MEDWIN PUBLISHERS
Committed to create value for Researchers

Physical Science & Biophysics Journal
ISSN: 2641-9165

A Member of the $r_T - X$ Family of Distributions Induced by a New V with Application to a Power Series Class of Distributions

Ampadu CB*

31 Carrolton Road, USA

***Corresponding author:** Clement Boateng Ampadu, 31 Carrolton Road, Boston, MA 02132-6303, USA, Email: drampadu@hotmail.com

Review Article

Volume 3 Issue 3

Received Date: September 09, 2019

Published Date: September 25, 2019

Abstract

The r_T - X family of distributions induced by V was introduced by Ampadu CB. In this paper a member of this class of distributions is introduced under a new weight, and its applicability is shown. The power series class of distributions first appeared, and in this paper, as an application of the $r_T - X$ family of distributions induced by a new V, we introduce a new power series class of distributions, and show its applicability. The concluding remarks, leaves the reader with further investigation.

Keywords: $r_T - X$ family of distributions induced by V; Power series class of distributions

Introduction

The cummulative distribution function (CDF) and the probability density function (PDF) of the r_T – X family induced by V appeared as follows

Theorem 3.1. The CDF of the $r_T - X$ family induced by V is given by Ampadu CB [1]

$$J(x) = R\left(V\left(F(x)\right)\right)$$

where the random variable $T \in [a,b]$, for $-\infty \le a < b \le \infty$, has *CDF* R, V is an appropriate weight, and
$F(x)$ is the *CDF* of any random variable X.

Theorem 3.2. The PDF of the $r_T - X$ family induced by V is given by Ampadu CB [1]

$$j(x) = r(V(F(x)))V'(F(x))f(x)$$

where the random variable $T \in [a,b]$, for $-\infty \le a < b \le \infty$, has *PDF* r, V is an appropriate weight $F(x)$ and $f(x)$ are the *CDF* and *PDF*, respectively, of any random variable X.

Remark 3.3. The choice of V depends on the support of T, In particular, when the support of T is [a, ∞), where $a \ge 0$, we can take V as follows.

(a) $V(x) = 1 - e^{-x}$

(b) $V(x) = \dfrac{x}{1+x}$

(c) $V(x) = \left[1 - e^{-x}\right]^{\frac{1}{\alpha}}$, where $\alpha > 0$

(d) $V(x) = \left[\dfrac{x}{1+x}\right]^{\frac{1}{\alpha}}$, where $\alpha > 0$

On the other hand, when the support of T is (−∞, ∞), we can take V as follows

(a) $V(x) = 1 - e^{-e^{x}}$

(b) $V(x) = \frac{e^{x}}{1+e^{x}}$

(c) $V(x) = \left[1-e^{-e^{x}}\right]^{\frac{1}{\alpha}}$, where $\alpha > 0$

(d) $V(x) = \left[\frac{e^{x}}{1+e^{x}}\right]^{\frac{1}{\alpha}}$, where $\alpha > 0$

Moreover, a so-called new Exponential-Weibull distribution arising from this broad class of distributions was shown to be a good fit to the breast cancer patients data [2], showing its practical significance.

The power series class of distributions was proposed and studied in Noack A [3]. This class of distributions includes binomial, geometric, logarithmic and Poisson distributions as special cases. However, these distributions may not be useful when a random variable takes the value of zero with high probability, that is, zero-inflated. In such situations, it is more appropriate to consider the distribution which is truncated at zero. More details on these distributions can be found in the context of univariate discrete distributions contained in Johnson NL, Kemp AW, Kotz S [4].

Power series distributions are usually motivated by the stochastic representations

$$Z = min\left(X_1, \cdots, X_N\right)$$

if the components are in series, or

$$Z = max\left(X_1, \cdots, X_N\right)$$

if the components are in parallel, and many lifetime data admit such representations, and the related distributions are very useful in modeling such data. We suppose $X_1, \ldots, X_N$ are independent and identically distributed random variables from a parent distribution with PDF $f(x)$ and CDF $F(x)$, and consider N to be a discrete random variable from a power series distribution (truncated at zero) and whose PDF is given by

$$P(N = n) = \frac{a_n \lambda^n}{C(\lambda)}, \text{ n=1,2,3,.....}$$

where $c(\lambda) = \sum_{n=1}^{\infty} a_n \lambda^n$, a_n depends on n, and $\lambda > 0$. $C(\lambda)$ is finite, and its first, second, and third derivatives with respect to λ are defined and given by $C'(\lambda)$, $C''(\lambda)$, and $C'''(\lambda)$, respectively. The Table 1 below represents some useful quantities including a_n $C(\lambda)$, $C^{-1}(\lambda)$, $C'(\lambda)$, $C''(\lambda)$, and $C'''(\lambda)$, respectively, for the Poisson, geometric, logarithmic and binomial (with m being the number of replicas) distributions which belong to the power series family of distributions.

