Original Article

STATISTICAL MODELING OF EXTREME VALUES: THE GOMPERTZ INVERSE PARETO DISTRIBUTION METHOD

Saima Gulzar Ahmed and Hasan Javed Malik

Department of Statistics, Forman Christian College a Chartered University Lahore Pakistan

Abstract: In real-life scenarios, classical probability distributions often fail to adequately capture the characteristics of empirical data. To address this limitation, researchers have introduced various distribution generators, each characterized by one or more parameters, offering enhanced flexibility in modeling data. Some notable generators include the Marshal-Olkin family (MO-G), the Beta-G, the Kumaraswamy-G (Kw-G), the McDonald-G (Mc-G), various types of gamma-G distributions, the log gamma-G, the Exponentiated generalized-G, Transformed-Transformer (T-X), Exponentiated (T-X), Weibull-G, and the Exponentiated half logistic generated family. Additionally, Ghosh et al. (2016) introduced the Gompertz-G generator, which extends continuous distributions with two extra parameters, further enriching the spectrum of available distribution generators. This paper explores the Gompertz-G generator and its general mathematical properties, contributing to the growing toolbox of distribution generators that offer more versatile modeling options for diverse data sets.

Keywords: Distribution generators, Gompertz-G generator, parameterized distributions, empirical data modeling, probability distributions.

1. Introduction:

In many real-life situations, the classical distributions do not provide adequate fit to some real data sets. Thus, researchers introduced many generators by introducing one or more parameters to generate new distributions. The new generated distributions are more flexible as compare to the classical distributions. Some well-known generators are

Marshal-Olkin generated family (MO-G) (Marshall and Olkin, 1997), the Beta-G by Eugene et al. (2002) and Jones (2004), Kumaraswamy-G (Kw-G for short) by Cordeiro and de Castro (2011) and McDonald-G (Mc-G) by Alexander et al. (2012), gamma-G (type 1) by Zografos and Balakrishnan (2009), gamma-G (type 2) by Risti´c and Balakrishnan (2012), gamma-G (type 3) by Torabi and Hedesh (2012) and log gamma-G by Amini et al. (2012), Exponentiated generalized-G by Cordeiro et al. (2011), Transformed-Transformer (T-X) by Alzaatreh et al. (2013) and Exponentiated (T-X) by Alzaghal et al. (2013), Weibull-G by Bourguignon et al. (2014) and Exponentiated half logistic generated family by Cordeiro et al. (2014).Ghosh et al. (2016) introduced a new generator of continuous distributions with two extra parameters called the Gompertz-G generator and studied some general mathematical properties of it.

In this article the Gompertz family of distribution is considered to develop a new model. It has been already used by Alizadeh et al. (2017), and Abdal-Hameed, Khaleel, Abdullah, Oguntunde, Adejumo and Oguntunde et al. (2018). The cumulative distribution function (cdf) and probability density function (pdf) of the Gompertz family of distributions is

F□x□ □1□e□□ □1□□1□G□x□□□□□□□; □□ 0, □□ 0. (1)

□□

1 □1□□1□G□x□□□□□ f □x□□□g□x□□□1□G□x□□□□□□ e □□□ □□; □□ 0, □□ 0. (2)

Where □and □are extra shape parameters and the cdf in eq. (1) and eq. (2) was developed using the following transformation:

□log□□1□G□x□□□

F□x□□ □ w(t)dt

0

w□t□is the probability density function (pdf) of the Gompertz distribution and t is a random variable.G□x□ and g□x□are the cdf and pdf of the baseline distribution. The probability density function (pdf) of the Pareto distribution is

□

f □x□ □□□x □□1 □□ 0,□□ 0 □□ x □ □. (3)

Where,□ is scale and □is shape parameter.

