

Optimal Designs for Generalized Pareto Model

Poonam Singh

Department of Statistics
University of Delhi, Delhi, India
E-mail: pbs_93@yahoo.co.in

Ashok Kumar

Department of Statistics
University of Lucknow, Lucknow, India
Corresponding author: stats.ashok@gmail.com

(Received February 27, 2019; Accepted June 2, 2019)

Abstract

The generalized Pareto model plays an important role in modelling extreme events. Hosking and Wallis (1987) discussed the parameter and quantile estimation for generalized Pareto distribution. Optimal experimental designs are used to accurately estimate the unknown parameters of the model. In this paper, locally D-, A- and E-optimal designs with two and three support points having equal and unequal weights for homoscedastic generalized Pareto regression model are obtained numerically. It is also proved that these designs are minimally supported. The results are illustrated through Norwegian fire insurance claim data.

Keywords- Fisher information matrix, Local optimality, D-optimality, minimally supported designs, Tchebycheff system.

1. Introduction

Pareto functions are very versatile and a variety of uncertainties can be usefully modeled by them such as life time models in actuarial sciences, survival analysis and growth models in economics, finance etc. Revankar et al., (1974) introduced the Pareto II or Lomax distribution for which the density function is given as:

$$f(x, c, a) = \frac{ac^a}{(x+c)^{a+1}} ; \quad x > 0, c > 0, a > 0$$

Pickands (1975) introduced the generalized Pareto distribution for which the density function is given by

$$f(x, c, k) = \frac{1}{k} \left(1 - \frac{cx}{k}\right)^{\frac{1-c}{c}} I \left(x > 0, \frac{cx}{k} < 1\right) ; k > 0, -\infty < c < \infty \quad (1)$$

The generalized Pareto model has applications in a number of fields, including reliability studies, in the modelling of large insurance claims, as a failure time distribution. Also it is frequently used in the study of income distribution and in the analysis of extreme events, e.g. for the analysis of the precipitation data, in the flood frequency analysis, in the analysis of the data of greatest wave heights or sea levels, maximum winds loads on buildings, in the maximum rain fall analysis. Hosking and Wallis (1987) discussed the parameter and quantile estimation for generalized Pareto

distribution. Castillo and Daoudi (2009) studied the profile-likelihood function of the generalized Pareto distribution. Hambuckers et al. (2016) derived the asymptotic properties of the resulting parameter estimators for a semiparametric model for generalized Pareto regression based on a dimension reduction assumption.

The choice of experimental design is very important in order to accurately estimate the unknown parameters of the model and efficiently improve the quality of statistical inferences. The strategy attempts to choose the design which maximize the accuracy of the resulting parameters estimate and the design is termed as optimal design. Optimal designs are widely studied and used in scientific investigation as they provide efficient plans for generating and analyzing data. Theory and applications of optimal designs are discussed by Silvey (1980), Pukelsheim (1993) and Atkinson et al. (2007).

Let χ be the design space and Ξ be the class of probability distributions on the borel set of χ , then any $\xi \in \Xi$ is called a design measure or an approximate design. Let θ denotes the column vector of k unknown parameters in the model and $M(\xi, \theta)$ denotes the Fisher information matrix induced by the design ξ , then $M(\xi, \theta)$ is a symmetric matrix of order k . Let ϕ be a real-valued function defined on $k \times k$ symmetric matrices and bounded above on $\{M(\xi, \theta): \xi \in \Xi\}$. An optimal design problem is concerned with finding ξ^* such that $\phi(M(\xi^*, \theta)) = \min_{\xi \in \Xi} \phi(M(\xi, \theta))$, which is called a ϕ -optimal design. Li and Majumdar (2008) obtained D-optimal designs for logistic models with three and four parameters. Widiharis et al. (2013) investigated the D-optimal design for weighted exponential and generalized exponential models. They considered two support point designs with equal weights to obtain D-optimal design. Singh and Kumar (2018) investigated D-optimal designs for exponentiated Pareto II and generalized exponentiated Pareto models. They considered design with two support points having equal weights to obtain D-optimal design. Yang (2008) discussed A-optimal designs for generalized linear models with two parameters. Dette and Haines (1994) obtained the E-optimal designs for linear and nonlinear models with two parameters. Heiligers (1994) obtained E-optimal designs for weighted polynomial regression. Dette et al. (2006) investigated locally E-optimal designs for exponential regression models.

In this paper, we obtain locally D-, A- and E-optimal designs for the homoscedastic generalized Pareto model (2) based on (1)

$$y = (1 - \theta_1 x)^{\theta_2} + \epsilon = \eta(x, \theta) + \epsilon \quad (2)$$

where $x > 0$ (if $\theta_1 < 0$) or $0 < x \leq 1/\theta_1$ (if $\theta_1 > 0$), $\theta_1 > 0$ and ϵ is a random error normally distributed with mean zero and variance σ^2 . Designs with two and three support points with equal and unequal weights are considered for this purpose.

