Advanced Risk Management

Week 3

Non-Normal Distributions; Extreme Value Theory

UNIVERSITY OF AMSTERDAM

AMSTERDAM BUSINESS SCHOOL Business

Copyright statement

Outline

- Visualizing non-normality: histograms and QQ plots
- Filtered historical simulation
- Standardized Student's t distribution
- Skewed t distribution
- Extreme value theory

Non-normality

- One of the stylized facts of daily financial returns is that their distribution is not normal: more peaked and longer-tailed, and skewed.
- To a lesser extent, the same applies to shocks $z_t = R_t/\sigma_t$ after fitting GARCH model.
- This can be quantified by the sample skewness and excess kurtosis: for a series $\{x_t\}_{t=1}^T$,

$$SK = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{x_t - \bar{x}}{s_x}\right)^3, \qquad EK = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{x_t - \bar{x}}{s_x}\right)^4 - 3,$$

  where $\bar{x}$ and s_x^2 are the sample mean and variance.
- Jarque-Bera test for $H_0 : x_t \sim N(\mu, \sigma^2)$:

$$JB = T\left(\frac{SK^2}{6} + \frac{EK^2}{24}\right) \overset{H_0}{\sim} \chi_2^2.$$

Non-normality – example

- Consider again the daily (2001–2010) S&P 500 returns R_t, and the shocks z_t after fitting a GJR-GARCH model.
- For the returns, we have

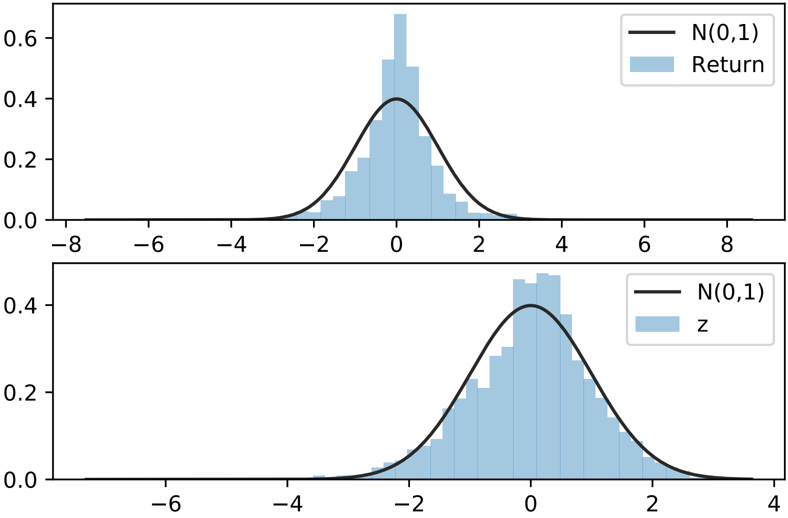
$$SK = -0.124, \qquad EK = 8.212, \qquad JB = 7071 > 5.99;$$

  for the shocks, we find

$$SK = -0.349, \qquad EK = 1.057, \qquad JB = 808 > 5.99.$$

- May be visualized using a histogram, and comparing it to the normal distribution.

Visualizing non-normality – example



QQ plot

- Alternative visualization is the QQ-plot.
- First, sort data in ascending order, so $\{x_i, i = 1, \ldots, T\}$ satisfy

