Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/020204

Expectation Identities of Upper Record Values from Generalized Pareto Distribution and a Characterization

Devendra Kumar¹ and Anamika Kulshrestha²

¹Department of Statistics, Amity Institute of Applied Sciences, Amity University, Noida-201301, India
²Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh-202 002, India

Received: 18 Mar. 2013, Revised: 20 Mar. 2013, Accepted: 26 Mar. 2013
Published online: 1 Jul. 2013

Abstract: In this paper we consider generalized Pareto distribution. Exact expressions and some recurrence relations for single and product moments of upper record values are derived. Further a characterization of this distribution based on conditional and recurrence relation of single moments of record values is presented.

Keywords: Record values; single moment; product moment; recurrence relations; generalized Pareto distribution; conditional expectation; Characterization.

1 Introduction

A random variable $X$ is said to have generalized Pareto distribution (Hall and Wellner [9]) if its probability density function (pdf) is of the form

$$f(x) = \frac{\beta (1 + \alpha)}{(\alpha x + \beta)^2} \left( \frac{\beta}{\alpha x + \beta} \right)^{1/\alpha}, \quad x > 0, \quad \alpha, \beta > 0$$

and the corresponding survival function is

$$\bar{F}(x) = \left( \frac{\beta}{\alpha x + \beta} \right)^{1+(1/\alpha)}, \quad x > 0, \quad \alpha, \beta > 0$$

Where $\bar{F}(x) = 1 - F(x)$, $\alpha > -1, \beta > 0$, then $F$ is said to be member of generalized Pareto distribution. It should be noted that for $\alpha > 0$ and $-1 < \alpha < 0$ this model is, respectively, a Pareto distribution and a Power distribution. Moreover the survival function (1.2) tends to the exponential survival function as $\alpha$ tends to zero. This model is a flexible one due to its properties, i.e. it has a linear mean residual life function its coefficient of variation of the residual life is constant and its hazard rate is the reciprocal of linear function.

For more details and some applications of this distribution one may refer to Hall and Wellner [9] and Johnson et al. [11]. Record values are found in many situations of daily life as well as in many statistical applications. Often we are interested in observing new records and in recording them: for example, Olympic records or world records in sport. Record values are used in reliability theory. Moreover, these statistics are closely connected with the occurrences times of some corresponding non homogeneous Poisson process used in shock models. The statistical study of record values started with Chandler [6], he formulated the theory of record values as a model for successive extremes in a sequence of independently and identically random variables. Feller [8] gave some examples of record values with respect to gambling problems. Resnick [22] discussed the asymptotic theory of records. Theory of record values and its distributional properties have been extensively studied in the literature, for example, see, Ahsanullah [1], Arnold et al. [2,3], Nevzorov [18] and Kamps [12] for reviews on various developments in the area of records.

We shall now consider the situations in which the record values (e.g. successive largest insurance claims in non-life insurance, highest water-levels or highest temperatures) themselves are viewed as “outliers” and hence the second or
third largest values are of special interest. Insurance claims in some non-life insurance can be used as one of the examples. Observing successive $k$ largest values in a sequence, Dziubdziela and Kopocinski [7] proposed the following model of $k$ record values, where $k$ is some positive integer.

Let $\{X_n, n \geq 1\}$ be a sequence of identically independently distributed (i.i.d) random variables with pdf $f(x)$ and distribution function $F(x)$. The $j$-th order statistics of a sample $(X_1, X_2, \ldots, X_n)$ is denoted by $X_{j:n}$. For a fix $k \geq 1$ we define the sequence $\{U_n^{(k)}, n \geq 1\}$ of $k$ upper record times of $\{X_n, n \geq 1\}$ as follows

$$U_1^{(k)} = 1,$$

$$U_{n+1}^{(k)} = \min \{j > U_n^{(k)} : X_j : j + k + 1 > X_{U_n^{(k)},U_n^{(k)}+k-1}\}.$$ 

The sequence $\{Y_n^{(k)}, n \geq 1\}$ with $Y_n^{(k)} = X_{U_n^{(k)},U_n^{(k)}+k-1}, n = 1, 2, \ldots$ are called the sequences of $k$ upper record values of $\{X_n, n \geq 1\}$.

