Quantile Estimation

Jussi Klemelä

April, 2016

Contents

1) Quantiles: Definition and applications

2) Estimation of Quantiles

   – empirical quantiles
   – parametric modeling: Pareto distributions

Part I: Quantiles: Definition and Applications

What is a Quantile?

- How high a barrage (dam, embankment) must be in order that the probability of the water level exceeding this year the barrage is smaller than $1/10,000$.

- How much liquid capital a bank must possess in order that the probability of running out of cash during the next month is smaller than $1/10,000$.

- Let Y be a real valued random variable.
  Let $0 < p < 1$ be a probability.
  The pth quantile is the smallest number $x \in \mathbf{R}$ so that $P(Y > x) \leq 1 - p$.

Figure 1: Flood of 1953.

Quantiles in Finance (Value-at-Risk)

- Regulatory officials want to ensure that systemic financial institutions (large banks, insurance companies) do not fall into a liquidity crisis: capital requirements.

- Investors want to control risk. If the 1% quantile of a monthly loss of S&P 500 returns is 11%, then an investor who owns S&P 500 index could expect to suffer 11% monthly loss once in every eight years.
  (100 months ≈ 8 years.)

What is a Quantile?

- The pth quantile is the smallest number $x \in \mathbf{R}$ so that $P(Y > x) \leq 1 - p$.
- We assume that Y has a continuous distribution.
  - The pth quantile is the number $x \in \mathbf{R}$ so that $P(Y > x) = 1 - p$.
  - The pth quantile is the number $x \in \mathbf{R}$ so that $P(Y \leq x) = p$.
  - Let the distribution function of Y be

$$F(x) = P(Y \leq x), \qquad x \in \mathbf{R}.$$

    The pth quantile is the number $x \in \mathbf{R}$ so that $F(x) = p$.
  - The pth quantile is

