Extremes (2012) 15:67–87
DOI 10.1007/s10687-011-0127-9

Generalized fiducial confidence intervals for extremes

Damian V. Wandler · Jan Hannig

Received: 1 December 2009 / Revised: 13 December 2010 /
Accepted: 4 January 2011 / Published online: 15 January 2011

Abstract The generalized Pareto distribution is relevant to many situations when modeling extremes of random variables. In particular, peaks over threshold data approximately follow the generalized Pareto distribution. We use a fiducial framework to perform inference on the parameters and the extreme quantiles of the generalized Pareto. This inference technique is demonstrated both when the threshold is a known and unknown parameter. Assuming the threshold is a known parameter resulted in fiducial intervals with good empirical properties and asymptotically correct coverage. Likewise, our simulation results suggest that the fiducial intervals and point estimates compare favorably to the competing methods seen in the literature. The proposed intervals for the extreme quantiles when the threshold is unknown also have good empirical properties regardless of the underlying distribution of the data. Comparisons to a similar Bayesian method suggest that the fiducial intervals have better coverage and are similar in length with fewer assumptions. In addition to simulation results, the proposed method is applied to a data set from the NASDAQ 100. The data set is analyzed using the fiducial approach and its competitors for both cases when the threshold is known and unknown. R code for our procedure can be downloaded at http://www.unc.edu/~hannig/.

Keywords Fiducial inference · Extreme quantile · Peaks over threshold · MCMC

AMS 2000 Subject Classifications 62G32 · 62G05 · 62F10 · 62P12 · 62P05 · 62P30

D. V. Wandler (✉)
Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA
e-mail: dvwandler@gmail.com

J. Hannig
Department of Statistics and Operations Research, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

1 Introduction

Extreme value theory is of practical interest in a variety of different fields (e.g. economics, hydrology, environmental science, insurance, etc.). It is well known that modeling data over a high threshold with the generalized Pareto distribution (GPD) is appropriate (Davison and Smith 1990). As stated by Hosking and Wallis (1987) the applications of the GPD include analysis in extreme events such as the modeling of large insurance claims and in situations that an exponential distribution might be used but robustness is required with heavy or light tailed alternatives.

The generalized Pareto was first introduced by Pickands (1975). Later Smith (1984, 1985), Davison (1984), and Monfort and Witter (1985) all showed interest in its application and theoretical properties. If $X \sim F$, Pickands (1975) showed that the limiting distribution of $(X - a)$ conditional on $X > a$ as $a \to \omega_F$ where ω_F is the right-hand endpoint of the distribution follows a generalized Pareto distribution. The density of the GPD is defined as

$$f(x) = \begin{cases} \dfrac{1}{\sigma}\left(1 + \dfrac{\gamma(x-a)}{\sigma}\right)^{-\frac{1}{\gamma}-1} & \gamma \neq 0 \\ \dfrac{1}{\sigma}\exp\{-(x-a)/\sigma\} & \gamma = 0 \end{cases}$$

for $x > a$ such that $1 + \gamma(x - a)/\sigma > 0$.

Estimators of the parameters using maximum likelihood, method of moments, and L-moments have been explored in Davison (1984), Hosking and Wallis (1987), and Smith (1984). Smith (1984) found the maximum likelihood estimators to be asymptotically normal and consistent when $\gamma > -\frac{1}{2}$. Hosking (1990) showed that procedures using L-moments and probability weighted moments are equivalent. Furthermore, L-moments are more robust than method of moments and are often more efficient than maximum likelihood estimates. A Bayesian solution to this problem has been explored in Castellanos and Cabras (2005).

We propose new confidence intervals for γ, σ, and the β-quantile for the cases when the threshold, a, is a known and unknown parameter. The proposed solution is based on generalized fiducial intervals of Hannig (2009b). Simulation results suggest that this inference technique performs well with small sample sizes, and, when a is known, the confidence intervals have asymptotically correct coverage. This fiducial method for calculating intervals for the extreme quantiles (return levels) of the GPD also compare favorably to the profile log-likelihood method described in Coles (2001) and the Bayesian method described in Castellanos and Cabras (2005). That is to say that the fiducial intervals have good empirical coverage and are often shorter than the comparable profile log-likelihood and Bayesian intervals. Furthermore, the point estimates for γ and σ using this fiducial approach have smaller bias when compared with estimators calculated using maximum likelihood, L-moments, and the aforementioned Bayesian methods. The bias for the estimate of the β-quantile is smaller than the estimates based on maximum likelihood and L-moments but slightly larger than those calculated by the Bayesian method.

We also developed fiducial methods when the threshold is unknown. As seen in Coles (2001), the threshold is generally chosen by some *ad hoc* procedure of looking at plots and fixing the threshold for all subsequent calculations. Other methods will test whether the generalized Pareto fits the data for various different thresholds as seen in Choulakian and Stephens (2001) and Dupuis (1999). Guillou and Hall (2001) investigate this problem of choosing a threshold by using a fixed number of the largest order statistics and Frigessi et al. (2002) used a weighting scheme with a mixture of a Weibull distribution and the GPD to model the data. These methods do not account for fact that the threshold is unknown in all practical applications. The unknown threshold will add variability to the estimates of the extreme quantiles. Our method, like the Bayesian methods developed in Cabras and Castellanos (2009) and Tancredi et al. (2006), assumes the threshold is another parameter that is unknown. As a result, the fiducial method will select likely values for the threshold based on the data. Using this method, we performed a simulation study for data that was generated from various distributions that could be seen in real-life settings. Based on the simulations, the fiducial framework produced intervals for the β-quantile that had reasonable frequentist coverage for all of the distributions and compared favorably to the Bayesian approach described in Cabras and Castellanos (2009).

This fiducial approach was also used to analyze a data set in both cases when the threshold is assumed to be known and unknown. The data set that was analyzed was the log-weekly losses of the NASDAQ 100 index. Our analysis produced fiducial intervals for the 0.99-quantile that were generally shorter than the intervals from the appropriate profile log-likelihood and Bayesian methods.

2 Generalized fiducial inference

2.1 Overview

The original idea for fiducial inference was developed by Fisher (1930) in an attempt to overcome what he perceived as a deficiency in the Bayesian framework. Namely, he was opposed to assuming a prior distribution when there was little or no information about the parameters available. Opposition to the fiducial framework arose when it was later discovered that some of the properties that Fisher had originally claimed were not actually true (Lindley 1958; Zabell 1992). More recently, fiducial inference has begun to gain more acceptance in the statistics community following the introduction of generalized inference by Weeranhandi (1993) and the work of Hannig et al. (2006) where a relationship between fiducial and generalized inference was established. More background on fiducial inference and discussion of the asymptotic and empirical properties can be found in Hannig (2009b).

The principle idea of generalized fiducial inference is similar to the likelihood function and “switches” the role of the data, $\mathbf{X}$, and the model parameter(s) ξ. We use a model and the observed data, $\mathbf{X}$, to gain information about the parameter(s) ξ. We use this function to define a probability measure on the parameter space, Ξ.

To formally describe fiducial inference we assume that a relationship between $\mathbf{X}$ and ξ exists in the form of

$$\mathbf{X} = G(\xi, \mathbf{U}) \tag{1}$$

where $G(\cdot, \cdot)$ is called the *structural equation* and $\mathbf{U}$ is a random vector with a completely known distribution and independent of any parameters. The parameter ξ and the random vector $\mathbf{U}$ will determine the distribution of $\mathbf{X}$. After $\mathbf{X}$ is observed the role of the data and the parameter can be switched and one can infer a distribution on ξ from what we know of the distribution of $\mathbf{U}$. If Eq. 1 can be inverted the inverse will be written as $G^{-1}(\cdot, \cdot)$. For an observed $\mathbf{x}$ and $\mathbf{u}$ we can calculate ξ from

$$\xi = G^{-1}(\mathbf{x}, \mathbf{u}). \tag{2}$$

Because of this inverse relationship we can generate a random sample of $\mathbf{u}_1', \mathbf{u}_2', \ldots, \mathbf{u}_M'$ and obtain a random sample for ξ: $\xi_1' = G^{-1}(\mathbf{x}, \mathbf{u}_1'), \xi_2' = G^{-1}(\mathbf{x}, \mathbf{u}_2'), \ldots, \xi_M' = G^{-1}(\mathbf{x}, \mathbf{u}_M')$. This sample is called a *fiducial sample* and can be used to calculate estimates and confidence intervals for the true parameter(s), ξ_0.

