Asimit, A. V., Li, D. & Peng, L. (2010). Pitfalls in using Weibull tailed distributions. Journal of Statistical Planning and Inference, 140(7), pp. 2018-2024. doi: 10.1016/j.jspi.2010.01.039

Pitfalls in using Weibull tailed distributions

Alexandru V. Asimit[1], Deyuan Li[2*] and Liang Peng[3]

[1]School of Mathematics, The University of Manchester, UK

[2]School of Management, Fudan University, China

[3]School of Mathematics, Georgia Institute of Technology, USA

October 25, 2009

Abstract

By assuming that the underlying distribution belongs to the domain of attraction of an extreme value distribution, one can extrapolate the data to a far tail region so that a rare event can be predicted. However, when the distribution is in the domain of attraction of a Gumbel distribution, the extrapolation is quite limited generally in comparison with a heavy tailed distribution. In view of this drawback, a Weibull tailed distribution has been studied recently. Some methods for choosing the sample fraction in estimating the Weibull tail coefficient and some bias reduction estimators have been proposed in the literature. In this paper, we show that the theoretical optimal sample fraction does not exist and a bias reduction estimator does not always produce a smaller mean squared error than a biased estimator. These are different from using a heavy tailed distribution. Further we propose a refined class of Weibull tailed distributions which are more useful in estimating high quantiles and extreme tail probabilities.

KEY WORDS: Asymptotic mean squared error, extreme tail probability, high quantile, regular variation, Weibull tail coefficient

1 Introduction

Suppose $X_1, \cdots, X_n$ are independent and identically distributed random variables with distribution function F, which has a Weibull tail coefficient θ. That is,

$$1 - F(x) = \exp\{-H(x)\} \quad \text{with} \quad H^{-}(x) = \inf\{t : H(t) \geq x)\} = x^{\theta} l(x), \tag{1.1}$$

*Corresponding address: School of Management, Fudan University, Room 736, Siyuan Building, 670 Guoshun Road, 200433 Shanghai, P.R. China. E-mail address: deyuanli@fudan.edu.cn (D. Li).

where $l(x)$ is a slowly varying function at infinity, i.e.,

$$\lim_{t\to\infty} l(tx)/l(t) = 1 \quad \text{for all} \quad x > 0.$$

This class of distributions includes some well-known light tailed distributions such as Weibull, Gaussian, gamma and logistic. Due to the applications of these distributions in insurance, estimating θ has attracted much attention recently. Accurate estimate of the probabilities associated with the extreme events contributes to a good understanding of the risk taken by the insurance company. In addition, estimates of certain risk measures can be obtained, such as the Value-at-Risk, which is a quantile function. This may be quite useful for risk management purposes, as it allows one to determine high quantiles of the insurance company losses and therefore enables one to obtain capital amounts that will be adequate with high probability.

There exist various estimators for θ in the literature; see Beirlant, Bouquiaux and Werker (2006), Gardes and Girard (2008), Girard (2004). A comparison study is given in Gardes and Girard (2006). Since the condition (1.1) is made asymptotically, each of these proposed estimators for θ can only involve a fraction of upper order statistics. How to choose this fraction plays an important role in practice. Motivated by similar studies on estimating extreme value index in Matthys and Beirlant (2003) and Mattys, Delafosse, Guillou and Beirlant (2004), Diebolt, Gardes, Girard and Guillou (2008a,b) proposed ways to choose the optimal fraction in estimating both θ and high quantiles of F. Moreover some bias reduction estimators for both θ and high quantiles are proposed in Diebolt, Gardes, Girard and Guillou (2008a,b), and Dierckx, Beirlant, de Waal and Guillou (2009).

