)}{\alpha_n}\right]^{-\left(\frac{1}{\xi_n}+1\right)} exp\left[-\left(1 + \frac{\xi_n(M_{i,n} - \beta_n)}{\alpha_n}\right)^{-1/\xi_n}\right]$$

if $\xi_n \neq 0$ where it has to hold that $1 + \frac{\xi_n(M_{i,n}-\beta_n)}{\alpha_n} > 0$, and

$$f(M_{i,n}) = \frac{1}{\alpha_n} exp\left[-\frac{M_{i,n} - \beta_n}{\alpha_n} - exp\left(-\frac{M_{i,n} - \beta_n}{\alpha_n}\right)\right]$$

if $\xi_n = 0$

Estimation

- The likelihood function is then

$$L(\alpha_n, \beta_n, \xi_n | M_{1,n}, \dots, M_{k,n}) = \prod_{i=1}^{k} f(M_{i,n})$$

- The estimates will be unbiased and asymptotically normal
- Estimations may also be made using regression or non-parametric techniques, see Tsay 3rd ed p.347-348
- On may also use ”gevfit” in matlab or some EV-package in R

Example

- Fitting the Gumbel distribution to the Decile1,2,9,10 data (from Tsay) using 21-day blocks and gevfit in Matlab gives

Data	Scale α_n	Location β_n
Decile1	0.0807	0.1291
Decile2	0.0633	0.0983
Decile9	0.0265	0.0787
Decile10	0.0286	0.0728

Checking model fit

- One can define residuals as (hats for parameter estimates are left out in what follows)

$$e_i = \left(1 + \frac{\xi_n\left(M_{i,n} - \beta_n\right)}{\alpha_n}\right)^{-1/\xi_n}$$

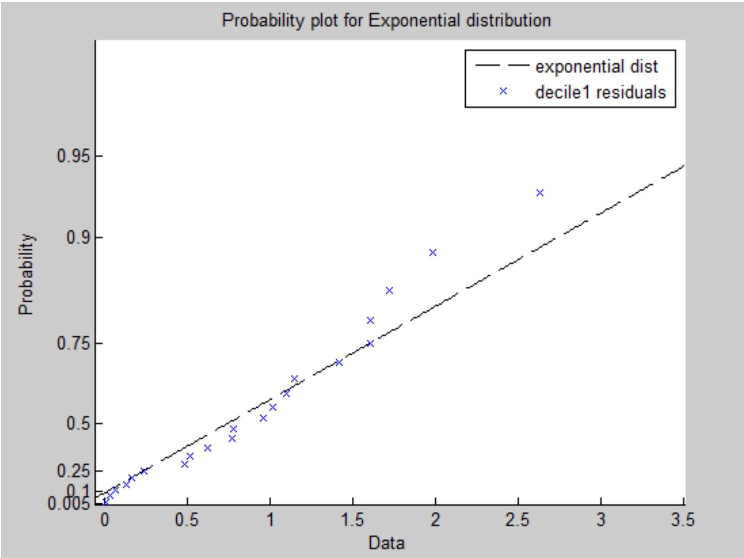
if $\xi_n \neq 0$ and

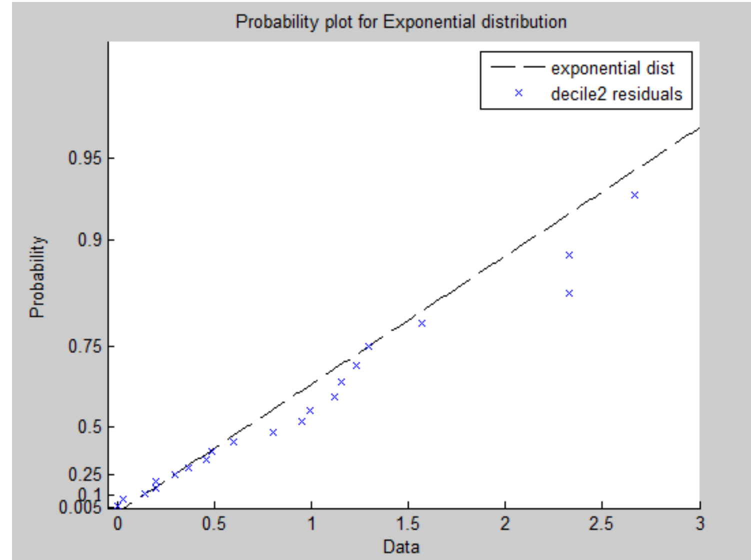
$$e_i = exp\left(-\frac{M_{i,n} - \beta_n}{\alpha_n}\right)$$