An observation is called a record values if its value is greater than (less than) all the preceding observations. Records values theory has wide application in the fields of studies such as climatology, sports, science, engineering, medicine, traffic, and industry, among others. For example, if we consider the weighing of objects on a scale missing its spring. An object is placed on this scale and its weight is measured. The needle indicates the correct value but does not return to zero when the object is removed. If various objects are placed on the scale and only the weights greater than the previous ones can be recorded. Then these recorded weights are the record value sequence. The development of the general theory of statistical analysis of record values began with the work of Chandler (1952). Further work done by, Foster and Stuart (1954), Renyi (1962), Resnick (1973), Nayak (1981), Dunsmore (1983), Gupta (1984), Houchens (1984), Ahsanullah (1978, 1979, 1980, 1981, 1982, 1987, 1988, 1991, 1995, 2004, 2006), Ahmadi et al. (2005), Ahsanullah and Aliev (2008) and Balakrishnan et al. (2009), Ahsanullah et al. (2010) and many more. The pdf of the sequence of upper record values □□ $X_{U□n□}$,n □1□□ is

f_n □x□ □ □□R□x□□□n□ 1 f □x□, □□ □ x □ □. (4)

□□n□

where, R□x□ □ □ln□□1□ F□x□□□.

2. Gompertz Inverse Pareto Distribution

In this section, we derived the inverse Pareto distribution using the pdf in eq. (3) first and then the Gompertz inverse Pareto distribution is developed. The pdf of the inverse Pareto (IP) is derived by transferring eq. (3) with pdf

g□x□ □□□x □□1, □□ 0,□□ 0, 0 □ x □ 1 . (5)

□

And the cdf of the IP distribution is

G□x□ □ □x□□□, □□ 0,□□ 0, 0 □ x □ 1 (6)

Original Article

□

The cdf and pdf of the GoIP distribution is derived by substituting eq. (5) and eq. (6) in eq. (1) and eq. (2),

□□□□□1□□1□□x□□□□□ □□

F□x□ □1□e (7)

f □x□□□□□□x□□1□□1□□x□□□□□□□□1e□□□□□1□□1□□x□□□□□□□, □□□□, , , □ 0, 0 □ x □□ 1 . (8)

Where,□ is scale and □□□, , are shape parameters. The graphs of the pdf and cdf of GoIP distribution have been shown in Figure 1 and 2.