It is known that there exists a theoretical optimal choice of the sample fraction in estimating the tail index of a heavy tailed distribution when the second order regular variation index is negative. In addition, a bias reduced estimator for the tail index produces a smaller order of asymptotic mean squared error than the corresponding biased tail index estimator theoretically. Since the estimation for the Weibull tail coefficient is partly motivated by the similar study in estimating the tail index of a heavy tailed distribution, one may conjecture that the bias reduction for estimating θ is always better. Although the above mentioned papers are in favor of bias-reduction estimation for θ, we show that bias reduction estimation is not always better in the sense of asymptotic mean squared error and the choice of sample fraction for a bias reduction estimator of θ becomes practically difficult. That is, a bias reduction estimator for θ is not particularly useful both theoretically and practically. These observations are in contrast to the case of tail index estimation. Finally, we propose a refined class of Weibull tailed distributions which are more useful in estimating high quantiles and extreme tail distributions.

We organize this paper as follows. Section 2 presents our main findings. A simulation study is given in Section 3. Some conclusions are given in Section 4.

2 Main results

Before giving our statements, we list some known estimators for θ and their asymptotic results.

Suppose $X_1, \cdots, X_n$ are independent and identically distributed random variables with distribution function F. Let $X_{n,1} \le \cdots \le X_{n,n}$ denote the order statistics of $X_1, \cdots, X_n$. Throughout we assume that F satisfies (1.1). Here we focus on the following estimators proposed in Diebolt, Gardes, Girard and Guillou (2008) and Dierckx, Beirlant, de Waal and Guillou (2009), respectively:

$$\hat{\theta}_H(k) = \frac{k^{-1}\sum_{i=1}^k \log(X_{n,n-i+1}/X_{n,n-k})}{k^{-1}\sum_{i=1}^k \log\log((n+1)/i) - \log\log((n+1)/(k+1))},$$

$$\hat{\theta}_{R,1}(k) = k^{-1}\sum_{i=1}^k i\log(n/i)\log(X_{n,n-i+1}/X_{n,n-i}),$$

$$\hat{\theta}_{R,2}(k) = k^{-1}\sum_{i=1}^k i\log(n/i)\log(X_{n,n-i+1}/X_{n,n-i}) - \frac{\sum_{j=1}^k (a_j-\bar{a})j\log(n/j)\log(X_{n,n-j+1}/X_{n,n-j})}{\sum_{j=1}^k (a_j-\bar{a})^2}\bar{a},$$

where $a_j = (\frac{\log(n/j)}{\log(n/k)})^{-1}$, $\bar{a} = k^{-1}\sum_{j=1}^k a_j$, and

$$\hat{\theta}_M(k) = \{1 - \frac{\sum_{j=1}^k \log\{\hat{m}(X_{n,n-j})/\hat{m}(X_{n,n-k-1})\}}{\sum_{j=1}^k \log(X_{n,n-j}/X_{n,n-k-1})}\}^{-1},$$

where $\hat{m}(X_{n,n-k}) = k^{-1}\sum_{i=1}^k X_{n,n-i+1} - X_{n,n-k}$. Note that $\hat{\theta}_{R,2}(k)$ and $\hat{\theta}_M(k)$ are bias-reduced estimators for θ. Here we want to compare these two bias-reduced estimators with the possibly biased estimators $\hat{\theta}_H(k)$ and $\hat{\theta}_{R,1}(k)$ in terms of asymptotic mean squared errors.

In order to derive the asymptotic limits of the above estimators, one needs the following stricter condition than (1.1): there exist $\rho \le 0$ and $b(x) \to 0$ (as $x \to \infty$) such that

$$\lim_{x\to\infty} b^{-1}(x)\log\frac{l(xy)}{l(x)} = \frac{y^\rho - 1}{\rho} \quad \text{for all} \quad y > 0. \tag{2.1}$$

From now on we assume that (1.1) and (2.1) hold and $k = k(n) \to \infty$ and $k/n \to 0$ as $n \to \infty$.