Two potential times that $G^{-1}(\cdot, \cdot)$ may not exist are listed in Hannig and Lee (2009). They occur when (i) there is no ξ that satisfies Eq. 2 or (ii) there is more than one ξ that satisfies Eq. 2. From Hannig (2009b) we will handle situation (i) by eliminating such $\mathbf{u}$'s and re-normalizing the sampling probabilities. This is reasonable because we know our data was generated using ξ_0 and $\mathbf{u}_0$. Consequently, we know that there is at least one solution for Eq. 2 when $\mathbf{u}_0$ is considered; we will only consider the $\mathbf{u}$'s that allow for $G^{-1}(\cdot, \cdot)$ to exist. Hannig (2009b) suggests that situation (ii) is handled by picking an ξ by some, possibly random, rule that satisfies the inverse in Eq. 2.

A more rigorous definition of the inverse is the set valued function of

$$Q(\mathbf{x}, \mathbf{u}) = \{\xi : \mathbf{x} = G(\xi, \mathbf{u})\}. \tag{3}$$

As we previously noted we know that our observed data was generated using some value of the model parameter, ξ_0, and random vector, $\mathbf{u}_0$. Thus, we know the distribution of $\mathbf{U}$ and that $Q(\mathbf{x}, \mathbf{u}_0) \neq \emptyset$. Using these two facts we can compute the *generalized fiducial distribution* from

$$V(Q(\mathbf{x}, \mathbf{U}^\star)) | \{Q(\mathbf{x}, \mathbf{U}^\star) \neq \emptyset\} \tag{4}$$

where $\mathbf{U}^\star$ is an independent copy of $\mathbf{U}$ and $V(S)$ is a random element for any measurable set, S, with support on the closure of S, $\bar{S}$ (i.e. $V(\cdot)$ is the random rule for selecting the possible ξ's). To simplify references to the generalized fiducial distribution in this manuscript, we will refer to a random vector that has a distribution given by Eq. 4 as $\mathcal{R}_\xi$.

For a more detailed discussion of the derivation of the generalized fiducial distribution see Hannig (2009b). From the distribution, we can also calculate the *generalized fiducial density* as proposed in Hannig (2009a, b). In these papers, there is also a theoretical justification of generalized fiducial inference. In particular, it is demonstrated that the confidence intervals based on the generalized fiducial density will have, under some regularity conditions, asymptotically correct coverage.

To compute the fiducial density, we will additionally assume that the structural equation (1) can be written as $X_i = g_i(\xi, U_i)$, $i = 1, \ldots, n$. Here $\mathbf{X} = (X_1, \ldots, X_n)$, $\mathbf{G} = (g_1, \ldots, g_n)$ and $\mathbf{U} = (U_1, \ldots, U_n)$ with U_i being independent random variables or vectors of known distribution. Note that ξ is a $p \times 1$ vector and let $\mathbf{X}_0 = (X_{1,} \ldots, X_p)$, $\mathbf{X}_c = (X_{p+1,} \ldots, X_n)$, $\mathbf{U}_0 = (U_1, \ldots, U_p)$, $\mathbf{U}_c = (U_{p+1}, \ldots, U_n)$, and assume that $\mathbf{G} = (\mathbf{G}_0, \mathbf{G}_c)$ where $\mathbf{X}_0 = \mathbf{G}_0(\xi, \mathbf{U}_0)$ and $\mathbf{X}_c = \mathbf{G}_c(\xi, \mathbf{U_c})$. Finally, we will assume that the functions $\mathbf{G}_0$ and $\mathbf{G}_c$ are one-to-one and differentiable. After establishing these relationships we can now follow the prescribed recipe to calculate the generalized fiducial density.

If we were to use the first p structural equations in $\mathbf{X}_0 = \mathbf{G}_0(\xi, \mathbf{U}_0)$ when we observe $\mathbf{x}_0$ and $\mathbf{u}_0$ the inverse with respect to $\mathbf{u}_0$ would be, $\mathbf{u}_0 = \mathbf{G}_0^{-1}(\xi, \mathbf{x_0})$. The resulting generalized fiducial density is

$$f_{\mathcal{R}_\xi}(\xi) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi) J_0(\mathbf{x}_0, \xi)}{\int_\Xi f_{\mathbf{X}}(\mathbf{x}|\xi') J_0(\mathbf{x}_0, \xi') d\xi'}$$

where

$$J_0(\mathbf{x}_0, \xi) = \left| \frac{\det\left(\frac{d}{d\xi}\mathbf{G}_0^{-1}(\mathbf{x}_0, \xi)\right)}{\det\left(\frac{d}{d\mathbf{u}_0}\mathbf{G}_0^{-1}(\mathbf{x}_0, \xi)\right)} \right|.$$

This depends on the choice of $\mathbf{G}_0$ and $\mathbf{G}_c$. Choosing $\mathbf{G}_0$ from the first p equations is rather arbitrary so we will average over all possible p combinations. Specifically, we will denote $\mathbf{X_i} = \mathbf{G}_{0,\mathbf{i}}(\xi, \mathbf{U_i})$ where $\mathbf{X_i} = (X_{i_1}, \ldots, X_{i_p})$ and $\mathbf{U_i} = (U_{i_1}, \ldots, U_{i_p})$ for all possible combinations of the indices $\mathbf{i} = (i_1, \ldots, i_p)$. This will produce the generalized fiducial density

$$f_{\mathcal{R}_\xi}(\xi) = \frac{f_{\mathbf{X}}(\mathbf{x}|\xi) J(\mathbf{x}, \xi)}{\int_\Xi f_{\mathbf{X}}(\mathbf{x}|\xi') J(\mathbf{x}, \xi') d\xi'} \tag{5}$$

where

$$J(\mathbf{x}, \xi) = \binom{n}{p}^{-1} \sum_{\mathbf{i}=(i_1,\ldots,i_p)} \left| \frac{\det\left(\frac{d}{d\xi}\mathbf{G}_{0,\mathbf{i}}^{-1}(\mathbf{x_i}, \xi)\right)}{\det\left(\frac{d}{d\mathbf{u_i}}\mathbf{G}_{0,\mathbf{i}}^{-1}(\mathbf{x_i}, \xi)\right)} \right| \tag{6}$$

is the mean of all subsets where $1 \leq i_1 < \cdots < i_p \leq n$ and the determinants in Eq. 6 are the appropriate Jacobians.

The generalized fiducial distribution describes our belief about possible values of the parameters. It can be used in the same way as the Bayesian posterior to define point estimators, confidence sets, etc.

3 Main results

3.1 Structural equation

3.1.1 Known threshold

If $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables from the GPD with parameters γ and σ then a structural equation, when a is known, can be defined as

$$X_i = a + \left(U_i^{-\gamma} - 1\right)\frac{\sigma}{\gamma}$$

where U_i for $i = 1, \ldots, n$ are independent random variables from the $U(0, 1)$ distribution.