For $k = 1$ and $n = 1, 2, \ldots$ we write $U_1^{(1)} = U_n$. Then $\{U_n, n \geq 1\}$ is the sequence of record times of $\{X_n, n \geq 1\}$. The sequence $\{Y_n^{(k)}, n \geq 1\}$, where $Y_n^{(k)} = X_{U_n^{(k)}}$ is called the sequence of $k$ upper record values of $\{X_n, n \geq 1\}$. For convenience, we shall also take $Y_0^{(k)} = 0$. Note that $k = 1$ we have $Y_n^{(1)} = X_{U_n}, n \geq 1$, which are record value of $\{X_n, n \geq 1\}$. Moreover $Y_1^{(k)} = \min \{X_1, X_2, \ldots, X_k = X_{1:k}\}$.

Let $\{X_{n}^{(k)}, n \geq 1\}$ be the sequence of $k$ upper record values then from (1.3). Then the pdf of $X_{n}^{(k)}, n \geq 1$ is given by

$$f_{X_{n}^{(k)}}(x) = \frac{k^n}{(n-1)![-ln(F(x))]^{n-1}[F(x)]^{k-1}f(x)} \text{ (1.3)}$$

and the joint pdf of $X_{m}^{(k)}$ and $X_{n}^{(k)}$, $1 \leq m < n, n \geq 2$ is given by

$$f_{X_{m}^{(k)}, X_{n}^{(k)}}(x,y) = \frac{k^n}{(n-1)!(n-m-1)![-ln(F(x))]^{m-1}}$$

$$\times [-lnF(y) + lnF(x)]^{n-m-1}[F(y)]^{k-1}f(x)f(y), x < y. \text{ (1.4)}$$

We shall denote

$$\mu_{n}^{(r)} = E((X_{U_{(n)}})^{r}), \quad 1, 2, \ldots,$$

$$\mu_{m,n,k}^{(r,s)} = E((X_{U_{(m)}})^{r},(X_{U_{(n)}})^{s}), \quad 1 \leq m \leq n - 1 \quad \text{and} \quad r,s = 1, 2, \ldots,$$

$$P_{m,n,k}^{(r)} = E((X_{U_{(m)}})^{r}) = \mu_{n}^{(r)}, \quad 1 \leq m \leq n - 1 \quad \text{and} \quad r = 1, 2, \ldots,$$

$$\mu_{m,n,k}^{(0,s)} = E((X_{U_{(n)}})^{s}) = \mu_{n}^{(s)}, \quad 1 \leq m \leq n - 1 \quad \text{and} \quad s = 1, 2, \ldots.$$ 

Recurrent relations are interesting in their own right. They are useful in reducing the number of operations necessary to obtain a general form for the function under consideration. Furthermore, they are used in characterizing the distributions, which in important area, permitting the identification of population distribution from the properties of the sample. Recurrence relations and identities have attained importance reduces the amount of direct computation and hence reduces the time and labour. They express the higher order moments in terms of order moments and hence make the evaluation of higher order moments easy and provide some simple checks to test the accuracy of computation of moments of order statistics.

Recurrent relations for single and product moments of $k$ record values from Weibull, Pareto, generalized Pareto, Burr, exponential and Gumble distribution are derived by Pawalas and Szymanl [19,20,21]. Sultan [24], Saran and Singh [23]. Kumar [16], Kumar and Khan [17] are established recurrence relations for moments of $k$ record values from modified Weibull, linear exponential, exponentiated log-logistic and generalized beta II distributions respectively. Balakrishnan and Ahsanullah [4,5] have proved recurrence relations for single and product moments of record values from generalized Pareto, Lomax and exponential distributions respectively. Recurrence relations for single and product moment of generalized exponential distribution are derived by Khan et al. [14] and Khan et al. [15] are characterized the distributions based on generalized order statistics. Kamps [13] investigated the importance of recurrence relations of
order statistics in characterization. In this paper, we established some explicit expressions and recurrence relations satisfied by the single and product moments of $k$ upper record values from the generalized Pareto distribution. A characterization of this distribution based on conditional expectation and recurrence relations of single moments of record values.