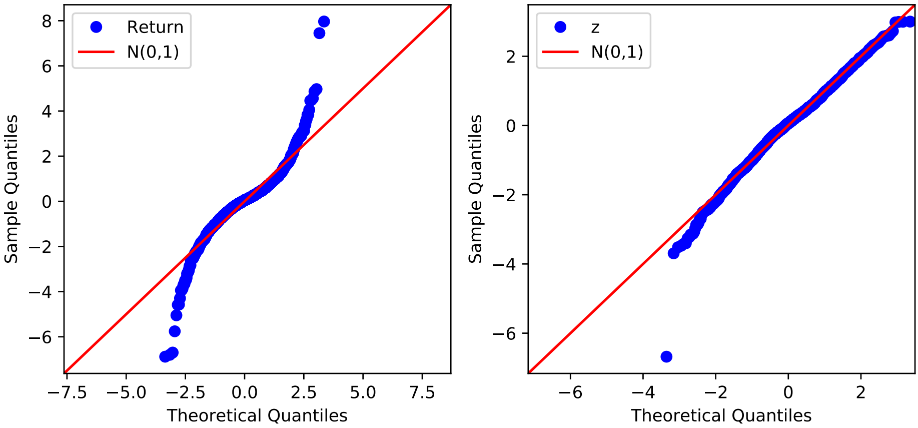
$$x_1 < x_2 < \ldots < x_{T-1} < x_T.$$

- QQ plot is a scatter plot with, for each $i = 1, \ldots, T$:
  - on the vertical axis: x_i;
  - on the horizontal axis: quantiles of theoretical $N(0, 1)$ distribution:

$$\Phi_{p_i}^{-1}, \qquad p_i = \frac{i - 0.5}{T};$$

  - under normality, we expect the points to lie close to the 45° line.
- Can also be applied to alternative theoretical distribution with cdf F: replace $\Phi_{p_i}^{-1}$ by $F_{p_i}^{-1}$.

QQ plot – example



Filtered historical simulation

- Consider again problem of obtaining Value at Risk in model

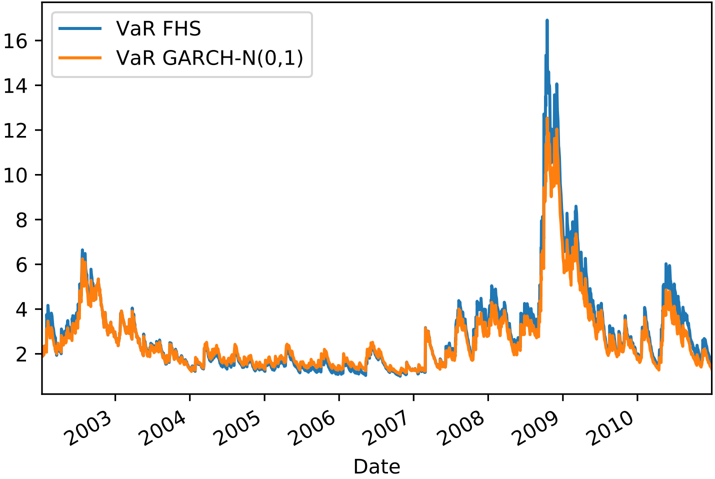
$$R_{PF,t+1} = \sigma_{PF,t+1} z_{t+1},$$

  with $\sigma^2_{PF,t+1}$ obtained from RiskMetrics, GARCH, or extensions; hence $\hat{z}_{t+1} = R_{PF,t+1}/\hat{\sigma}_{PF,t+1}$.
- If $\hat{z}_{t+1}$ has no autocorrelation (acf of $\hat{z}_{t+1}$) or volatility clustering (acf of $\hat{z}^2_{t+1}$), we may assume it to be i.i.d.
- Then we can approximate the distribution of $\hat{z}_t$ by its histogram over the past m days, so

$$VaR^p_{t+1} = -\hat{\sigma}_{PF,t+1} Percentile(\{\hat{z}_{t-j}\}_{j=0}^{m-1}, 100p).$$

- So for $p = 0.01$, replace $\Phi^{-1}_{0.01} = -2.33$ by its vertical coordinate in QQ-plot of $\{\hat{z}_{t-j}\}_{j=0}^{m-1}$.
- Extension to Expected Shortfall: $-\hat{\sigma}_{PF,t+1} \times$ average over pm exceedances $\hat{z}_{t-j}$.