2. Locally Optimal Designs

2.1 Optimal Designs for Linear Models

Consider the linear regression model

$$y = f^T(x)\theta + \epsilon \quad (3)$$

where, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)^T$ is a vector of k unknown parameters, $f(x) = (1, x, x^2, \dots, x^{k-1})^T$ is vector of regressors and ϵ is a random error normally distributed with mean zero and variance σ^2 . Let approximate design measure ξ which puts weight w_i on the distinct point x_i for $i = 1, 2, \dots, m$ be denoted by

$$\xi = \left\{ \begin{matrix} x_1, x_2, \dots, x_m \\ w_1, w_2, \dots, w_m \end{matrix} \right\} \quad (4)$$

where $w_i = \frac{n_i}{n}$, $n_i =$ number of observations at the point x_i , $n = \sum n_i$, $0 < w_i < 1$; $\sum w_i = 1$. Then the Fisher information matrix $M(\xi, \boldsymbol{\theta})$ corresponding to the approximate design ξ is given by

$$M(\xi, \boldsymbol{\theta}) = \sum_{i=1}^m w_i f(x_i) f^T(x_i) \quad (5)$$

A design is said to be D-, A- and E-optimal if it minimize, the determinant, the trace and the maximum of eigen value of the inverse of the Fisher information matrix $M(\xi, \boldsymbol{\theta})$, a symmetric matrix of order k .

2.2 Optimal Designs for Non-Linear Models

Consider the non-linear model

$$y = \eta(x, \boldsymbol{\theta}) + \epsilon \quad (6)$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)^T$ and ϵ is a random error normally distributed with mean zero and variance σ^2 .

We linearized non-linear model (6) by method of Taylor series expansion around the point $\boldsymbol{\theta}_0 = (\theta_{01}, \theta_{02}, \dots, \theta_{0k})^T$ as follows:

$$\begin{aligned} E[y] &= \eta(x, \boldsymbol{\theta}) = \eta(x, \boldsymbol{\theta}_0) + \frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \\ &= c + \left[\frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} \boldsymbol{\theta} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \end{aligned}$$

$$y_{\boldsymbol{\theta}_0}(x) = f_{\boldsymbol{\theta}_0}^T(x) \boldsymbol{\theta}_0 + \epsilon$$

$$\text{where } f_{\boldsymbol{\theta}_0}(x) = \left(\frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_1}, \frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_2}, \dots, \frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_k} \right)_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}^T.$$

In general (6) can be written as

$$y = f^T(x) \boldsymbol{\theta} + \epsilon \quad (7)$$

where $f(x) = \left(\frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_1}, \frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_2}, \dots, \frac{\partial \eta(x, \boldsymbol{\theta})}{\partial \theta_k} \right)^T$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)^T$ is a vector of k unknown parameters and ϵ is a random error normally distributed with mean zero and variance σ^2 . Then the Fisher information matrix $M(\xi, \boldsymbol{\theta})$ for model (7) is same as (5).

The Fisher information matrix for non-linear models depend on the unknown parameters, therefore, obtaining optimal designs in this case is difficult. The most widely used approach to deal with this problem is local optimality approach introduced by Chernoff (1953). Chernoff (1953) adopted an initial guess $\theta = \theta_0$ for unknown parameter vector and maximize the criteria function evaluated at guessed value of the parameters. The resulting design is termed as locally optimal design.

Definition 1. [Atkinson et al., (2007)]. *D-optimal criterion is a criterion which is obtained by minimizing $\phi(M(\xi, \theta))$, where $\phi(M(\xi, \theta)) = \log|M^{-1}(\xi, \theta)| = -\log|M(\xi, \theta)|$.*

Let the design measure $\bar{\xi}$ put the unit mass at the point x and let the measure ξ^α be given by $\xi^\alpha = (1 - \alpha)\xi + \alpha\bar{\xi}$, ($0 \leq \alpha \leq 1$) then, $M(\xi^\alpha, \theta) = (1 - \alpha)M(\xi, \theta) + \alpha M(\bar{\xi}, \theta)$ and the derivative of ϕ in the direction $\bar{\xi}$ is

$$\psi(x, \xi) = \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} [\phi\{(1 - \alpha)M(\xi, \theta) + \alpha M(\bar{\xi}, \theta)\} - \phi(M(\xi, \theta))] \quad (8)$$

The following general equivalence theorem is a useful tool for checking whether the design is optimal or not (Atkinson et al., 2007).

Theorem 1.

The generalized equivalence theorem states the equivalence of the following three condition on ξ^ :*

- (i) *The design ξ^* minimizes $\phi(M(\xi, \theta))$.*
- (ii) *The design ξ^* maximizes the minimum over χ of $\psi(x, \xi)$.*
- (iii) *The minimum over χ of $\psi(x, \xi^*)$ is equal to the zero, this minimum occurring at the points of support of the design.*

As a consequence of (iii), we obtain the further condition:

- (iv) *For any non-optimum design ξ the minimum over χ of $\psi(x, \xi) < 0$.*

For **D-optimal** criterion, we minimize

$$\phi(\xi, \theta) = -\log|M(\xi, \theta)| \quad (9)$$

\therefore Equation (8) becomes

$$\psi(x, \xi) = k - \frac{\partial f(x)}{\partial \theta} M^{-1}(\xi, \theta) \frac{\partial f(x)}{\partial \theta^T} = k - d(x, \xi) \quad (10)$$

where

$$d(x, \xi) = \frac{\partial f(x)}{\partial \theta} M^{-1}(\xi, \theta) \frac{\partial f(x)}{\partial \theta^T}$$

is dispersion function (standardized variance) associated with the design measure ξ .