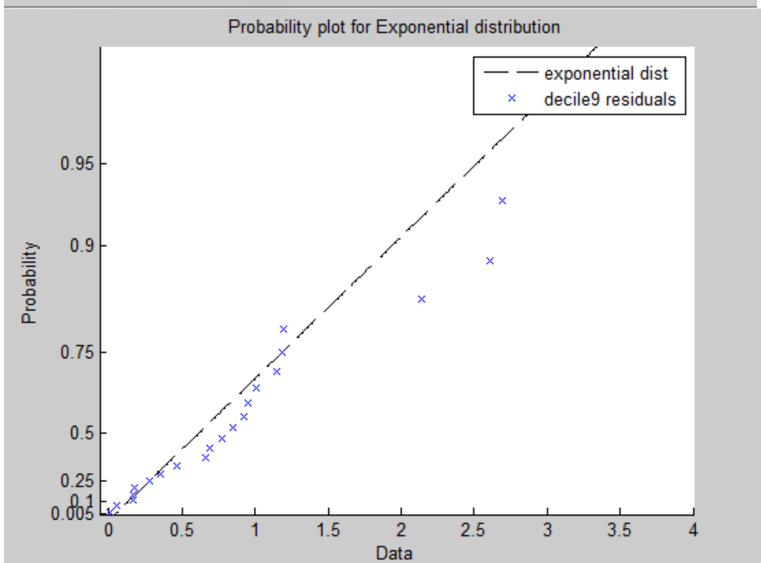
if $\xi_n = 0$

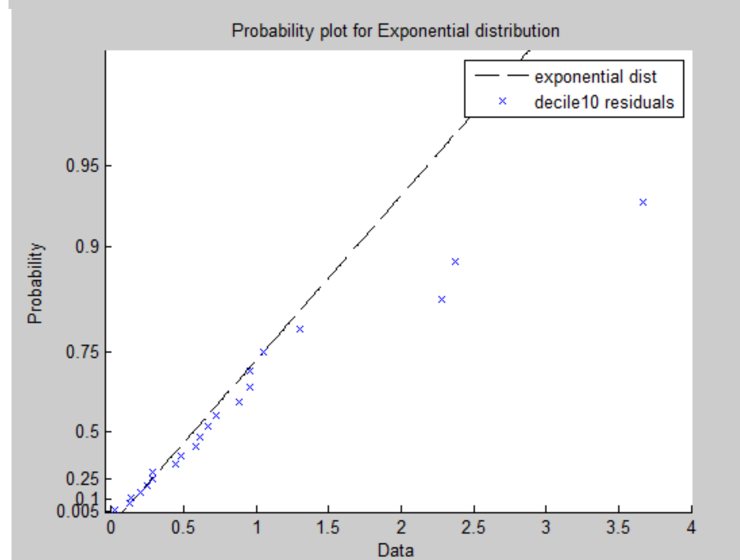
- Residuals should follow an exponential distribution if the model is correctly specified

PP-plots (Matlab ”probplot” ML decile)









Alternative Simple Method, Gumbel

- If a random variable has the Gumbel distribution function

$$exp\{-exp\{(x-\beta_n)/\alpha_n\}\}$$

it can be shown that the expected value is $\beta_n + \alpha_n\gamma$, where $\gamma \approx 0.5772$ (Euler's constant), and that the variance is $(\pi\alpha_n)^2/6$

- This means that we may estimate α_n and β_n using

$$\hat{\alpha}_n = \frac{\sqrt{6}}{\pi} S_{M_n} \quad \text{and} \quad \hat{\beta}_n = \bar{M}_n - \gamma\hat{\alpha}_n$$

where $\bar{M}_n$ and S_{M_n} are the mean and standard deviation of our observed block maxima.

