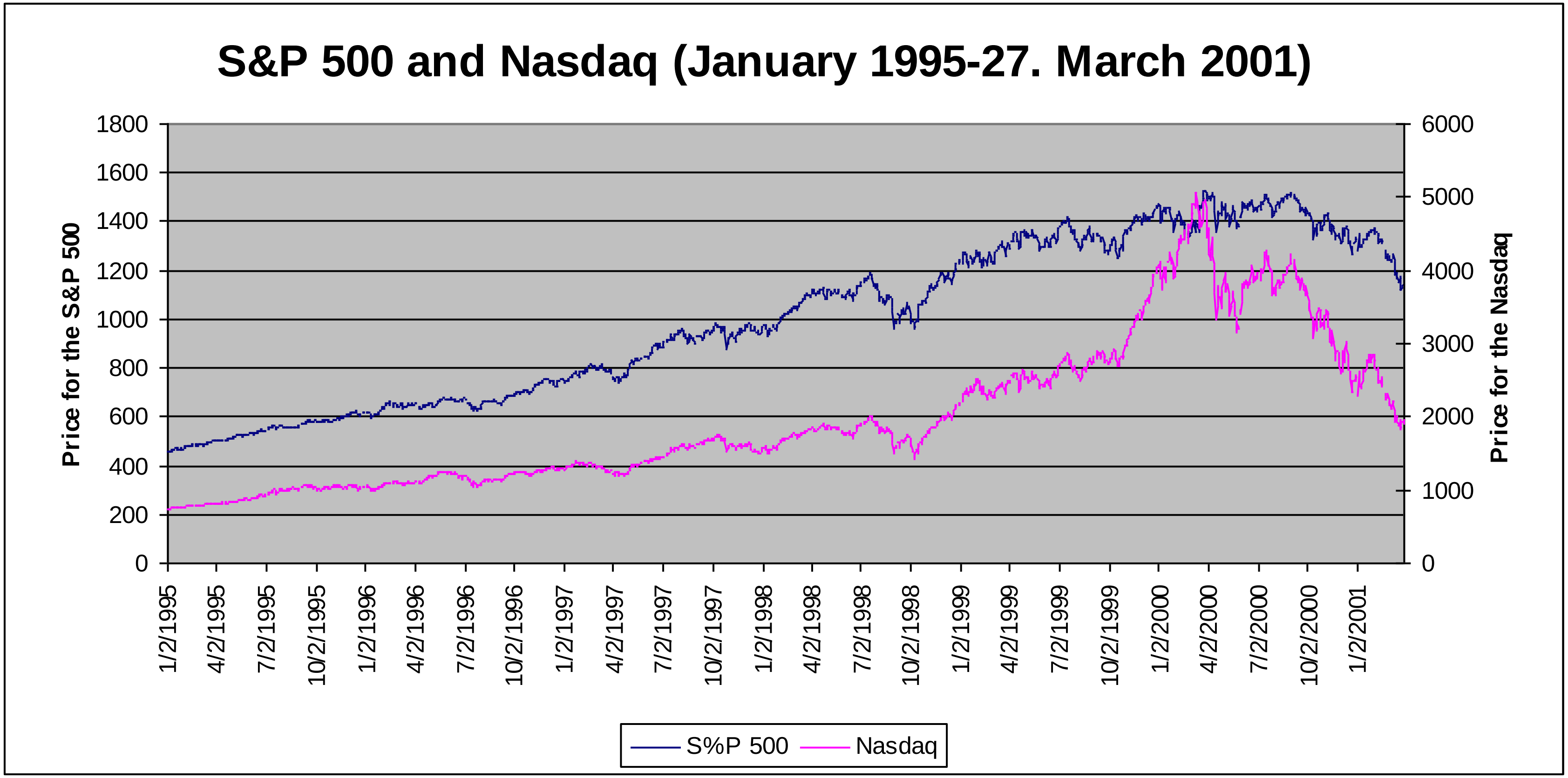
Tail Risk – Measuring

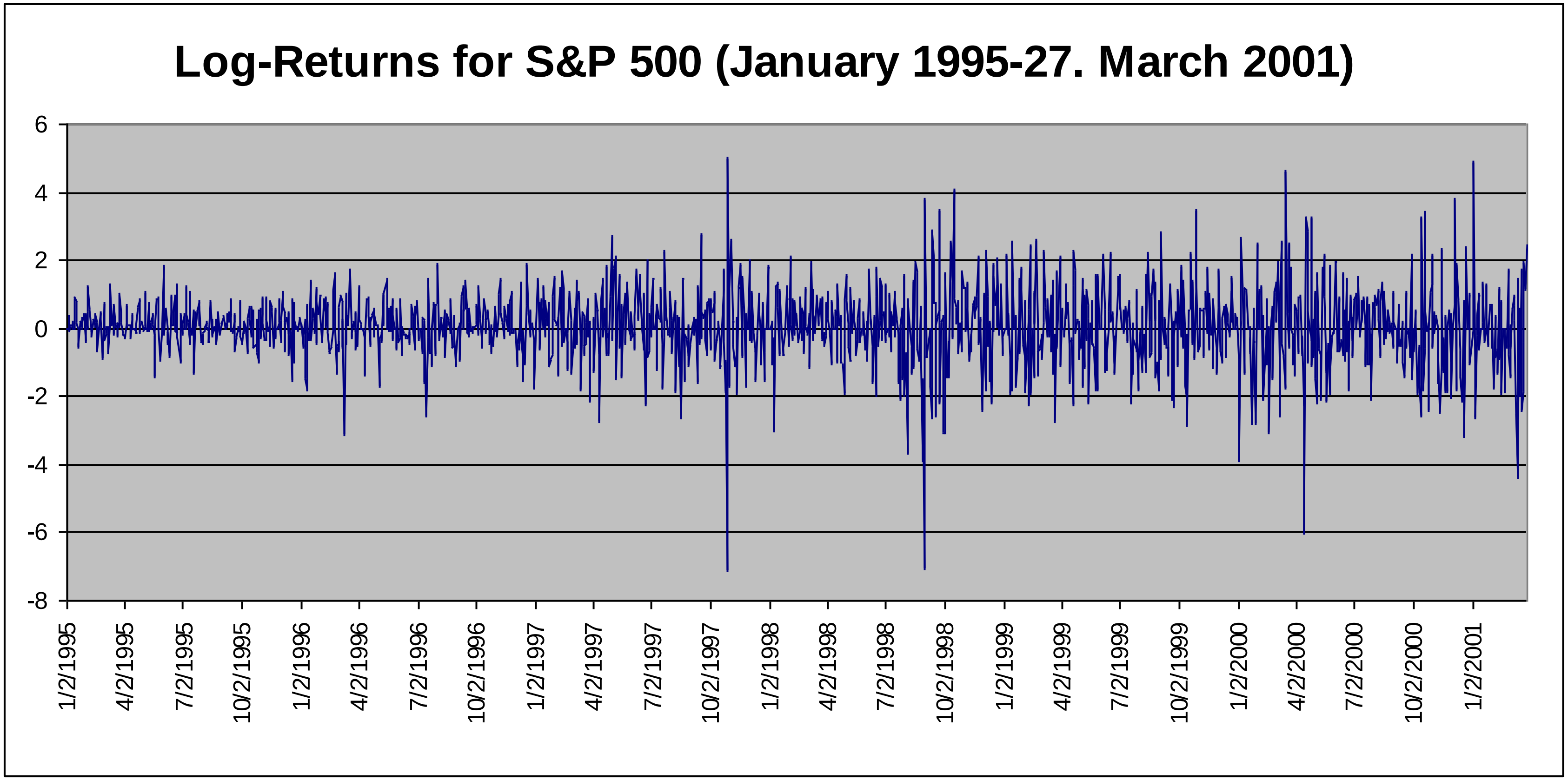
Presented by: Claus Madsen

Quant-Network: 18. April 2012