From Theorem 1 condition 3, $\psi(x, \xi) \geq 0$, therefore $d(x, \xi) \leq k$

For **A-optimal** criterion, we minimize

$$\phi(\xi, \theta) = \text{tarce}(M^{-1}(\xi, \theta)) \quad (11)$$

For **E-optimal** criterion, we minimize the maximum of eigenvalue of

$$\phi(\xi, \theta) = M^{-1}(\xi, \theta) \quad (12)$$

3. Optimal Design for Generalized Pareto Model

Keeping in mind the sufficient conditions of Li and Majumdar (2008), we are giving the following sufficient conditions for obtaining A- and E- optimal designs alongwith sufficient conditions of D-optimal designs (Li and Majumdar, 2008) for non-linear regression models.

Theorem 2.

- (i) For $\chi_0 = (-\infty, \infty)$, if $\forall \xi \in \Xi, \exists \epsilon > 0$ such that every function in $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ has at most $2k + 1$ roots in the design space and the D-, A- and E-optimal design exists, then the D-, A- and E-optimal designs must be minimally supported and unique.
- (ii) Let χ be one of the following two forms: $\chi_1 = [a, \infty)$ or $\chi_2 = (-\infty, b]$. If $\forall \xi \in \Xi, \exists \epsilon > 0$ such that every function in $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ has at most $2k$ roots in the design space χ_1 or χ_2 and the D-, A- and E-optimal designs exists, then the D-, A- and E-optimal designs must be minimally supported and unique. In addition, if $\forall \xi \in \Xi, \exists \epsilon > 0$ such that every function in $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ has at most $2k - 1$ roots in the design space and the D-, A- and E-optimal designs exists, then a (for χ_1) or b (for χ_2) is one of the support point of the D-, A- and E-optimal designs.
- (iii) For $\chi_3 = [a, b]$, if $\forall \xi \in \Xi, \exists \epsilon > 0$ such that every function in $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ has at most $2k - 1$ roots in the design space χ_3 , then the D-, A- and E-optimal design must be minimally supported and unique and atleast one of the boundary points is a support point of the D-, A- and E-optimal designs. In addition, if $\forall \xi \in \Xi, \exists \epsilon > 0$ such that every function in $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ has at most $2k - 2$ roots in the design space χ_3 , then both a and b are support points of the D-, A- and E-optimal designs.

To verify sufficient conditions given in Theorem 2, it is most important to verify the maximum number of roots of $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ is according to Theorem 2, where k is number of parameters. The theory of Tchebycheff system or T-system (Karlin and Studden, 1966), chapter 11; theorem 102) plays an important role in the derivation of this result. We apply Theorem 2 to obtain D-, A- and E-optimal designs for generalized Pareto model given in (2).

The Fisher information matrix is

$$\begin{aligned} M(\xi, x) &= \left(\frac{\partial \eta(x, \theta)}{\partial \theta} \right) \left(\frac{\partial \eta(x, \theta)}{\partial \theta^T} \right) = f(x) f^T(x) \\ &= \begin{bmatrix} \theta_2^2 x^2 (1 - \theta_1 x)^{2\theta_2 - 2} & -(1 - \theta_1 x)^{2\theta_2 - 1} \theta_2 x \log(1 - \theta_1 x) \\ -(1 - \theta_1 x)^{2\theta_2 - 1} \theta_2 x \log(1 - \theta_1 x) & (1 - \theta_1 x)^{2\theta_2} (\log(1 - \theta_1 x))^2 \end{bmatrix} \\ &= \begin{bmatrix} f_{11} & f_{12} \\ f_{12} & f_{22} \end{bmatrix} \end{aligned}$$

Consider a two point design with equal weight

$$\xi = \begin{Bmatrix} x_1 & x_2 \\ 0.5 & 0.5 \end{Bmatrix} \quad (13)$$

then the Fisher information matrix will be

$$M(\xi, \theta) = \sum_{i=1}^2 w_i f^T(x) f(x) = \begin{bmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{bmatrix} \quad (14)$$

where

$$\begin{aligned} M_{11} &= 0.5f_{11}|_{x=x_1} + 0.5f_{11}|_{x=x_2} \\ M_{12} &= 0.5f_{12}|_{x=x_1} + 0.5f_{12}|_{x=x_2} \\ M_{22} &= 0.5f_{22}|_{x=x_1} + 0.5f_{22}|_{x=x_2}. \end{aligned}$$

The Fisher information matrix for three point design can also be obtained, similarly. Now, we apply Theorem 2 to see the existence of minimally supported designs and their uniqueness. Let m_{ij} denote the $(i, j)^{th}$ element of $M^{-1}(\xi, \theta)$, then

$$\begin{aligned} d(x, \xi) &= \frac{\partial \eta(x, \theta)}{\partial \theta^T} M^{-1}(\xi, \theta) \frac{\partial \eta(x, \theta)}{\partial \theta} \\ \therefore \frac{d(x, \xi)}{(1 - \theta_1 x)^{2\theta_2}} &= \frac{m_{11} \theta_2^2 x^2}{(1 - \theta_1 x)^2} - \frac{2m_{12} \theta_2 x \log(1 - \theta_1 x)}{1 - \theta_1 x} + m_{22} (\log(1 - \theta_1 x))^2. \end{aligned}$$

Now, $\{d(x, \xi) - k + c : 0 < c < \epsilon\}$ is a linear combination of

$$\left\{ \frac{\theta_2^2 x^2}{(1 - \theta_1 x)^2}, \frac{\theta_2 x \log(1 - \theta_1 x)}{1 - \theta_1 x}, (\log(1 - \theta_1 x))^2, \frac{1}{(1 - \theta_1 x)^{2\theta_2}} \right\} \quad (15)$$

therefore by Karlin and Studden (1966), (15) is a T-system.