Following the recipe from Eq. 5 we were able to calculate the fiducial density for $\xi = (\gamma, \sigma)$. The generalized fiducial density when a is known is

$$\begin{aligned} f_{\mathcal{R}_\xi}(\xi) \propto \frac{1}{\sigma^n}\prod_{i=1}^{n}\left(1+\frac{\gamma\,(x_i-a)}{\sigma}\right)^{-\frac{1}{\gamma}-1} I_{(a,\infty)}(x_{(1)})\, I_{(0,\infty)}(\sigma)\, I_{(-\frac{\sigma}{x_{(n)}-a},\infty)}(\gamma) \\ \times \binom{n}{2}^{-1}\frac{1}{\gamma^2}\sum_{i<j}\Bigg|(x_i-a)\left(1+\frac{\gamma\,(x_j-a)}{\sigma}\right)\log\left(1+\frac{\gamma\,(x_j-a)}{\sigma}\right) \\ -\left(x_j-a\right)\left(1+\frac{\gamma\,(x_i-a)}{\sigma}\right)\log\left(1+\frac{\gamma\,(x_i-a)}{\sigma}\right)\Bigg| \end{aligned} \tag{7}$$

where $\mathcal{R}_\xi = (\mathcal{R}_\gamma, \mathcal{R}_\sigma)$ is the fiducial random variable for (γ, σ) and $x_{(i)}$ is the order statistic for $i = 1, \ldots, n$.

To find the fiducial density for the β-quantile (return level) a transformation on Eq. 7 is needed. Namely, we need to find the distribution of

$$\mathcal{R}_q = a + \frac{\mathcal{R}_\sigma}{\mathcal{R}_\gamma}\left((1-\beta)^{-\mathcal{R}_\gamma} - 1\right) \qquad \beta \in (0, 1) \tag{8}$$

where $\mathcal{R}_q$ is the fiducial random variable associated with the β-quantile. The fiducial density for the β-quantile is

$$f_{\mathcal{R}_q}(q) \propto \int f_{\mathcal{R}_{\gamma q}}(\gamma, q)\, d\gamma \tag{9}$$

where $f_{\mathcal{R}_{\gamma q}}(\gamma, q)$ is the joint density of $\left(\mathcal{R}_\gamma, \mathcal{R}_q\right)$ using the transformation in Eq. 8.

3.1.2 Unknown threshold

In most practical applications the threshold is unknown and the GPD does not fit the tail of the distribution exactly. As a result, we consider the following model as an approximation:

$$X_i = I_{(0,p)}(U_i)\,(aW_i) + I_{(p,1)}(U_i)\left(a + \left(W_i^{-\gamma} - 1\right)\frac{\sigma}{\gamma}\right) \tag{10}$$

where U_i and W_i for $i = 1, \ldots, n$ are independent random variables from the $U(0, 1)$ distribution. Furthermore, p is chosen so the density based on Eq. 10 is continuous at a. Notice that Eq. 10 is a structural equation of the form given in Eq. 1 with $\mathbf{U} = \{(U_i, W_i), i = 1, \ldots, n\}$. For computational stability, we also assume that it is known a priori that at least B observations are above the threshold and at least one is below the threshold, a.

As we are only interested in the data above the threshold the model below the threshold serves as a way of introducing a penalty. The penalty helps to ensure that the threshold is not forced far into the tail of the distribution. The uniform distribution, while not the correct distribution of the data below the threshold, seems to introduce the correct penalty for selecting the threshold as demonstrated by our simulation studies.

Using this structural equation, the recipe described earlier produces the generalized fiducial density as

$$f(\xi) \propto \sum_{i=1}^{n-B} \left\{ \frac{I_{(x_{(i)}, x_{(i+1)})}(a) J\left(\mathbf{x}_{i+1:n}, \xi\right)}{(a+\sigma)^n} \prod_{j=i+1}^{n} \left[\left(1 + \frac{\gamma\left(x_{(j)} - a\right)}{\sigma} \right)^{-\frac{1}{\gamma}-1} \right] \right\} \tag{11}$$

where

$$\begin{aligned} J\left(\mathbf{x}_{i+1:n}, \xi\right) = & \binom{n-i-1}{3}^{-1} \frac{1}{\gamma^2} \\ & \times \sum_{1 \le i < j < k < l \le n} \Bigg| (x_j - x_l) \left(1 + \frac{\gamma(x_k - a)}{\sigma}\right) \log\left(\frac{\gamma(x_k - a)}{\sigma} + 1\right) \\ & - (x_k - x_l) \left(1 + \frac{\gamma(x_j - a)}{\sigma}\right) \log\left(\frac{\gamma(x_j - a)}{\sigma} + 1\right) \\ & - (x_j - x_k) \left(1 + \frac{\gamma(x_l - a)}{\sigma}\right) \log\left(\frac{\gamma(x_l - a)}{\sigma} + 1\right) \Bigg|, \end{aligned}$$

$\mathbf{x}_{j:n} = \left(x_{(j)}, x_{(j+1)}, \ldots, x_{(n)}\right)$ is a vector of order statistics ($x_{(0)} = 0$) and $B < n$. Note that B forces a certain number of values in the tail to be fit with a GPD to ensure that the threshold is not selected too close to $x_{(n)}$. In our simulations (discussed later) we chose $B = 10$.

Finally we remark that the $(a + \sigma)^{-n}$ portion is a product of using the $U(0, a)$ distribution below the threshold and down-weights the generalized fiducial density for large threshold values. We have chosen the $U(0, a)$ below the threshold because it gives good empirical properties and leads to a simple form of the penalty term, $(a + \sigma)^{-n}$. Other choices for the distribution below the threshold are possible and would lead to a more complicated penalty term, which we do not investigate here.

3.2 Confidence intervals, coverage, and point estimates when a is known

3.2.1 *Confidence intervals and coverage*

Based on the fiducial densities in Eqs. 7 and 9 we defined point estimates and confidence intervals for the parameters. The point estimates are defined as the median of the marginal distributions of Eqs. 7 and 9. We constructed the equal tailed confidence region for the true parameters (γ_0, σ_0) and one and two sided intervals for the true high quantile, q_{β_0}. First, one sided lower and upper tailed intervals for q_{β_0} are defined as (c_1, ∞) and $(0, c_2)$ respectively. The values c_1 and c_2 are the α and $1-\alpha$ quantiles of Eq. 9. Two tailed intervals were calculated in two different ways. A symmetric $(1-\alpha)\,100\%$ interval is obtained by combining two $1-\frac{\alpha}{2}$ one tailed intervals to get (c_1, c_2). The second two tailed interval is defined as

$$\left\{(d_1, d_2) : \arg\min_{d_1,d_2}\left\{(d_2-d_1),\ \int_{d_1}^{d_2} f_{\mathcal{R}_Q}(q)dq = 1-\alpha\right\}\right\}. \tag{12}$$

The interval (c_1, c_2) will be referred to as the "fiducial symmetric interval" and the interval in Eq. 12 will be referred to as the "fiducial shortest interval". Likewise, we define the equal tailed joint confidence region for the true parameters (γ_0, σ_0) as

$$\begin{aligned} C(\mathbf{X}) = \Bigg\{(\gamma,\sigma):\ & A_1 = \int_0^\infty \int_{\frac{-\sigma}{x_{(n)}-a}}^{d_1} f_{\mathcal{R}_\xi}(\gamma,\sigma)d\gamma d\sigma, \\ & A_2 = \int_0^{d_2} \int_{\frac{-\sigma}{x_{(n)}-a}}^{\infty} f_{\mathcal{R}_\xi}(\gamma,\sigma)d\gamma d\sigma,\ A_3 = \int_0^\infty \int_{d_3}^\infty f_{\mathcal{R}_\xi}(\gamma,\sigma)d\gamma d\sigma, \\ & A_4 = \int_{d_4}^\infty \int_{\frac{-\sigma}{x_{(n)}-a}}^{\infty} f_{\mathcal{R}_\xi}(\gamma,\sigma)d\gamma d\sigma,\ \text{where } d_1, d_2, d_3, d_4 \text{ satisfy} \\ & A_1 = A_2 = A_3 = A_4,\ \int_{d_2}^{d_4}\int_{d_1}^{d_3} f_{\mathcal{R}_\xi}(\gamma,\sigma)d\gamma d\sigma = 1-\alpha\Bigg\}. \end{aligned} \tag{13}$$