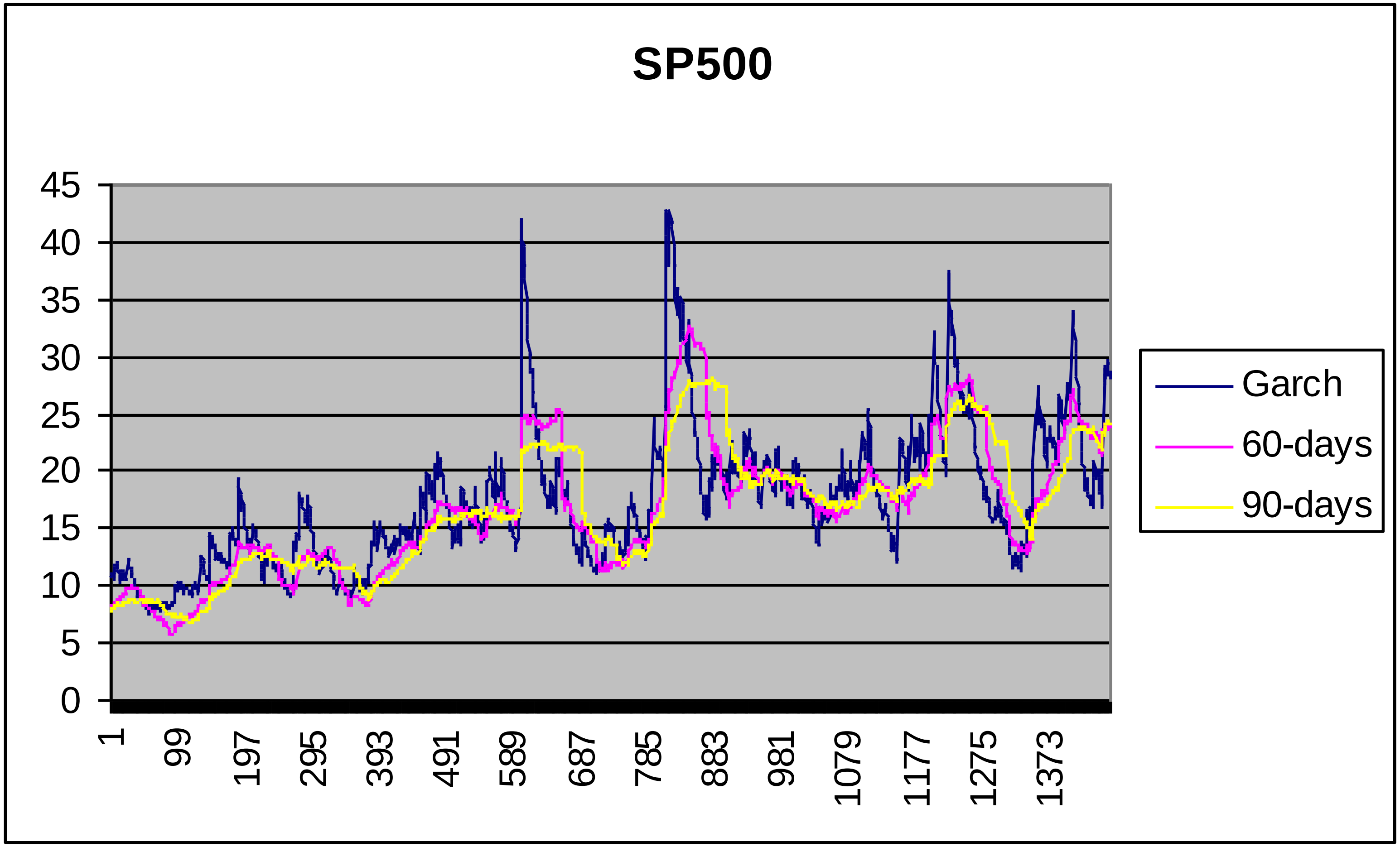
Volatility - I



Volatility - II



Volatility Estimation



How to calculate Tail Risk - I

- The question can be re-phrased: How to Calculate Risk?