Hence, $\{d(x, \xi) - k + c\}$ has at most four roots and from Theorem 2, locally A-, D- and E-optimal designs for $[a, \infty); a > 0$ are minimally supported if D-, A- and E-optimal design exists in the corresponding design space.

To determine the support points for D-, A- and E-optimal designs, we minimize the determinant, trace and maximum eigenvalue of the inverse of Fisher information matrix $M(\xi, \theta)$ given in (14), respectively. Since the determinant, trace and maximum eigenvalue of the inverse of Fisher information matrix are not in explicit form, therefore, we obtain optimal support points numerically for some known values of the parameters θ_1 and θ_2 . Theorems 3, 4 and 5 give conditions on support points for designs to be D-, A- and E-optimal, respectively.

Theorem 3.

- (i) For design support $A_0 = [a, \infty), a > 0$, D-optimal design is supported on (x_1^*, x_2^*) .
- (ii) Consider $A_0 = [a, b)$
 - If $a < x_1^* < x_2^* < b$, then the D-optimal design is supported on (x_1^*, x_2^*)
 - If $a \geq x_1^*, b > x_2^*$, then the D-optimal design is supported on (x_a^*, x_2^*) , where x_a^* is solution of minimized (9) with $x_1^* = a$.
 - If $a < x_1^*, b \leq x_2^*$, then the D-optimal design is supported on (x_1^*, x_b^*) , where x_b^* is solution of minimized (9) with $x_2^* = b$.
 - If $a \geq x_1^*, b \leq x_2^*$, then the D-optimal design is supported on (a, b) .

For some values of θ_1 and θ_2 the D-optimal design with two and three support points having equal weights are listed in Table 1 and Table 2 respectively and can be shown to satisfy generalized equivalence theorem as $d(x, \xi) \leq 2$. For equal and unequal weights, D-optimal design support points are same.

Table 1. D-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.5, w_2 = 0.5$

(θ_1, θ_2)	Design region	Design points		(θ_1, θ_2)	Design region	Design points	
		x_1	x_2			x_1	x_2
(0.25, 2)	[0.1, 4]	1.2749	3.0877	(0.5, 1.5)	[0.1, 2]	0.8467	1.7850
	[0.5, 3.5]	1.2749	3.0877		[0.5, 1.8]	0.8497	1.7834
	[1.5, 3.5]	1.5000	3.1386		[1, 2]	1.0000	1.8027
	[0.1, 2.5]	1.0793	2.5000		[0.1, 1.75]	0.7182	1.7500
	[1.5, 2.5]	1.5000	2.5000		[1, 1.75]	1.0000	1.7500
(0.25, 3)	[0.1, 4]	0.8490	2.3745	(0.5, 2.5)	[0.1, 2]	0.5097	1.3457
	[0.5, 3.5]	0.8490	2.3745		[0.5, 1.5]	0.5097	1.3457
	[1, 3.5]	1.0000	2.4267		[0.75, 2]	0.7500	1.4204
	[0.1, 2]	0.7597	2.0000		[0.1, 1]	0.4118	1.0000
	[1, 2]	1.0000	2.0000		[0.75, 1]	0.7500	1.0000

Table 2. D-optimal designs for generalized Pareto model with three support points having weights $w_1 = 1/3, w_2 = 1/3, w_3 = 1/3$

(θ_1, θ_2)	Design region	Design points			(θ_1, θ_2)	Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
(0.25, 2)	[0.1, 4]	1.2749	1.2749	3.0877	(0.5, 1.5)	[0.1, 2]	0.8467	0.8467	1.7850
	[0.5, 3.5]	1.2749	1.2749	3.0877		[0.5, 1.8]	0.8497	0.8497	1.7834
	[1.5, 3.5]	1.5000	1.5000	3.1386		[1, 2]	1.0000	1.0000	1.8027
	[0.1, 2.5]	1.0793	1.0793	2.5000		[0.1, 1.75]	0.7182	0.7182	1.7500
	[1.5, 2.5]	1.5000	1.5000	2.5000		[1, 1.75]	1.0000	1.0000	1.7500
(0.25, 3)	[0.1, 4]	0.8490	0.8490	2.3745	(0.5, 2.5)	[0.1, 2]	0.5097	0.5097	1.3457
	[0.5, 3.5]	0.8490	0.8490	2.3745		[0.5, 1.5]	0.5097	0.5097	1.3457
	[1, 3.5]	1.0000	1.0000	2.4267		[0.75, 2]	0.7500	0.7500	1.4204
	[0.1, 2]	0.7597	0.7597	2.0000		[0.1, 1]	0.4118	0.4118	1.0000
	[1, 2]	1.0000	1.0000	2.0000		[0.75, 1]	0.7500	0.7500	1.0000

Theorem 4.

