

NUMERICAL METHOD

Introduction to Algorithmic Trading Strategies Lecture 8

Risk Management

Haksun Li

haksun.li@numericalmethod.com

www.numericalmethod.com

Outline

- ▶ Value at Risk (VaR)
- ▶ Extreme Value Theory (EVT)



References

- ▶ AJ McNeil. Extreme Value Theory for Risk Managers. 1999.
- ▶ Blake LeBaron, Ritirupa Samanta. Extreme Value Theory and Fat Tails in Equity Markets. November 2005.



Risks

- ▶ Financial theories say:
 - ▶ the most important single source of profit is risk.
 - ▶ profit \propto risk.
- ▶ *I personally do not agree.*



What Are Some Risks? (1)

- ▶ **Bonds:**

- ▶ duration (sensitivity to interest rate)
- ▶ convexity
- ▶ term structure models

- ▶ **Credit:**

- ▶ rating
- ▶ default models



What Are Some Risks? (2)

- ▶ **Stocks**

- ▶ volatility
- ▶ correlations
- ▶ beta

- ▶ **Derivatives**

- ▶ delta
- ▶ gamma
- ▶ vega



What Are Some Risks? (3)

- ▶ **FX**
 - ▶ volatility
 - ▶ target zones
 - ▶ spreads
 - ▶ term structure models of related currencies



Other Risks?

- ▶ Too many to enumerate...
 - ▶ natural disasters, e.g., earthquake
 - ▶ war
 - ▶ politics
 - ▶ operational risk
 - ▶ regulatory risk
 - ▶ wide spread rumors
 - ▶ alien attack!!!
- ▶ Practically infinitely many of them...



VaR Definition

- ▶ Given a loss distribution, F , quintile $1 > q \geq 0.95$,
- ▶ $\text{VaR}_q = F^{-1}(q)$



Expected Shortfall

- ▶ Suppose we hit a big loss, what is its expected size?
- ▶ $ES_q = E[X|X > VaR_q]$



VaR in Layman Term

- ▶ VaR is the maximum loss which can occur with certain confidence over a holding period (of n days).
- ▶ Suppose a daily VaR is stated as \$1,000,000 to a 95% level of confidence.
- ▶ There is only a 5% chance that the loss the next day will *exceed* \$1,000,000.



Why VaR?

- ▶ Is it a true way to measure risk?
 - ▶ NO!
- ▶ Is it a universal measure accounting for most risks?
 - ▶ NO!
- ▶ Is it a good measure?
 - ▶ NO!
- ▶ Only because the industry and regulators have adopted it.
 - ▶ It is a widely accepted standard.



VaR Computations

- ▶ Historical Simulation
- ▶ Variance-CoVariance
- ▶ Monte Carlo simulation



Historical Simulations

- ▶ Take a historical *returns* time series as the returns distribution.
- ▶ Compute the loss distribution from the historical returns distribution.



Historical Simulations Advantages

- ▶ Simplest
- ▶ Non-parametric, no assumption of distributions, no possibility of estimation error



Historical Simulations Dis-Advantages

- ▶ As all historical returns carry equal weights, it runs the risk of over-/under- estimate the recent trends.
- ▶ Sample period may not be representative of the risks.
- ▶ History may not repeat itself.
- ▶ Cannot accommodate for new risks.
- ▶ Cannot incorporate subjective information.



Variance-CoVariance

- ▶ Assume all returns distributions are Normal.
- ▶ Estimate asset variances and covariances from historical data.
- ▶ Compute portfolio variance.
 - ▶ $\sigma_P^2 = \sum_{i,j} \rho_{ij} \omega_i \omega_j \sigma_i \sigma_j$



Variance-CoVariance Example

- ▶ 95% confidence level (1.645 stdev from mean)
- ▶ Nominal = \$10 million
- ▶ Price = \$100
- ▶ Average return = 7.35%
- ▶ Standard deviation = 1.99%
- ▶ The VaR at 95% confidence level = $1.645 \times 0.0199 = 0.032736$
- ▶ The VaR of the portfolio = $0.032736 \times 10 \text{ million} = \$327,360$.