- There are a lot of competing methods – disregarding the standard normal-model:
  - Monte Carlo Simulation
    - Dimension Reduction, possible Scenario Simulation
  - Historical Simulation
    - Potential with Stochastic Volatility
  - Non-Normal Distribution
    - t – Distribution, mixed-normal distribution
  - EVT
  - Copula

Monte Carlo issues...

- The following things needs to be addressed:
  - How to handle a large number of factors – or put differently how to handle the dimension of the correlation matrix
    - PCA (SVD)
  - How to handle the fact that the standard deviation only decreases as $O\left(N^{-\frac{1}{2}}\right)$
  - Other questions:
    - Which processes to simulate
    - Which probability distribution to select

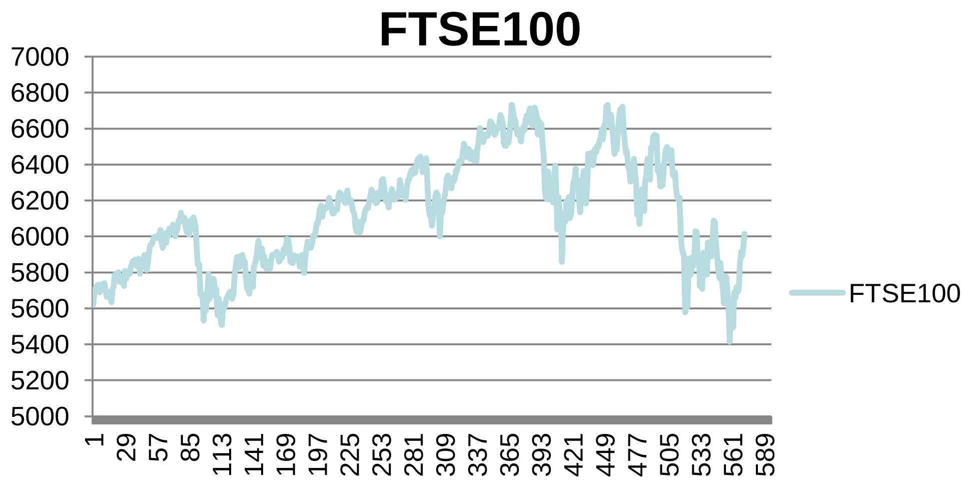
Monte Carlo contra Historical

- Historical simulation is a special case of the algorithm in which case the scenarios are drawn directly from historical data
- This approach is sometimes defended due to the fact that it is easy to explain to a nontechnical audience, but the historical distribution inevitable has gaps – especially in the tails

- Remark: Fitting a theoretical distribution (like for example the Generalization of Tukeys lambda distribution) and then applying Monte Carlo gets around this artificial feature of the purly data-driven implementation

Mixed-Normal Distribution

- The mixture setting is designed to capture regime switch – different regimes
- Example: FTSE 100



- We get:

Estimated Parameters for the Normal Mixture Distribution (annualized)

	Weight	Mean 1	Sigma 1	Mean 2	Sigma 2
FTSE100	0,340027	-0,1243%	1,6754%	0,0821%	0,6022%

- Risk:

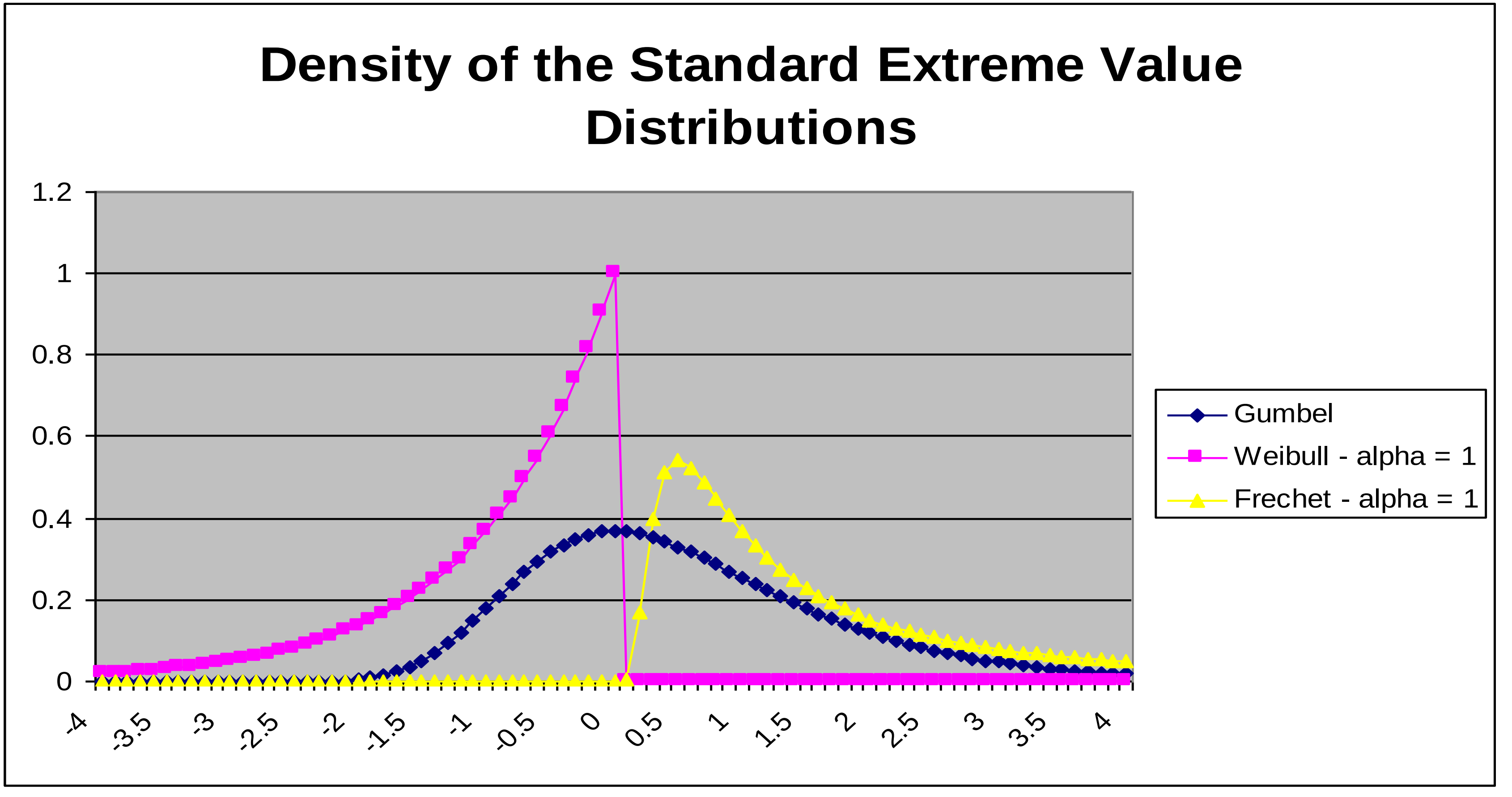
Risk	Normal	Mixture Normal
FTSE100 VaR	16,14	20,07
FTSE100 ETL	18,49	24,11