Filtered historical simulation – example



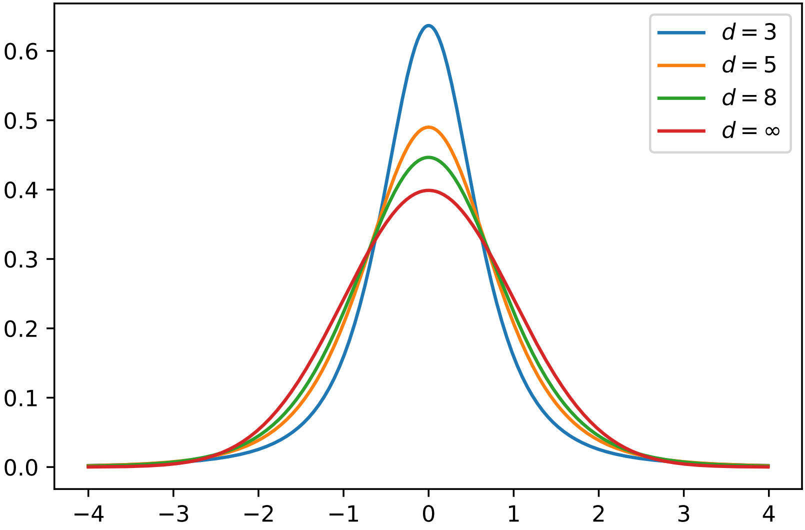
Standardized Student's t distribution

- Instead of using the histogram, we may fit a more flexible distribution to shocks z_t.
- Student's $t(d)$ distribution is symmetric around zero; has excess kurtosis but no skewness.
- For $d > 2$, its variance is $d/(d-2)$; hence standardized $\tilde{t}(d)$ distribution, with density:

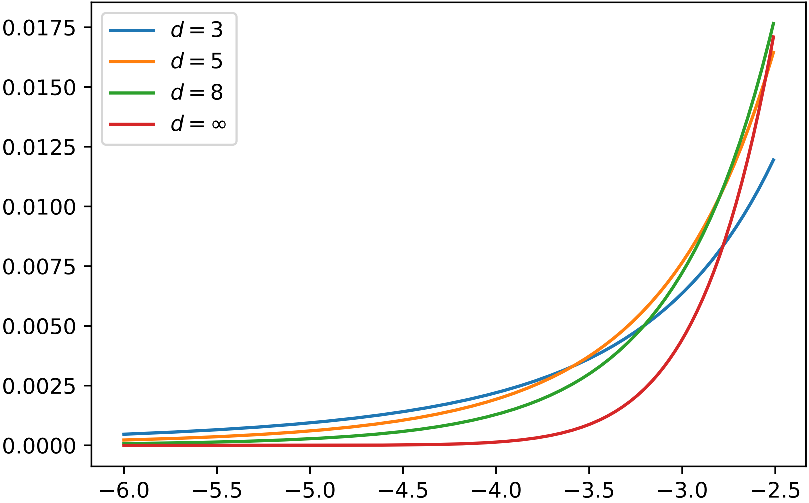
$$f_{\tilde{t}(d)}(z;d) = C(d)\left(1 + \frac{z^2}{(d-2)}\right)^{-(d+1)/2}, \qquad C(d) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)\sqrt{\pi(d-2)}}.$$

- Distribution has mean 0 and variance 1, skewness 0 and (if $d > 4$) excess kurtosis $6/(d-4)$.
- Hence easy estimate: $\hat{d} = 6/EK + 4$.

Standardized $\tilde{t}(d)$ densities



Standardized $\tilde{t}(d)$ densities – left tails



Standardized Student's t distribution – estimation

- Alternative to easy estimator of d: maximum likelihood.
- Two-step method: first estimate GARCH parameters by QMLE; then estimate d from

$$\max_{d} \sum_{t=1}^{T} \ln \left(f_{\tilde{t}(d)} \left(\hat{z}_t; d \right) \right).$$

- One-step ML estimation, joint with GARCH parameters:

$$\max_{\omega,\alpha,\beta,d} \sum_{t=1}^{T} \ln \left(\frac{1}{\sigma_t} f_{\tilde{t}(d)} \left(\frac{R_t}{\sigma_t}; d \right) \right), \qquad \sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta \sigma_t^2 .$$

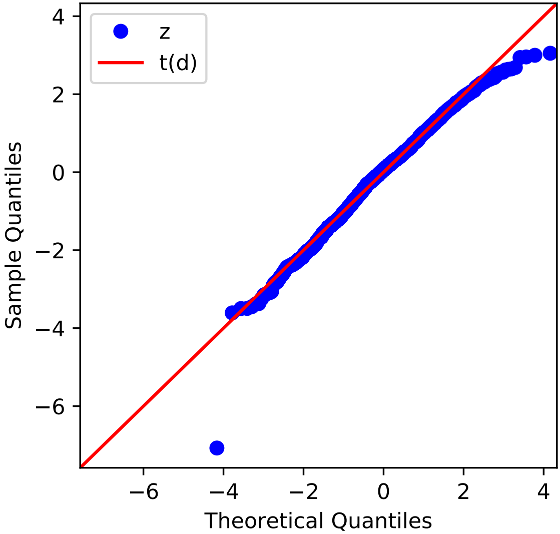
- Python implementation:

```
am = arch_model(R, p=1, o=1, q=1, dist='StudentsT')
```