$$Q(p) = F^{-1}(p), \qquad p \in (0, 1).$$

Definition of a Quantile

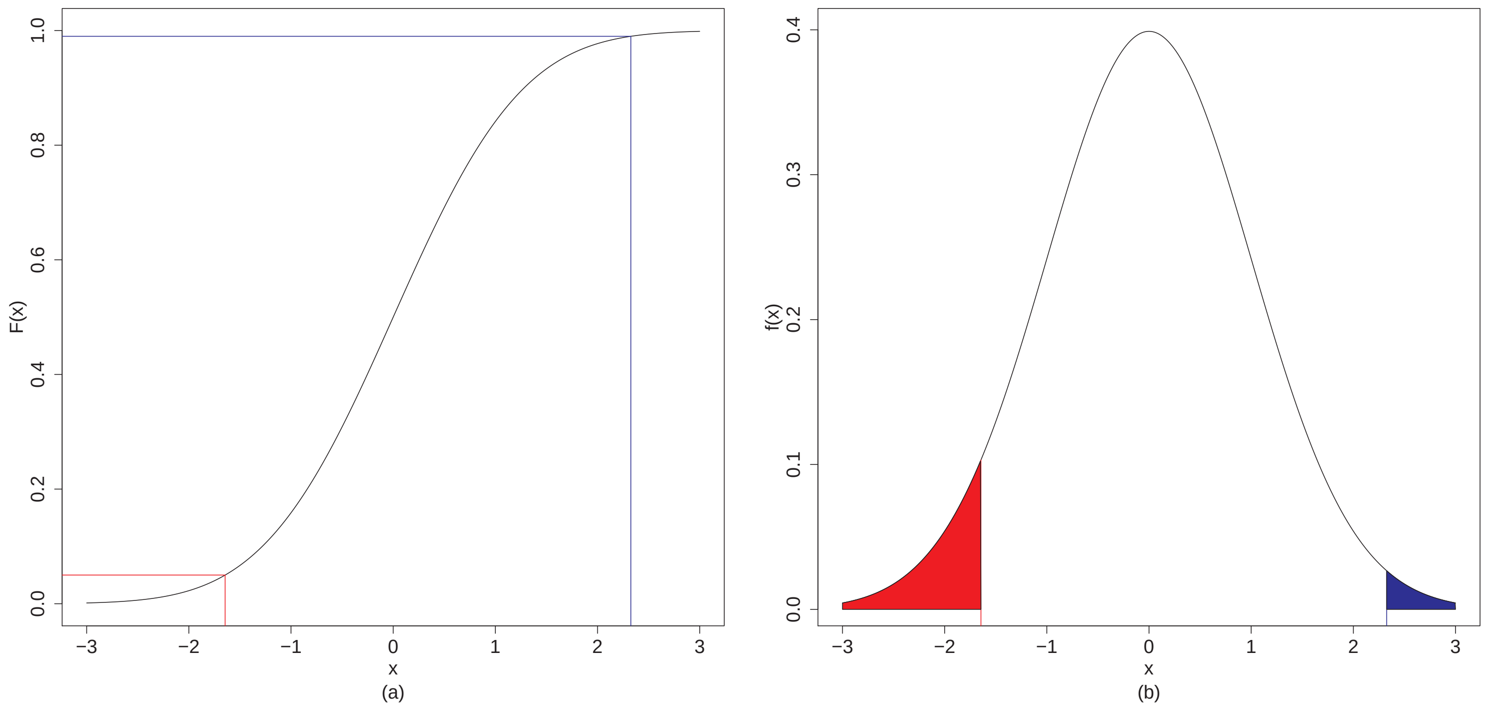

Figure 2: (a) Distribution function. (b) Density function. The 5% quantile with red. The 99% quantile with blue.

Extreme Quantiles

- We are interested in the quantiles $Q(p)$, where p is close to zero or close to one. For example, $p = 0.0001$ or $p = 0.9999$.

- By definition, there are few observations in tail areas. This makes estimation in tail areas difficult.

- In order to estimate extreme quantiles we make special models for tail areas of the distribution, and ignore the central area of the distribution.

Part II: Estimation of Quantiles

Empirical Quantiles

- We observe $Y_1, \ldots, Y_n$ with a common distribution.
  We want to estimate the pth quantile $Q(p)$, $0 < p < 1$.
  The pth quantile is such x that $P(Y \le x) = p$.

- (1) For any x, we can estimate $P(Y \le x)$ by calculating the frequencies.
  (2) The estimate of $P(Y \le x)$ is

$$\hat{p} = \frac{\#\{Y_i : Y_i \le x\}}{n}.$$

  (3) The empirical quantile is such x that $\hat{p}$ is closest to p.

- Computation of an empirical quantile:
  (1) Let m be such integer that $p \approx m/n$ (round pn to the closest integer).
  (2) The estimate of $Q(p)$ is the mth largest observation.