EVT- I

- Extreme Value Theory - the main tools:
  - Distribution of maxima - The generalised Extreme Value Distribution (GEV)
  - Distribution of the tails - The generalised Pareto Distribution (GPD)

Remark: The Method is not often used in Finance but widely used in Insurance - which has a long history of modelling extreme events. In Finance, on the other hand, focus is generally on modelling expected events.

EVT - II



Density of the Standard Extreme Value Distributions
1.2
1
0.8
0.6
0.4
0.2
0
-4
-3.5
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3
3.5
4
Gumbel
Weibull - alpha = 1
Frechet - alpha = 1

Mean excess function, and excess distribution function

- For some threshold y we are interested in the conditional excess loss function:

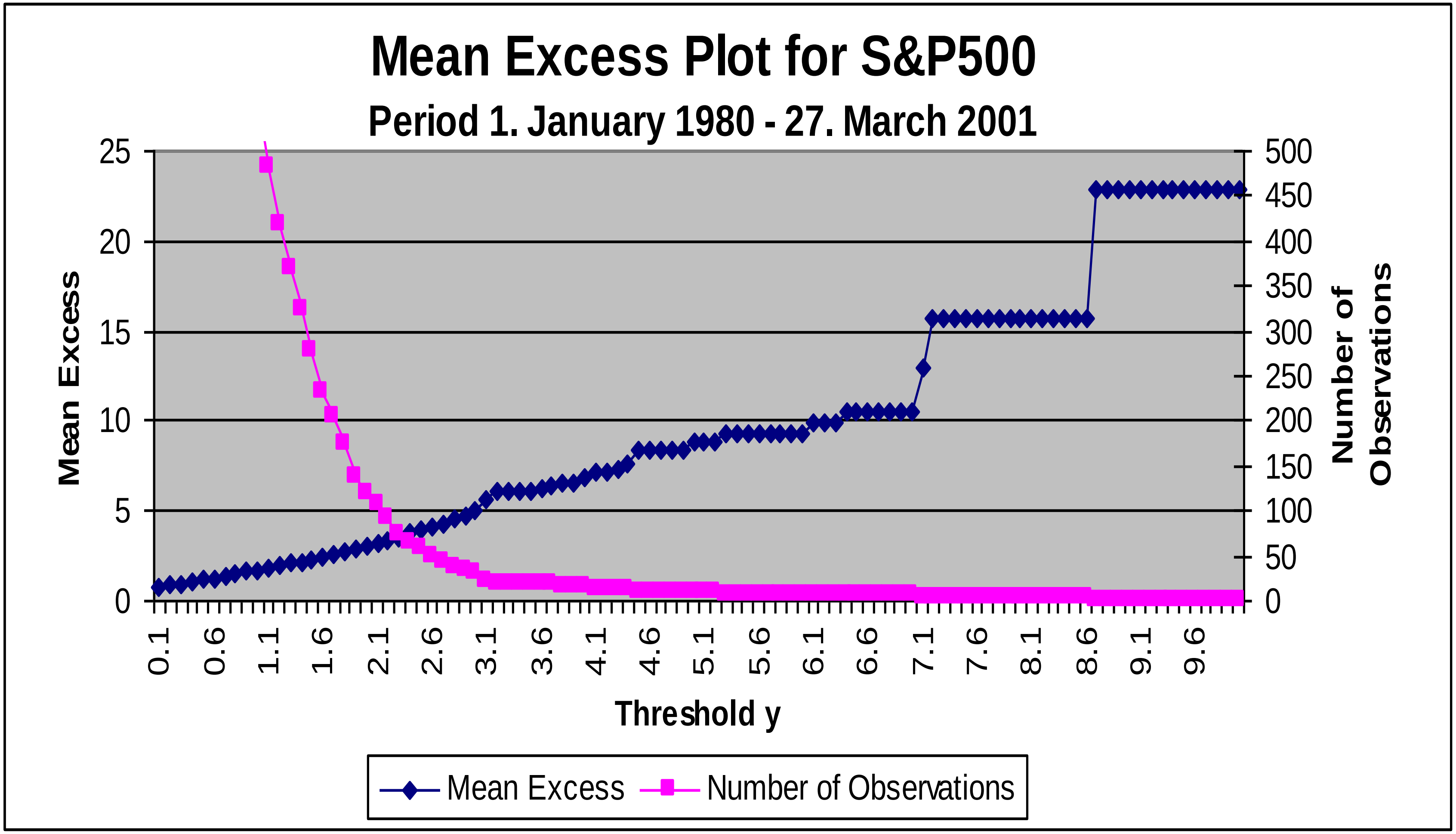
$$e(y) = E(X - y \mid \mathrm{X} > \mathrm{y})$$