Result 1 (Theorem 1 of Gardes and Girard (2008)). If

$$k^{1/2}b(\log n) \to \lambda \in (-\infty, \infty) \quad \text{and} \quad k^{1/2}/\log n \to 0, \tag{2.2}$$

then

$$\sqrt{k}\{\hat{\theta}_H(k) - \theta\} \xrightarrow{d} N(\lambda, \theta^2).$$

Result 2 (Theorem 2.2 of Diebolt, Gardes, Girard and Guillou (2008a)). If

$$|kb(k)| \to \infty, \quad k^{1/2} b(\log(n/k)) \to \lambda \in (-\infty, \infty) \quad \text{and} \quad \log k/\log n \to 0 \quad \text{when} \quad \lambda = 0, \tag{2.3}$$

then

$$\sqrt{k}\{\hat{\theta}_{R,1}(k) - \theta - b(\log(n/k))k^{-1}\sum_{j=1}^{k} a_j^{-\rho}\} \xrightarrow{d} N(0, \theta^2).$$

Result 3 (Theorem 3.1 of Diebolt, Gardes, Girard and Guillou (2008a)). If

$$\begin{array}{l} |kb(k)| \to \infty, \quad \frac{\sqrt{k}}{\log(n/k)} b(\log(n/k)) \to \Lambda \in (-\infty, \infty) \quad \text{and,} \\ \frac{\log^2 k}{\log(n/k)} \to 0 \quad \text{and} \quad \frac{\sqrt{k}}{\log(n/k)} \to \infty \quad \text{when} \quad \Lambda = 0, \end{array} \tag{2.4}$$

then

$$\frac{\sqrt{k}}{\log(n/k)}\{\hat{\theta}_{R,2}(k) - \theta\} \xrightarrow{d} N(0, \theta^2).$$

Result 4 (Theorem 2.3 of Dierckx, Beirlant, de Waal and Guillou (2009)). If $x^{-\rho}|b(x)|$ is a normalized slowly varying function and

$$k^{1/2}/\log(n/k) \to \infty \quad \text{and} \quad \log^2 k/\log n \to 0, \tag{2.5}$$

then

$$\frac{\sqrt{k}}{\log(n/k)}\{\hat{\theta}_M(k) - \theta - (1+\rho)b(\log(n/k)) - \frac{\theta - \theta^2}{\log(n/k)}\} \xrightarrow{d} N(0, \theta^2).$$

Now, using the above results, we can articulate our statements as follows.

Statement 1: no theoretical optimal k. Recently, Diebolt, Gardes, Girard and Guillou (2008a) proposed to choose k to minimize the following estimated asymptotic mean squared error

$$A\hat{M}SE(k) = k^{-1}\hat{\theta}^2_{R,1}(k) + \{\frac{\sum_{j=1}^{k}(a_j - \bar{a})j\log(n/j)\log(X_{n,n-j+1}/X_{n,n-j})}{\sum_{j=1}^{k}(a_j-\bar{a})^2}k^{-1}\sum_{j=1}^{k} a_j\}^2. \tag{2.6}$$

Now the question is whether the minimum exists. Note that the theoretical asymptotic mean square error of $\hat{\theta}_{R,1}(k)$ is $AMSE(k) = k^{-1}\theta^2 + \{b(\log(n/k))k^{-1}\sum_{j=1}^{k} a_j^{-\rho}\}^2$. Since b is a regular variation with index ρ, (2.3) implies that $\limsup_{n\to\infty}\sqrt{k}\{\log(n/k)\}^{\rho-\epsilon} < \infty$ for any $\epsilon > 0$, i.e., $\sqrt{k} = O(\{\log(n/k)\}^{-\rho+\epsilon}) = O(\{\log n\}^{-\rho+\epsilon})$. Thus,

$$\log k = o(k^{\frac{1}{2\epsilon - 2\rho}}) = o(\log n),$$

which implies that

$$\left\{ \begin{array}{l} \lim_{n\to\infty}\log(n/k)/\log n = 1 - \lim_{n\to\infty}\log k/\log n = 1 \\ \lim_{n\to\infty} b(\log(n/k))/b(\log(n)) = \lim_{n\to\infty}(\frac{\log(n/k)}{\log n})^{\rho} = 1. \end{array} \right. \tag{2.7}$$

Write $a_j = (1 - \log(j/k)/\log(n/k))^{-1}$. For any $t > 0$, we have

$$\begin{aligned} 1 &\geq k^{-1}\sum_{i=1}^k a_i^{-\rho} \\ &\geq k^{-1}\sum_{i=1}^k (1 - \tfrac{\log(i/k)}{t})^{\rho} \\ &\to \int_0^1 (1 - \tfrac{\log x}{t})^{\rho}\,dx. \end{aligned}$$