Variance-CoVariance Advantages

- ▶ Widely accepted approach in banks and regulations.
- ▶ Simple to apply; straightforward to explain.
- ▶ Datasets immediately available
 - ▶ very easy to estimate from historical data
 - ▶ free data from RiskMetrics
 - ▶ <http://www.jpmorgan.com>
- ▶ Can do scenario tests by twisting the parameters.
 - ▶ sensitivity analysis of parameters
 - ▶ give more weightings to more recent data



Variance-CoVariance Disadvantages

- ▶ Assumption of Normal distribution for returns, which is known to be not true.
- ▶ Does not take into account of fat tails.
- ▶ Does not work with non-linear assets in portfolio, e.g., options.



Monte Carlo Simulation

- ▶ You create your own returns distributions.
 - ▶ historical data
 - ▶ implied data
 - ▶ economic scenarios
- ▶ Simulate the joint distributions many times.
- ▶ Compute the empirical returns distribution of the portfolio.
- ▶ Compute the (e.g., 1%, 5%) quantile.



Monte Carlo Simulation Advantages

- ▶ Does not assume any specific models, or forms of distributions.
- ▶ Can incorporate any information, even subjective views.
- ▶ Can do scenario tests by twisting the parameters.
 - ▶ sensitivity analysis of parameters
 - ▶ give more weightings to more recent data
- ▶ Can work with non-linear assets, e.g., options.
- ▶ Can track path-dependence.



Monte Carlo Simulation Disadvantages

- ▶ **Slow.**
 - ▶ To increase the precision by a factor of 10, we must make 100 times more simulations.
 - ▶ Various variance reduction techniques apply.
 - ▶ antithetic variates
 - ▶ control variates
 - ▶ importance sampling
 - ▶ stratified sampling
- ▶ Difficult to build a (high) multi-dimensional joint distribution from data.