S&P 500 Prices and Returns

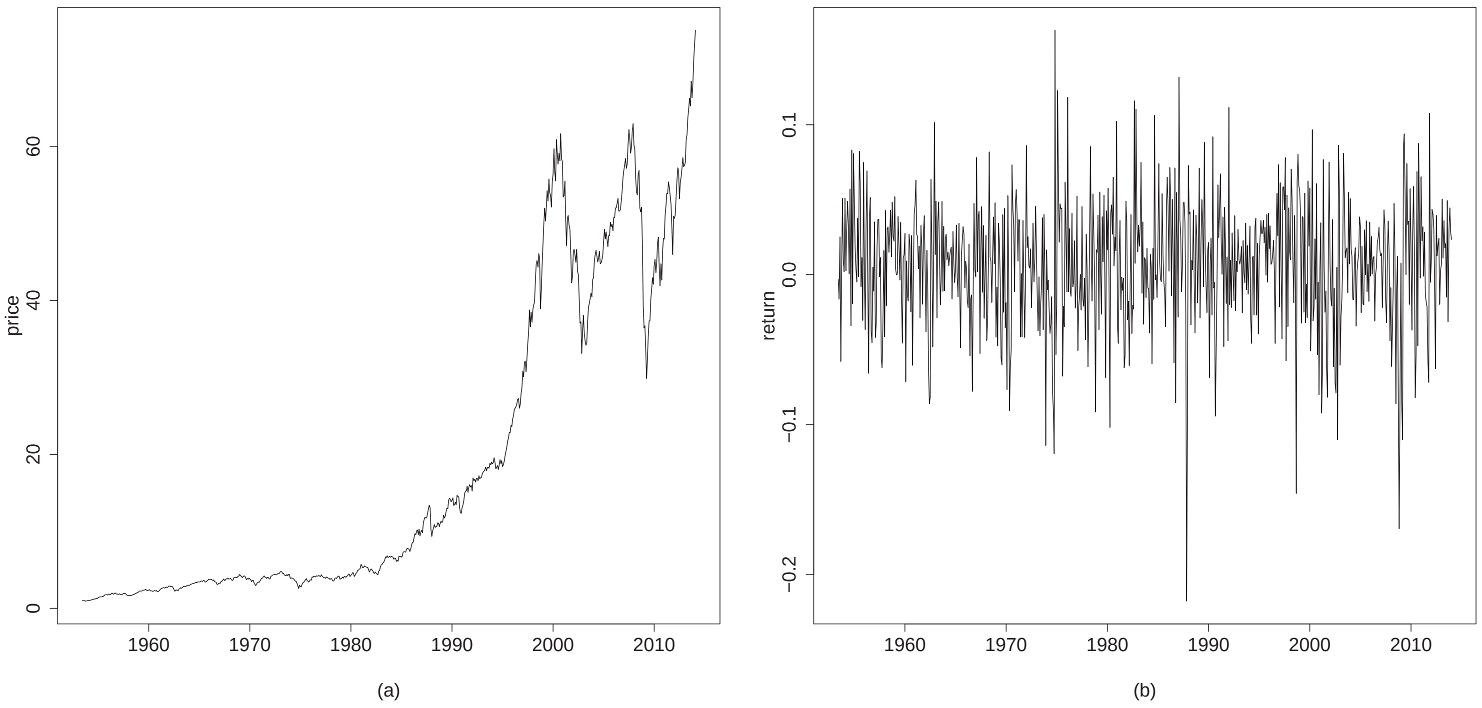

Figure 3: (a) Prices of S&P 500. (b) Returns $Y_t = (S_t - S_{t-1})/S_{t-1}$. There are 728 monthly observations from May 1953 to December 2013.

