

Quantile Estimation – sample quantiles versus kernel quantiles



PillarOne Conference, St. Gallen, September 10, 2010

Jessika Walter, PhD, Intuitive Collaboration

Markus Stricker, PhD, Intuitive Collaboration



- VaR, TVaR/CVaR is one of the most frequently used key figures in risk management
 - Solvency II: VaR 99.5%
 - SST: TVaR 99%
- Most partial-internal and internal risk management models are Monte-Carlo simulation models
 - VaR has to be estimated from simulated samples

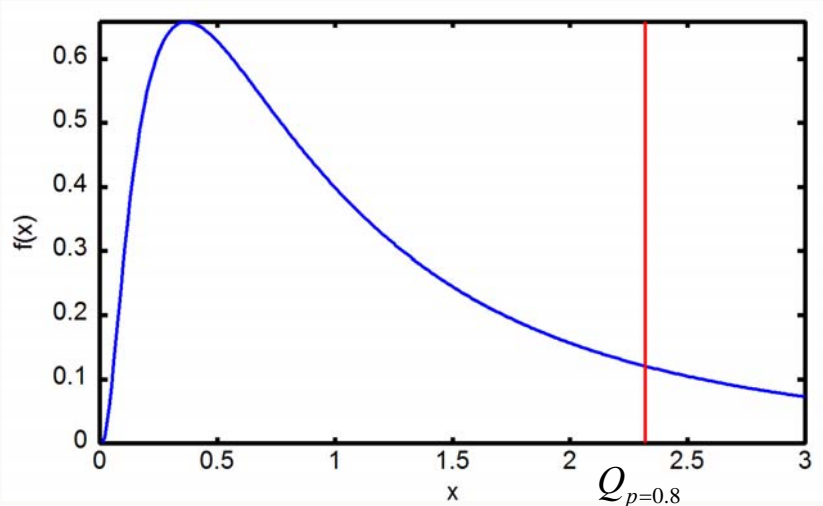
Everybody talks about modeling, but

- Are the key figures derived from the samples stable?
- Are they biased?
- How do they converge?

Ultimate goal

- Derive confidence intervals for VaR
- Derive criteria for required sample sizes, i.e. required number of iterations

- Sample quantile estimators:
 - Definitions
 - Asymptotic considerations
- Kernel quantile estimators:
 - Kernel density estimation
 - Selecting the appropriate bandwidth
 - Sample quantiles versus kernel estimators technique
- Conclusion and outlook



Probability function $f(x)$:

$$\int f(x)dx = 1, f \geq 0$$

Cumulative distribution $F(x)$:

$$F(x) = \int_{-\infty}^x f(y)dy$$

Interpretation: $X \propto f(x)dx \Rightarrow \mathbf{P}[X \in (a,b)] = F(b) - F(a)$

Quantile function Q_p is defined as the generalized inverse of F

$$Q_p = \inf\{x \in \mathbf{R} : \mathbf{P}[X \leq x] \geq p\}$$

- Assumption:**
1. X_1, X_2, \dots, X_N i.i.d. random variables according to $X \propto f(x)dx$
 2. Let x_1, \dots, x_N denote a realization of $\{X_n\}_n$ ordered by increasing size (order statistics)

Prop: Under certain assumptions, as $N \rightarrow \infty$ the following is satisfied:

1. The stationary density $\hat{f}_N(x)$ of $\{x_n\}_n$ converges to $f(x)$ in $L^1(dx)$.
2. If \hat{Q}_p denotes a sample quantile, then \hat{Q}_p converges pointwise to the real quantile Q_p .

Popular choice: $\hat{Q}_p := x_{[Np]+1}$

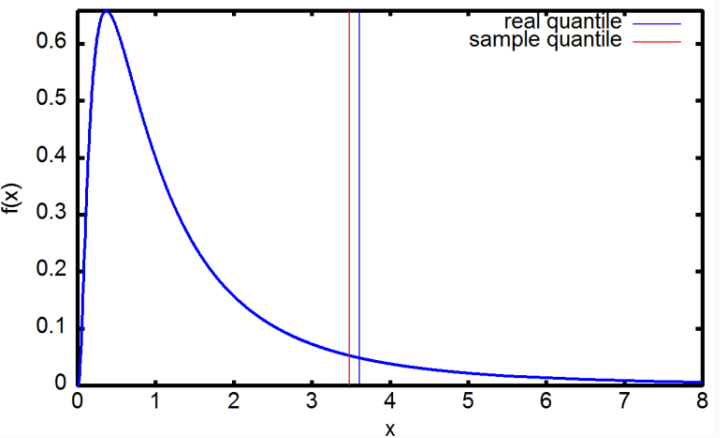
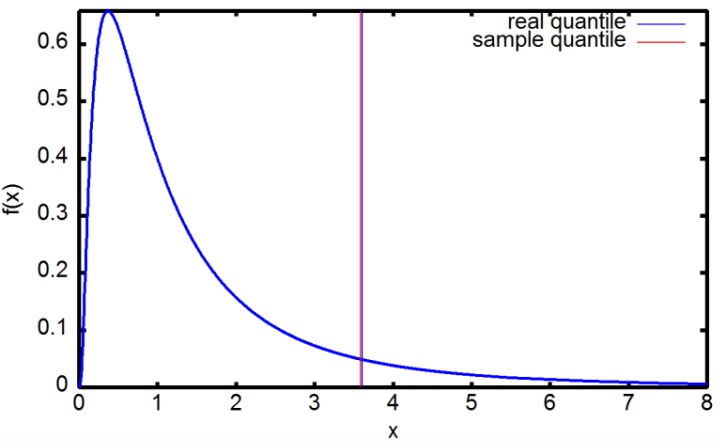
Note: There is not a unique definition for sample quantiles and magnitude of bias highly depends on the specific choice and the underlying distribution.

Sample quantiles are biased, but asymptotically unbiased:

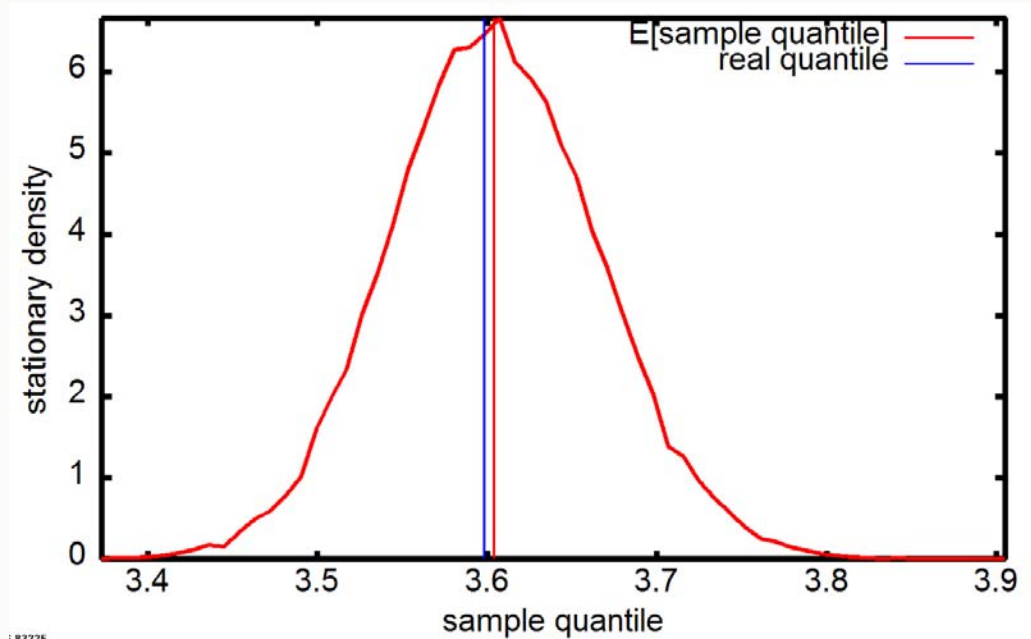
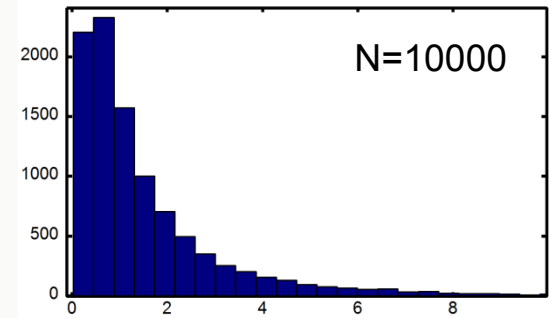
$$\mathbf{E}[\hat{Q}_{p,N}] \neq Q_p \text{ for all } N, \quad \lim_{N \rightarrow \infty} \mathbf{E}[\hat{Q}_{p,N}] = Q_p$$