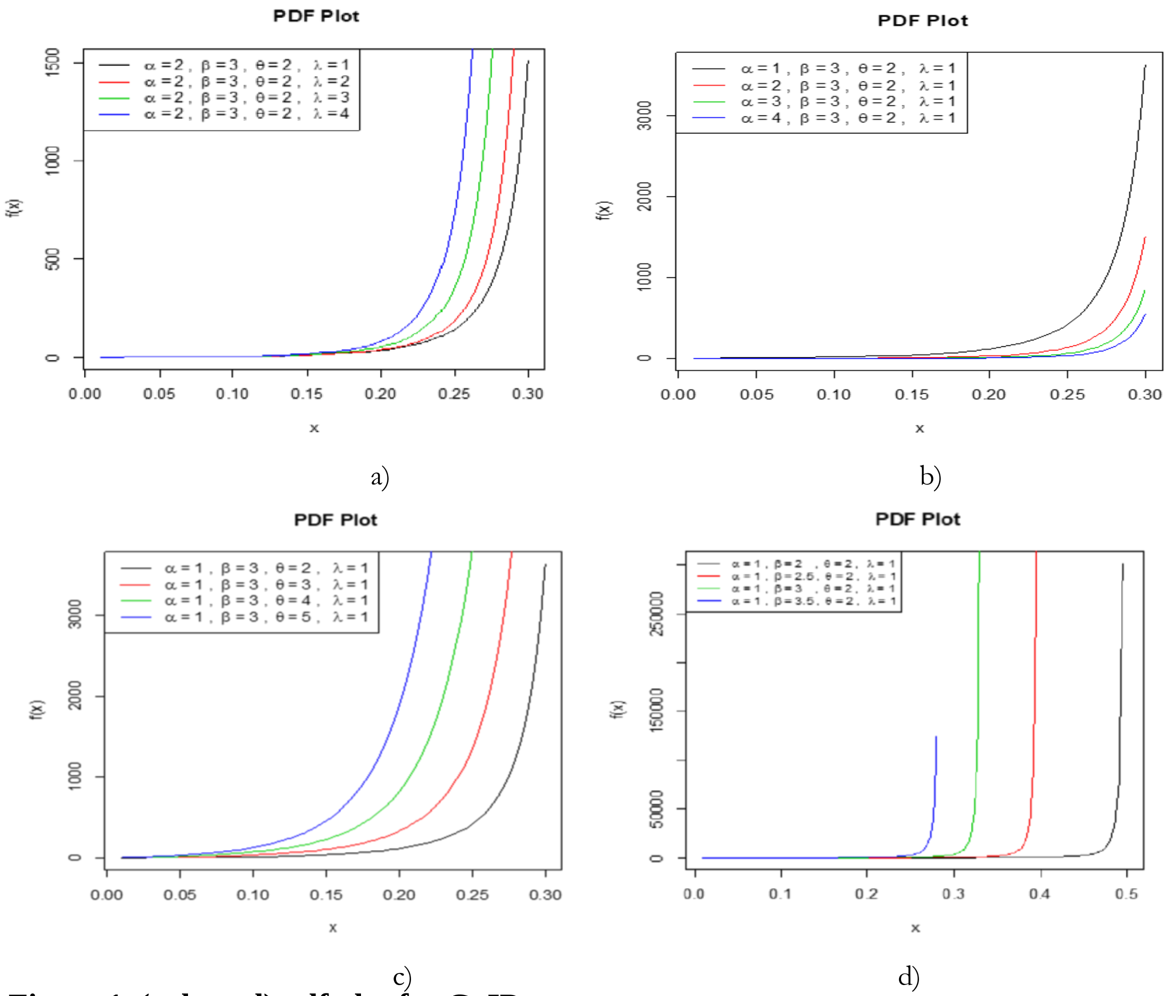


Figure 1. (a, b, c, d) pdf plot for GoIP

2.1. Some Basic Properties of the Gompertz Inverse Pareto Distribution

In this section some reliability measures of the GoIP have been derived. The reliability function of the GoIP distribution is

R□x□□ e□□□1□□1□□x□□□□□ □□ (9) □

The hazard rate function of the GoIP distribution is

□□□1 h□x□□□□□□x□□1□□1□□x□□□□ (10)