Empirical Quantiles for S&P 500

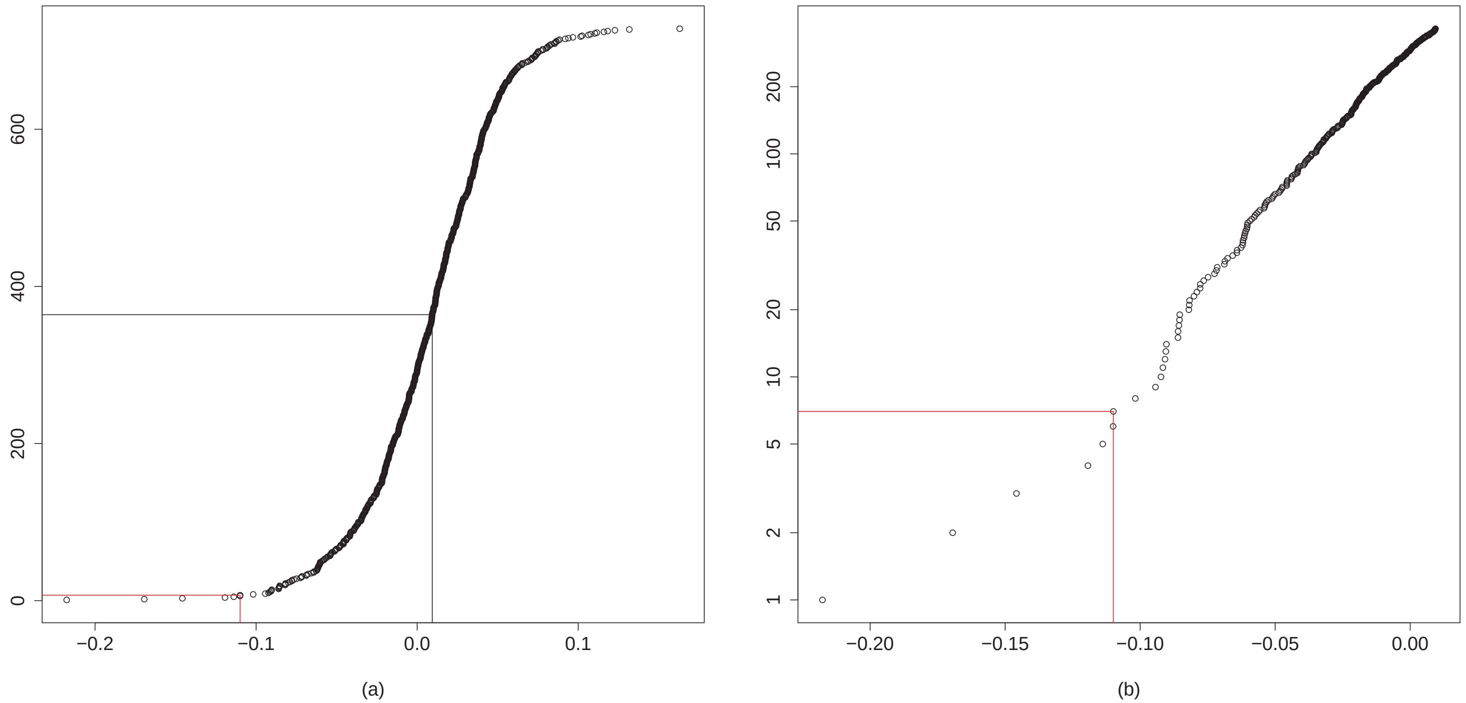

Figure 4: (a) The empirical distribution function. (b) The left part of the empirical distribution function, with logarithmic y-axis. Red: the 1% empirical quantile. There are 728 observations. What if we want to estimate the pth quantile when $p = 1/1000$?.

Model Fitting for S&P 500

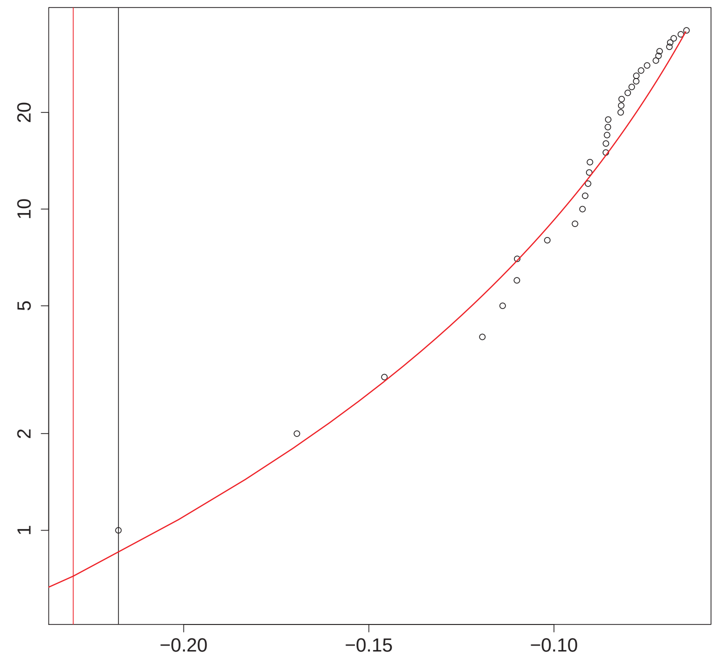

Figure 5: The left part of the empirical distribution function. Estimation of the pth quantile with $p = 1/1000 = 0.1\%$. Blue: Pareto fit and quantile estimate. Black: empirical quantile.

Part III: A Research Problem: Conditional Quantiles

- Expectations and conditional expectations:
  E(temperature) = $+5°C$
  E(temperature | it is January) = $-10°C$
  E(temperature | it is July) = $+15°C$.