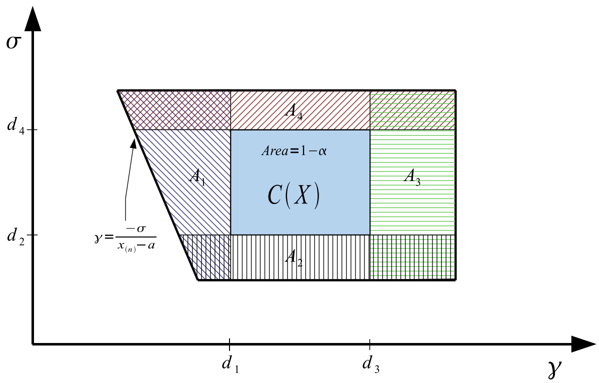


Fig. 1 Equal tailed confidence region for (γ_0, σ_0)

Table 1 Simulation values when a is known

γ_0	-0.2	0	0.2	0.4	0.6
σ_0	1				
n	20	50	200		

See Fig. 1 for illustration.

The traditional way to assess the coverage of confidence intervals is to choose a desired confidence level (e.g. 95%), simulate data, then check the frequency in which the true parameter(s) is/are captured by the constructed interval(s). Alternatively, we use a graphic device demonstrated in Hannig (2009b), which allows us to check the coverage at all confidence levels simultaneously. To accomplish this for the one tailed intervals for the β-quantile set $Q_\beta(\mathbf{X}, q_{\beta_0}) = P(\mathcal{R}_q < q_{\beta_0}|\mathbf{X})$. This is essentially the smallest coverage level of an upper tailed confidence interval that will contain the true quantile, q_{β_0}. If the confidence interval for q_{β_0} were exact at all confidence levels then $Q_\beta(\mathbf{X}, q_{\beta_0})$ (which can be thought of as a p-value) would follow the $U(0, 1)$ distribution.

We generated 1,000 data sets from a generalized Pareto distribution with the parameter values seen in Table 1. An MCMC algorithm was used to draw a sample from Eq. 7. Each generated data set produced one $Q_\beta(\mathbf{X}, q_{\beta_0})$ value which we used to construct $U(0, 1)$ QQ-plots. To assess the coverage look at the nominal coverage and then note the corresponding actual coverage that coincides with the simulated line. For example, the dotted line in the first two plots of Fig. 2 reflect that the 0.95 lower and upper tailed intervals have actual coverage of 0.948 and 0.962 respectively. Figures 3 and 4 are additional QQ-plots for specified parameter values. In general, intervals with exact coverage would follow the diagonal line in the plots. Because there is variation due to simulation we provide confidence bands (dashed lines). If the observed $Q_\beta(\mathbf{X}, q_{\beta_0})$ values (simulated line) stay within the dashed lines (95% confidence bands) then they cannot be distinguished from a sample of the $U(0, 1)$ distribution and we claim good coverage properties.

A similar calculation can be done with the two tailed interval (c_1, c_2), the interval in Eq. 12, and the joint confidence region in Eq. 13. The coverage for the intervals of the 0.99-quantile and the joint confidence region are seen in Figs. 2, 3, and 4. The figures with the titles "Upper tailed", "Lower tailed", and "Symmetric" coincide with the previously described intervals of $(0, c_2)$, (c_1, ∞), and (c_1, c_2) respectively. The

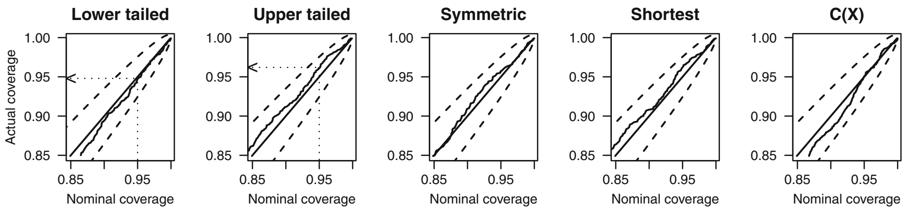


Fig. 2 QQ-plots when $\gamma_0 = -0.2$, $\sigma_0 = 1$, and $n = 50$

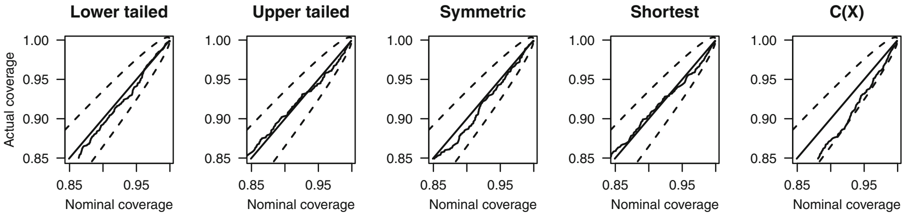


Fig. 3 QQ-plots when $\gamma_0 = 0$ and $\sigma_0 = 1$, and $n = 50$

figures titled "Shortest" and "$C(X)$" coincide with the intervals defined in Eqs. 12 and 13, respectively. As the plots illustrate all of the intervals are very close to achieving the nominal coverage rate. This behavior was also seen for all of the γ_0, σ_0, and sample size combinations in Table 1.

3.2.2 Asymptotic properties

For the parametric model with a known threshold we proved that the intervals described previously have asymptotically correct frequentist coverage. If the threshold is not known the relevant model is no longer parametric. Asymptotic properties of the generalized fiducial distribution for non-parametric and semi-parametric problems are still not fully understood. See Hannig and Lee (2009) for an early result in this direction.

We have verified that the confidence region in Eq. 13 has asymptotically correct coverage using Theorem 5.1 from Hannig (2009a). Similarly, because the β-quantile is a differentiable function of γ and σ it follows directly from the delta method that the intervals for the β-quantile are also asymptotically correct. These calculations have been relegated to Appendix A.

3.2.3 Interval comparisons

Our assessment of the confidence intervals for the β-quantile also involved a comparison to intervals constructed using the profile log-likelihood described in Coles (2001) and the Bayesian approach using Jefferys prior described in Castellanos and Cabras (2005). The coverage for two tailed 95 and 99% intervals based on the γ_0 values in Table 1 for the 0.99-quantile are seen in Fig. 5. Each dot represents the

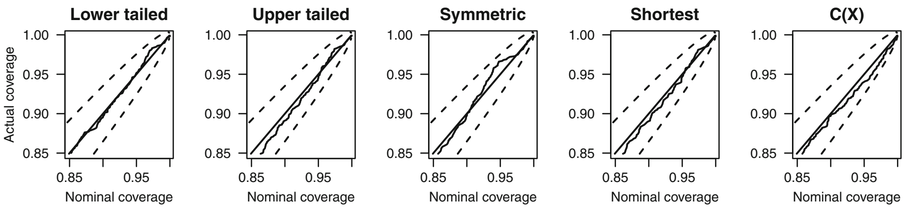


Fig. 4 QQ-plots when $\gamma_0 = 0.4$ and $\sigma_0 = 1$, and $n = 50$

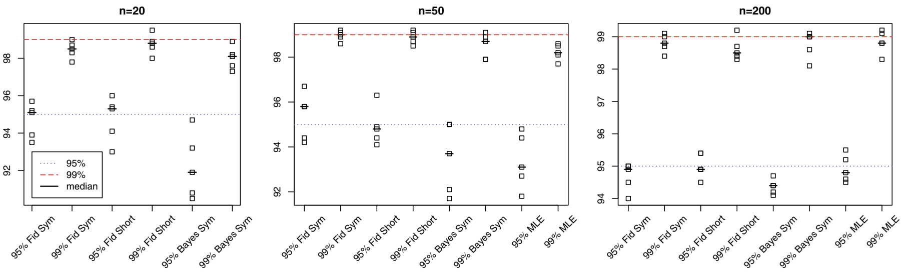


Fig. 5 Coverage of the 95 and 99% confidence intervals for the 0.99-quantile with various γ_0 values and $\sigma_0 = 1$

coverage for a simulation using a particular γ_0 value. Due to convergence problems the profile log-likelihood could not always be calculated for the small sample sizes. As the plots illustrate, the fiducial intervals are very close to the desired coverage rate while the Bayesian and profile log-likelihood methods tend to be slightly liberal. Figure 6 demonstrates the lengths of the intervals when $\gamma_0 > 0$. The mean (denoted by the triangle) and the median length of the fiducial shortest interval described in Eq. 12 is less than its competitors for all sample sizes. When $\gamma_0 \leq 0$ similar results are seen. However, the mean and median lengths of the fiducial intervals are slightly longer than the competitors when $n = 20$ and 50. The fiducial shortest interval is shorter than its competitors when $n = 200$.