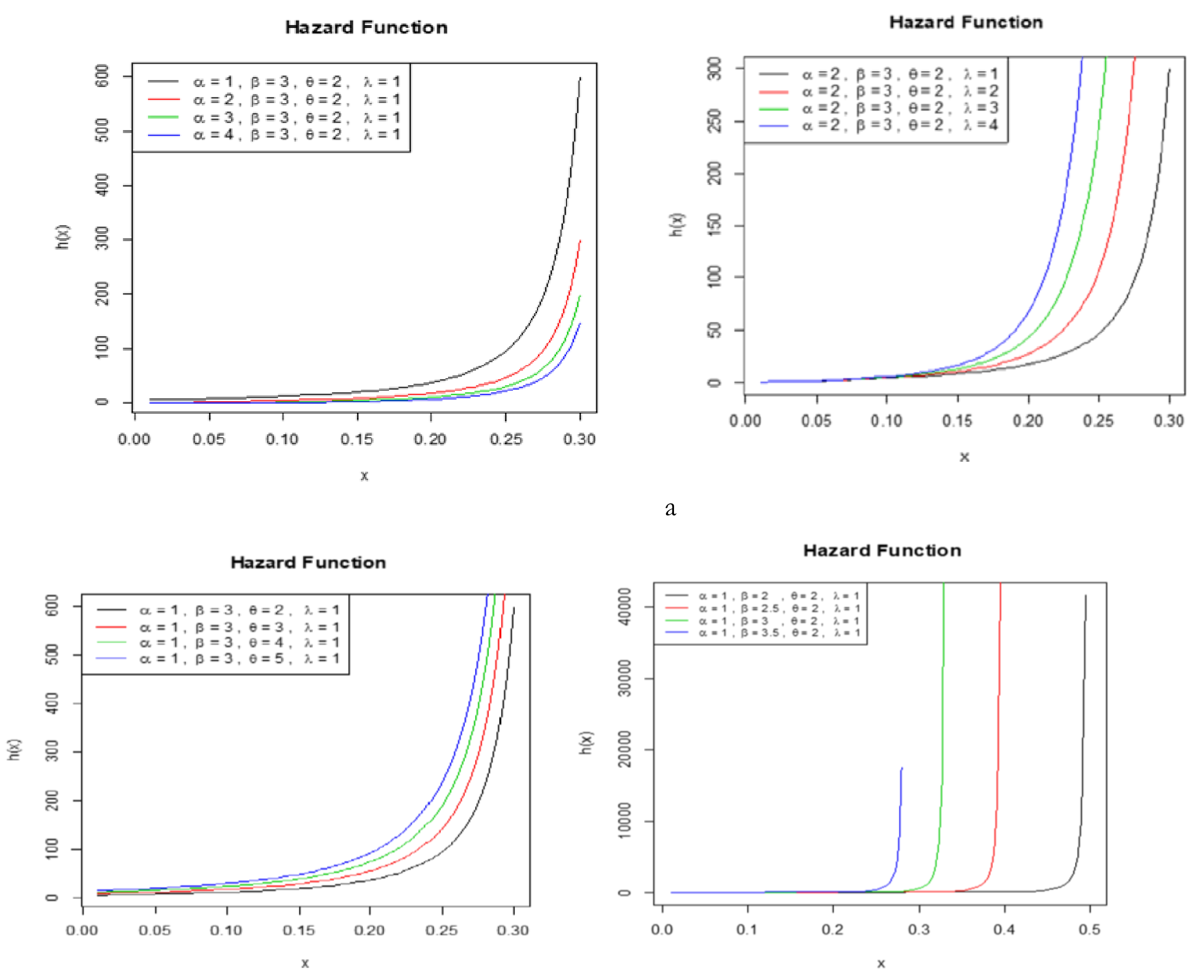
Original Article

The graphs of the reliability function and hazard rate function of the GoIP are given in figure 3 and 4. The reversed hazard rate function of the GoIP distribution is

r□x□□□□□□x□□1 □11□□□e x□□□ □1□□□□1□□□□x□□□□□□1□□□□ e□□□□□ 1□□1□□x□□□□ □□□ (11)

The odds function of the GoIP distribution is

$$O(x) = \frac{1 - e^{\square}}{e^{\square}} \qquad (12)$$



c) b)

Figure 2. (a, b, c, d). Plots for hazard rate function of GoIP

2.2. Quantile Function and Median

In this section the median and quantile function is derived. Q□u□□F□1□u□, where U is Uniform (0,1). The quantile function of the GoIP distribution is

1

Original Article

Q□u□ □ 1 □□1□□□1□□ln□1□u□□□□□ 1 □□□ (13)

—

—

□□ □ □ □ □

□ □

Random numbers for GoIP distribution can be generated using eq. (13). The median of the GoIP distribution is

1

1 – –

1 □ □1□□□1□□ln□0.5□□□□□ □□□ (14) median □

—

—

□□ □ □ □ □

□ □

2.3. Estimation

The method of maximum likelihood estimation (MLE) is used to estimate the parameters of the GoIP distribution. Let $x_1,x_2,....,x_n$ be the random samples distributed GoIP with pdf given in eq. (8),

n □□□1 □□1□□1□□x□□□□ □

L□$x_1,x_2,...,x_n$;□□□□, , , □□□□□□□x□□1□□1□□x□□□□□ e□□□ □□

i□1

n

ln L□□□□, , ,□□ □ nlog□□ nlog□□ n□log□□□□□1□□log x_i □

i□1

n n □□ (15)