Taking $t \to \infty$, we have

$$\lim_{k\to\infty} k^{-1}\sum_{i=1}^k a_i^{-\rho} = 1. \tag{2.8}$$

By (2.7) and (2.8), we have

$$AMSE(k) = \{k^{-1}\theta^2 + b^2(\log n)\}\{1 + o(1)\}.$$

Apparently the minimum of AMSE(k) is achieved when $k = n$. Hence, the theoretical optimal k in terms of minimizing the asymptotic mean squared error of $\hat{\theta}_{R,1}$ does not exist at all. So, the method in choosing k in Diebolt, Gardes, Girard and Guillou (2008a) is not mathematically sound. Similar thing happens for the way of choosing k in estimating high quantiles proposed in Diebolt, Gardes, Girard and Guillou (2008b). These are not surprising since similar study exists in estimating an extreme value index γ, where the case of $\gamma = 0$ is excluded in considering the optimal choice of sample fraction.

Statement 2: no need to reduce bias when $\sqrt{k}b(\log(n/k)) \to \lambda \in (-\infty, \infty)$. It follows from Results 1-4 that biased estimators $\hat{\theta}_H(k)$ and $\hat{\theta}_{R,1}(k)$ have a faster rate of convergence than the bias-reduced estimators $\hat{\theta}_{R,2}(k)$ and $\hat{\theta}_M(k)$. Hence, when one employs the same k such that $\sqrt{k}b(\log(n/k)) \to \lambda \in (-\infty, \infty)$, the biased estimators have a smaller order of mean squared error than the bias-reduced estimators. This is different from the study for a heavy tailed distribution.

Statement 3: bias reduction is useful only when a large sample fraction is employed. Now let's compare the bias estimator $\hat{\theta}_{R,1}(m)$ with the bias reduction estimator $\hat{\theta}_{R,2}(k)$ when m and k satisfy (2.3) with $\lambda \neq 0$ and (2.4) with $\Lambda \neq 0$, respectively. By (2.7) and (2.8), Results 2 and 3 imply that the asymptotic mean squared errors for $\hat{\theta}_{R,1}(m)$ and $\hat{\theta}_{R,2}(k)$ are $b^2(\log n)\{1 + \theta^2\lambda^{-2}\}$ and $b^2(\log n)\theta^2\Lambda^{-2}$, respectively. Hence, $\hat{\theta}_{R,2}(k)$ has a smaller asymptotic mean squared error than $\hat{\theta}_{R,1}(m)$ only when $\Lambda^2 \geq \lambda^2\theta^2/(\lambda^2 + \theta^2)$. That is, when the sample fraction k in the bias-reduced estimator $\hat{\theta}_{R,2}(k)$ is not large enough, i.e., Λ is not large enough, the bias-reduced estimator $\hat{\theta}_{R,2}(k)$ has a larger asymptotic mean squared error than the biased estimator $\hat{\theta}_{R,1}(m)$. On the other hand, how large a sample fraction in a bias-reduced estimator should be chosen becomes practically difficult. This is different from tail index estimation, where a bias reduction tail index estimator has a smaller order of asymptotic mean squared error than a biased one.

Statement 4: not enough for estimating an extreme tail probability. It is known that heavy tailed distributions can be employed to estimate both high quantiles and extreme tail probabilities. Although model (1.1) has been employed to estimating high quantiles, it is doubtful that it can be used to estimate an extreme tail probability. Suppose

$$1-F(x)\sim x^{\alpha}\exp\{-cx^{1/\theta}\}=\exp\{-cx^{1/\theta}+\alpha\log x\}$$

as $x\to\infty$, which satisfies (1.1). As in Diebolt, Gardes, Girard and Guillou (2008b), estimating a high quantile for (1.1) is based on the inverse function of $\exp\{-cx^{1/\theta}\}$ and estimators for c and θ. However, the factor x^{α} is not negligible in estimating an extreme tail probability, i.e., estimating $1-F(x_n)$ where $x_n\to\infty$ as $n\to\infty$. Therefore a more refined model than (1.1) is needed for estimating an extreme tail probability. A possible class is