3.2.4 *Point estimates comparisons*

Comparisons of the point estimates of our method, the Bayesian estimates, the MLE estimates, and estimates based on L-moments were also performed. The results were similar for all γ_0 values, we report these comparisons in Fig. 7. Again, the MLE estimates at small sample sizes occasionally had convergence issues so we limit the MLE comparisons to the sample sizes of $n = 50$ and 200. The fiducial estimates for γ_0 and σ_0 have smaller bias than all of the competitors. The Bayesian estimate for the 0.99-quantile is slightly less biased. All of the methods have similar variability amongst the estimates.

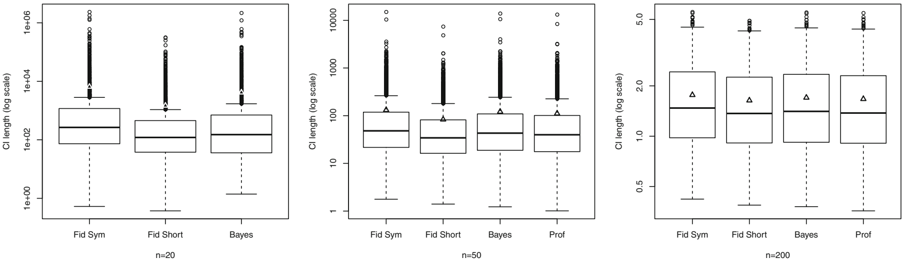


Fig. 6 Length of the fiducial, Bayesian, and profile log-likelihood confidence intervals for the 0.99-quantile when $\gamma_0 > 0$

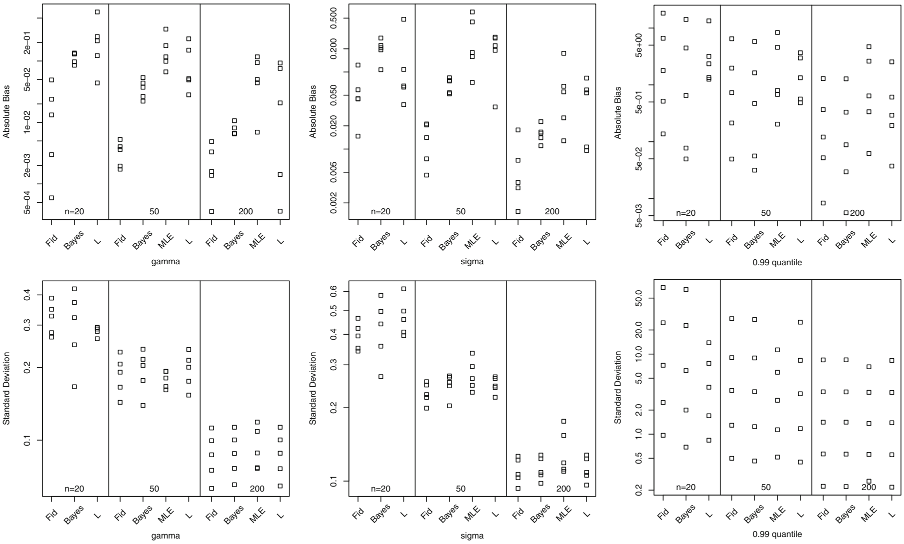


Fig. 7 Absolute bias and standard deviation of the point estimates using distributions with parameter values given in Table 1

3.3 Confidence intervals and coverage when a is unknown

3.3.1 Confidence intervals and coverage

Similar analysis can be done when the threshold, a, is unknown. We have derived the generalized fiducial density when a is treated as an unknown parameter in Eq. 11. We assume the data came from a mixture of the $U(0, a)$ and the GPD beyond a. As a result, all computations were done on the transformed data set of $X' = X - X_{(1)}$ where $X_{(1)}$ is the minimum and then back transformed to the original scale. Using the transformation in Eq. 8 we can calculate the generalized fiducial density for the β-quantile of the X' data set. A Metropolis-Hastings algorithm allowed us to draw a sample from Eq. 11 and calculate intervals for the parameters in the same manner that was discussed earlier.

To assess the usefulness of our method we applied it to general data sets. We generated 1,000 data sets from the distributions listed in Table 2 with a sample size of 1,000 and assessed the confidence intervals for the β-quantile. Figures 8, 9, and 10 reflect the coverage for the 0.99 and 0.999 quantiles when the underlying distributions of the data are Exp(1), $t(5) + 10$, and $N(10, 100)$. Our intervals are very close to achieving the nominal coverage rate in those scenarios with the exception of the

Table 2 Distributions of the simulated data

Exp(1)	$t(5) + 10$
Exp(5)	$t(10) + 10$
Gamma(5, 1)	$N(10, 100)$

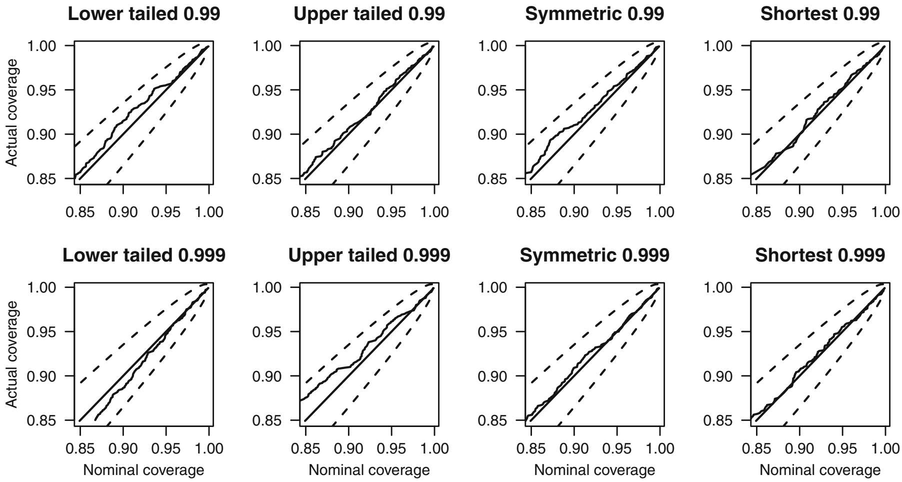


Fig. 8 QQ-plots for the 0.99 and 0.999 quantile when $X \sim \text{Exp}(1)$ using the fiducial method

two tailed intervals for the 0.99-quantile in Fig. 9. Those intervals are slightly liberal. Similar results were seen when the data was generated from the other distributions.