□ □

□□□1□□i□1 log□□1□□x□□□□□ □□□i□1 □□ 1□□1□□x□□□□ □ □ □

—

□L□□□□□, , , □ □ n □ 1 □n □□ 1□□1□□x□□□□□□□□□ (16)

—— — —

□□ □ □ i□1 □

□L□□□□□, , , □ n nlog□□□n □ □n □x^i□□□ log□x□□

□ □ log x_i □□□ 1□ □ (17)

—— —

——

□□ □ i□1 i□1 □□ 1□□x□□ □□

□L□□□□□, , , □ □ □ □□ □n □ □1□□x□□□□□□□log□x□□□□i□n1

□□□1□□1□□x□□□□□□□□□□□□□ □□i□n1 log□□1□□x□□ □□□ (18)

□□

—— —

□□ □2 □ i□1

□ˆ □ 1 (19)

x n

3. Order Statistics

The pdf of the rth order statistics from the GoIP distribution is

n□r□1□□ r□ 1

f r:n □x□□ □r □1□!n□!n□r□! □□□□x□□1 □□1□□x□□□□□□□1 e□□ □ □□1□□1□□x□ □□□□□□ □□ □□□1□e□□□□□1□□1□□x□□□□□□□□□ □□□□ , 0 □ x □□ 1 . (20)

□

The pdf of minimum and maximum order statistics from GoIP distribution is

n□□ □ □□□

f1:n □x□□ n□□□□x□□1 □□1□□x□□□□□□□□1 e□□□1□□1□□x□□ □ □□, 0 □ x □□ 1 . (21)

□□ □□ n□1

□□

□□□1 □1□□1□□x□□□□ □□ □ □□1□□1□□x□□□□ □□ □1

□

fn:n □x□□ n□□□□x□□1 □□1□□x□□□□□ e□□ □□□1□e □□ □ □□□ , 0 □ x □□ . (22)

4. Record Values

If the upper record values X_U□1□, X_U□2□,......, X_U□n□arise from GoIP distribution then the pdf of the upper record values from Gompertz inverse Pareto (UR-GoIP) distribution is derived using eq. (8) in eq. (4), we get

n □ □□□1□□ n□1 □□□1□ 1□□x□□□ □□□□□□ 0 1

fn □x□□□□□□□n□1□□ n□ x□□1□□1□□x□□□□□□ □□□□1□□x□□□□ □1□□□ exp□□□□ □ □ □ □□□ □ x □□. (23)

The cdf of the UR-GoIP distribution is

F_n □x□ □ □□

1 n□ □□n, x □□□□□□ (24)

□□

The survival function the UR-GoIP distribution is

Original Article

S_n □x□ □ 1 n□ □□□□n, x □□□□□□ (25)
□□

The hazard rate function of the UR-GoIP distribution is

h_n □x□ □ □□□□ ^{n} □ x□□^{1} □1□□x□ □ □ □□□^{1} □1□□x□ □ □□ □1□ $^{n□1}$ exp □□ □1□□1□□x□ □ □□ □□ / □□n, x□ □ (26)

x □

Where, x□ □1□□1□□x□□□□□□, □□a, x□□□t$^{a□1}$e$^{□t}$dt and □□a,x□□□t$^{a□1}$e$^{□t}$dt are the lower
0 x

incomplete gamma and upper incomplete gamma functions respectively. The relationship between pdf and cdf of GoIP is

□□□1

f □x□ □□□□□x$^{□□1}$□1□□x□□□□ □□1□ F□x□□□ (27)

and,

f □x□

□□1□ F□x□□□ □□□□□x □□$_1$□1□□x□□□□ □□□1 (28)

Theorem 1: If a sequence of upper record values $x_{U□1□}$, $x_{U□2□}$,......, $x_{U□n□}$, n > 1, arise from the GoIP distribution given in eq. (8), then