### 2 Relations for Single moment

First of all, we may note that for the generalized Pareto distribution in (1.1)

$$F(x) = \frac{\alpha x + \beta}{(1 + \alpha)} f(x).$$  \hspace{1cm} (2.1)

The relation in (2.1) will be exploited in this paper to derive recurrence relations for the moments of record values from the generalized Pareto distribution.

We shall first establish the explicit expression for single moment $k$ record values $E((X^{(k)}_u)^r)$. Using (1.3), we have

$$\mu^{(r)}_{nk} = \left(\frac{k^n}{(n-1)!}\right) \int_0^\infty x^r[-\ln(F(x))]^{n-1}[\bar{F}(x)]^{k-1}f(x).$$ \hspace{1cm} (2.2)

By setting $t = [\bar{F}(x)]^{\alpha/(1+\alpha)}$ in (2.2), we get

$$\mu^{(r)}_{nk} = \left(\frac{1+\alpha}{\alpha}n\right) r \sum_{p=0}^{r} (-1)^p \left(\frac{r}{p}\right) \int_0^1 l^{(1+\alpha)/n+1-p-r}[-\ln(l^{(1+\alpha)/\alpha})]^{n-1} dt.$$

Again by putting, $w = -\ln(l^{(1+\alpha)/\alpha}$, we obtain

$$\mu^{(r)}_{nk} = \frac{l^{(1+\alpha)}^{r} \sum_{p=0}^{r} (-1)^p \left(\frac{r}{p}\right)}{(1+\alpha)k + \alpha(p-r)}.$$  \hspace{1cm} (2.3)

**Remark 2.1:** For $k = 1$ in (2.3) we deduce the explicit expression for single moments of upper record values from the generalized Pareto distribution.

Recurrence relations for single moments of $k$ upper record values from $df$ (1.2) can be derived in the following theorem.

**Theorem 2.1:** For a positive integer $k \geq 1$ and for $n \geq 1$ and $r = 0, 1, 2, \ldots$,

$$\left(1 - \frac{\alpha r}{(1+\alpha)k}\right)\mu^{(r)}_{nk} = \mu^{(r)}_{n-1,k} + \frac{\beta r}{(1+\alpha)k}\mu^{(r-1)}_{nk}.$$  \hspace{1cm} (2.4)

**Proof** We have from equations (1.3)

$$\mu^{(r)}_{nk} = \frac{k^n}{(n-1)!} \int_0^\infty x^r[-\ln(F(x))]^{n-1}[\bar{F}(x)]^{k-1}f(x)dx.$$  \hspace{1cm} (2.5)

Integrating by parts treating $[\bar{F}(x)]^{k-1}f(x)$ for integration and the rest of the integrand for differentiation, we get

$$\mu^{(r)}_{nk} = \mu^{(r)}_{n-1,k} + \frac{rk^n}{(n-1)!} \int_0^\infty x^{r-1}[-\ln(F(x))]^{n-1}[\bar{F}(x)]^{k}f(x)dx$$

the constant of integration vanishes since the integral considered in (2.5) is a definite integral. On using (2.1), we obtain

$$\mu^{(r)}_{nk} = \mu^{(r)}_{n-1,k} + \frac{rk^n}{(1+\alpha)k(n-1)!} \left\{\alpha \int_0^\infty x^{r-1}[-\ln(F(x))]^{n-1}[\bar{F}(x)]^{k-1}f(x)dx \right.$$ \hspace{1cm} (2.6)

and hence the result given in (2.4).