- (i) For design support $\Lambda_0 = [a, \infty), a > 0$, A-optimal design is supported on (x_1^*, x_2^*) .
- (ii) Consider $\Lambda_0 = [a, b)$
 - If $a < x_1^* < x_2^* < b$, then the A-optimal design is supported on (x_1^*, x_2^*)
 - If $a \geq x_1^*, b > x_2^*$, then the A-optimal design is supported on (x_a^*, x_2^*) , where x_a^* is solution of minimized (11) with $x_1^* = a$.
 - If $a < x_1^*, b \leq x_2^*$, then the A-optimal design is supported on (x_1^*, x_b^*) , where x_b^* is solution of minimized (11) with $x_2^* = b$.
 - If $a \geq x_1^*, b \leq x_2^*$, then the A-optimal design is supported on (a, b) .

For some values of θ_1 and θ_2 the A-optimal design with two support points having equal and unequal weights and three support points having equal and unequal weights are listed in Tables 3-6 respectively and can be shown to satisfy generalized equivalence theorem as $d(x, \xi) \leq 2$.

Table 3. A-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.5, w_2 = 0.5$

(θ_1, θ_2)	Design region	Design points		(θ_1, θ_2)	Design region	Design points	
		x_1	x_2			x_1	x_2
(0.25, 1.5)	[0.1, 4]	1.3121	3.7159	(0.5, 1.5)	[0.1, 2]	0.6571	1.8263
	[0.5, 3.75]	1.3121	3.7159		[0.5, 1.9]	0.6554	1.8562
	[1.5, 3.75]	1.5000	3.7094		[0.75, 2]	0.7500	1.8531
	[0.5, 2.75]	1.1461	2.7500		[0.5, 1.5]	0.6125	1.5000
	[1.5, 2.75]	1.5000	2.7500		[0.75, 1.5]	0.7500	1.5000
(0.25, 2.5)	[0.1, 4]	0.7253	2.9106	(0.5, 2.5)	[0.1, 2]	0.3623	1.4545
	[0.5, 3.5]	0.7253	2.9106		[0.25, 1.5]	0.3623	1.4545
	[1.5, 3.5]	1.5000	2.9757		[0.5, 2]	0.5000	1.4494
	[0.5, 2.5]	0.7468	2.5000		[0.25, 1]	0.3425	1.0000
	[1.5, 2.5]	1.5000	2.5000		[0.5, 1]	0.5000	1.0000

Table 4. A-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.35, w_2 = 0.65$

(θ_1, θ_2)	Design region	Design points		(θ_1, θ_2)	Design region	Design points	
		x_1	x_2			x_1	x_2
(0.25, 1.5)	[0.1, 5]	1.4368	3.7652	(0.5, 1.5)	[0.1, 2]	0.7181	1.8809
	[0.1, 4]	1.4368	3.7652		[0.5, 2]	0.7181	1.8809
	[1.5, 4]	1.5000	3.7628		[0.75, 2]	0.7500	1.8798
	[0.5, 3.5]	1.4466	3.5000		[0.5, 1.5]	0.6513	1.5000
	[1.5, 3.5]	1.5000	3.5000		[0.75, 1.5]	0.7500	1.5000
(0.25, 2.5)	[0.1, 4]	0.8077	3.0029	(0.5, 2.5)	[0.1, 2]	0.4036	1.5005
	[0.5, 3.5]	0.8077	3.0029		[0.25, 2]	0.4036	1.5005
	[1, 3.5]	1.0000	2.9908		[0.5, 2]	0.5000	1.4946
	[0.5, 2.5]	0.8196	2.5000		[0.25, 1.5]	0.3674	1.0000
	[1.5, 2.5]	1.5000	2.5000		[0.5, 1]	0.5000	1.0000

Table 5. A-optimal designs for generalized Pareto model with three support points having weights $w_1 = 1/3, w_2 = 1/3, w_3 = 1/3$

(θ_1, θ_2)	Design region	Design points			(θ_1, θ_2)	Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
(0.25, 1.5)	[0.1, 4]	1.1587	1.1587	3.6617	(0.5, 1.5)	[0.1, 2]	0.5784	0.5785	1.8291
	[0.5, 3.75]	1.1587	1.1587	3.6617		[0.5, 1.9]	0.5784	0.5785	1.8291
	[1.5, 3.75]	1.5000	1.5000	3.6539		[0.75, 2]	0.7500	0.7500	1.8258
	[0.5, 3.5]	1.1825	1.1825	3.5000		[0.5, 1.5]	0.5556	0.5556	1.5000
	[1.5, 3.5]	1.5000	1.5000	3.5000		[0.75, 1.5]	0.7500	0.7500	1.5000
(0.25, 2.5)	[0.1, 4]	0.6292	0.6292	2.8182	(0.5, 2.5)	[0.1, 2]	0.3142	0.3142	1.4084
	[0.5, 3.5]	0.6292	0.6292	2.8182		[0.25, 1.5]	0.3142	0.3142	1.4084
	[1.5, 3.5]	1.5000	1.5000	2.9204		[0.5, 2]	0.5000	0.5000	1.4089
	[0.5, 2.5]	0.6517	0.6517	2.5000		[0.25, 1]	0.3069	0.3069	1.0000
	[1.5, 2.5]	1.5000	1.5000	2.5000		[0.5, 1]	0.5000	0.5000	1.0000

Table 6. A-optimal designs for generalized Pareto model with three support points having weights $w_1 = 0.15, w_2 = 0.25, w_3 = 0.60$