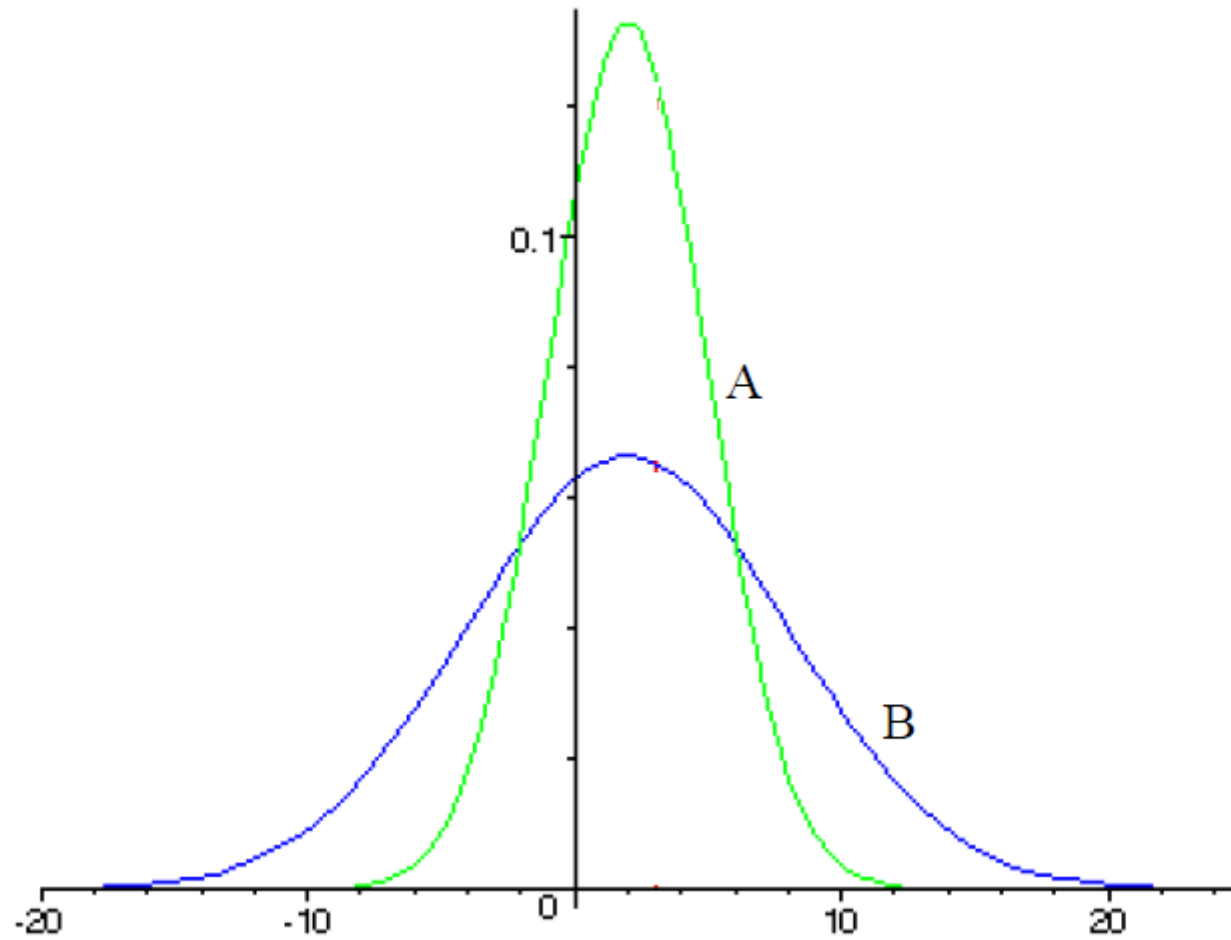


100-Year Market Crash

- ▶ How do we incorporate rare events into our returns distributions, hence enhanced risk management?
- ▶ Statistics works very well when you have a large amount of data.
- ▶ How do we analyze for (very) small samples?