**Remark 2.2** Setting $k = 1$ in (2.4) we deduce the recurrence relation for single moments of upper record values from the generalized Pareto distribution.
3 Relations for Product moment

On using (1.4), the explicit expression for the product moments of \( k \) record values \( \mu_{m,n,k}^{(r,s)} \) can be obtained

\[
\mu_{m,n,k}^{(r,s)} = \frac{k^n}{(m-1)!(n-m-1)!} \int_0^\infty x^\left( -\ln(\bar{F}(x)) \right)^m \frac{f(x)}{\bar{F}(x)} G(x) \, dx,
\]

where

\[
G(x) = \int_x^\infty y^{\left[ -\ln(\bar{F}(y)) + \ln(\bar{F}(x)) \right]^{n-m-1} \left[ \frac{\bar{F}(x)}{\bar{F}(y)} \right]^{k-1}} \, f(y) \, dy.
\]

By setting \( w = \ln(\bar{F}(x)) - \ln(\bar{F}(y)) \) in (3.2), we obtain

\[
G(x) = (1 + \alpha)^{n-m} \left( \frac{\beta}{\alpha} \right)^r \sum_{p=0}^r \sum_{q=0}^s (-1)^{p+q} \left( \frac{s}{p} \right) \left( \frac{s}{q} \right) \frac{[\bar{F}(x)]^{k+p}(p-s)\alpha/(1+\alpha)\Gamma(n-m).}{(1+\alpha)k + \alpha(p-s)}
\]

On substituting the above expression of \( G(x) \) in (3.1) and simplifying the resulting equation, we obtain

\[
\mu_{m,n,k}^{(r,s)} = \left( 1 + \alpha \right)^{\alpha_m-m} \left( \frac{\beta}{\alpha} \right)^r \sum_{p=0}^r \sum_{q=0}^s (-1)^{p+q} \left( \frac{s}{p} \right) \left( \frac{s}{q} \right) \frac{1}{\left( 1 + \alpha \right)k + \alpha(p+q-r-s)}
\]

Remark 3.1 Setting \( k = 1 \) in (3.3) we deduce the explicit expression for product moments of record values from the generalized Pareto distribution.

Making use of (2.1), we can drive recurrence relations for product moments of \( k \) upper record values

Theorem 3.1: For \( 1 \leq m \leq n-2 \) and \( r,s = 1,2,\ldots \),

\[
\left( 1 - \frac{\alpha s}{(1+\alpha)k} \right) \mu_{m,n,k}^{(r,s)} = \mu_{m,n-1,k}^{(r,s)} + \frac{\beta s}{(1+\alpha)k} \mu_{m,n,k-1}^{(r,s)}.
\]

Proof: From equation (1.4) for \( 1 \leq m \leq n-2 \) and \( r,s = 0,1,2,\ldots \),

\[
\mu_{m,n,k}^{(r,s)} = \frac{k^n}{(m-1)!(n-m-1)!} \int_0^\infty x^\left( -\ln(\bar{F}(x)) \right)^m \frac{f(x)}{\bar{F}(x)} I(x) \, dx,
\]

where

\[
I(x) = \int_x^\infty x^{\left[ -\ln(\bar{F}(y)) + \ln(\bar{F}(x)) \right]^{n-m-1} \left[ \frac{\bar{F}(x)}{\bar{F}(y)} \right]^{k-1} f(y) \, dy.
\]

Integrating \( I(x) \) by parts treating \( [\bar{F}(x)]^{k-1} f(y) \) for integration and the rest of the integrand for differentiation, and substituting the resulting expression in (3.5), we get

\[
\mu_{m,n,k}^{(r,s)} = \mu_{m,n-1,k}^{(r,s)} + \frac{s k^n}{(m-1)!(n-m-1)!} \int_0^\infty \int_x^\infty x^\left[ -\ln(\bar{F}(x)) \right]^{m-1} \frac{f(x)}{\bar{F}(x)} f(y) \, dy \, dx,
\]

the constant of integration vanishes since the integral in \( I(x) \) is a definite integral. On using the relation (2.1), we obtain

\[
\mu_{m,n,k}^{(r,s)} = \mu_{m,n-1,k}^{(r,s)} + \frac{s k^n}{k(1+\alpha)(m-1)!(n-m-1)!} \left[ \alpha \int_0^\infty \int_x^\infty x^\left[ -\ln(\bar{F}(x)) \right]^{m-1} \frac{f(x)}{\bar{F}(x)} f(y) \, dy \right] dx
\]

and hence the result given in (3.4).