(θ_1, θ_2)	Design region	Design points			(θ_1, θ_2)	Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
(0.25, 1.5)	[0.1, 4]	1.3955	1.3955	3.7483	(0.5, 1.5)	[0.1, 2]	0.6973	0.6973	1.8725
	[1, 4]	1.3955	1.3955	3.7483		[0.5, 2]	0.6973	0.6973	1.8725
	[1.5, 4]	1.5000	1.5000	3.7444		[0.75, 2]	0.7500	0.7500	1.8706
	[0.5, 3.5]	1.4109	1.4109	3.5000		[0.5, 1.5]	0.6393	0.6393	1.5000
	[1.5, 3.5]	1.5000	1.5000	3.5000		[0.75, 1.5]	0.7500	0.7500	1.5000
(0.25, 2.5)	[0.1, 4]	0.7799	0.7799	2.9702	(0.5, 2.5)	[0.1, 2]	0.3897	0.3897	1.4842
	[0.5, 3.5]	0.7799	0.7799	2.9702		[0.25, 1.5]	0.3897	0.3897	1.4842
	[1.5, 3.5]	1.5000	1.5000	3.0178		[0.5, 2]	0.5000	0.5000	1.4782
	[0.5, 2.5]	0.7962	0.7962	2.5000		[0.25, 1]	0.3597	0.3597	1.0000
	[1.5, 2.5]	1.5000	1.5000	2.5000		[0.5, 1]	0.5000	0.5000	1.0000

Theorem 5.

- (i) For design support $A_0 = [a, \infty), a > 0$, E-optimal design is supported on (x_1^*, x_2^*) .
- (ii) Consider $A_0 = [a, b)$
 - If $a < x_1^* < x_2^* < b$, then the E-optimal design is supported on (x_1^*, x_2^*)
 - If $a \geq x_1^*, b > x_2^*$, then the E-optimal design is supported on (x_a^*, x_2^*) , where x_a^* is solution of minimized (12) with $x_1^* = a$.
 - If $a < x_1^*, b \leq x_2^*$, then the E-optimal design is supported on (x_1^*, x_b^*) , where x_b^* is solution of minimized (12) with $x_2^* = b$.
 - If $a \geq x_1^*, b \leq x_2^*$, then the E-optimal design is supported on (a, b) .

For some values of θ_1 and θ_2 the E-optimal design with two support points haing equal and unequal weights and three support points having equal and unequal weights are listed in Tables 7-10 respectively and can be shown to satisfy generalized equivalence theorem as $d(x, \xi) \leq 2$.

Table 7. E-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.5, w_2 = 0.5$

(θ_1, θ_2)	Design region	Design points		(θ_1, θ_2)	Design region	Design points	
		x_1	x_2			x_1	x_2
(0.25, 2)	[0.1, 4]	0.9324	3.2931	(0.5, 2)	[0.1, 2]	0.4648	1.6455
	[0.5, 3.5]	0.9324	3.2931		[0.25, 1.75]	0.4648	1.6455
	[1.5, 3.5]	1.5000	3.2992		[0.5, 2]	0.5000	1.6425
	[0.5, 2.5]	0.898	2.5000		[0.1, 1.5]	0.4781	1.5000
	[1.5, 2.5]	1.5000	2.5000		[0.5, 1.5]	0.5000	0.5000
(0.25, 3)	[0.1, 4]	0.5937	2.5918	(0.5, 3)	[0.1, 2]	0.2965	1.2955
	[0.5, 3]	0.5937	2.5918		[0.25, 1.5]	0.2965	1.2955
	[1, 3]	1.0000	2.6023		[0.5, 2]	0.5000	1.3008
	[0.5, 2]	0.6017	2.0000		[0.1, 1]	0.3007	1.0000
	[1, 2]	1.0000	2.0000		[0.5, 1]	0.5000	1.0000

Table 8. E-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.35, w_2 = 0.65$

(θ_1, θ_2)	Design region	Design points		(θ_1, θ_2)	Design region	Design points	
		x_1	x_2			x_1	x_2
(0.25, 2)	[0.1, 4]	1.0327	3.3733	(0.5, 2)	[0.1, 2]	0.5152	1.6857
	[0.5, 3.5]	1.0327	3.3733		[0.25, 1.75]	0.5152	1.6857
	[1.5, 3.5]	1.5000	3.3696		[0.75, 2]	0.7500	1.6840
	[0.5, 3]	1.0518	3.0000		[0.1, 1.5]	0.5252	1.5000
	[1.5, 3]	1.5000	3.0000		[0.75, 1.5]	0.7500	1.5000
(0.25, 3)	[0.1, 4]	0.6633	2.6874	(0.5, 3)	[0.1, 2]	0.3314	1.3432
	[0.5, 3]	0.6633	2.6874		[0.25, 1.5]	0.3314	1.3432
	[1, 3]	1.0000	2.6852		[0.5, 2]	0.5000	1.3422
	[0.5, 2]	0.6553	2.0000		[0.1, 1]	0.3275	1.0000
	[1, 2]	1.0000	2.0000		[0.5, 1]	0.5000	1.0000

Table 9. E-optimal designs for generalized Pareto model with three support points having weights $w_1 = 1/3, w_2 = 1/3, w_3 = 1/3$