Where this represents the expected return above the level y, conditional on the return being greater than y. e(y) + y is the expected loss, if every loss is greater than y.

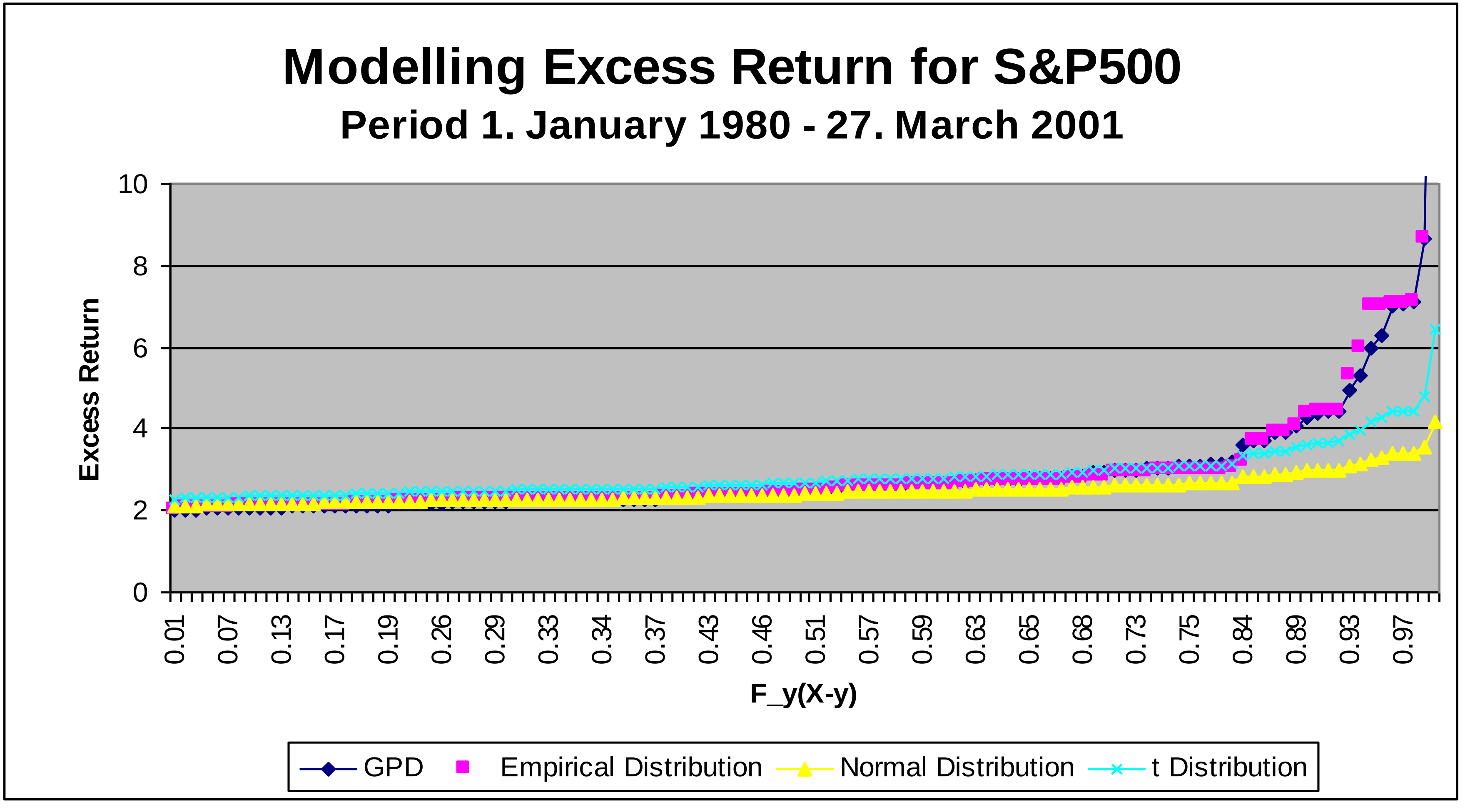
- The conditional excess distribution is:

$$F_y(x) = \Pr[X - y \leq x \mid X > y], \text{for} \quad x \geq 0$$

Mean Excess Plot for S&P500



Estimating the tail distribution for S&P500 (Threshold (y) = 2) - I



Estimating the tail distribution for S&P500 (Threshold (y) = 2) - II

- To summarize lets us look at the probability for observing a return at 4% or higher for the GPD, Normal distribution and the t distribution:
- For the Normal distribution the probability is: 0.000116%
  - Approximately 1 every 34 years
- For the t distribution the probability is: 0.00116%
  - Approximately 1 every 3.5 years
- For the GPD the probability is: 0.00263%
  - Approximately 1 every 1.5 years