3.3.2 Comparisons

In addition to checking the coverage we also compared our method to a similar Bayesian method described in Cabras and Castellanos (2009). This method used a mixture of a normal, Weibull, or nonparametric model below the threshold and a GPD above the threshold. Obviously, this method depends on which central model is

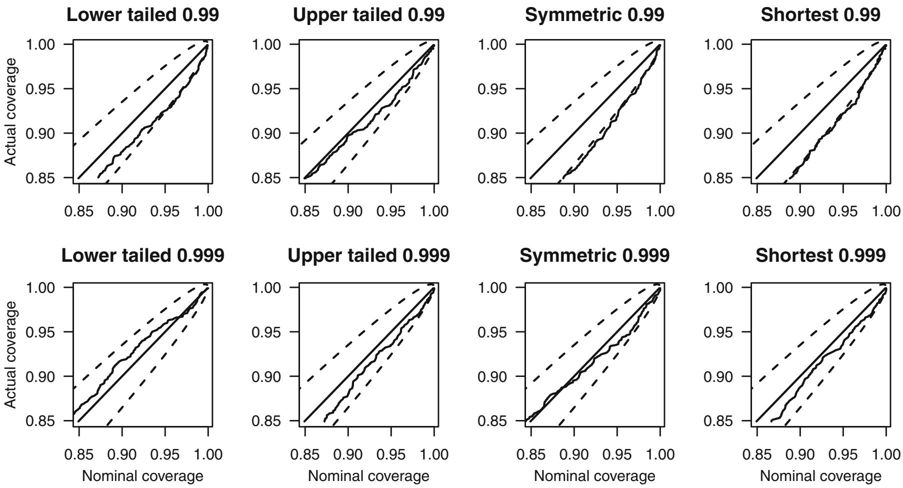


Fig. 9 QQ-plots for the 0.99 and 0.999 quantile when $X \sim t(5) + 10$ using the fiducial method

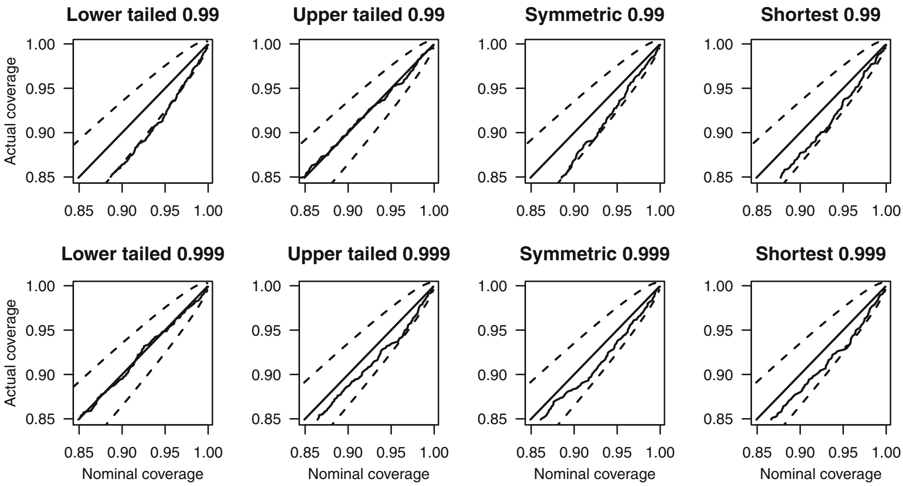


Fig. 10 QQ-plots for the 0.99 and 0.999 quantile when $X \sim N(10, 100)$ using the fiducial method

chosen. We found that the nonparametric model often chose a threshold far into the tail and estimated the extreme quantiles using the prescribed nonparametric distribution. As a result, there was very little variation in the MCMC chain which caused very liberal confidence intervals that rarely contained the true quantile. When the data came from a $t(5)$+10 the nonparametric method does a better job but is still rather liberal in its two tailed intervals, seen in Fig. 11. Using a normal central model worked well when the data was normal and when it was very different from a normal distribution, seen in Figs. 12 and 13. When the data came from a t distribution the normal central model attempted to fit the bulk of the data and forced the threshold into the tail. This caused an underestimation of the quantiles and produced very liberal upper tailed and symmetric intervals, seen in Fig. 14. The Weibull central model

Fig. 11 QQ-plots for the 0.99 and 0.999 quantile when $X \sim t(5) + 10$ using the Bayesian method with a nonparametric central model

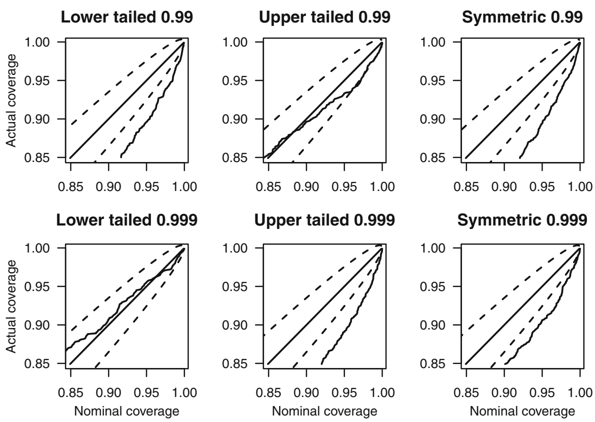


Fig. 12 QQ-plots for the 0.99 and 0.999 quantile when $X \sim \text{Exp}(1)$ using the Bayesian method with a normal central model

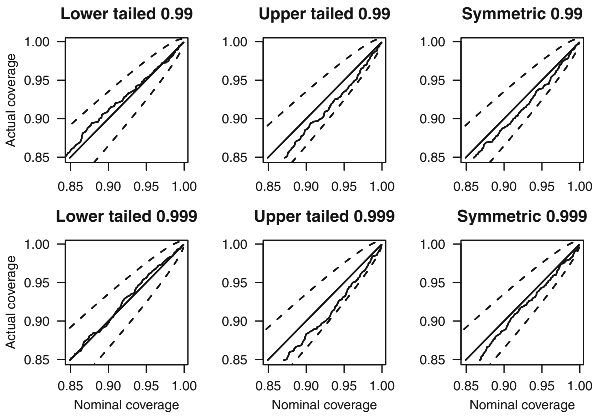


Fig. 13 QQ-plots for the 0.99 and 0.999 quantile when $X \sim N(10, 100)$ using the Bayesian method with a normal central model

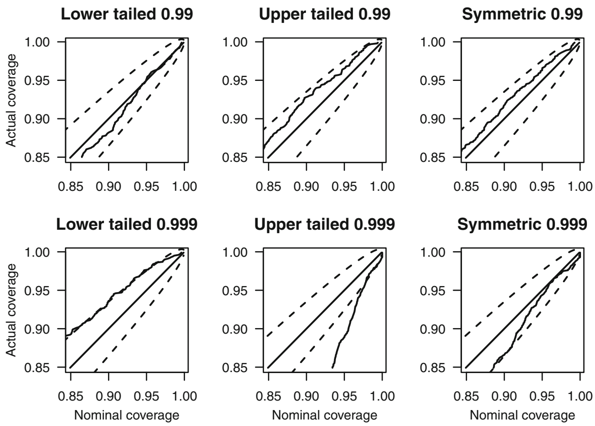
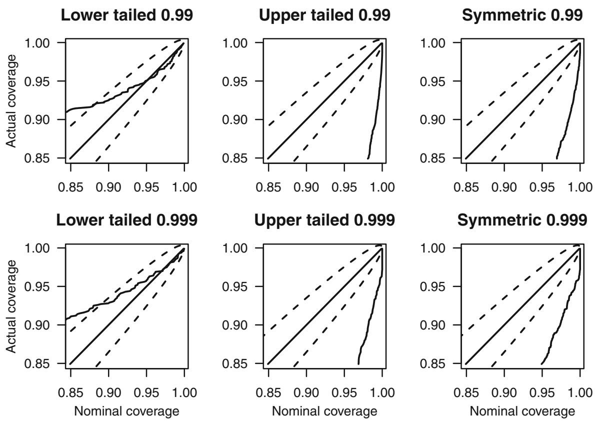