E□□ XU□□□nr□1□ □1□□XU□n□1□□□□□□□□□□□ □ E□□XU□□□nr□ □1□□XU□n□1□□□□□□□□□□□ □□□□r□ □□E□XUr □n□ □□E□XU□□□nr□ □□□□□1 □□□□□r □□□□□□□ (29)

□ □ □ □□□ □

Proof: Consider the pdf of UR-GoIP in eq. (23),

1 / 1 □

E□□□XUr □n□ □1□□XU□n□□□□□□□□□□ □□n□ □0 xr □1□□x□□□ □□□□ln□1□F□x□□□□n□1 f □x□dx

Using the relation of given in eq. (27), then in eq. (28) and simplifying it the results in eq. (3) is obtained.

4.1. Simulations: Random numbers of size 50 are generated taking a sample of 15, using the R software. From these results the upper records have been noted and we get the mean, median, geometric mean (G.M), harmonic mean (H.M), variance, standard deviation (S.D), mean deviation (M.D), and coefficient of variation (C.V) of the UR-GoIP distribution.

Table 1: descriptive measures for UR-GoIP distribution

Measures for n□15,□□1.5,□□ 0.195,□□ 0.5,□□1.25							
Mean	Median	G.M	H.M	Variance	S.D	M.D	C.V
9.878532	10.009666	9.874452	9.870271	0.07844	0.2801	0.2210	2.835%

5. Conclusion

In this article a new form four parameter Pareto distribution named ‘Gompertz Inverse Pareto (GoIP) distribution’ is developed using Gompertz family G generator. Some properties of the newly derived model including cdf, survival function, hazard rate function, reversed hazard rate function, odds function median, quantile function have been derived. Parameters of the GoIP distribution are estimated by MLE. Order statistics for GoIP distribution have been introduced. Graphs of the pdf and hazard rate function of the GoIP distribution are presented. From figure 1(a, b, c, d), it can be seen that the shape of distribution is extremely left skewed. From figure 2 (a, b, c, d) it can be seen that the shape of the hazard rate function of the GoIP distribution is increasing bathtub (IBT) shape. Moreover, the upper record values have developed form GoIP distribution. Properties of the UR-GoIP distribution including cdf, survival function, hazard rate function, and recurrence relation for single moments for the UR-GoIP distribution have been derived. Finally, a simulation study has been done. Random numbers of size 50 has been generated with a sample of size 15. The upper records have been noted and some measures have been calculated numerically.

Reference

Abdal-Hameed, M., et al. (2018). Parameter estimation and reliability, hazard functions of Gompertz Burr Type XII distribution. Tikrit Journal for Administration and Economics Sciences, 1(41 part 2), 381–400.

Ahsanullah, M. (1978). Record values and the exponential distribution. Annals of Institute of Statistical Mathematics, 30, A, 429-433.

Ahsanullah, M. (1979). Characterization of the exponential distribution by record values, Sankhya, 41, B, 116-121.

Ahsanullah, M. (1982). Characterizations of the exponential distribution by some properties of record values, StatisticheHefte, 23, 326-332.

Ahsanullah, M. (1988). Introduction to Record Statistics, Ginn Press, Needham Height, MA.

Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, Annals of Institute of Statistical Mathematics, 32, A, 363-368.

Ahsanullah, M. (1981). Record values of the exponentially distributed random variables, StatisticheHefte, 22, 121127.

Ahsanullah, M. (1987). Two characterizations of exponential distribution, Communications in Statistics TheoryMethods, 16, 375-381.

Ahsanullah, M. (1991). Some characteristic properties of the record values from the exponential distribution, Sankhya, 53, B, 403-408.

Ahsanullah, M. (1995). Record Statistics, Nova Science Publishers Inc., New York, NY.

Ahsanullah, M. (2004). Record Values-Theory and Applications, University Press of America, Lanham, MD.

Original Article

Ahsanullah, M. (2006). The generalized order statistics from exponential distribution by record values, Pakistan Journal of Statistics, 22, 121-128.