Illustrative Example: variability of sample quantile

p=0.9



$$\hat{Q}_p := x_{[Np]+1}$$

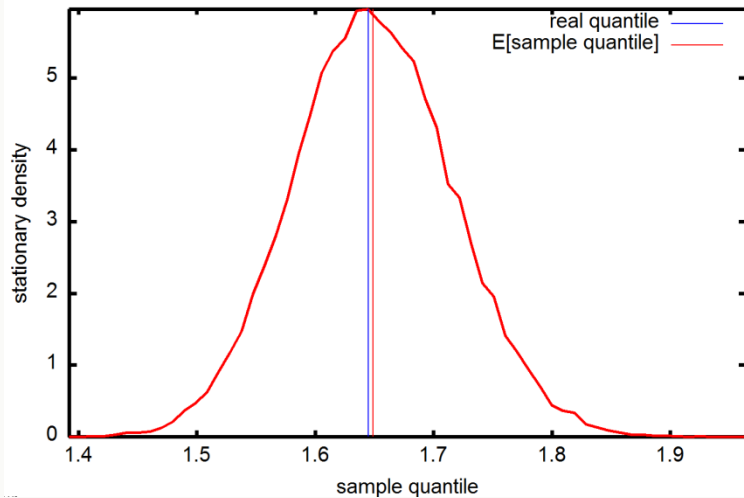


$f(x)$ standard lognormal

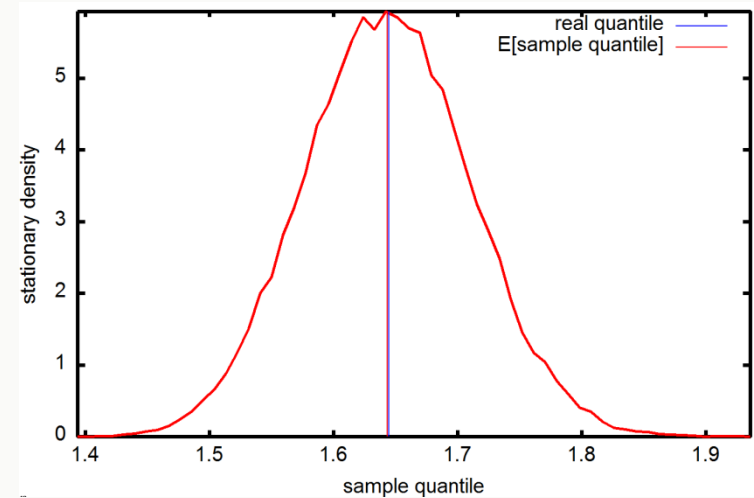
Distribution of sample quantile

$f(x)$ is standard gaussian

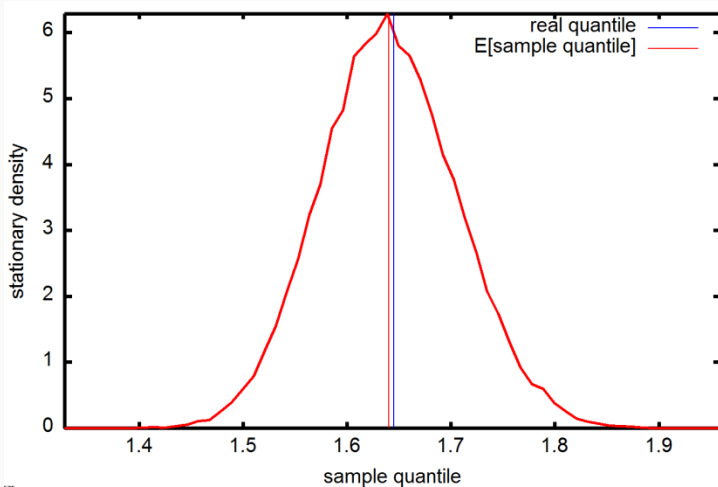
N=1000



N=1010



N=1022



$$\hat{Q}_p := x_{[Np]+1}$$

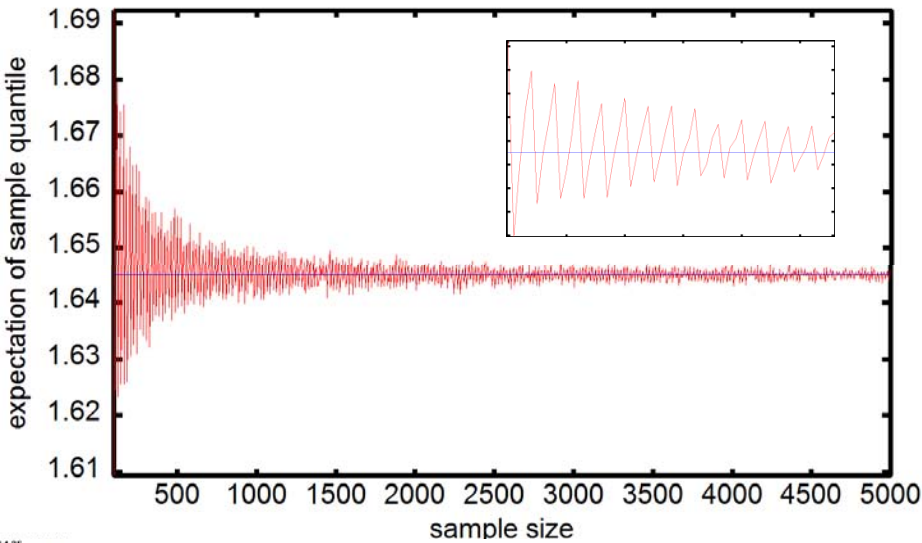
here: $p=0.95$

N=1000: $Np=950$

N=1010: $Np=959.50$

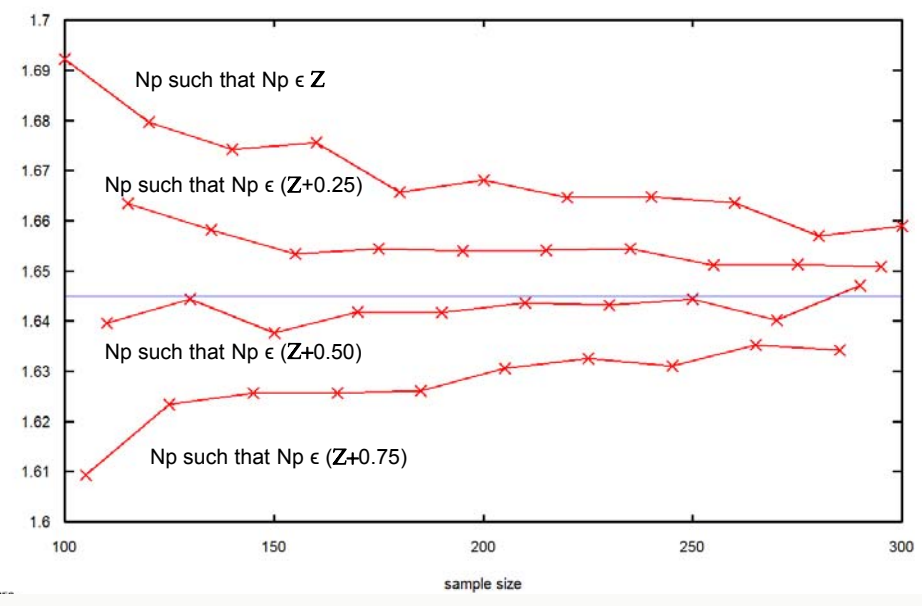
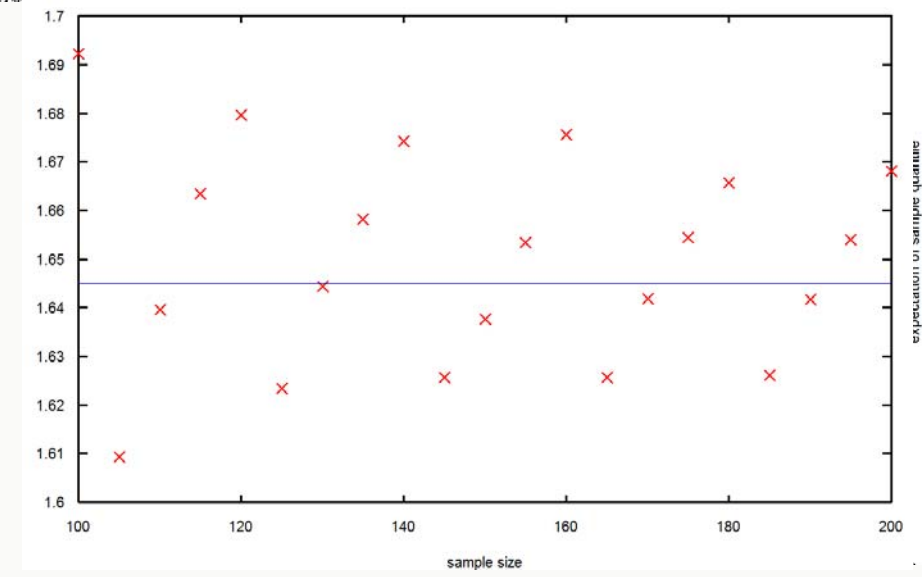
N=1022: $Np=970.90$