Fat Tails

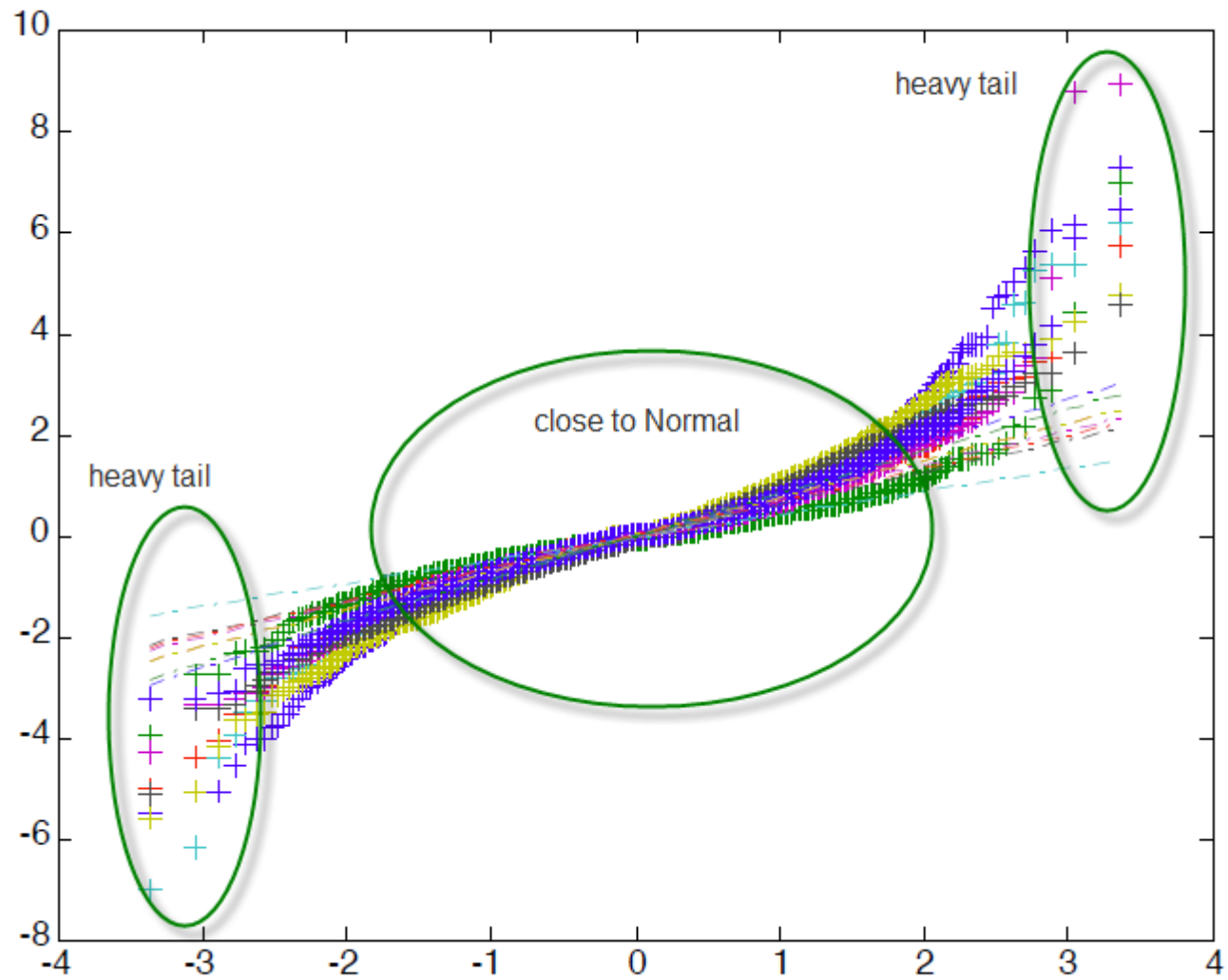


QQ

- ▶ A QQ plots display the quintiles of the sample data against those of a standard normal distribution.
- ▶ This is the first diagnostic tool in determining whether the data have fat tails.



QQ Plot



Asymptotic Properties

- ▶ The (normalized) mean of a the sample mean of a large population is normally distributed, *regardless of the generating distribution.*
- ▶ What about the sample maximum?



Intuition

- ▶ Let X_1, \dots, X_n be i.i.d. with distribution $F(x)$.
- ▶ Let the sample maxima be $M_n = X_{(n)} = \max_i X_i$.
- ▶ $P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x)$
- ▶ $= \prod_{i=1}^n P(X_i \leq x) = F^n(x)$
- ▶ What is $\lim_{n \rightarrow \infty} F^n(x)$?



Convergence

- ▶ Suppose we can scale the maximums $\{c_n\}$ and change the locations (means) $\{d_n\}$.
- ▶ There may exist non-negative sequences of these such that
 - ▶ $c_n^{-1}(M_n - d_n) \rightarrow Y$, Y is not a point
 - ▶ $H(x) = \lim_{n \rightarrow \infty} P(c_n^{-1}(M_n - d_n) \leq x)$
 - ▶ $= \lim_{n \rightarrow \infty} P(M_n \leq c_n x + d_n)$
 - ▶ $= \lim_{n \rightarrow \infty} F^n(c_n x + d_n)$



Example 1 (Gumbel)