Remark 3.2 Setting \( k = 1 \) in (3.4) we deduce the recurrence relation for product moments of upper record values from the generalized Pareto distribution.

One can also note that Theorem 2.1 can be deduced from Theorem 3.1 by putting \( s = 0 \).
4 Characterization

Let \( \{X_n, n \geq 1\} \) be a sequence of i.i.d continuous random variables with \( df \) \( F(x) \) and \( pdf \) \( f(x) \). Let \( X_{U(n)} \) be the \( n \)-th upper record values, then the conditional \( pdf \) of \( X_{U(n)} \) given \( X_{U(m)} = x, 1 \leq m < n \) in view of (1.3) and (1.4), is

\[
f(X_{U(n)}|X_{U(m)} = x) = \frac{1}{(n - m - 1)!}[-\ln F(y) + \ln f(x)]^{n-m-1} \frac{f(y)}{F(x)}, \quad x > y.
\]  

(Theorem 4.1) Let \( X \) be an absolutely continuous random variable with \( df \) \( F(x) \) and \( pdf \) \( f(x) \) on the support \((0, \infty)\), then for \( m < n \),

\[
E[X_{U(n)}|X_{U(m)} = x] = \frac{[\alpha x + \beta](1 + \alpha)^{n-m} - \beta}{\alpha}
\]  

(4.2) if and only if

\[
F(x) = \left(\frac{\beta}{\alpha x + \beta}\right)^{1+(1/\alpha)}, \quad x > 0, \quad \alpha, \beta > 0.
\]  

Proof: From (4.1), we have

\[
E[X_{U(n)}|X_{U(m)} = x] = \frac{1}{(n - m - 1)!} \int_x^\infty \left[\ln \left(\frac{F(y)}{F(x)}\right)\right]^{n-m-1} \frac{f(y)}{F(x)} dy
\]  

(4.3)

By setting \( t = \ln \left(\frac{F(x)}{F(y)}\right) = \ln \left(\frac{\alpha x + \beta}{\alpha y + \beta}\right)^{(1+1/\alpha)} \) from (1.2) in (4.3), we obtain

\[
E[X_{U(n)}|X_{U(m)} = x] = \frac{1}{\alpha(n - m - 1)!} \int_0^\infty \left[(\alpha x + \beta) e^{\alpha t/(1+\alpha)} - \beta\right]^{n-m-1} e^{-t} dt
\]

(4.4)

\[
E[X_{U(n)}|X_{U(m)} = x] = \frac{(\alpha x + \beta)}{\alpha(n - m - 1)!} \int_0^\infty e^{-1+\alpha/(1+\alpha)} t^{n-m-1} dt - \frac{\beta}{\alpha(n - m - 1)!} \int_0^\infty e^{-t^{n-m-1}} dt
\]

Simplifying the above expression, we derive the relation given in (4.2).

To prove sufficient part, we have from (4.1) and (4.2)

\[
\frac{1}{(n - m - 1)!} \int_x^\infty y\left[-\ln F(y) + \ln f(x)\right]^{n-m-1} f(y) dy = F(x) H_r(x),
\]  

where

\[
H_r(x) = \frac{[\alpha x + \beta](1 + \alpha)^{n-m} - \beta}{\alpha}
\]

(4.5)

Differentiating (4.4) both sides with respect to \( x \), we get

\[
- \frac{1}{(n - m - 2)!} \int_x^\infty y\left[-\ln F(y) + \ln f(x)\right]^{n-m-2} f(y) dy = - f(x) H_r(x) + F(x) H_r'(x)
\]