Asymptotic behaviour of expectation

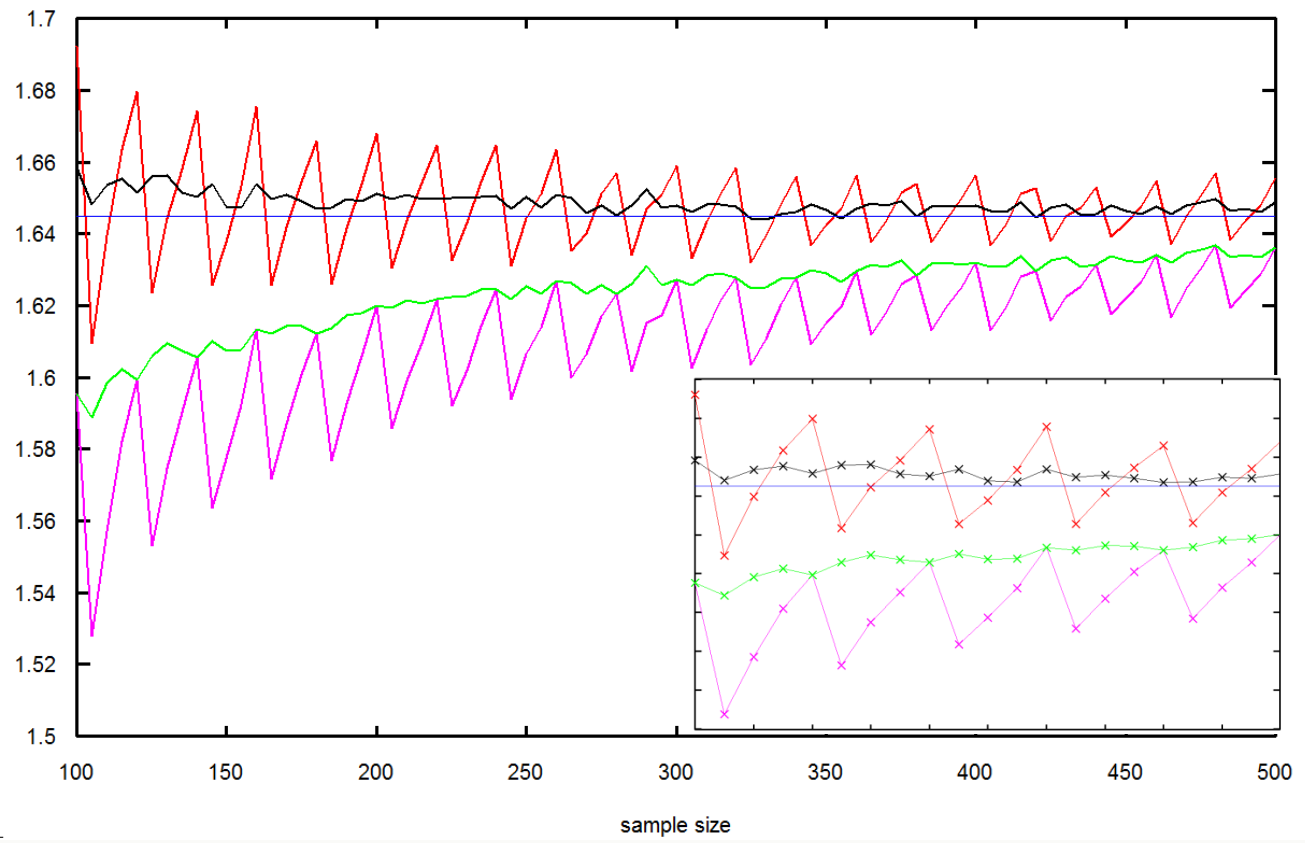


$N=[100:5:5000]$, $p=0.95$

$$\hat{Q}_p := x_{[Np]+1}$$



Comparison of sample quantiles (L-estimators)



$N=[100:5:500], \quad p=0.95$

$$\hat{Q}_U(p) := x_{\lfloor Np \rfloor + 1}$$

$$\hat{Q}_L(p) := x_{\lfloor Np \rfloor}$$

linear interpolation between
 $\hat{Q}_U := x_{\lfloor Np \rfloor}$ and $\hat{Q}_L := x_{\lfloor Np \rfloor + 1}$

Estimator of Hyndman and Fan lies between low side and high side

$$\hat{Q}_{HF}(p) := (1 - \gamma)x_g + \gamma x_{g+1}$$

$$g = \lfloor (N + 1/3)p + 1/3 \rfloor, \quad \gamma = (N + 1/3)p + 1/3 - g$$

L-Estimator of Hyndman and Fan:

$$\hat{Q}_{HF}(p) := (1 - \gamma)x_g + \gamma x_{g+1}$$

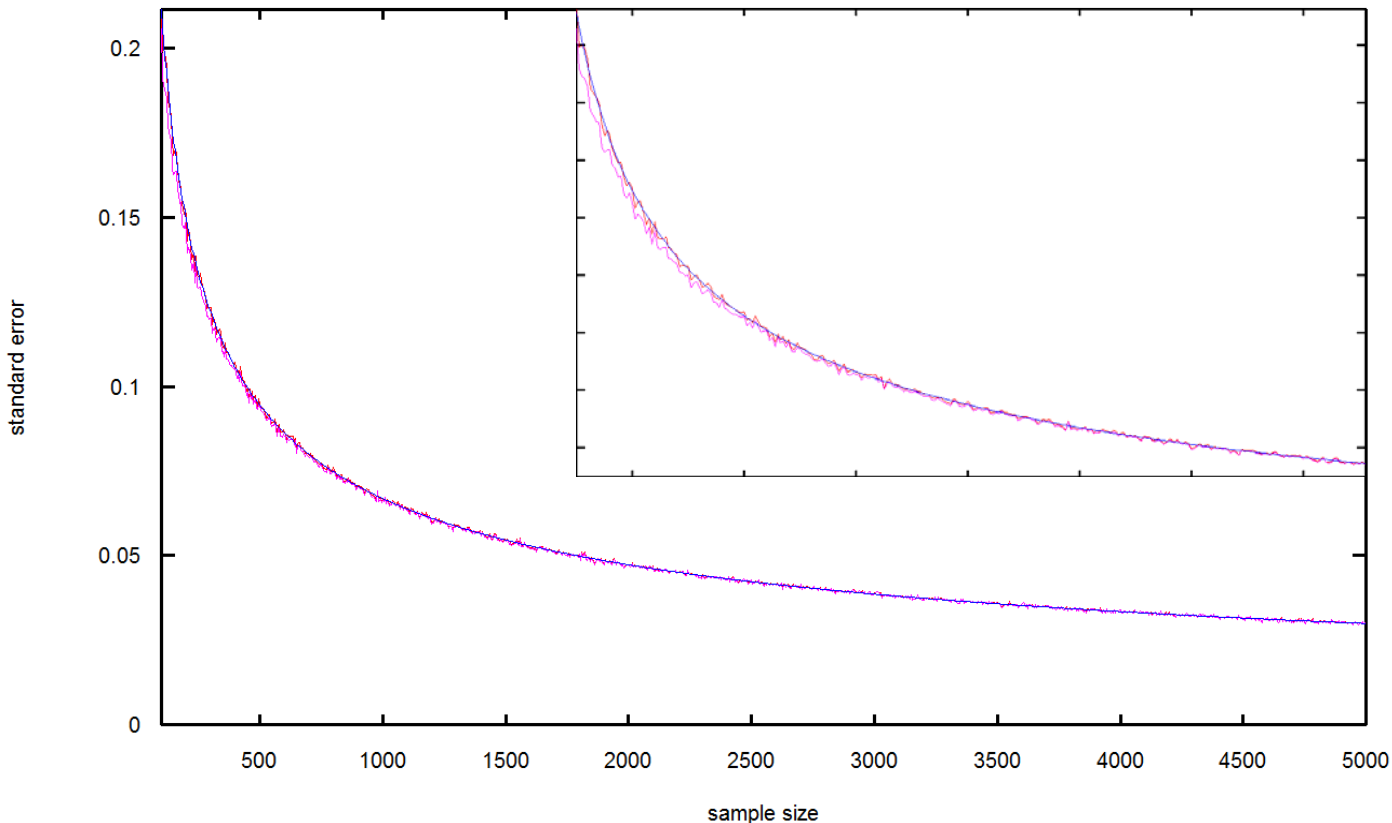
$$g = \lfloor (n+1/3)p + 1/3 \rfloor, \quad \gamma = (n+1/3)p + 1/3 - g$$