Ahsanullah, M., and Aliev, F. (2008). Some characterizations of exponential distribution by record values, Journal of Statistical Research, 2, 11-16.

Ahsanullah, M., Hamedani, G.G. and Shakil, M. (2010). Expanded version of 'On record values of Univariate exponential distribution', Technical Report, MSCS, Marquette University.

Alexander, C., Cordeiro, G.M., Ortega, E.M.M., and Sarabia, J. M. (2012). Generalized beta- generated distributions. Computational Statistics & Data Analysis, 56, 1880-1897

Alizadeh, M., Cordeiro, G. M., Bastos Pinho, L. G., & Ghosh, I. (2017). The Gompertz-g family of distributions. Journal of Statistical Theory and Practice, 11(1), 179–207. doi:10.1080/15598608.2016.1267668

Alzaatreh, A., Lee, C. and Famoye, F. (2013). A new method for generating families of continuous distributions. Metron, 71, 63-79.

Alzaghal, A., Famoye, F. and Lee, C. (2013). Exponentiated T-X family of distributions with some applications. International Journal of Statistics and Probability, 2, 1-31.

Amini, M., Mir Mostafaee, S.M.T.K. and Ahmadi, J. (2012). Log-gamma-generated families of distributions. Statistics, 1, 1-20.

Balakrishnan, N., Doostparast, M. and Ahmadi, J. (2009). Reconstruction of past records, Metrika, 70, 89-109.

Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull-G family of probability distributions. Journal of Data Science, 12, 53–68.

Chandler, K.N. (1952). The distribution and frequency of record values, Journal of

Cordeiro, G. M., Alizadeh, M., and Ortega, E.M. (2014). The Exponentiated half-logistic family of distributions: Properties and applications. Journal of Probability and Statistics, 2014, Article ID 864396. doi:10.1155/2014/864396

Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81, 883-898.

Cordeiro, G.M., Ortega, E.M.M. and Silva, G.O. (2011). The Exponentiated generalized gamma distribution with application to lifetime data. Journal of statistical computation and simulation, 81, 827-842

Dunsmore, I.R. (1984). The future occurrence of records, Annals of Institute of Statistical Mathematics, 35, 267-277.

Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics - Theory and Methods, 31, 497-512.

Foster, F.G. and Stuart, A. (1954). Distribution free tests in time series based on the breaking of records, Journal of Royal Statistical Society, B,16, 1-22

Ghosh et al. (2016). The Gompertz-G family of distributions, Journal of Statistical Theory and Practice 11(1), 179– 207. http://dx.doi.org/10.1080/15598608.2016.1267668

Gupta, R.C. (1984). Relationships between order statistics and record values and some characterization results, Journal of Applied Probability, 21, 425-430.

Houchens, R.L. (1984). Record Value Theory and Inference, PhD Dissertation, University of California, Riverside, California.

Jones, M.C (2004). Families of distributions arising from distributions of order statistics. Test, 13, 1-43.

Marshall, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84, 641-652.

Nayak, S.S. (1981). Characterizations based on record values, Journal of Indian Statistical Association, 19, 123-127.

Renyi, A. (1962). Theorie des elements saillantsd'unesuit d'observations, Colloquium Combinatorial Methods of Probability Theory, Aarhus University, 104-115.

Resnick, S.I. (1973). Record values and maxima, Annals of Probability, 1, 650-662.

Risti´c, M.M. and Balakrishnan, N. (2012). The gamma -Exponentiated exponential distribution. Journal of Statistical Computation and Simulation, 82, 1191-1206.

Royal Statistical Society, B, 14, 220-228.

Torabi, H. and Hedesh, N. M. (2012). The gamma-uniform distribution and its applications. Kybernetika, 1, 16-30.

Zografos, K. and Balakrishnan, N. (2009). On families of beta- and generalized gamma- generated distributions and associated inference. Statistical Methodology, 6, 344-362.