Creating QQ-plots

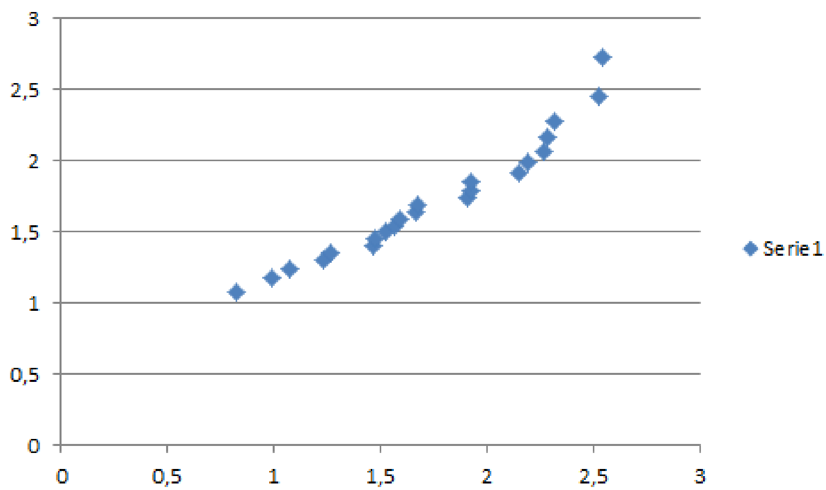
- If we want QQ-plots to check the fit of a GEV we may use our ordered (increasing values) observations $M_{(i),n}$ for $i = 1, \dots, m$ and the quantiles

$$M^*{}_{i,n} = \begin{cases} \beta_n - \dfrac{\alpha_n}{\xi_n}\left\{1 - \left[-\ln\left(\dfrac{i}{m+1}\right)\right]^{-\xi_n}\right\} \\ \beta_n - \alpha_n \ln\left[-\ln\left(\dfrac{i}{m+1}\right)\right] \end{cases}$$

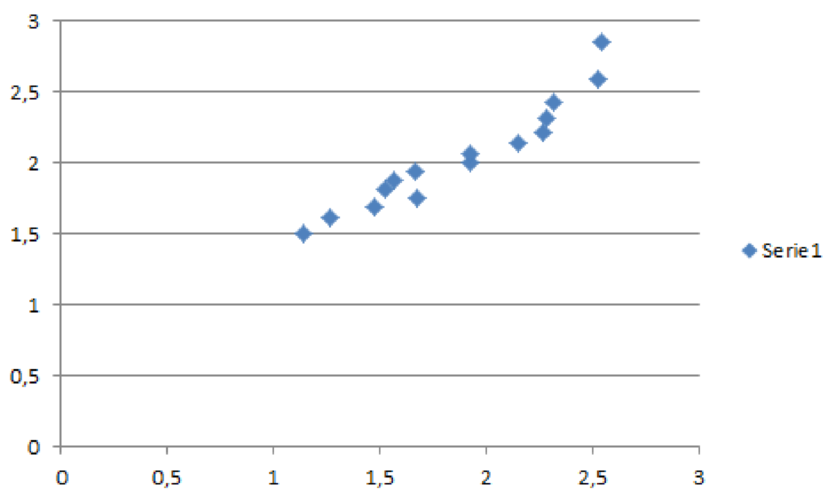
- The pairs $\left(M^*{}_{i,n}, M_{(i),n}\right)$ should form a straight line

QQ-plots (Excel) for Gumbel fit to standardized negated FB log-returns

- 21-day



- 30-day



Using Block-Maxima for VaR

- In VaR we are interested in quantiles

- Using GEV distributions and assuming that we have negated returns so that a high return is a big loss we let q be the (small) probability of a great loss and write

$$1 - q = \begin{cases} exp\left\{-\left[1 + \frac{\xi_n (M^*{}_n - \beta_n)}{\alpha_n}\right]^{-1/\xi_n}\right\} \\ exp\left\{-exp\left[-\frac{M^*{}_n - \beta_n}{\alpha_n}\right]\right\} \end{cases}$$