Example with $p=0.95$: Then $p/3+1/3 = 0.65$

1. $N = 100 \Rightarrow Np = 95 \Rightarrow \hat{Q}_{HF} = 0.35x_{95} + 0.65x_{96}$
2. $N = 105 \Rightarrow Np = 99.75 \Rightarrow \hat{Q}_{HF} = 0.60x_{100} + 0.40x_{101}$
3. $N = 110 \Rightarrow Np = 104.5 \Rightarrow \hat{Q}_{HF} = 0.85x_{105} + 0.15x_{106}$
4. $N = 115 \Rightarrow Np = 109.25 \Rightarrow \hat{Q}_{HF} = 0.10x_{109} + 0.90x_{110}$

Standard error of sample quantiles

$N=[100:5:5000]$, $p=0.95$



$$\hat{Q}_p := x_{[Np]+1}$$

$$\hat{Q}_p := x_{[Np]}$$

Bahadur representation: $\hat{Q}_p \xrightarrow{d} N(Q_p, \frac{\sigma^2}{N})$ with $\sigma = \frac{\sqrt{p(1-p)}}{f(Q_p)}$

Let x_1, x_2, \dots, x_N denote a sample of size N from a random variable with density f

Kernel density estimate:

$$\hat{f}_N(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - x_i}{h}\right)$$

Kernel K is nonnegative, bounded, Lipschitz function with

$$\int K(y)dy = 1, \quad \int yK(y)dy = 0, \quad \int y^2 K(y)dy = \sigma^2 > 0$$

Gaussian Kernel :

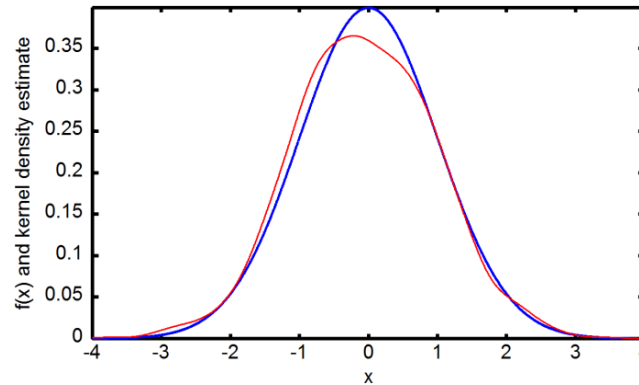
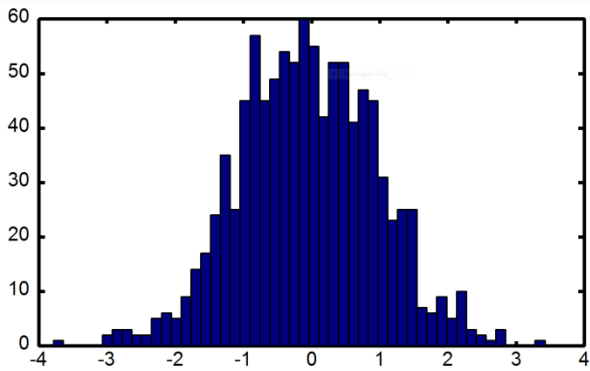
$$K(y) = \frac{1}{\sqrt{2\pi}} \exp(-y^2 / 2)$$

Convergence assumption :

$$\lim_{N \rightarrow \infty} h = 0, \quad \lim_{N \rightarrow \infty} (Nh) = \infty$$

What is the optimal bandwidth h ?

$N=1000, f(x)=N(0,1)$



slight oversmoothing
for $h=0.27$

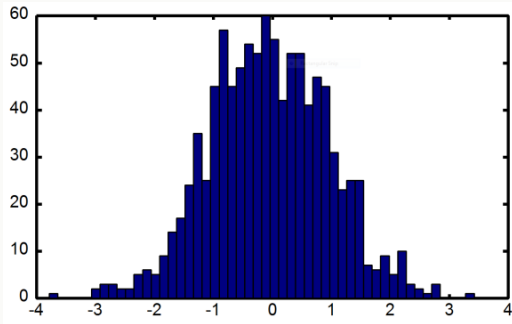
$$h_{opt} = h_{opt}(I(K^2), I((f'')^2), N) = O(N^{-1/5})$$

The optimal bandwidth minimizes the asymptotic mean integrated squared error

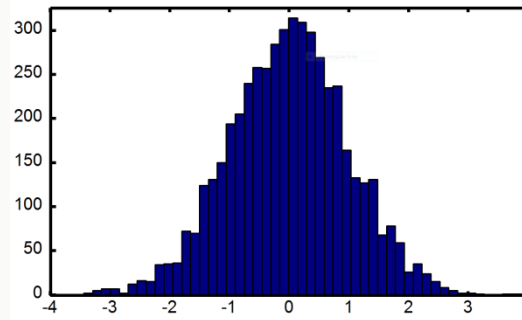
$$\text{MISE} = \int \text{Bias}(\hat{f}_h(y))^2 dy + \int \text{Var}(\hat{f}_h(y)) dy$$

Illustrative Example: Optimal global bandwidth for standard Gaussian

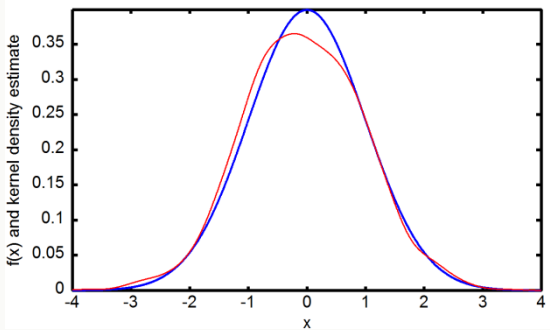
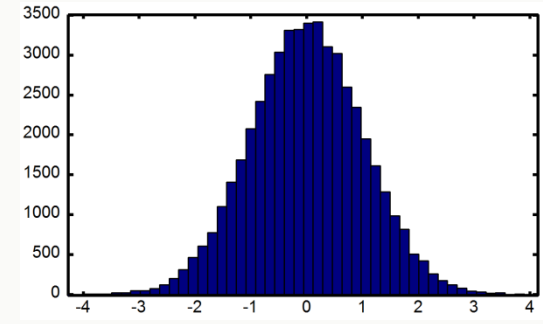
N=1000