- ▶ $F(x) = 1 - e^{-\lambda x}, x > 0.$
- ▶ Let $c_n = \lambda^{-1}, d_n = \lambda^{-1} \log n.$
- ▶ $P(\lambda(M_n - \lambda^{-1} \log n) \leq x)$
- ▶ $= P(M_n \leq \lambda^{-1}(x + \log n))$
- ▶ $= (1 - e^{-(x+\log n)})^n$
- ▶ $= \left(1 - \frac{e^{-x}}{n}\right)^n$
- ▶ $\rightarrow e^{-e^{-x}} = e^{-e^{-x}} 1_{\{x>0\}}$



Example 2 (Fre'chet)

- ▶ $F(x) = 1 - \frac{\theta^\alpha}{(\theta+x)^\alpha} = 1 - \frac{1}{\left(1+\frac{x}{\theta}\right)^\alpha}, x > 0.$
 - ▶ Let $c_n = \theta n^{\frac{1}{\alpha}}, d_n = 0.$
 - ▶ $P(\vartheta^{-1}n^{-1/\alpha}M_n \leq x)$
 - ▶ $= P(M_n \leq \vartheta n^{1/\alpha}x)$
 - ▶ $= \left(1 - \frac{1}{(1+n^{1/\alpha}x)^\alpha}\right)^n \sim \left(1 - \frac{1}{(n^{1/\alpha}x)^\alpha}\right)^n$
 - ▶ $= \left(1 - \frac{x^{-\alpha}}{n}\right)^n$
 - ▶ $\rightarrow e^{-x^{-\alpha}} \mathbf{1}_{\{x>0\}}$
-



Fisher-Tippett Theorem

- ▶ It turns out that H can take only one of the three possible forms.
- ▶ Fréchet
 - ▶ $\Phi_\alpha(x) = e^{-x^{-\alpha}} 1_{\{x>0\}}$
- ▶ Gumbel
 - ▶ $\Lambda(x) = e^{-e^{-x}} 1_{\{x>0\}}$
- ▶ Weibull
 - ▶ $\Psi_\alpha(x) = e^{-(-x)^\alpha} 1_{\{x<0\}}$



Maximum Domain of Attraction

- ▶ **Fre'chet**

- ▶ Fat tails
- ▶ E.g., Pareto, Cauchy, student t,

- ▶ **Gumbel**

- ▶ The tail decay exponentially with all finite moments.
- ▶ E.g., normal, log normal, gamma, exponential

- ▶ **Weibull**

- ▶ Thin tailed distributions with finite upper endpoints, hence bounded maximums.
- ▶ E.g., uniform distribution



Why Fre'chet?

- ▶ Since we care about fat tailed distributions for financial asset returns, we rule out Gumbel.
- ▶ Since financial asset returns are theoretically unbounded, we rule out Weibull.
- ▶ So, we are left with Fre'chet, the most common MDA used in modeling extreme risk.



Frechet Shape Parameter

- ▶ α is the shape parameter.
- ▶ Moments of order r greater than α are infinite.
- ▶ Moments of order r smaller than α are finite.
 - ▶ Student t distribution has $\alpha \geq 2$. So its mean and variance are well defined.



Frechet MDA Theorem

- ▶ $F \in \text{MDA } H$, H Frechet if and only if
- ▶ the complement cdf $\bar{F}(x) = x^{-\alpha}L(x)$
- ▶ L is slowly varying function
 - ▶ $\lim_{x \rightarrow \infty} \frac{L(tx)}{L(x)} = 1, t > 0$
- ▶ This restricts the maximum domain of attraction of the Frechet distribution quite a lot, it consists only of what we would call heavy tailed distributions.