$$1-F(x)=cx^{\alpha}\exp\{-dx^{1/\theta}\}\{1+O(x^{-\beta})\} \tag{2.9}$$

as $x\to\infty$, where $c>0,\alpha\in R,d>0,\beta>0$ and $\theta>0$. Note that the class of the distributions satisfying (2.9) is a sub-class of Weibull tailed distributions defined in (1.1). One example which is Weibull tailed distribution but not satisfying (2.9) is $1-F(x)=\exp\{-x^{\alpha}(\log x)^{\beta}\}$ for some $\alpha,\beta>0$ and large x.

Since Theorem 1.2.6 of de Haan and Ferreira (2006) implies that (2.9) is in the domain of attraction of the Gumbel distribution, one may wonder how useful (2.9) is in estimating high quantiles in comparison with the way developed in extreme value theory.

Statement 5: model (2.9) is useful in estimating very high quantiles. Let's consider estimating the high quantile x_p defined by $p=1-F(x_p)$, where $p=p(n)\to 0$. A proposed estimator based on a Weibull tailed distribution in the literature is $\tilde{x}_p(k)=X_{n,n-k+1}\{\frac{\log(1/p)}{\log(n/k)}\}^{\hat{\theta}_H(k)}$ and it follows from Diebolt, Gardes, Girard and Guillou (2008b) that

$$\tilde{x}_p(k)/x_p-1=O(\log(\frac{\log(1/p)}{\log(n/k)})/\sqrt{k}) \tag{2.10}$$

when

$$\text{(2.9) holds and}\quad \sqrt{k}\frac{\log\log n}{\log n}\to\lambda<\infty,\quad \liminf_{n\to\infty}\frac{\log(1/p)}{\log(n/k)}>1. \tag{2.11}$$

Since (2.9) implies that F is in the domain of attraction of the Gumbel distribution, x_p can be estimated by some known methods in extreme value theory; see Section 4.3 of de Haan and Ferreira (2006). Since the extreme value index is zero, we can estimate x_p by

$$\hat{x}_p(k)=X_{n,n-k}+log(\frac{k}{np})X_{n,n-k}k^{-1}\sum_{i=1}^{k}\log\frac{X_{n,n-i+1}}{X_{n,n-k}},$$

which is slightly different from the estimator for x_p given in Section 4.3.1 of de Haan and Ferreira (2006). Denote the inverse function of $1/(1-F(t))$ by $U(t)$. Then (2.9) implies that

$$\begin{aligned} U(x) &= d^{-\theta}(\log x)^{\theta}\{1+a_1\tfrac{\log\log x}{\log x}+a_2\tfrac{1}{\log x}+a_3\tfrac{(\log\log x)^2}{(\log x)^2} \\ &\quad +a_4\tfrac{\log\log x}{(\log x)^2}+a_5\tfrac{1}{(\log x)^2}+O\big(\tfrac{(\log\log x)^2}{(\log x)^3}+(\log x)^{-\beta\theta-1}\big)\} \end{aligned} \tag{2.12}$$

as $x \to \infty$, where

$$a_1=\alpha\theta^2,\quad a_2=\theta\log c-\alpha\theta^2\log d,\quad a_3=-\frac{1-\theta}{2\theta}a_1^2,$$

$$a_4=\alpha\theta a_1-\frac{a_1a_2(1-\theta)}{\theta}\quad \text{and}\quad a_5=\alpha\theta a_2-\frac{1-\theta}{2\theta}a_2^2.$$