Using Block-Maxima for VaR

- Solving for the quantile ${M^*}_n$ we get

$$
{M^*}_n = \begin{cases} \beta_n - \frac{\alpha_n}{\xi_n}\left\{1 - [-\ln(1-q)]^{-\xi_n}\right\} \\ \beta_n - \alpha_n \ln[-\ln(1-q)] \end{cases}
$$

- Now this is the quantile for the number (n) of observations in each the block so we have to transform it to use it for one-day VaR

Using Block-Maxima for VaR

- Under the assumption of independent returns we may use that

$$1 - q = P\left(M_{i,n} \leq {M^*}_n\right) = [P(r_t \leq {M^*}_n)]^n$$

- So if we want $P(r_t \leq {M^*}_n) = 1 - q$ we get

$$VaR_{1-q} = \begin{cases} \beta_n - \dfrac{\alpha_n}{\xi_n}\left\{1 - [-n\ln(1-q)]^{-\xi_n}\right\} \\ \beta_n - \alpha_n \ln[-n\ln(1-q)] \end{cases}$$

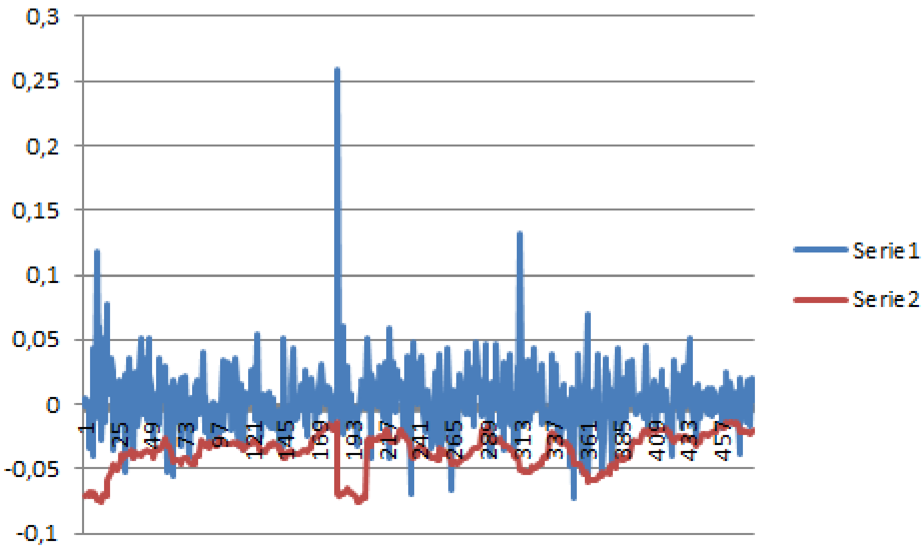
- It should be noted that r_t should be replaced by z_t using the dynamic scheme proposed in lecture 1.

Using Block-Maxima for VaR

- Using the decile data and the Gumbel model for block maxima of negated returns (without standardization which yields a constant VaR) fit with ML using Matlab (gevfit) we get

Data	95% VaR	99% VaR
Decile1	0.1231	0.2546
Decile2	0.0936	0.1968
Decile9	0.0767	0.1199
Decile10	0.0707	0.1173

Gumbel VaR for FB in Excel



Using the simple scheme and Gumbel for the 21-day block maxima also underestimates VaR as there are approximately 6% exceedances (not seen from the graph but easy to get in Excel). As we shall see in lecture 4, this underestimation can be taken care of by taking into account dependencies in the data.