Generalized Extreme Value Distribution (GEV)

▶ $H_\tau(x) = e^{-(1+\tau x)^{-\frac{1}{\tau}}}, \tau \neq 0$

▶ $H_\tau(x) = e^{-e^{-x}}, \tau = 0$

▶ $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^{-n} = e^{-x}$

▶ tail index $\tau = \frac{1}{\alpha}$

▶ Frechet: $\tau > 0$

▶ Gumbel: $\tau = 0$

▶ Weibull: $\tau < 0$



Generalized Pareto Distribution

- ▶ $G_\tau(x) = 1 - (1 + \tau x)^{-\frac{1}{\tau}}$
- ▶ $G_0(x) = 1 - e^{-x}$
 - ▶ simply an exponential distribution
- ▶ Let $Y = \beta X$, $X \sim G_\tau$.
- ▶ $G_{\tau,\beta} = 1 - \left(1 + \tau \frac{y}{\beta}\right)^{-\frac{1}{\tau}}$
- ▶ $G_{0,\beta} = 1 - e^{-\frac{y}{\beta}}$



The Excess Function

▶ Let u be a tail cutoff threshold.

▶ The excess function is defined as:

▶ $F_u(x) = 1 - \bar{F}_u(x)$

▶ $\bar{F}_u(x) = P(X - u > x | X > u) = \frac{P(X > u+x)}{P(X > u)} = \frac{\bar{F}(x+u)}{\bar{F}(u)}$



Asymptotic Property of Excess Function

- ▶ Let $x_F = \inf\{x: F(x) = 1\}$.
- ▶ For each τ , $F \in \text{MDA}(H_\tau)$, if and only if
 - ▶ $\lim_{u \rightarrow x_F^-} \sup_{0 < x < x_F - u} |F_u(x) - G_{\tau, \beta(u)}(x)| = 0$
- ▶ If $x_F = \infty$, we have
 - ▶ $\lim_{u \rightarrow \infty} \sup_x |F_u(x) - G_{\tau, \beta(u)}(x)| = 0$
- ▶ Applications: to determine τ , u , etc.