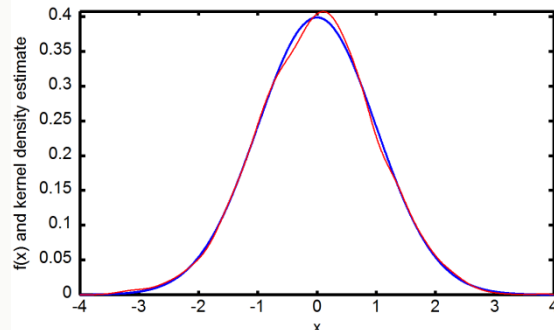
N=5000



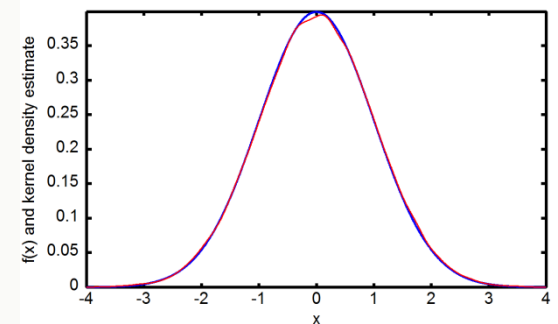
N=50000



$h=0.27$



$h=0.19$



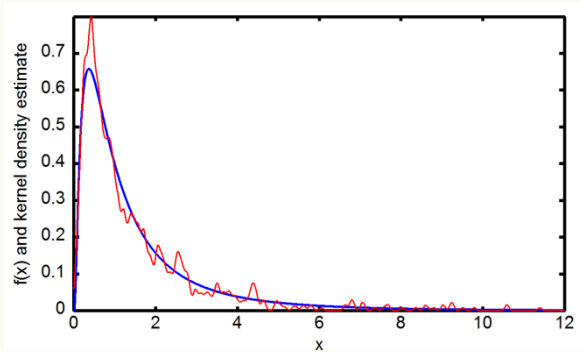
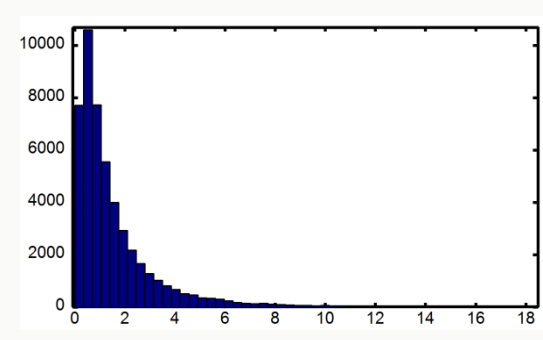
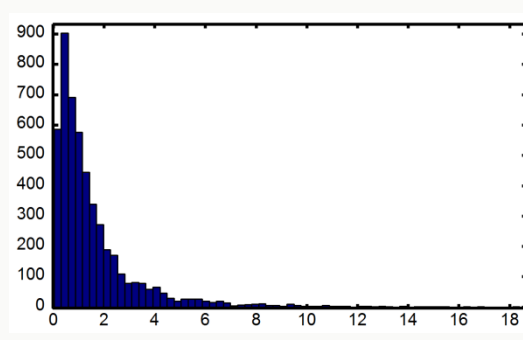
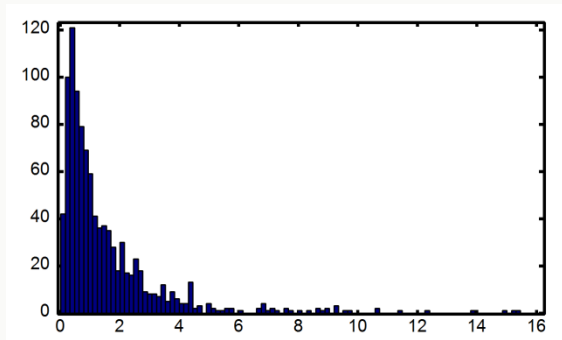
$h=0.12$

Illustrative Example: Optimal global bandwidth for standard lognormal distribution

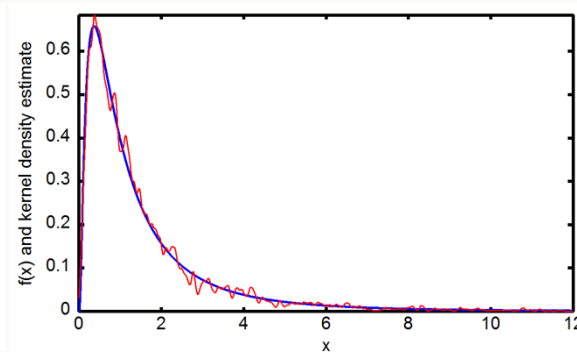
N=1000

N=5000

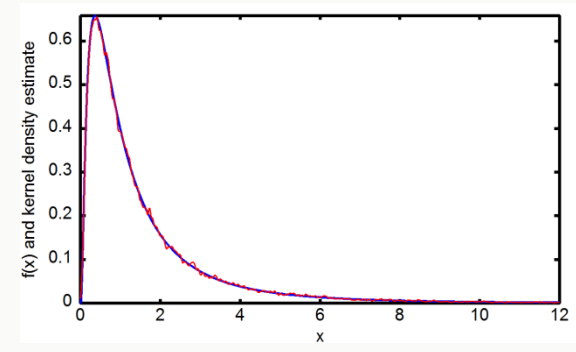
N=50000



h=0.05

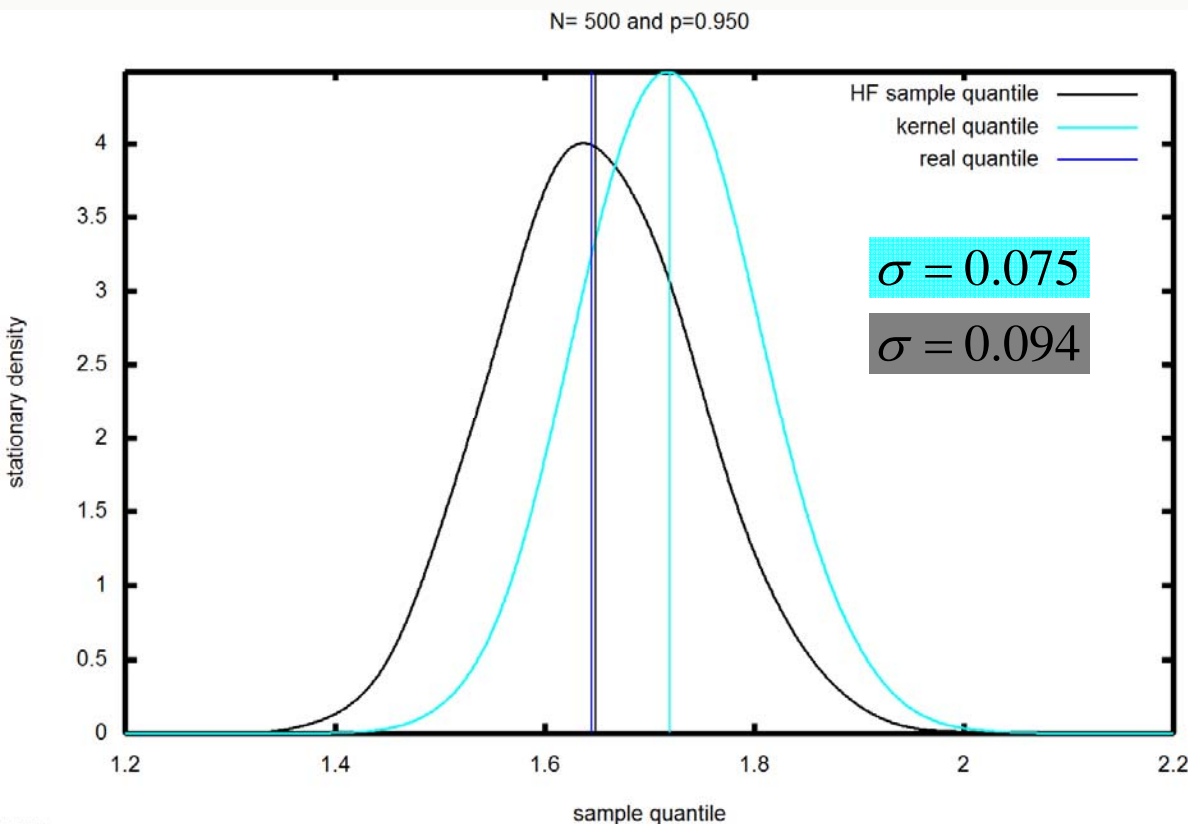


h=0.036



h=0.023

Distribution of kernel quantile estimator with global bandwidth



Bad news:

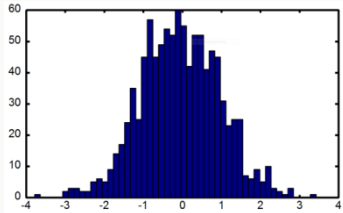
Optimal bandwidth for minimizing error in Lebesgue space is not a good candidate for quantile estimation

Good news: There exists a locally adapted optimal bandwidth for quantile estimation

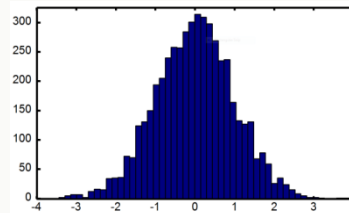
$$h_p = h_p(f(Q_p), f'(Q_p), N) = O(N^{-1/3})$$

Illustrative Example: Optimal local bandwidth for $p=0.9$ applied to standard Gaussian sample

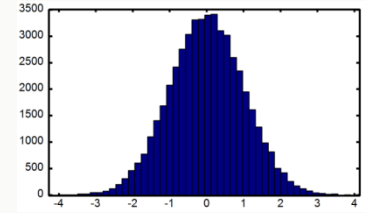
N=1000



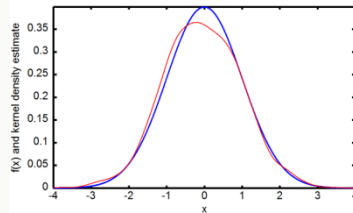
N=5000



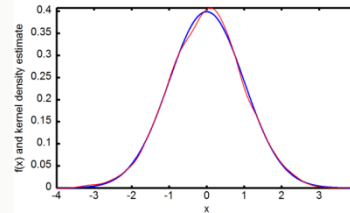
N=50000



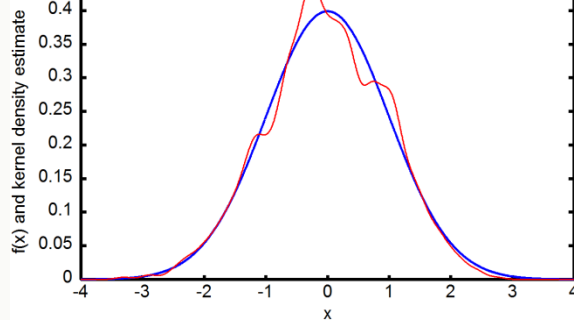
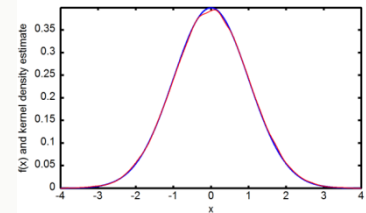
$h_{opt} = 0.27$



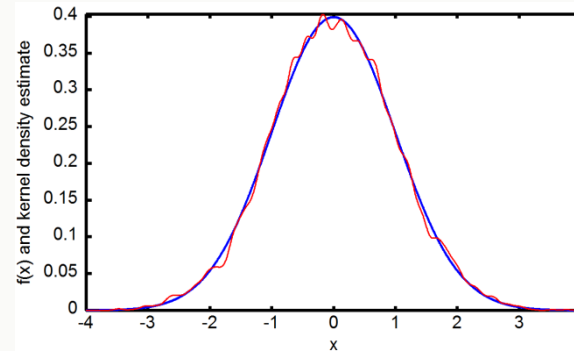
$h_{opt} = 0.19$



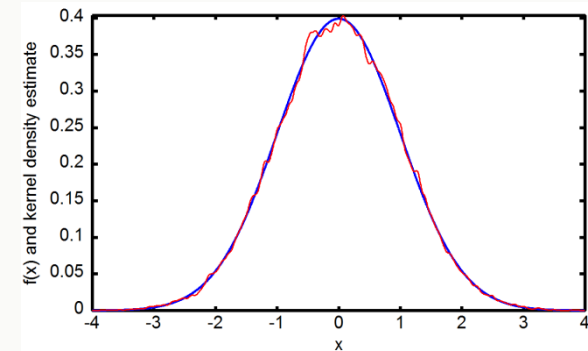
$h_{opt} = 0.12$



$h_p = 0.125$



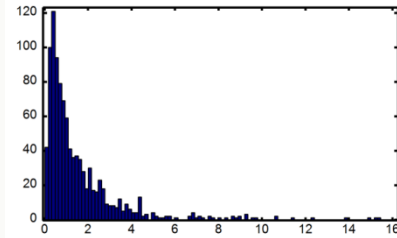
$h_p = 0.073$



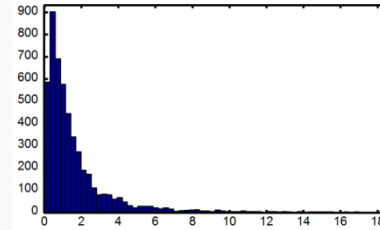
$h_p = 0.034$

Illustrative Example: Optimal local bandwidth for $p=0.9$ applied to standard Lognormal sample

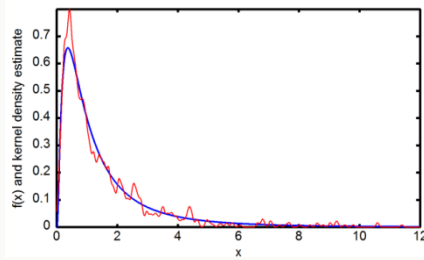
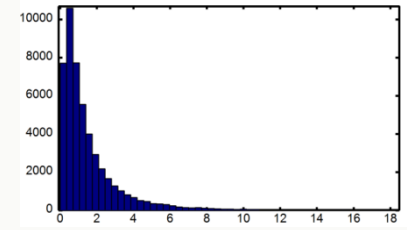
N=1000