It follows from (2.12) that

$$\begin{aligned} U(tx) &= d^{-\theta}(\log t)^{\theta}\{1+a_1\tfrac{\log\log t}{\log t}+a_2\tfrac{1}{\log t}+a_3\tfrac{(\log\log t)^2}{(\log t)^2} \\ &\quad +a_4\tfrac{\log\log t}{(\log t)^2}+a_5\tfrac{1}{(\log t)^2}+\log x\big(\tfrac{a_1-a_2+\theta a_2}{(\log t)^2}+\tfrac{a_1(\theta-1)\log\log t}{(\log t)^2}+\tfrac{\theta}{\log t}\big) \\ &\quad +\tfrac{(\log x)^2}{2}\tfrac{\theta(\theta-1)}{(\log t)^2}+O\big(\tfrac{(\log\log t)^2}{(\log t)^3}+(\log t)^{-\beta\theta-1}\big)\} \end{aligned} \tag{2.13}$$

for any $x>0$ as $t\to\infty$. Hence, when $\beta\theta>1$ and $\theta\neq1$,

$$\lim_{t\to\infty}\frac{\frac{U(tx)-U(t)}{a(t)}-\log x}{A(t)}=\frac{(\log x)^2}{2}$$

for $x>0$, where

$$a(t)=d^{-\theta}(\log t)^{\theta}\{\frac{a_1-a_2+\theta a_2}{(\log t)^2}+\frac{a_1(\theta-1)\log\log t}{(\log t)^2}+\frac{\theta}{\log t}\}$$

and $A(t)=d^{-\theta}\theta(\theta-1)(\log t)^{\theta-2}/a(t)$. Similar to the proof of Theorem 4.3.1 of de Haan and Ferreira (2006), we have

$$\hat{x}_p(k)/x_p-1=O_p(\frac{a(n/k)\{\log(k/(np))\}^2}{x_p\sqrt{k}}) \tag{2.14}$$

when

$$(2.9)\quad \text{holds and}\quad \sqrt{k}A(n/k)\to\lambda\in(-\infty,\infty),\quad \log(np)=o(\sqrt{k}),\quad np=o(k). \tag{2.15}$$

Note that the condition $\sqrt{k}A(n/k)\to\lambda\in(-\infty,\infty)$ in (2.15) and the formula for $A(t)$ imply that $\sqrt{k}/\log(n/k)$ converges to a finite number, i.e., $\sqrt{k}/\log n$ converges to a finite number. Combining this with the condition $\log(np)=o(\sqrt{k})$ in (2.15), we conclude that (2.15) implies that $\lim_{n\to\infty}\log(np)/\log n=0$. It is easy to check that (2.11) implies that $\lim_{n\to\infty}\{-\log(np)\}/\log n>0$. Hence model (2.9) works for a much higher quantile than the standard high quantile estimation developed in extreme value theory. This is exactly what we need to cope with the extrapolation limitation of using the condition of domain of attraction of the Gumbel distribution. Is it possible to have a high quantile estimator work for the case $\lim_{n\to\infty}\log(np)/\log n\geq0$? Since the high quantile estimator $\tilde{x}_p(k)$ is only based on the first order in (2.12), the model approximation error becomes large when x_p is small. This explains why $\tilde{x}_p$ only works for a very high quantile. It is of interest to study a high quantile estimator based on (2.12) and estimators for c, α, d, θ under the setup of (2.9). We conjecture that this new high quantile estimator works when $\lim_{n\to\infty} \log(np)/\log n \geq 0$. If this is true, then the model (2.9) becomes more practically useful than the methods based on either (1.1) or the domain of attraction of the Gumbel distribution since one does not need to worry whether the target quantile is high enough.

3 A Simulation Study

Here, we perform a simulation study to support Statements 1, 2, 3 and 5. We simulate 1000 random samples of size $n = 1000$ from the Gamma distribution with shape parameter 1.2 and scale parameter 1.

First, for each sample we determined the optimal value $k \in [2, n-1]$ such that the $AMSE$ in (2.6) is minimized. Figure 1 plots the $AMSE$ evaluated at each optimal k ($Kopt$) against the optimal k. This figure shows that most of the optimal k values are near the sample size $n = 1000$, which supports Statement 1.