Tail Index Estimation by Quantiles

- ▶ Hill, 1975
- ▶ Pickands, 1975
- ▶ Dekkers and DeHaan, 1990

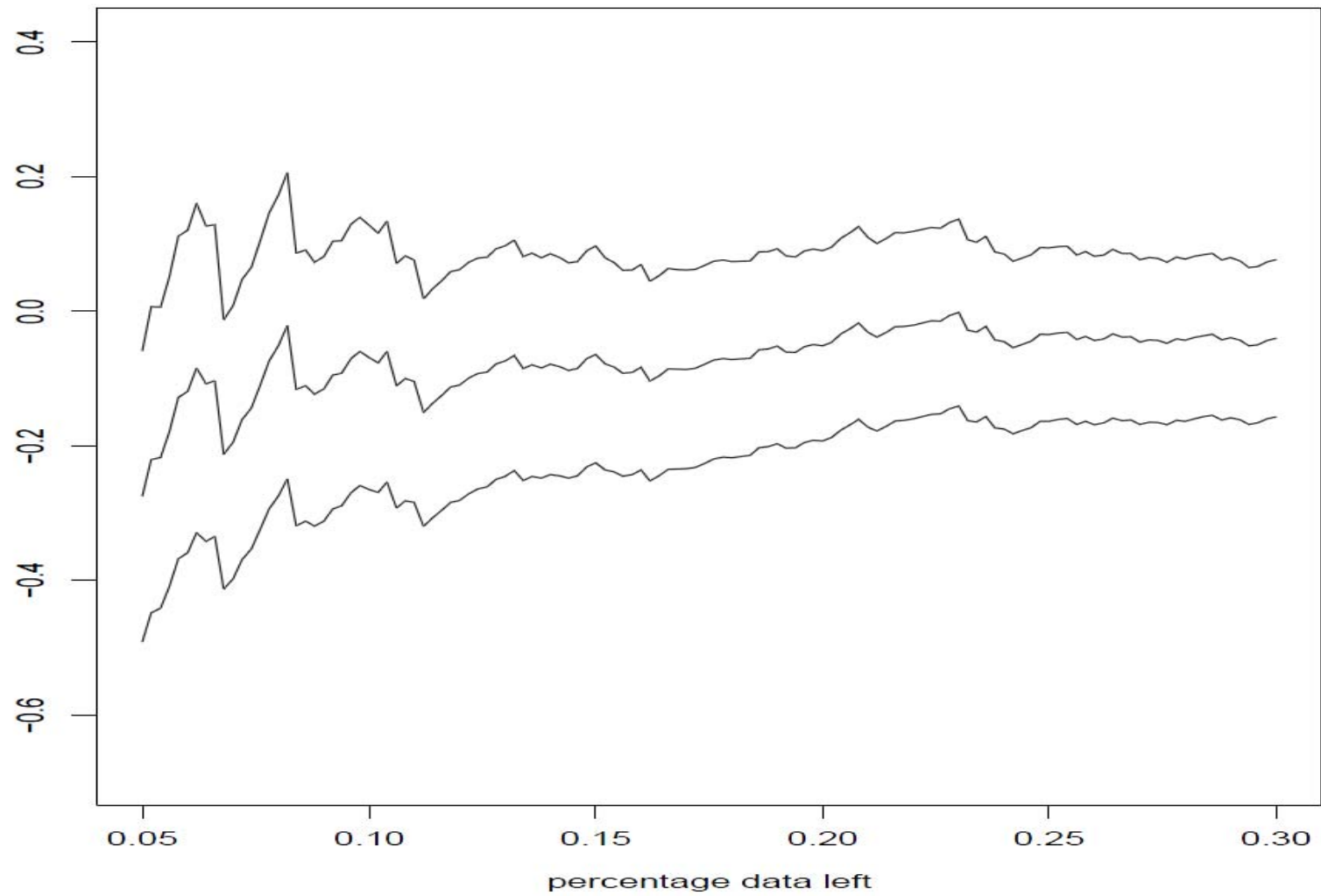


Hill Estimator

- ▶ $\tau_{n,m}^H = \frac{1}{m-1} \sum_{i=1}^{m-1} (\ln X^*_i - \ln X^*_{n-m,n})$
- ▶ X^* : the order statistics of observations
- ▶ m : the number of observations in the (left) tail
- ▶ Mason (1982) shows that $\tau_{n,m}^H$ is a consistent estimator, hence convergence to the true value.
- ▶ Pictet, Dacorogna, and Muller (1996) show that in finite samples the expectation of the Hill estimator is biased.
- ▶ In general, bigger (smaller) m gives more (less) biased estimator but smaller (bigger) variance.



POT Plot



Pickands Estimator

▶ $\tau_{n,m}^P = \frac{\ln(X_m^* - X_{2m}^*) / (X_{2m}^* - X_{4m}^*)}{\ln 2}$



Dekkers and DeHaan Estimator

- ▶ $\tau_{n,m}^D = \tau_{n,m}^H + 1 - \frac{1}{2} \left(1 - \frac{(\tau_{n,m}^H)^2}{\tau_{n,m}^{H2}} \right)^{-1}$
- ▶ $\tau_{n,m}^{H2} = \frac{1}{m-1} \sum_{i=1}^{m-1} (\ln X_i^* - \ln X_m^*)^2$



VaR using EVT

- ▶ For a given probability $q > F(u)$ the VaR estimate is calculated by inverting the excess function. We have:

- ▶
$$\widehat{\text{VaR}}_q = u + \frac{\hat{\beta}}{\hat{\tau}} \left(\left(\frac{n}{m} (1 - q) \right)^{-\hat{\tau}} - 1 \right)$$

- ▶ Confidence interval can be computed using profile likelihood.