- What is a conditional quantile?
  The quantile is $Q(p) = F_Y^{-1}(p)$, where $F_Y(x) = P(Y \leq x)$ is the distribution function.
  The conditional quantile is $Q(p \mid X) = F_{Y|X}^{-1}(p)$, where $F_{Y|X}(x) = P(Y \leq x \mid X)$ is the conditional distribution function.

Conditional Quantiles: GARCH Estimation

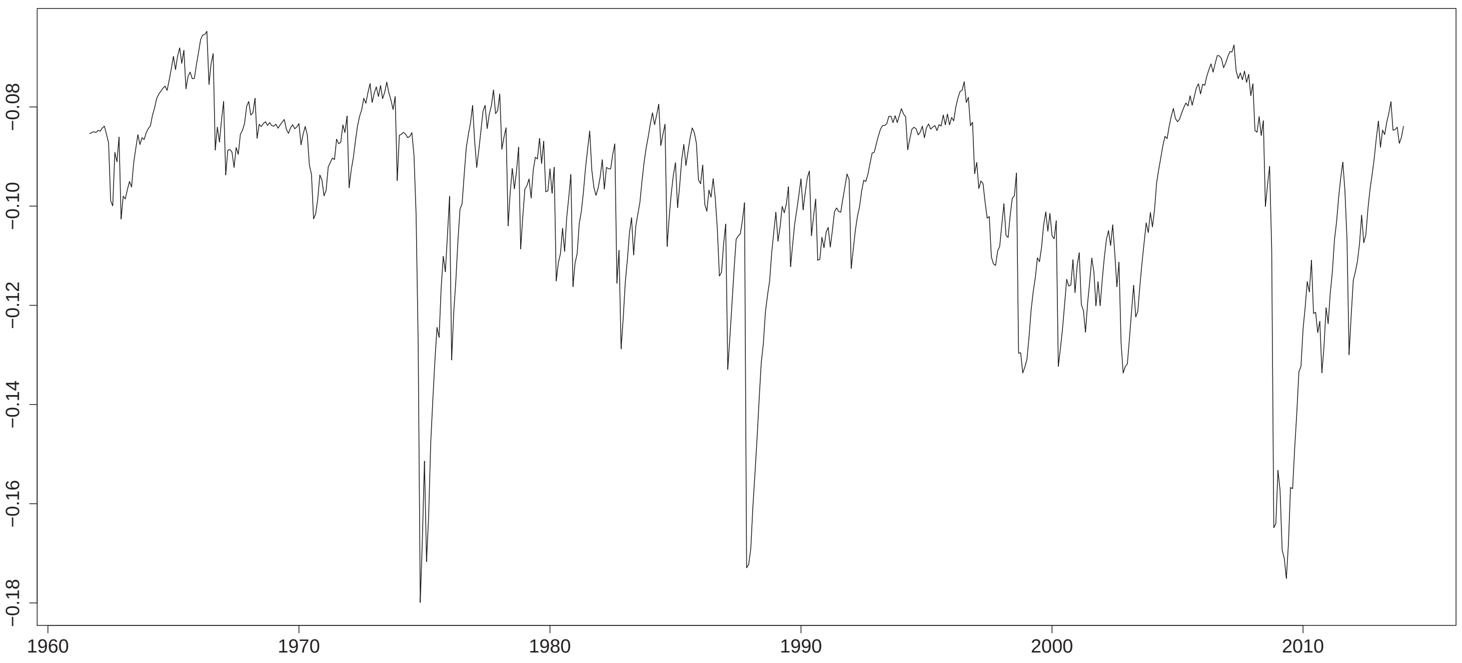

Figure 6: Conditional quantiles with level $p = 1\%$. $Q_{Y|X}(p) = \mu_{Y|X} + \sigma_{Y|X}\Phi^{-1}(p)$.

Conclusion

- Quantile estimation is a classical problem with many important applications.
- In quantile estimation we are facing the problem of having only few observations.
- There are many open research problems. For example, how to estimate conditional quantiles?