

I want to determine the distribution that a random variable is below an estimate quantile when estimated with a limited number of observations. The quantile of interest is the  $\alpha$ -quantile. It is assumed that  $X_t \sim N(\mu, \sigma^2)$  and  $X_1, \dots, X_T$  are drawn from a random sample. This  $\alpha$ -quantile is estimated by in following way:

1. draw  $T$  observations from  $N(\mu, \sigma^2)$
2. determine  $k = \text{roundup}(\alpha \cdot T)$
3. determine the estimated quantile by taking out the  $k^{th}$  lowest value out of the  $T$  observations, this value is the estimated quantile and is denoted by  $X_{k:T}$ .

For example, for 525 observations and  $\alpha = 0.01$ , the  $\alpha$ -quantile is estimated by taking out the 6<sup>th</sup> lowest loss. But since  $6/525 = 0.0114 \neq 0.01 = \alpha$ , there is an estimation error.

The goal is to find following probability  $Pr(X_t < X_{k:T})$ . According to Bain and Engelhardt (1991, p. 244) if  $\lim_{T \rightarrow \infty} \frac{k}{T} = \alpha$  (and for a continuous distribution with pdf  $f(x)$  that is continuous and nonzero at the  $\alpha$ -quantile) then the following holds:

$$\sqrt{T}(X_{k:T} - X_\alpha) \sim N(0, c^2/T), \quad (1)$$

where  $c^2 = \frac{\alpha(1-\alpha)}{[f(x_\alpha)]^2}$  and  $f(x)$  is the normal probability density function. Then, the probability of a VaR exceedance equals:

$$Pr(X_t < X_{k:T}) = Pr(X_t - X_{k:T} < 0) = Pr(Q < 0), \quad (2)$$

To solve above equation, the distribution of  $Q$  has to be derived:

$$Q \sim N(\mu, \sigma^2) - N(X_\alpha, c^2/T^2) = N(\mu - X_\alpha, \sigma^2 + c^2/T^2) \quad (3)$$

So, the probability is computed by the following (by using the Central Limit Theorem):

$$Pr(Q < 0) = Pr\left(\frac{X_t - X_{k:T} - \mu + X_\alpha}{\sqrt{\sigma^2 + c^2/T^2}} < \frac{-\mu + X_\alpha}{\sqrt{\sigma^2 + c^2/T^2}}\right) = \Phi_{0,1}\left(\frac{-\mu + X_\alpha}{\sqrt{\sigma^2 + c^2/T^2}}\right) \quad (4)$$

But in really,  $\mu$  and  $\sigma^2$  are unknown. But the distributions of these parameters are known:

$$\sqrt{T}(\hat{\mu} - \mu) \sim N(0, \sigma^2) \quad (5)$$

$$\sqrt{T}(\hat{\sigma} - \sigma) \sim N(0, 2\sigma^4) \quad (6)$$

The questions:

- Is above derivation correct?
- How to determine probability  $Pr(X_t < X_{k:T})$  with the distributions  $\hat{\mu}$  and  $\hat{\sigma}$ ?