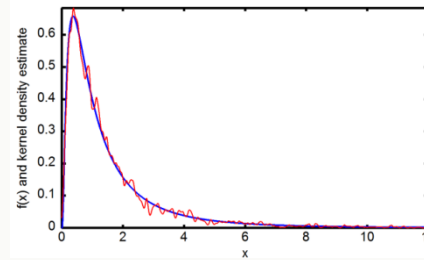
N=5000



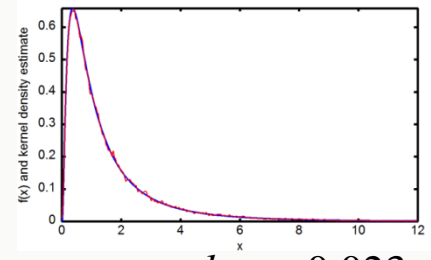
N=50000



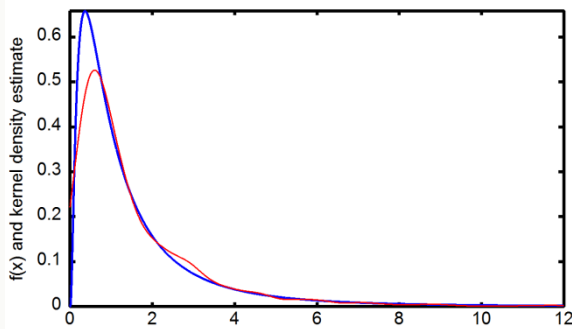
$h_{opt} = 0.05$



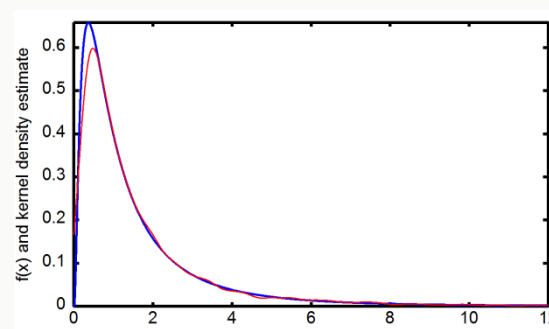
$h_{opt} = 0.036$



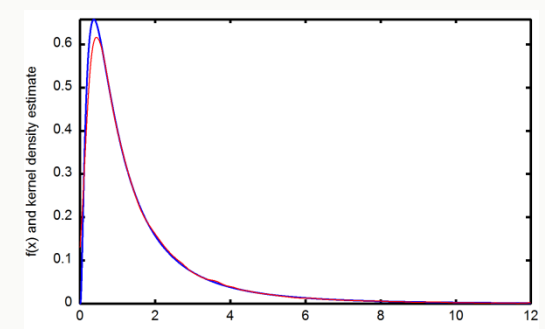
$h_{opt} = 0.023$



$h_p = 0.306$

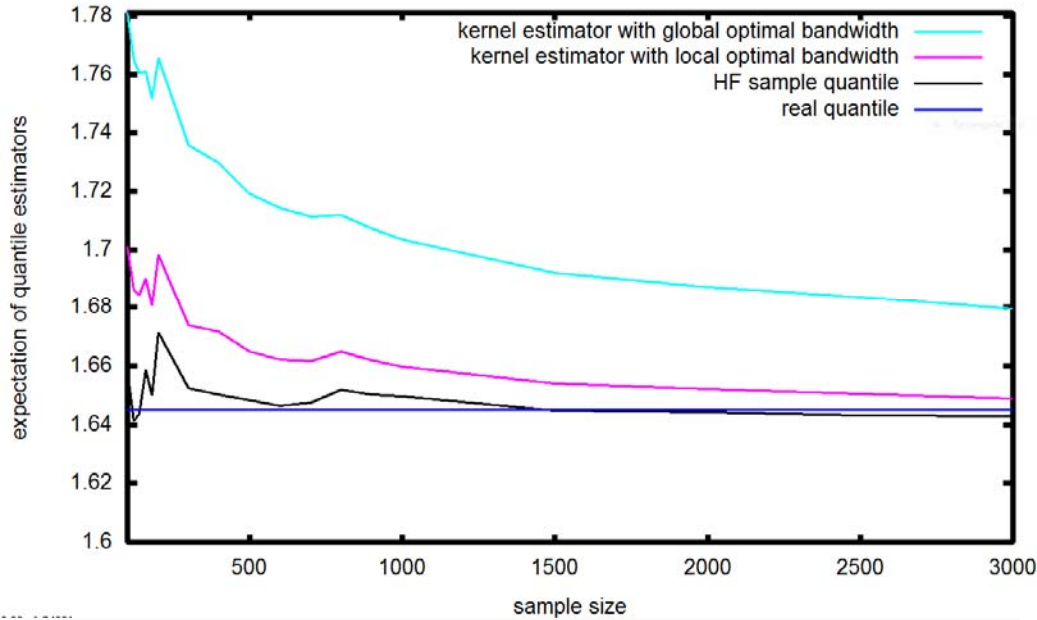


$h_p = 0.180$



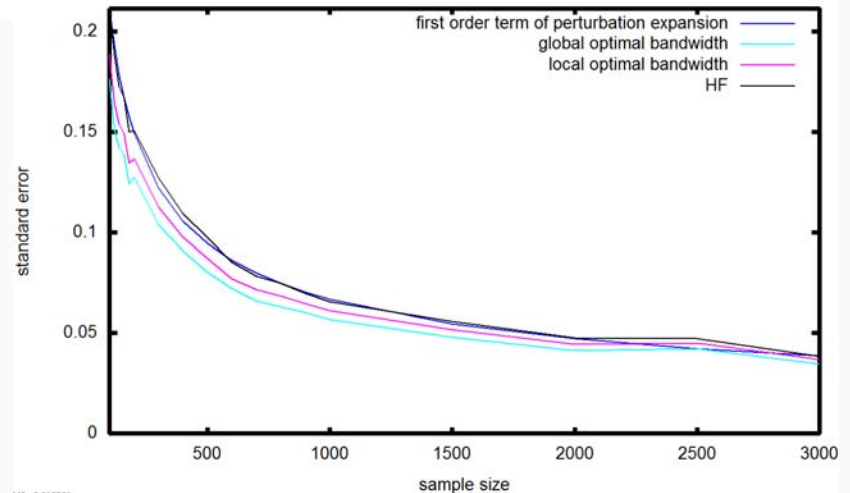
$h_p = 0.130$

Asymptotic behaviour of kernel quantile estimators versus sample quantiles



$f(x)$ standard Gaussian, $p=0.95$

$$\hat{Q}_{kernel} \xrightarrow{d} N(Q_p, \frac{\sigma^2}{N}), \quad \sigma = \frac{\sqrt{p(1-p)}}{f(Q_p)}$$



- Sample quantiles are biased and magnitude and direction depends on the selected estimator
- To reduce systematic errors use L-estimators that use at least two observations
- There are L-estimators that include three or more observations: one of them mixes L-estimation with kernel density estimation

Kernel smoothing the order statistics :

$$\hat{Q}_{p,N} := \sum_{i=1}^N x_i \int_{\frac{i-1}{n}}^{\frac{i}{n}} \frac{1}{h} K\left(\frac{p-x}{h}\right) dx$$

strongly depends on choosing the optimal bandwidth

one of the next tasks !!!

- Derivation of stopping criteria for required sample size:

Let ε and confidence level $1-\alpha$ be given :

N_0 such that for all $N \geq N_0$ we have $\hat{Q}_p \in [Q_{p-\varepsilon}, Q_{p+\varepsilon}]$

Here, the crucial problem is to estimate the function value at the quantile

All of the results will be considered for the implementation in Pillar One!