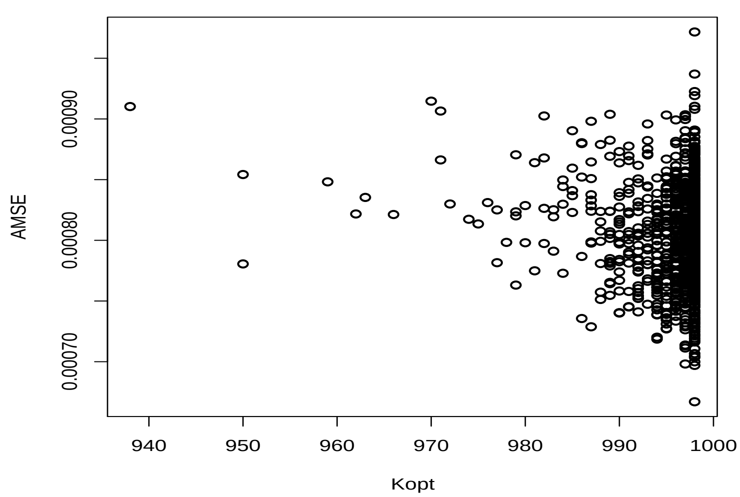

Figure 1: Plots of $AMSE(Kopt)$ against $Kopt$ for $Gamma(1.2, 1)$.

Next, to support Statements 2 and 3, a simulation is performed in which the biased and bias-reduced estimators, $\hat{\theta}_{R,1}(m)$ and $\hat{\theta}_{R,2}(k)$, are compared. In Figure 2, we plot the mean squared errors of these two estimators against difference choices of $k = m$. From this figure, we observe that the MSE of $\hat{\theta}_{R,1}(m)$ is smaller than that of $\hat{\theta}_{R,2}(k)$ when $k = m \leq 500$, which supports Statement 2. When m is around 200, one really needs a very large k to ensure that the MSE of $\hat{\theta}_{R,2}(k)$ is smaller than that of $\hat{\theta}_{R,1}(m)$. This observation supports Statement 3.

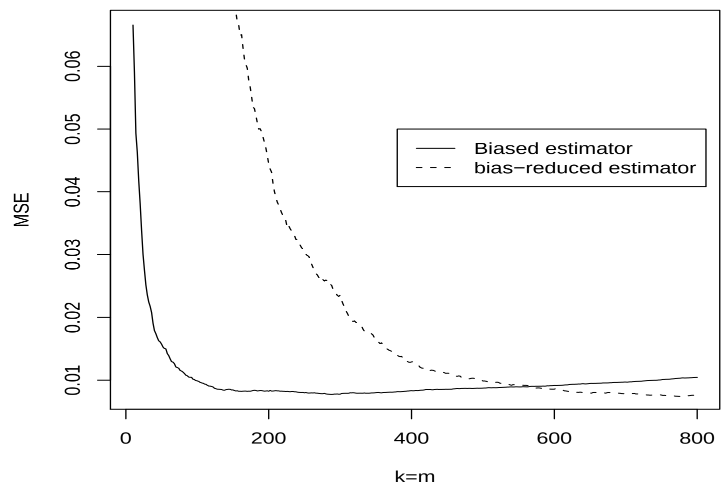

Figure 2: MSEs of $\hat{\theta}_{R,1}(m)$ and $\hat{\theta}_{R,2}(k)$ are plotted for $Gamma(1.2, 1)$.

Finally, we calculate the MSEs of $\tilde{x}_p(k)$ and $\hat{x}_p(k)$ for $p = 10^{-2}, 10^{-4}, 10^{-6}$. The first 50 smallest MSEs of these two estimators are plotted in Figure 3, which shows that $\tilde{x}_p(k)$ works much better than $\hat{x}_p(k)$ when p becomes smaller.

4 Conclusions

Unlike tail index estimation, the theoretical optimal sample fraction in estimating the Weibull tail coefficient does not exist, and a bias reduction estimator only shows an advantage when a large sample fraction is employed. There is no theory to guide the choice of a large sample fraction which still satisfies some necessary conditions such that $\frac{\sqrt{k}}{\log(n/k)} b(\log(n/k)) \to \Lambda \in (-\infty, \infty)$. Therefore, it should be extremely cautious to employ any adaptive estimation and bias reduction estimation for the Weibull tail coefficient in practice due to the lack of theoretical support! Weibull tailed distributions are more useful in estimating a higher quantile than the standard high quantile estimation by assuming the condition of domain of attraction of the Gumbel distribution. The proposed refined class of Weibull tailed distributions is necessary for estimating extreme tail probabilities and may be more practical in estimating high quantiles.