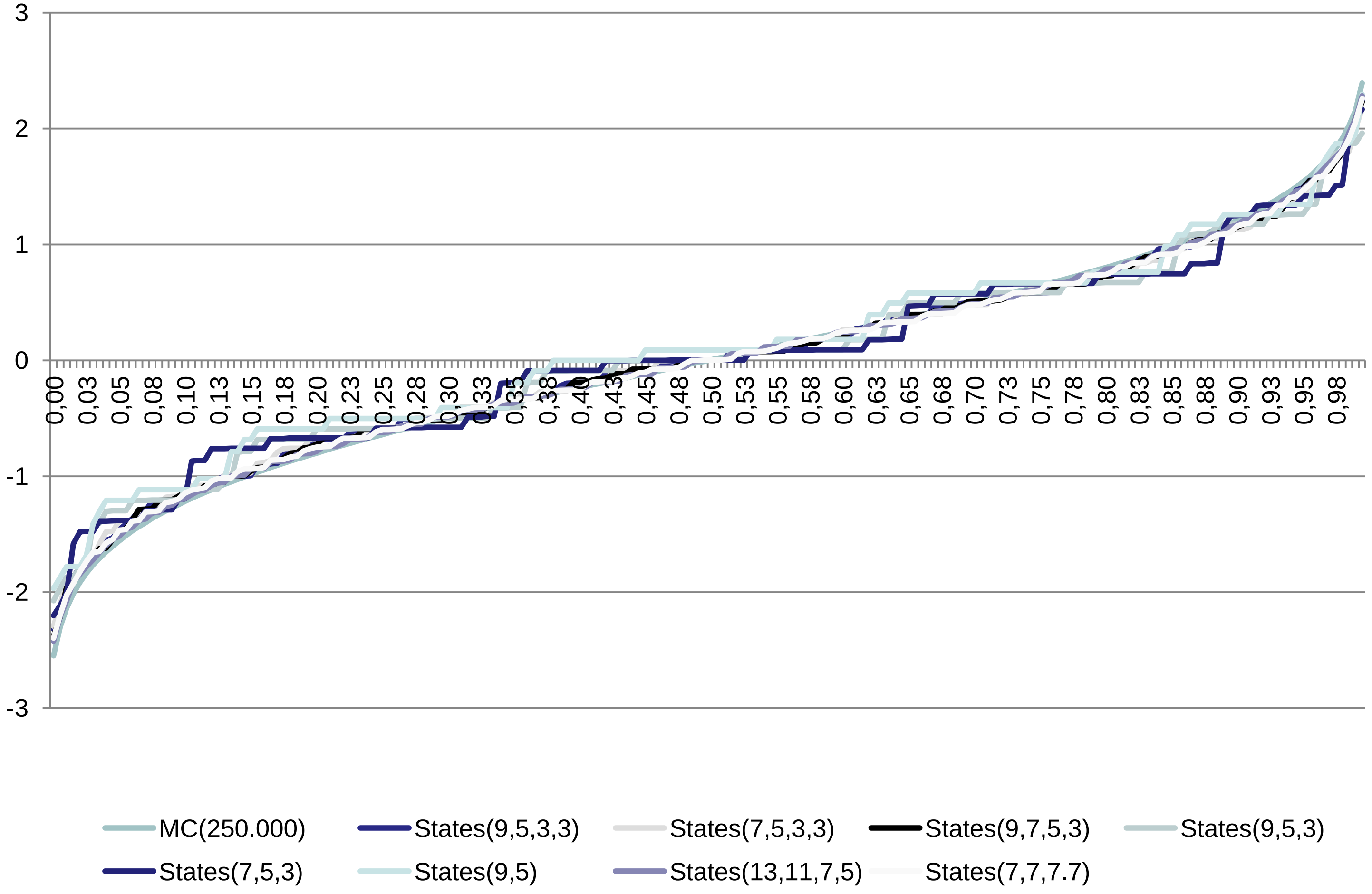
- **Remark: For the actual data set we have the an event equal to or higher than 4% for around 14 times - is approximately = 0.00261%**

Copulas

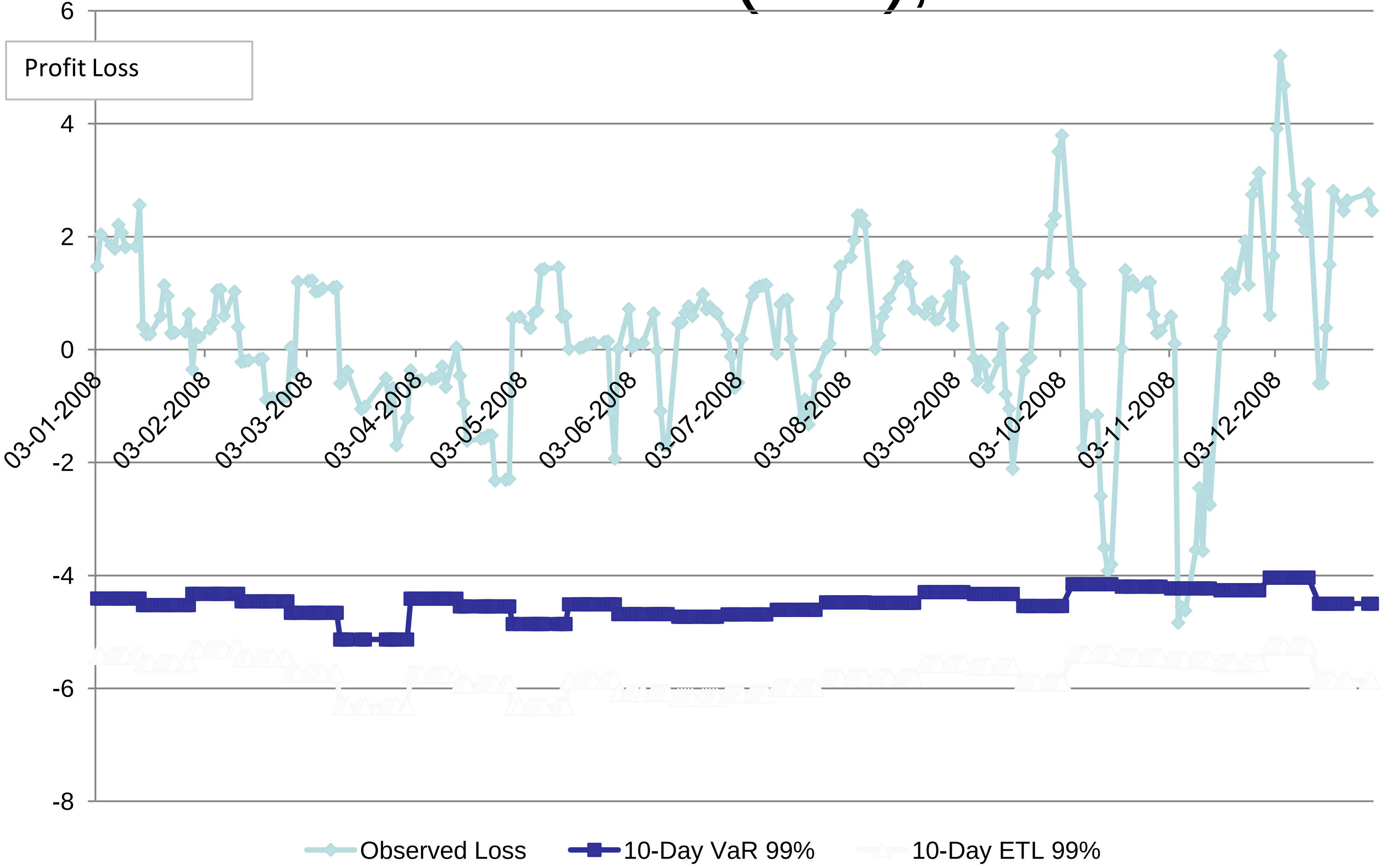
- A way to allow for heterogeneous marginal distributions one can model the dependency using a Copula distribution
  - Copulas are multivariate distributions with uniform marginals
- More general dependency – instead of the traditional linear association