(θ_1, θ_2)	Design region	Design points			(θ_1, θ_2)	Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
(0.25, 2)	[0.1, 4]	0.8134	0.8134	3.2105	(0.5, 2)	[0.1, 2]	0.4053	0.4053	1.6043
	[0.5, 3.5]	0.8134	0.8134	3.2105		[0.25, 1.75]	0.4053	0.4053	1.6043
	[1.5, 3.5]	1.5000	1.5000	3.2369		[0.5, 2]	0.5000	0.5000	1.5985
	[0.5, 3]	0.8367	0.8367	3.0000		[0.1, 1.5]	0.4171	0.4171	1.5000
	[1.5, 3]	1.5000	1.5000	3.0000		[0.5, 1.5]	0.5000	0.5000	1.5000
(0.25, 3)	[0.1, 4]	0.5132	0.5132	2.4978	(0.5, 3)	[0.1, 2]	0.2562	0.2562	1.2485
	[0.5, 3]	0.5132	0.5132	2.4978		[0.25, 1.5]	0.2562	0.2562	1.2485
	[1, 3]	1.0000	1.0000	2.5321		[0.5, 2]	0.5000	0.5000	1.2658
	[0.5, 2]	0.5293	0.5293	2.0000		[0.1, 1]	0.2644	0.2644	1.0000
	[1, 2]	1.0000	1.0000	2.0000		[0.5, 1]	0.5000	0.5000	1.0000

Table 10. E-optimal designs for generalized Pareto model with three support points having weights $w_1 = 0.15, w_2 = 0.25, w_3 = 0.60$

(θ_1, θ_2)	Design region	Design points			(θ_1, θ_2)	Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
(0.25, 2)	[0.1, 4]	0.9990	0.9990	3.3452	(0.5, 2)	[0.1, 2]	0.4983	0.4983	1.6716
	[0.5, 3.5]	0.9990	0.9990	3.3452		[0.25, 1.75]	0.4983	0.4983	1.6716
	[1.5, 3.5]	1.5000	1.5000	3.3440		[0.75, 2]	0.7500	0.7500	1.6715
	[0.5, 3]	1.0217	1.0217	3.0000		[0.1, 1.5]	0.5100	0.5100	1.5000
	[1.5, 3]	1.5000	1.5000	3.0000		[0.75, 1.5]	0.7500	0.7500	1.5000
(0.25, 3)	[0.1, 4]	0.6398	0.6398	2.6534	(0.5, 3)	[0.1, 2]	0.3196	0.3196	1.3262
	[0.5, 3]	0.6398	0.6398	2.6534		[0.25, 1.5]	0.3196	0.3196	1.3262
	[1, 3]	1.0000	1.0000	2.6546		[0.5, 2]	0.5000	0.5000	1.3270
	[0.5, 2]	0.6383	0.6383	2.0000		[0.1, 1]	0.3190	0.3190	1.0000
	[1, 2]	1.0000	1.0000	2.0000		[0.5, 1]	0.5000	0.5000	1.0000

4. A Real Data Application

Beirlant and Geogebur (2003) have studied the fitting of generalized Pareto model to the Norwegian fire claim data for the period 1988-1991. The data is available in R software library in the package *ReIns* as *norwegianfire*, it contains 9181 observations on two variables *size* and *year*

for the period 1971-1992. We use this data to fit the model (2), by taking claim size an dependent variable and time as independent variable, we get $\theta_1 = -0.0001754$ and $\theta_2 = -3.902088$. For these parameters of the model, we have obtain D-, A- and E-optimal two and three support point designs having equal and unequal weights. It is found that support points for A- and E-optimal designs are same for all the cases, therefore we are showing D- and A-optimal design support points for all cases. The obtained results are listed in the Tables 11-14.

Table 11. D-, A- and E-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.5, w_2 = 0.5$

	Design region	Design points			Design region	Design points	
		x_1	x_2			x_1	x_2
D-optimal	[500, 465365]	921.8780	4295.3240	A-optimal or E-optimal	[500, 465365]	572.6625	5068.4895
	[800, 5365]	921.8783	4295.3247		[550, 6365]	572.6625	5068.4895
	[1000, 5365]	1000.0000	4372.6100		[800, 6365]	800.0000	5022.0290
	[500, 4000]	897.8773	4000.0000		[550, 5000]	575.3937	5000.0000
	[1000, 4000]	1000.0000	4000.0000		[800, 5000]	800.0000	5000.0000

Table 12. D-, A- and E-optimal designs for generalized Pareto model with two support points having weights $w_1 = 0.35, w_2 = 0.65$

	Design region	Design points			Design region	Design points	
		x_1	x_2			x_1	x_2
D-optimal	[500, 465365]	921.8781	4295.3223	A-optimal or E-optimal	[500, 465365]	653.7900	5508.8220
	[800, 5365]	921.8782	4295.3249		[600, 6365]	653.7900	5508.8220
	[1000, 5365]	1000.0000	4372.6100		[800, 6365]	800.0000	5456.8240
	[500, 4000]	897.8774	4000.0000		[600, 5000]	672.3323	5000.0000
	[1000, 4000]	1000.0000	4000.0000		[800, 5000]	800.0000	5000.0000

Table 13. D-, A- and E-optimal designs for generalized Pareto model with three support points having weights $w_1 = 1/3, w_2 = 1/3, w_3 = 1/3$