Acknowledgment. We thank a reviewer for his/her helpful comments. Li's research was partially

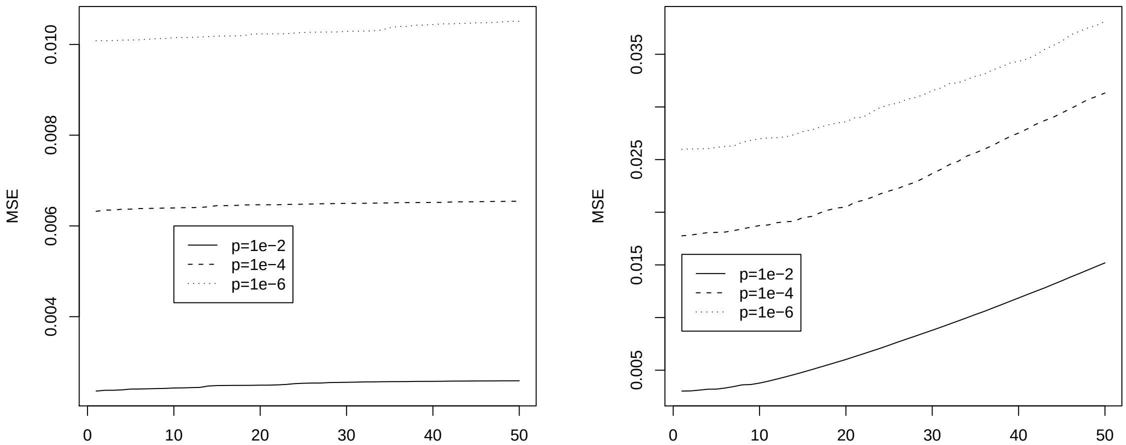

Figure 3: The first 50 smallest MSEs of $\tilde{x}_p(k)$ and $\hat{x}_p(k)$ are plotted in left and right panels, respectively, for $Gamma(1.2, 1)$.

supported by NNSFC grant 10801038. Peng's research was supported by NSF grant SES-0631608.

References

[1] J. Beirlant, C. Bouquiaux and B.J.M. Werker (2006). Semiparametric lower bounds for tail index estimation. *J. Statist. Plann. and Inference 136, 705–729.*

[2] L. de Haan and A. Ferreira (2006). Extreme Value Theory, an Introduction. *Springer, New York.*

[3] J. Diebolt, L. Gardes, S. Girard and A. Guillou (2008a). Bias-reduced estimators of the Weibull tail-coefficient. *Test 17, 311–331.*

[4] J. Diebolt, L. Gardes, S. Girard and A. Guillou (2008b). Bias-reduced extreme quantile estimators of Weibull tail-distributions. *J. Statist. Plann. and Inference 138, 1389–1401.*

[5] G. Dierckx, J. Beirlant, D. de Waal and A. Guillou (2009). A new estimation method for Weibull-type tails based on the mean excess function. *J. Statist. Plann. and Inference 139, 1905–1920.*

[6] L. Gardes and S. Girard (2006). Comparison of Weibull tail-coefficient estimators. *REVSTAT 4, 163–188.*

[7] L. Gardes and S. Girard (2008). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *J. Statist. Plann. and Inference 138, 1416–1427.*

[8] S. Girard (2004). A Hill type estimator of the Weibull Tail-Coefficient. *Communications in Statistics, Theory and Methods 33, 205–234.*

[9] G. Matthys and J. Beirlant (2003). Estimating the extreme value index and high quantiles with exponential regression models. *Statist. Sinica 13, 853–880.*

[10] G. Mattys, E. Delafosse, A. Guillou and J. Beirlant (2004). Estimating catastrophic quantile levels for heavy-tailed distributions. *Insurance Math. Econom. 34, 517–537.*