Return Level

- We may be interested in what levels losses are expected to exceed within in a certain time frame

- We refer to this as the return level $L_{n,k}$ where k denotes the number of periods of length n and

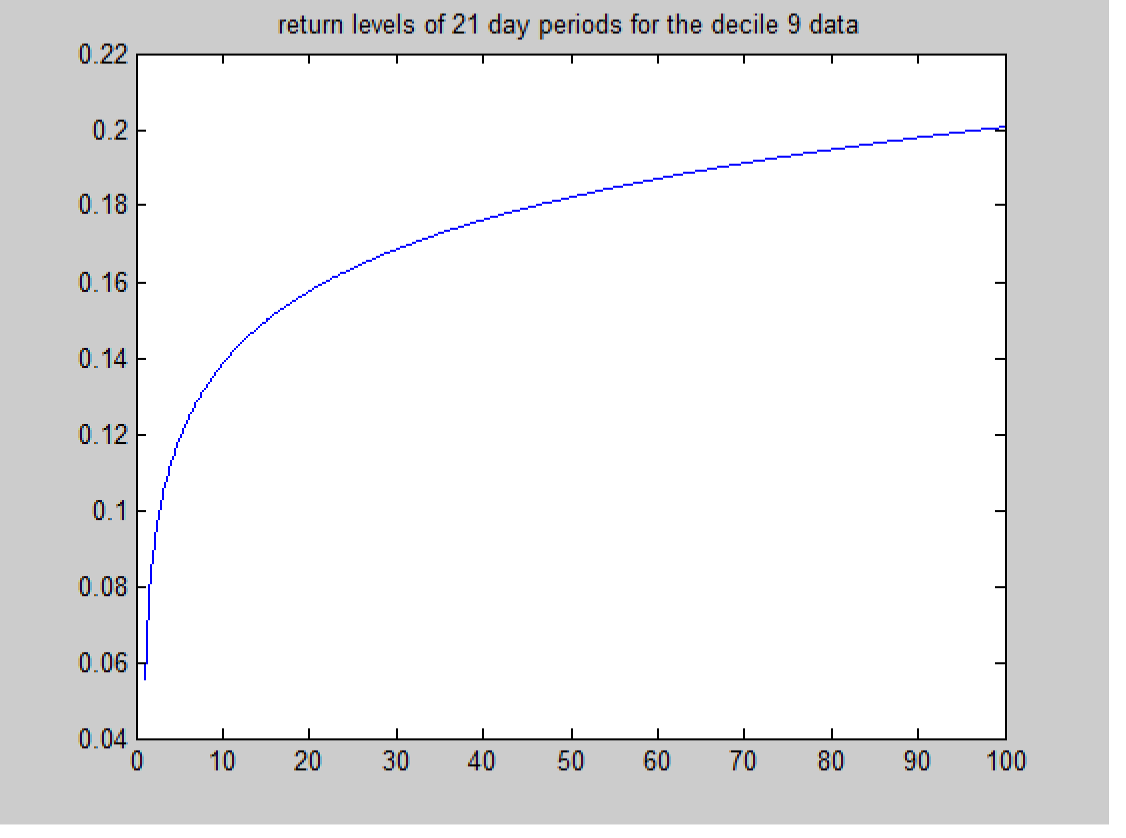
$$P\left(M_{i,n} > L_{n,k}\right) = \frac{1}{k}$$

Return Level

- If n is large enough so that the distribution of maxima is GEV we have that

$$L_{n,k} = \begin{cases} \beta_n - \frac{\alpha_n}{\xi_n}\left\{1 - \left[-\ln\left(1 - \frac{1}{k}\right)\right]^{-\xi_n}\right\} \\ \beta_n - \alpha_n \ln\left[-\ln\left(1 - \frac{1}{k}\right)\right] \end{cases}$$

Return Level



Summary and about tech proj 1

- We have seen that it is possible to compute and forecast Gumbel VaR using simple methods
- Next time we will look into fitting GPD distributions to data
- What has been covered today will be included in technical project 1 and in particular one must know how to create a dynamic Gumbel VaR series under i.i.d. assumption starting from an arbitrary set of stock prices in order to get a grade higher than 3/G
- Hence it is recommended that you download some data set, preferrably two years of prices, from e.g. yahoo finance and starting playing around with it.
- Note that file downloaded from yahoo finance are csv-files that you may transfer to xlsx by using ”text to columns” under ”data” in Excel
- To get started check out the xlsx-file available on facebook and at the course web page