- A very attractive class of distribution are the Archimedian Copulas…..
- Can be combined with MC…which is neat

Monte Carlo and Scenario Simulation



The Financial Crisis
PCA, Stochastic Volatility, Scenario Simulation (MC), HJM



Closing Remarks - I

- Nassim Taleb:
  - “It is often said that [one] ‘is wise who can see things coming.’ Perhaps the wise one is the one who knows that he cannot see things far away.”
- Nassim Taleb:
  - ”No amount of observations of white swans can allow the inference that all swans are white, but the observation of a single black swan is sufficient to refute that conclusion”
- Sun Tzu ”The art of War”, chapter VIII ”Variation on Tactics”, rule 11
  - The art of war teaches us to rely not on the likelihood of the enemy’s not coming, but on our readiness to receive him; not on the chance of his not attacking, but rather on the fact that we have made our position unassailable
- The 8th commandment from ”The Rogue Warrior's Ten Commandments of SpecWar”
  - Thou shalt never assume

Closing Remarks - II

- *“There is always going to be an element of doubt, as one is extrapolating into areas one doesn’t know about. But what EVT is doing is making the best of whatever data-sets you have about extreme phenomena”,* Smith from the preface to EKM (1997)

- There is no magical technique which yields reliable results for free - for that reason one could formulate as in finance: “*There is no free lunch when it comes to high quantile estimation!”*

Appendix

GEV - The distribution of maxima - I

- Due to the Fisher-Tippet theorem we have (3.2.3 in EKM (1979)):
  - For any sample distribution the maxima are asymptotically distributed as one of the following 3, where α is the tail parameter:
    - Fat tails: Frechet - with cdf equal to: $\Phi_\alpha(x) = e^{-x^{-\alpha}}, \text{for } x > 0$
      - (for $\alpha > 0$) $\Phi_\alpha(\mathrm{x}) = 0, \text{for } x \geq 0$
    - Short tails: Weibull - with cdf equal to: $\Psi_\alpha(x) = e^{-(-x)^{-\alpha}}, \text{for } x \leq 0$
      - (for $\alpha > 0$) $\Psi_\alpha(\mathrm{x}) = 1, \text{for } x > 0$
    - Light tails: Gumbel - with cdf equal to: $\Lambda(x) = e^{-e^{-x}}, \text{for all } x$

Remark: GEV is the limit distribution for centred and normalised maxima

GEV - The distribution of maxima - II

- Generalising the GEV:

$$H_{\tau;\mu,\sigma}(x) = e\left\{-\left(1+\tau\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\tau}}\right\}, \text{ for } 1+\tau\frac{x-\mu}{\sigma} > 0$$

$$\Phi_{\alpha}\left(1+\frac{x-\mu}{\alpha\sigma}\right), \text{ for } x > \mu-\alpha\sigma, \tau = \frac{1}{\alpha} > 0$$

$$= \Psi_{\alpha}\left(-\left(1-\frac{x-\mu}{\alpha\sigma}\right)\right), \text{ for } x < \mu+\alpha\sigma, \tau = \frac{-1}{\alpha} < 0$$

$$\Lambda\left(\frac{x-\mu}{\sigma}\right), \text{ for } x \in \Re, \tau = 0$$

- where μ = location parameter, σ = scale parameter, τ is the tail parameter and α is the shape parameter