Fig. 14 QQ-plots for the 0.99 and 0.999 quantile when $X \sim t(10) + 10$ using the Bayesian method with a normal central model



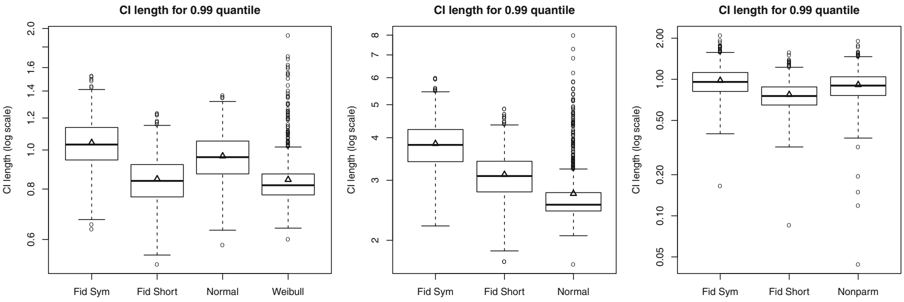


Fig. 15 Length of the two tailed fiducial and Bayesian intervals for the 0.99-quantile when the data was generated from Exp(1), Normal(10, 100), and $t(5) + 10$ respectively

produced reasonable coverage when the data came from an exponential distribution. The other cases were not appropriate for a Weibull central model.

Figure 15 illustrates the lengths of the two tailed intervals when the methods had reasonable coverage rates. The first plot reflects the Bayesian approach with a Weibull central model is the shortest. The fiducial shortest interval is only slightly longer in this case when the data came from an Exp(1). In the second plot the data was generated from a $N(10, 100)$ and, as expected, the Bayesian method that fits a normal model is shorter than both of our fiducial methods. In the third plot the data came from a $t(5) + 10$ and our fiducial shortest interval was the best. The fiducial symmetric interval was only slightly longer than the Bayesian method using a nonparametric central model.

4 NASDAQ 100 data set

We analyzed a data set of the log-weekly losses of the NASDAQ 100 index . The data consists of 1222 weeks from October 1985 to March 2009.

Using the ad hoc approaches like the mean residual life plot in Fig. 16, suggested in Coles (2001), a fixed threshold ranging from 0.03 to 0.05 would be appropriate. Figure 16 also shows some fit diagnostics for the GPD when $a = 0.04$. Neither plot would suggest that the GPD is not appropriate.

The estimates and confidence intervals for the 0.99-quantile can be seen in Table 3. The estimates for the 0.99-quantile are all very similar and the fiducial shortest interval is equivalent or shorter than the Bayesian interval and the profile log-likelihood interval. Considering that the Bayesian and profile log-likelihood methods were both slightly liberal and the fiducial intervals were very close to exact the fiducial interval shows improvement.

Like Bayesian analysis we could also look at the density of any parameter. Figure 17 illustrates the fiducial density and Bayesian posterior for γ when $a = 0.04$. As the plots show, both methods produce a very similar density.

When the threshold is unknown the Bayesian method using the normal central model was slightly shorter than the fiducial intervals. As demonstrated in the

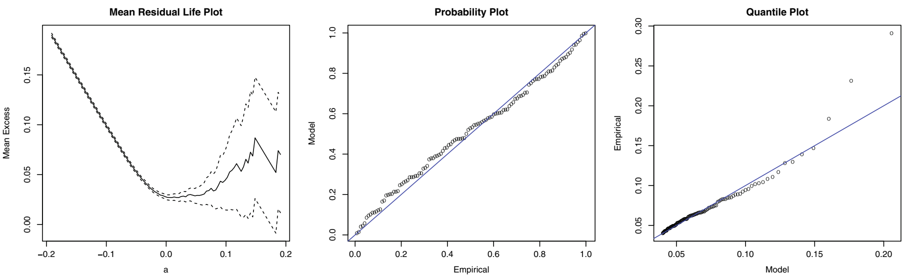


Fig. 16 Mean residual life plot for the NASDAQ 100 data set and probability and quantile plots when $a = 0.04$

simulations, when the data came from a normal distribution the Bayesian method with a normal central model produced the shortest intervals. However, as normal QQ-plots would suggest, it is not likely that these data came from a normal distribution. As a result, the coverage of the Bayesian interval with a normal central model may be called into question. The fiducial shortest interval was equal in length to the Bayesian method that used a nonparametric central model and the symmetric interval was slightly longer. When the threshold was unknown the fiducial method still

Table 3 Estimates and confidence intervals for the 0.99-quantile of the NASDAQ 100 data set

Method	a (fixed)	$q_{0.99}$ estimate	$q_{0.99}$ 95% CI
Fiducial symmetric interval	0.03	0.109	(0.096, 0.128)
	0.04	0.109	(0.097, 0.128)
	0.05	0.107	(0.095, 0.125)
Fiducial shortest interval	0.03	0.109	(0.095, 0.126)
	0.04	0.109	(0.096, 0.125)
	0.05	0.107	(0.094, 0.123)
Bayesian interval	0.03	0.109	(0.097, 0.128)
	0.04	0.110	(0.097, 0.129)
	0.05	0.108	(0.095, 0.125)
Profile log-likelihood interval	0.03	0.108	(0.096, 0.126)
	0.04	0.109	(0.097, 0.126)
	0.05	0.106	(0.094, 0.124)
Method	**a (estimate)**	**$q_{0.99}$ estimate**	**$q_{0.99}$ 95% CI**
Fiducial symmetric interval	0.038	0.110	(0.096, 0.129)
Fiducial shortest interval	0.038	0.110	(0.095, 0.126)
Bayesian normal central model	0.016	0.110	(0.099, 0.129)
Bayesian nonpar central model	0.029	0.109	(0.096, 0.127)

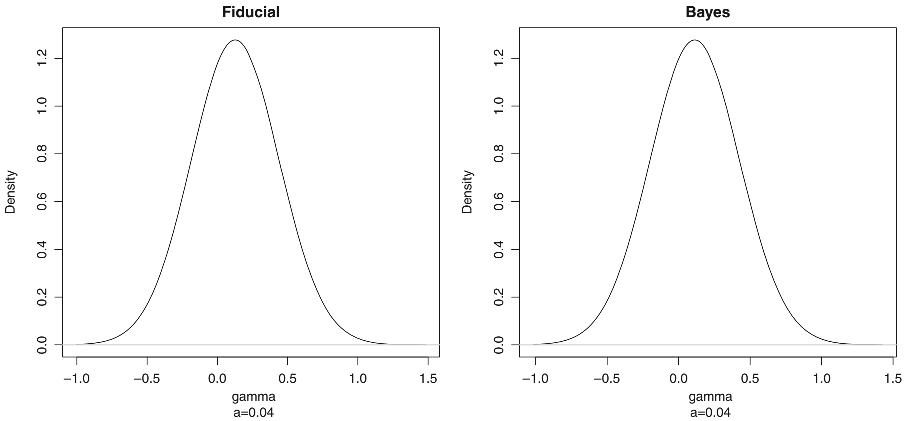


Fig. 17 Fiducial density and Bayesian posterior for γ when $a = 0.04$

produced confidence intervals for the β-quantile with good coverage. Thus, we can be confident that the reported intervals have good empirical properties. Even though similar intervals are produced, our fiducial method is more flexible as it can account for different distributions while still maintaining good coverage properties.

In the case when the threshold is unknown, our method is computationally intensive. As a result, analysis can take longer than the commonly used profile log-likelihood analysis. The benefit of our method comes from the fact that the threshold is an unknown parameter in any real life setting. Assuming that the threshold is known, as is done with the profile log-likelihood, may not be appropriate for all situations.

5 Conclusion

There has been substantial interest in peaks over threshold modeling. Various estimation techniques have been developed and methods continue to improve. The challenge in modeling peaks over threshold data come from estimating the extreme quantiles and also estimating the threshold. We developed a fiducial approach to both problems.