	Design region	Design points				Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
D-optimal	[500, 465365]	921.878	921.878	4295.324	A-optimal or E-optimal	[500, 465365]	500.000	500.000	4668.251
	[800, 5365]	921.878	921.878	4295.324		[500, 6365]	500.000	500.000	4668.251
	[1000, 5365]	1000.000	1000.000	4372.610		[600, 6365]	600.000	600.000	4634.187
	[800, 4000]	897.877	897.877	4000.000		[500, 5000]	504.520	504.520	4000.000
	[1000, 4000]	1000.000	1000.000	4000.000		[600, 5000]	600.000	600.000	4000.000

Table 14. D-, A- and E-optimal designs for generalized Pareto model with three support points having weights $w_1 = 0.15, w_2 = 0.35, w_3 = 0.60$

	Design region	Design points				Design region	Design points		
		x_1	x_2	x_3			x_1	x_2	x_3
D-optimal	[500, 465365]	921.878	921.878	4295.324	A-optimal or E-optimal	[500, 465365]	625.896	625.895	5346.316
	[800, 5365]	921.878	921.877	4295.322		[600, 6365]	625.895	625.895	5346.316
	[1000, 5365]	1000.000	1000.000	4372.610		[800, 6365]	800.000	800.000	5293.034
	[800, 4000]	897.877	897.877	4000.000		[600, 5000]	639.101	639.101	5000.000
	[1000, 4000]	1000.000	1000.000	4000.000		[800, 5000]	800.000	800.000	5000.000

5. Conclusion

We obtained locally D-, A- and E-optimal designs for generalized Pareto model with equal and unequal weights numerically. It is observed that in A- and E-optimal designs with three design points, two points are same but these are different from two support points design. For D-optimal designs, in designs with three support points, two points are same. It is also observed that in D-optimal designs the distinct points for two and three support point designs are same while in A- and E- optimal designs distinct points for two and three support point designs are different. D-optimal designs are same for equal and unequal weights. As we increase the values of parameters, the design points lie closer to center of the design region. For Norwegian fire claim data, the D-, A- and E-optimal designs with two and three support points having equal and unequal weights are obtained. It is found that D-optimal design with two and three support points are almost same for equal and unequal weights. It is also observed that the support points for A- and E-optimal designs are same for all the cases. The designs obtained in the paper can be used in practice.

Conflict of Interest

The authors confirm that this article contents have no conflict of interest.

Acknowledgment

The authors would like to thank the editor and the referees for their constructive comments that helped to improve the quality of the paper.

References

- Atkinson, A.C., Donev, A.N., & Tobias, R.D. (2007). *Optimum experimental designs, with SAS*. Oxford University Press. New York.
- Beirlant, J., & Geoghebur, Y. (2003). Regression with response distributions of Pareto-type. *Computational Statistics & Data Analysis*, 42(4), 595-619.
- Castillo, J.D., & Daoudi, J. (2009). Estimation of generalized Pareto distribution. *Statistics and Probability Letters*, 79(5), 684-688.
- Chernoff, H. (1953). Locally optimal designs for estimating parameters, *The Annals of Mathematical Statistics*, 24(4), 586–602.
- Dette, H., & Haines, L.M. (1994). E-optimal designs for linear and nonlinear models with two parameters, *Biometrika*, 81(4), 739-754.
- Dette, H., Melas, V.B. & Wong, W.K. (2006). Locally D-optimal designs for exponential regression, *Statistica Sinica*, 16(3), 789-803.
- Hambuckers, J., Heuchenne, C., & Lopez, O. (2016). A semiparametric model for Generalized Pareto regression based on a dimension reduction assumption, 1-01362314.
URL: <https://hal.archives-ouvertes.fr/hal-01362314>.
- Heiligers, B. (1994). E-optimal designs in weighted polynomial regression, *Annals of Statistics*, 22(2), 917-929.
- Hosking, J.R.M., & Wallis, J.R. (1987). Parameter and quantile estimation for the generalized pareto distribution. *Technometrics*, 29(3), 339-349.

- Karlin, S. & Studden, W.J. (1966). *Tchebycheff systems: with applications in analysis and statistics*. John Wiley & Sons, New York.
- Li, G., & Majumdar, D. (2008). D-optimal designs for logistic models with three and four parameters, *Journal of Statistical Planning and Inference*, 138(7), 1950–1959.
- Pickands, J. (1975). Statistical inference using extreme order statistics, *The Annals of Statistics*, 3(1), 119–131.
- Pukelsheim, P. (1993). *Optimal design of experiments*. New York Wiley.
- Revankar, N.S., Hartley, M.J., & Pagano, M. (1974). A characterization of the Pareto distribution, *The Annals of Statistics*, 2(3), 599–601.
- Silvey, S.D. (1980). *Optimal designs*. London: Chapman Hall.
- Singh, P., & Kumar, A. (2018). D-optimal designs for exponentiated Pareto II and generalized exponentiated Pareto Models, *Journal of Statistical Theory and Application* (accepted for publication).
- Widiharih, T., Haryatmi, S., & Gunardi (2013). D-optimal designs for weighted exponential and generalized exponential models, *Applied Mathematical Sciences*, 7(22), 1067–1079.
- Yang, M. (2008). A-optimal designs for generalized linear model with two parameters. *Journal of Statistical Planning and Inference*, 138(3), 624–641.