Contacts



Dr. Markus Stricker; Partner, Managing Director

+41 44 926 10 88; +41 76 423 26 67

markus.stricker@intuitive-collaboration.com

Dr. Jessika Walter; Consultant, Actuarial Consulting

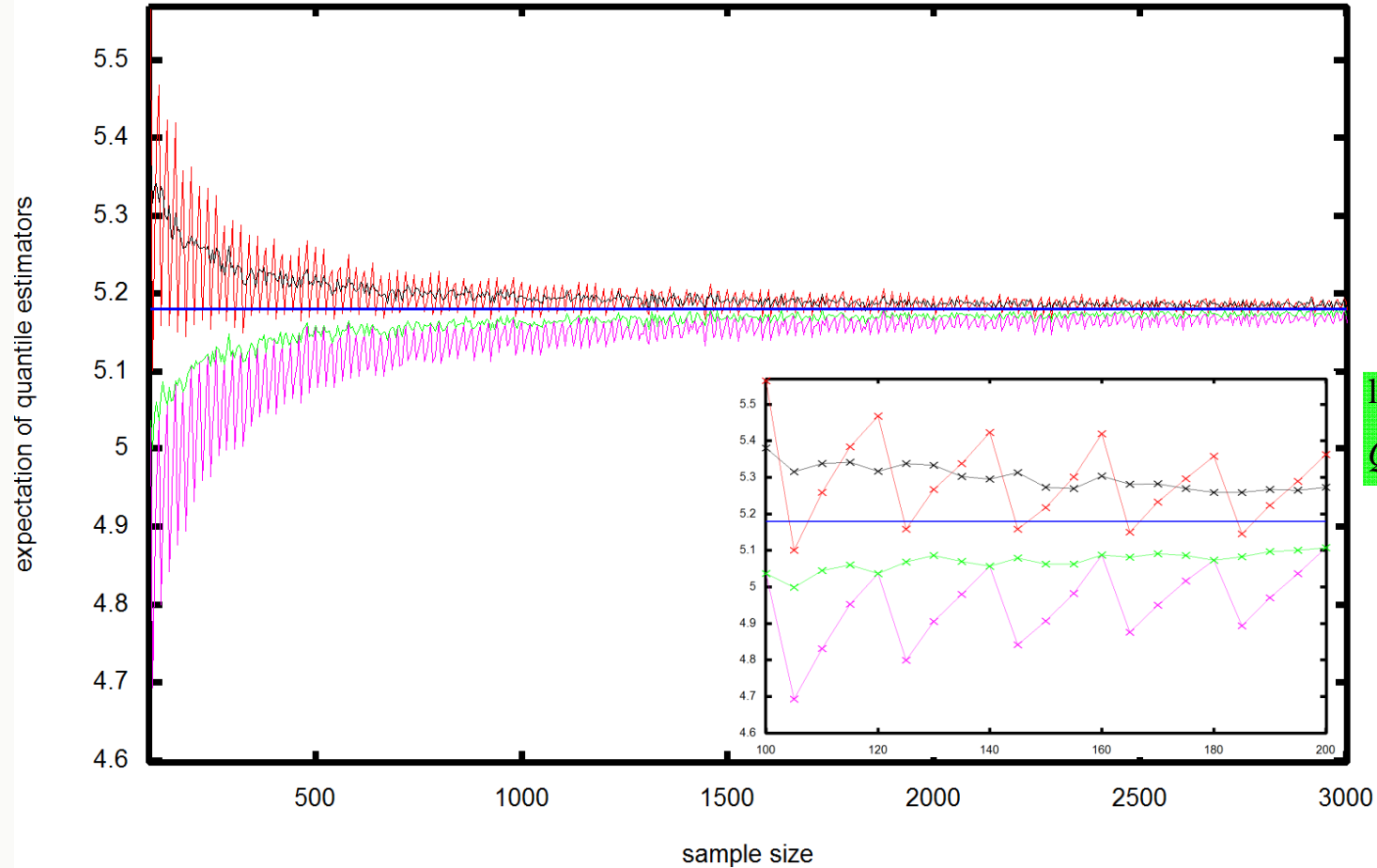
+41 44 926 14 07; +41 79 364 83 62

jessika.walter@intuitive-collaboration.com

Intuitive Collaboration AG; Seestrasse 16; CH-8712 Staefa

Appendix: asymptotic results for lognormal

Asymptotic behaviour of expectations of sample quantiles for standard Lognormal



$$\hat{Q}_U(p) := x_{\lfloor Np \rfloor + 1}$$

$$\hat{Q}_L(p) := x_{\lfloor Np \rfloor}$$

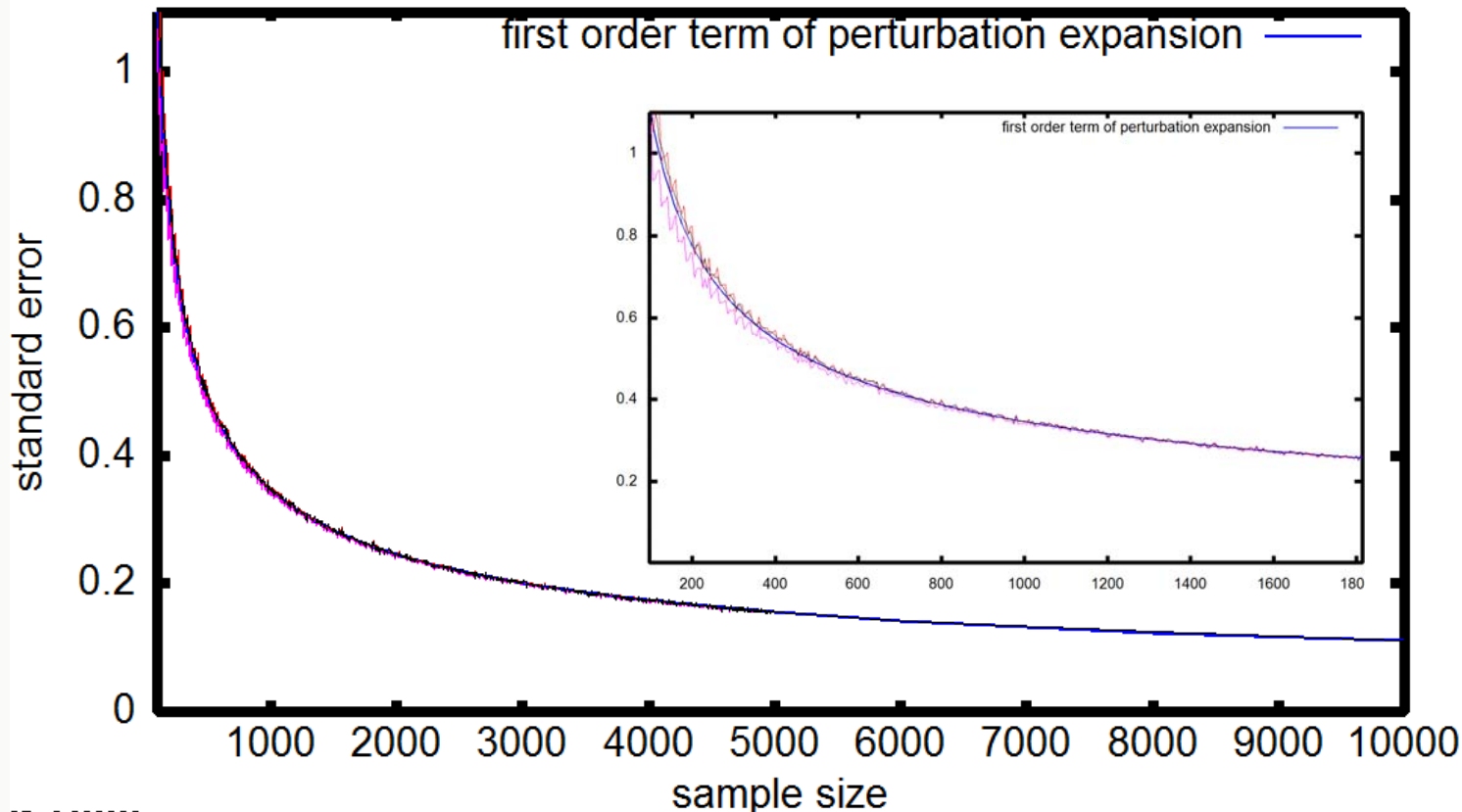
linear interpolation between $\hat{Q}_U := x_{\lfloor Np \rfloor}$ and $\hat{Q}_L := x_{\lfloor Np \rfloor + 1}$

$$\hat{Q}_{HF}(p) := (1 - \gamma)x_g + \gamma x_{g+1}$$

$$g = \lfloor (N+1/3)p + 1/3 \rfloor, \quad \gamma = (N+1/3)p + 1/3 - g$$

Standard error of sample quantiles for standard lognormal

$N=[100:5:5000\ 6000\ 8000\ 10000]$, $p=0.95$

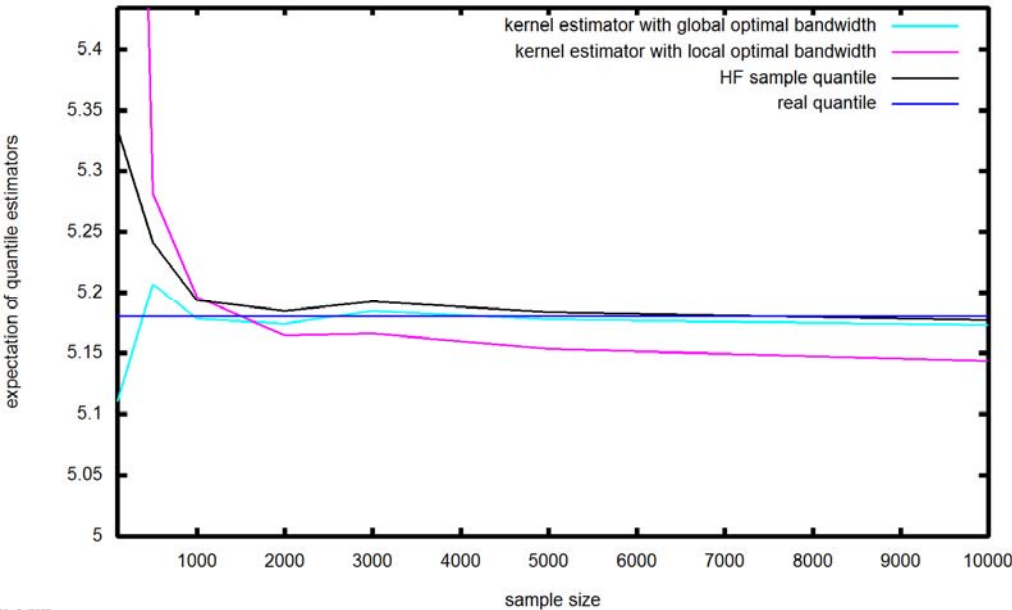


$$\hat{Q}_p := x_{[Np]+1}$$

$$\hat{Q}_p := x_{[Np]}$$

Bahadur representation: $\hat{Q}_p \xrightarrow{d} N(Q_p, \frac{\sigma^2}{N})$ with $\sigma = \frac{\sqrt{p(1-p)}}{f(Q_p)}$

Asymptotic behaviour of kernel quantile estimators versus sample quantiles



$f(x)$ standard Lognormal, $p=0.95$

$$\hat{Q}_{kernel} \xrightarrow{d} N(Q_p, \frac{\sigma^2}{N}), \quad \sigma = \frac{\sqrt{p(1-p)}}{f(Q_p)}$$