GPD - The tail distribution - I

- The generalized Pareto distribution (GPD) is defined as:

$$G_{\tau;\mu,\sigma}(x) = 1-\left(1+\tau\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\tau}}, \text{for } \tau \neq 0$$

$$1-e^{-x}, \text{for } \tau = 0$$

If $\tau > 0$, $x \geq 0$ - then the tails are Frechet as $x \to \infty$

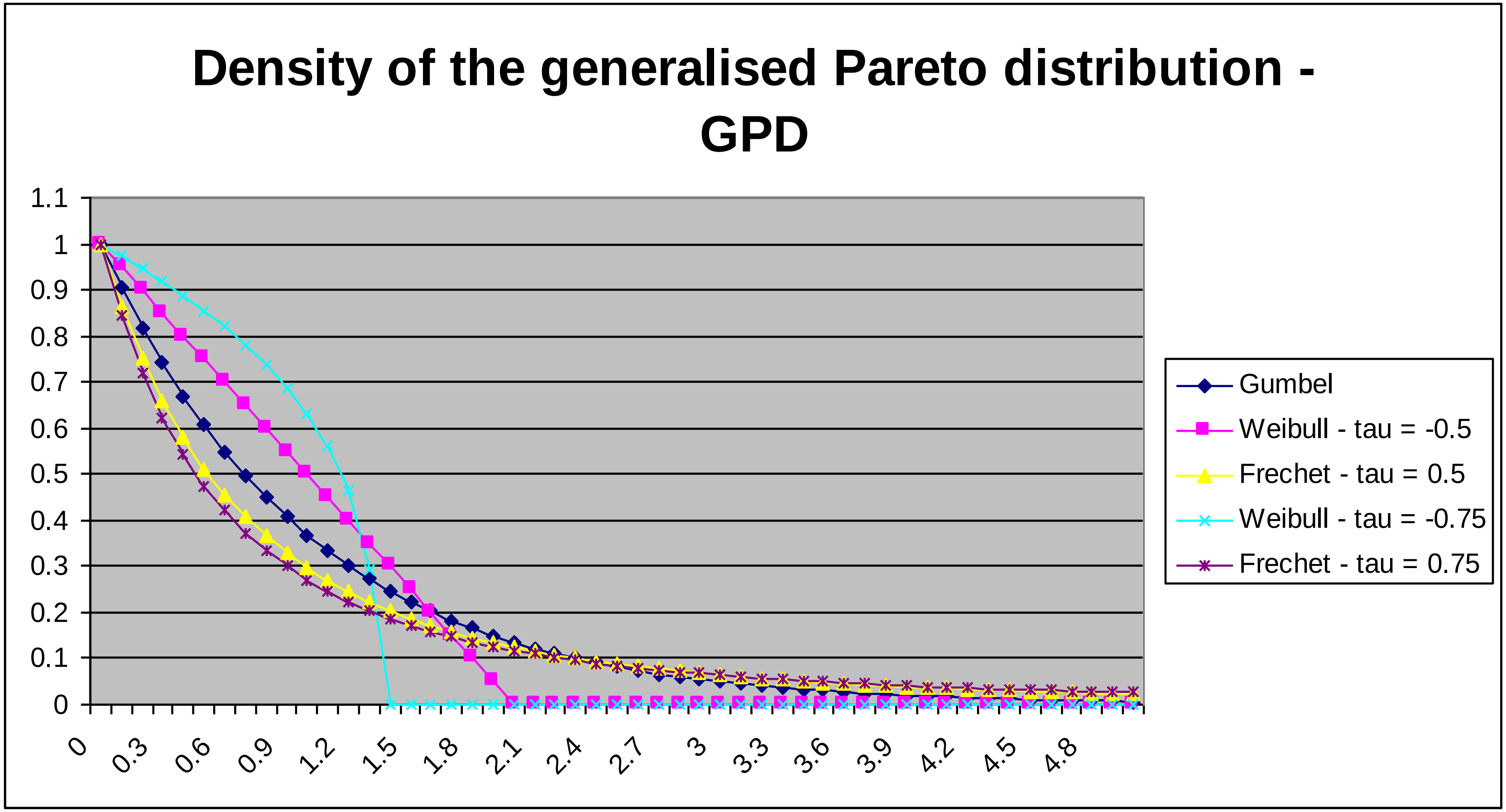
If $\tau < 0$, $0 \leq x \leq -1/\tau$ - then the tails are Weibull as $x \to 1/\tau$

If $\tau = 0$ - then the tails are Gumbel as $x \to \infty$

Remark 1: Sometimes referred to as the POT-model - peaks-over-threshold

Remark 2: For conditional processes add the extremal index parameter θ - see EKM (1979) - chapter 8. θ = 1 = weak dependency or where the dependency structure is not "too strong".

GPD - The tail distribution - II



GPD - The tail distribution - III

- For any large enough threshold y, the conditional excess distribution is allways GPD, for some τ, μ and σ
  - this is true for the whole tail - not just the tail of the distribution of maxima - which makes GPD ideal for data-sparse tails.
- The GPD is the best distribution to model any extreme tail and it can use all the available data
- For the GPD we need 3 parameters:
  - The tail parameter $\tau = 1/\alpha$ (where α is the shape parameter)
  - The location parameter μ
  - The scale parameter σ
- The tail-parameter is the most important, as it describes the extent of the fatness of the tail

GPD - The tail distribution - IV

- An interesting observation is that in general we are most interested in the case $\tau > 0$ (the Frechet distribution) as this corresponds with fat-tails

- In the Frechet distribution the shape parameter $\alpha = 1/\tau$ represents the maximal order of finite moments. For example if $\alpha > 1$ - then the mean exist, if $\alpha > 2$ - then the variance is finite, if $\alpha > 3$ - then the skewness is well defined and so on.
- The shape parameter is an intrinsic parameter of the distribution of returns and does not depend on the number of returns n from which the minimal return is selected. The shape parameter corresponds to the degrees of freedom for the t distribution and to the characteristics exponent of a stable Paretian distribution

GPD - What have we got?

- Finding the probability of a given event (loss), even though no such event (loss) has never occurred
- Finding the level of a loss at a given probability, even when there is insufficient data to calculate the level empirically