GARCH estimation with t distribution – example

```
================================================================================
Vol Model:                            GJR-GARCH   Log-Likelihood:           -3646.84
Distribution:      Standardized Student's t   No. Observations:             2514
                                  Mean Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
mu             0.0237  1.767e-02      1.344      0.179 [-1.090e-02,5.839e-02]
                               Volatility Model
=============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
-----------------------------------------------------------------------------
omega      9.0209e-03  3.573e-03      2.525  1.158e-02  [2.018e-03,1.602e-02]
alpha[1]       0.0000  1.779e-02      0.000      1.000 [-3.487e-02,3.487e-02]
gamma[1]       0.1277  1.888e-02      6.767  1.316e-11    [9.073e-02,  0.165]
beta[1]        0.9276  1.946e-02     47.666      0.000      [  0.889,  0.966]
                              Distribution
========================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
------------------------------------------------------------------------
nu            10.7201      2.652      4.043  5.283e-05 [  5.523, 15.917]
========================================================================
```

QQ plot of shocks versus $t(d)$ distribution



Standardized Student's t distribution – VaR and ES

- If $t_p^{-1}(d)$ is the pth quantile of the (non-standardized) $t(d)$ distribution, then

$$\tilde{t}_p^{-1}(d) = \sqrt{\frac{d-2}{d}}\, t_p^{-1}(d).$$

- $t_p^{-1}(d)$ is available from programming languages (e.g. `stats.t.ppf()`); hence

$$VaR_{t+1}^{p} = -\sigma_{PF,t+1}\sqrt{\frac{d-2}{d}}\, t_p^{-1}(d).$$

- Expected shortfall (some typing errors in book):

$$ES_{t+1}^{p} = \sigma_{PF,t+1}\left[\frac{C(d)}{p}\left(\frac{d-2}{d-1}\right)\left(1+\frac{t_p^{-1}(d)^2}{d}\right)^{(1-d)/2}\right],$$

with $C(d)$ as defined before.

Skewed t distribution

- To allow for skewness, the t distribution can be generalized in various ways.
- Asymmetric or skewed t distribution of Section 6.7 obtained by pasting together two t densities:

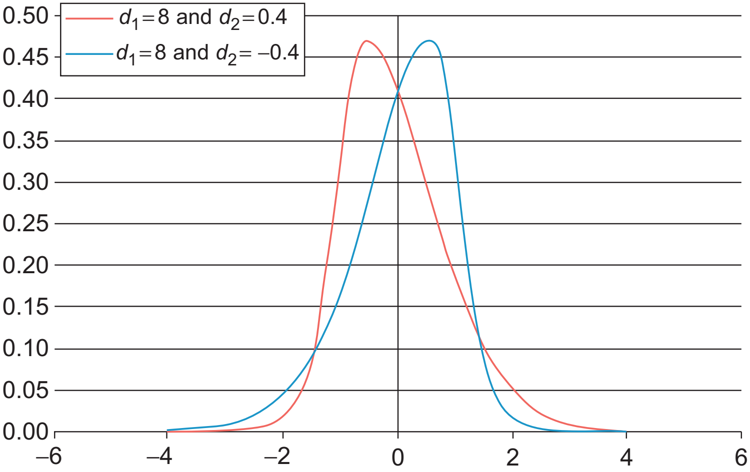
$$f_{asyt}(z; d, \lambda) = \begin{cases} B \times f_{\tilde{t}(d)}\left(\dfrac{B(z-m)}{1-\lambda}; d\right), & \text{if } z < m, \\ B \times f_{\tilde{t}(d)}\left(\dfrac{B(z-m)}{1+\lambda}; d\right), & \text{if } z \geq m, \end{cases}$$

  with:
    - $\lambda \in (-1, 1)$ is the skewness parameter: $\lambda = 0$ gives $\tilde{t}(d)$;
    - m is the mode of the distribution (at which point density has maximum);
    - B is scale parameter needed to make variance equal to 1.