When the threshold is assumed to be known our fiducial approach constructs intervals and point estimates for the true parameters (γ_0, σ_0) and the β-quantile that have good small sample properties. Furthermore, all of the proposed intervals have asymptotically correct coverage. Comparisons with a Bayesian method and the profile log-likelihood approach suggest that the fiducial intervals were closer to achieving the nominal coverage rate for the 0.99-quantile and one of the fiducial intervals was shortest. Likewise, the point estimates for the shape and scale parameters using the fiducial method had the smallest bias. The point estimate for the 0.99-quantile was slightly better using the Bayesian approach.

When the threshold is unknown the proposed fiducial intervals for the β-quantile had good frequentist coverage. Our method worked well on the different data types that could be seen in real life while the competing Bayesian method did not universally work well for all data types. First, a central model had to be chosen from either a Weibull, normal, or nonparametric distribution. Even after choosing an adequate central model it was not assured that the coverage for the β-quantiles would be reasonably close to exact. When the Bayesian intervals had an acceptable coverage rate the recommended fiducial intervals were either close to the same length or shorter.

We analyzed a data set from the log-weekly losses of the NASDAQ 100 index. Our analysis demonstrated that the intervals for the fiducial method were similar in length (sometimes shorter) when compared to the competing methods in both cases where the threshold was known and unknown. Because we demonstrated that the coverage for the fiducial intervals were reasonably close to exact for any data type, we can be confident that these intervals have adequate coverage while the competitors tended to have lower than their stated coverage.

Based on our findings the fiducial approach to the generalized Pareto distribution is a viable alternative to modeling peaks over threshold data sets. The good small sample properties and the asymptotic results make this an attractive solution to a difficult problem. R code for our procedure can be downloaded at http://www.unc.edu/~hannig/.

Acknowledgments We would like to thank Johan Segers for his helpful suggestions in the early stages of this project and the referees for their comments. All of the comments and suggestions were very beneficial to help us better explain our methods.

Appendix A: Asymptotic calculation

From Hannig (2009a) we can show that the intervals, when the threshold is known, are asymptotically correct from a frequentist prospective. The assumption are easily verifiable.

First, the structural equation must be set up as

$$X_i = F^{-1}(\xi, U_i)$$

for $i = 1, \ldots, n$ and U_i are i.i.d. $U(0, 1)$. This is exactly how we set up our structural equation so this has been verified.

Second, we need $F(x, \xi)$ to be continuously differentiable in ξ for all x. This is true by definition of the GPD.

For each $(x_1, \ldots, x_p)$ the map of $(F(x_1, \xi), \ldots, F(x_p, \xi)) = (u_1, \ldots, u_p)$ as a function of ξ is one to one. To accomplish this we will show that there is only one nonzero root for γ when considering the function

$$f(\gamma) = X_2 U_1^{-\gamma} - X_1 U_2^{-\gamma} + X_1 - X_2 = 0$$

where $f(\gamma)$ is a result of solving one of the structural equations for σ and substituting into the other, without loss of generality we will also assume $X_1 > X_2$. This function has one minimum at

$$\gamma = -\frac{\log\left(\frac{X_2\log(U_1)}{X_1\log(U_2)}\right)}{\log\left(\frac{U_2}{U_1}\right)}$$

and the derivative is

$$\frac{df(\gamma)}{d\gamma} = X_1 U_2^{-\gamma}\log(U_2) - X_2 U_1^{-\gamma}\log(U_1).$$

The derivative is negative when γ is small and positive for large γ. Also, $f(\gamma)$ approaches $X_1 - X_2$ as $\gamma \to -\infty$, which is always positive. Since there is still only one minimum and the function goes through the origin there can be only one other root. Thus satisfying the assumption.

Finally, we will assume that there exists a fixed grid in $\mathbb{R}$ of the form

$$(-\infty, c_1], (c_1, c_2), \ldots, (c_n, \infty)$$

and that each observed value $x_j \in (c_j, c_{j+1}]$. This means that X_j is only observed to the resolution of the grid. This assumption is very natural in any real setting. Every observation is measured to some measurement resolution and is not an exact value.

After these assumptions are satisfied we can directly apply Theorem 5.1 from Hannig (2009a) to demonstrate that the interval in Eq. 13 and intervals for the β-quantile are asymptotically correct.

References

Cabras, S., Castellanos, M.E.: An objective Bayesian approach for threshold estimation in the peaks over the threshold model. Tech. Rep. Department of Mathematics, University of Cagliari, Italy (2009)

Castellanos, M.E., Cabras, S.: A default Bayesian procedure for the generalized Pareto distribution. J. Stat. Plan. Inference **137**, 473–483 (2005)

Choulakian, V., Stephens, M.A.: Goodness-of-fit tests for the generalized Pareto distribution. Technometrics **43**, 478–484 (2001)

Coles, S.: An introduction to statistical modeling of extreme values. Springer (2001)

Davison, A.C.: Statistical extremes and applications, chap. modelling excesses over high thresholds with an application, pp. 461–482. Reidel, Dordrecht (1984)

Davison, A.C., Smith, R.L.: Models for excedances over high thresholds. J. R. Stat. Soc., Ser. B **52**, 393–442 (1990)

Dupuis, D.J.: Exceedances over high thresholds: a guide to threshold selection. Extremes **1**, 251–261 (1999)

Fisher, R.A.: Inverse probability. Proc. Camb. Philos. Soc. **xxvi**, 528–535 (1930)

Frigessi, A., Haug, O., Rue, H.: A dynamic mixture model for unsupervised tail estimation without threshold selection. Extremes **5**, 219–235 (2002)

Guillou, A., Hall, P.: A diagnostic for selecting the threshold in extreme values theory. J. R. Stat. Soc., Ser. B **63**, 293–305 (2001)

Hannig, J.: On asymptotic properties of generalized fiducial inference for discretized data. Tech. Rep. UNC/STOR/09/02, Department of Statistics and Operations Research, The University of North Carolina, Chapel Hill (2009a)

Hannig, J.: On generalized fiducial inference. Stat. Sin. **19**, 491–544 (2009b)

Hannig, J., Iyer, H.K., Patterson, P.: Fiducial generalized confidence intervals. J. Am. Stat. Assoc. **101**, 254–269 (2006)

Hannig, J., Lee, T.C.: Generalized fiducial inference for wavelet regression. Biometrika **96**, 847–860 (2009)

Hosking, J.R.M.: L-moments: analysis and estimation of distributions using linear combinations of order statistics. J. R. Stat. Soc., Ser. B **52**, 105–124 (1990)

Hosking, J.R.M., Wallis, J.R.: Parameter and quantile estimation for the generalized Pareto distribution. Technometrics **29**, 339–349 (1987)

Lindley, D.V.: Fiducial distributions and Bayes' theorem. J. R. Stat. Soc., Ser. B **20**, 102–107 (1958)

Monfort, M.A.J.V., Witter, J.V.: Testing exponentiality against generalized Pareto distribution. J. Hydrol. **78**, 305–315 (1985)

Pickands, J.: Statistical inference using extreme order statistics. Ann. Stat. **3**, 119–131 (1975)

Smith, R.L.: Statistical Extremes and Applications, Chap. Threshold Methods for Sample Extremes, pp. 621–638. Reidel, Dordrecht (1984)

Smith, R.L.: Maximum likelihood estimation in a class of nonregular cases. Biometrika **72**, 67–90 (1985)

Tancredi, A., Anderson, C., O'Hagan, A.: Accounting for threshold uncertainty in extreme value estimation. Extremes **9**, 87–1006 (2006)

Weeranhandi, S.: Generalized confidence intervals. J. Am. Stat. Assoc. **88**, 899–905 (1993)

Zabell, S.L.: R.A. Fisher and the fiducial argument. Stat. Sci. **7**, 369–387 (1992)