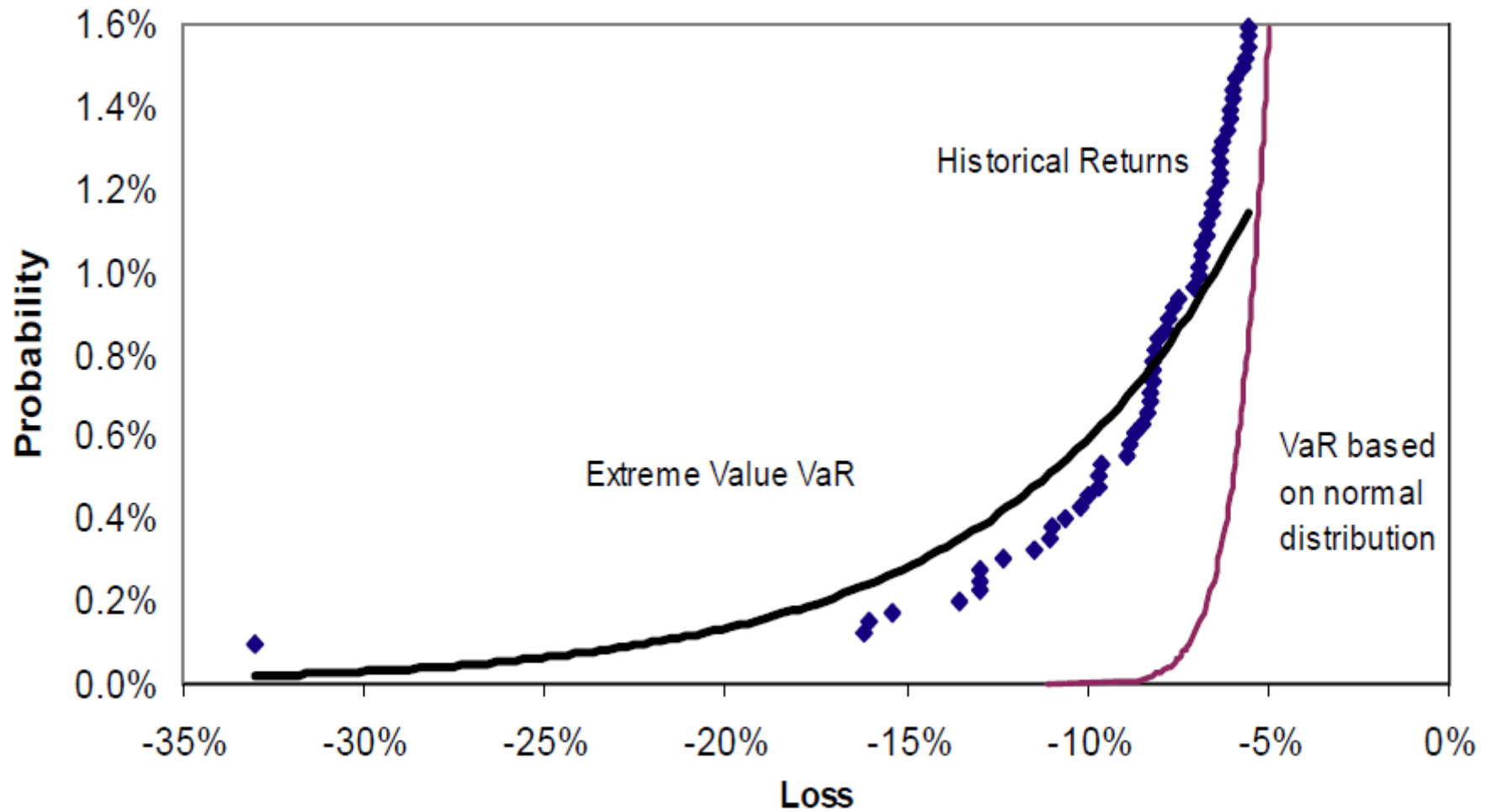


ES using EVT

$$\triangleright \widehat{\text{ES}}_q = \frac{\widehat{\text{VaR}}_q}{1-\hat{\tau}} + \frac{\hat{\beta} - \hat{\tau}u}{1-\hat{\tau}}$$



VaR Comparison



► http://www.fea.com/resources/pdf/a_evt_1.pdf