(4.6)

\[
\frac{f(x)}{F(x)} = \frac{H_r'(x)}{H_{r+1}(x) - H_r(x)} = \frac{(1 + \alpha)}{(\alpha x + \beta)}
\]

which proves that

\[
F(x) = \left(\frac{\beta}{\alpha x + \beta}\right)^{1+(1/\alpha)}, \quad x > 0, \quad \alpha, \beta > 0.
\]  

(Theorem 4.2) Let \( k \geq 1 \) be a fix positive integer, \( r \) be a non- negative integer and \( X \) be an absolutely continuous random variable with \( pdf \) \( f(x) \) and \( cdf \) \( F(x) \) on the support \((0, \infty)\), then

\[
\left(1 - \frac{\alpha r}{(1 + \alpha)k}\right) \mu_{w:k}^{(r)} = \mu_{w+1:k}^{(r)} + \frac{\beta r}{(1 + \alpha)k} \mu_{w:k}^{(r-1)}
\]  

(4.7)

if and only if

\[
F(x) = \left(\frac{\beta}{\alpha x + \beta}\right)^{1+(1/\alpha)}, \quad x > 0, \quad \alpha, \beta > 0.
\]  

© 2013 NSP
Natural Sciences Publishing Cor.
Proof: The necessary part follows immediately from equation (2.4). On the other hand if the recurrence relation in equation (4.5) is satisfied, then on using equation (1.3), we have

\[
\frac{k^n}{(n-1)!} \int_0^\infty x^r [-\ln(\bar{F}(x))]^{n-1}[\bar{F}(x)]^{k-1} f(x)dx
\]

\[
= \frac{k_n}{(n-2)!} \int_0^\infty x^r [-\ln(\bar{F}(x))]^{n-2}[\bar{F}(x)]^{k-1} f(x)dx
\]

\[
+ \frac{x_n}{(n-1)!} \int_0^\infty x^r [-\ln(\bar{F}(x))]^{n-1}[\bar{F}(x)]^{k-1} f(x)dx
\]

\[
+ \frac{\alpha r k_n}{(1+\alpha)(n-1)!} \int_0^\infty x^{r-1} [-\ln(\bar{F}(x))]^{n-1}[\bar{F}(x)]^{k-1} f(x)dx.
\]  

(4.6)

Integrating the first integral on the right hand side of equation (4.6), by parts and simplifying the resulting expression, we find that

\[
\frac{r k_n}{k(n-1)!} \int_0^\infty x^{r-1} [-\ln(\bar{F}(x))]^{n-1}[\bar{F}(x)]^{k-1}
\]

\[
\times \left\{ \bar{F}(x) - \left( \frac{\alpha x}{1+\alpha} + \frac{\beta}{1+\alpha} \right) f(x) \right\} dx = 0
\]  

(4.7)

Now applying a generalization of the Müntz-Szász Theorem (Hwang and Lin, [10] to equation (4.7), we get

\[
\frac{f(x)}{\bar{F}(x)} = \frac{1+\alpha}{\alpha x + \beta}
\]

which proves that

\[
\bar{F}(x) = \left( \frac{\beta}{\alpha x + \beta} \right)^{1+(1/\alpha)}, \quad x > 0, \quad \alpha, \beta > 0.
\]

5 Applications

The results established in this paper can be used to determine the mean, variance and coefficients of skewness and kurtosis. The moments can also be used for finding best linear unbiased estimator (BLUE) for parameter and conditional moments. Some of the results are then used to characterize the distribution.

6 Conclusion

In this study some exact expressions and recurrence relations for single and product moments of record values from the generalized Pareto distribution have been established. Further, conditional expectation and recurrence relation of single moments of record values has been utilized to obtain a characterization of the generalized Pareto.

Acknowledgement

The authors are grateful to the referee for giving valuable comments that led to an improvement in the presentation of the study.
References