  B and m are chosen as function of λ, so if $\lambda = 0$, then $m = 0$ and $B = 1$.
- Notation in the book: $d = d_1$, $\lambda = d_2$, $m = -A/B$.

Skewed t distribution

Figure 6.4 The asymmetric t distribution.



Extreme value theory

- EVT is branch of statistics concerned with distribution of extremes (minima and maxima).
- Dutch contributions, historically related to analysis of required height of sea dikes.
- Requires i.i.d. observations, hence applied to shocks z_t, not returns R_t.
- Let $x_t = -z_t$, and define $u > 0$. We want to characterize right tail of distribution, for $x_t > u$:

$$F_u(y) = \Pr(x_t \leq u + y | x_t > u) = \frac{\Pr(u < x_t \leq u + y)}{\Pr(x_t > u)}, \qquad y > 0.$$

- Pickands-Balkema-de Haan theorem: for large class of distributions of x_t,

$$\lim_{u \to \infty} F_u(y) = GPD(y; \xi, \beta) = \begin{cases} 1 - (1 + \xi y/\beta)^{-1/\xi}, & \text{if } \xi \neq 0, \\ 1 - \exp(-y/\beta), & \text{if } \xi = 0. \end{cases},$$

  where β is a scale parameter, and ξ is the tail index.
- $\xi > 0$ implies heavy tails (e.g., Student's $t(d) \Rightarrow \xi = 1/d$).

Hill estimator

- One approach to estimating ξ (and β) is by MLE, using only observations $x_t = -z_t$ satisfying $x_t > u$.
- Simpler estimator uses approximation of distribution of x_t in the right tail (only valid for $\xi > 0$):

$$F(x) = \Pr(x_t \leq x) \approx 1 - cx^{-1/\xi}.$$

- MLE for this distribution is the Hill estimator:

$$\hat{\xi} = \frac{1}{T_u} \sum_{i=1}^{T_u} \ln\left(\frac{x_i}{u}\right),$$

where $x_i = -z_i, i = 1, \ldots, T_u$ are the largest T_u observations, all exceeding u.

- In practice, u is chosen to keep fraction T_u/T of data characterizing the tail;
  - at the minimum, T_u/T should be larger than p if we need VaR^p_{t+1}.

VaR and ES based on EVT

- Simpler approximate distribution with $F(u) = 1 - T_u/T$ imposed gives

$$\Pr(z_t < -x) = \Pr(x_t > x) = \frac{T_u}{T}\left(\frac{x}{u}\right)^{-1/\xi};$$

  setting this equal to p and solving for x gives

$$VaR^p_{t+1} = \sigma_{t+1} u \left(\frac{p}{T_u/T}\right)^{-\xi}.$$

- Using same approximation,

$$\frac{ES^p_{t+1}}{VaR^p_{t+1}} = \frac{1}{1-\xi}.$$

VaR and ES based on t distribution and EVT – example

- Student's $t(d)$: estimate $\hat{d} = 10.7$ implies

$$\begin{aligned} VaR^{0.01}_{t+1} &= \sigma_{PF,t+1} \times 2.46, \\ ES^{0.01}_{t+1} &= \sigma_{PF,t+1} \times 2.98. \end{aligned}$$

  The fraction of exceedances is 1.11%.

- EVT: estimate $\hat{\xi} = 0.19$ implies

$$\begin{aligned} VaR^{0.01}_{t+1} &= \sigma_{PF,t+1} \times 2.54, \\ ES^{0.01}_{t+1} &= \sigma_{PF,t+1} \times 3.13. \end{aligned}$$

  The fraction of exceedances is 0.88%.

Copyright statement

Figure 4.6 on Slide 18 has been extracted from:

> Christoffersen, P. F. (2012), *Elements of Financial Risk Management*, Second Edition. Waltham (MA): Academic Press. ISBN 978-0-12-374448-7.

Back to first slide