Distribution	$C(\lambda)$	$C'(\lambda)$	$C''(\lambda)$	$C'''(\lambda)$	$C^{-1}(\lambda)$	a_n	Parameter Space
Poisson	$e^{\lambda} - 1$	e^{λ}	e^{λ}	e^{λ}	$\log(1 + \lambda)$	$(n!)^{-1}$	$(0, \infty)$
Geometric	$\lambda(1 - \lambda)^{-1}$	$(1 - \lambda)^{-2}$	$2(1 - \lambda)^{-3}$	$6(1 - \lambda)^{-4}$	$\lambda(1 + \lambda)^{-1}$	1	(0, 1)
Logarithmic	$-\log(1 - \lambda)$	$(1 - \lambda)^{-1}$	$(1 - \lambda)^{-2}$	$2(1 - \lambda)^{-3}$	$1 - e^{-\lambda}$	n^{-1}	(0, 1)
Binomial	$(1+\lambda)^{m} - 1$	$\frac{m}{(1+\lambda)^{1-m}}$	$\frac{m(m-1)}{(1+\lambda)^{2-m}}$	$\frac{m(m-1)(m-2)}{(1+\lambda)^{3-m}}$	$(\lambda+1)^{\frac{1}{m}} - 1$	$\binom{m}{n}$	$(0, \infty)$

Table 1: Useful Quantities for Some Power Series Distributions.

If $X_{(1)} = min\left(X_1, \cdots, X_N\right)$, then the conditional CDF of $X(1)|N = n$ is given by

$$1 - \left[1 - F\left(x\right)\right]n$$

The CDF of the power series class of distributions associated with some parent distribution with PDF $f(x)$ and CDF $F(x)$ is the marginal CDF of $X_{(1)}$ which is given by

$$1 - \frac{C(\lambda(1-F(x))}{C(\lambda)}$$

In this paper we will adopt the following terminology

Definition 3.4. A power series class of distributions associated with some parent distribution whose *PDF* is *f(x)* and whose *CDF* is *F(x)*, and defined by the marginal *CDF* of $X_{(1)}$ which is given by

$$1-\frac{C(\lambda(1-F(x))}{C(\lambda)}$$

will be called a Minimal Family of Power Series Distributions

The $\left(\frac{1}{e}\right)^{\alpha}$ $PT-G$ Family of Distributions of Ampadu CB [5] was further explored in Anafo AY [6]. In particular, by modifying the parameter space for α in the $\left(\frac{1}{e}\right)^{\alpha}$ PT - Standard Uniform family and using it in Remark 3.3 [7], we introduced a new weight as follows

$$W^{*}(x,\alpha)=-\log\left(\frac{1-e^{\alpha-\alpha x}}{1-e^{\alpha}}\right)$$

where $\alpha \in$ R and $\alpha \neq 0$ and $x \in [0, 1]$.

By inverting $W^{*}(x;\alpha)$, we introduce a new weight in class of distributions proposed in Ampadu CB [1], and hence a new family in the next section. In Section 3, we consider a power series application of the new family. Section 4 illustrates applicablity of the new families. In Section 5, we compare the new families via some goodness of fit measures. The last section is devoted to some further recommendations.

The New Family

By finding the inverse of $W^{*}(x;\alpha)$, we define the new weight as

$$V^{*}(x;\alpha)=\frac{\log\left(\frac{e^{\alpha+x}}{e^{\alpha}+e^{x}-1}\right)}{\alpha}$$

where $\alpha \in$ R and $\alpha \neq 0$ and $x \in [0, 1]$.

The new weight can be obtained by solving the following equation for V*(x; α):

$$x=-\log\left(\frac{1-e^{\alpha-\alpha V^{*}(x;\alpha)}}{1-e^{\alpha}}\right)$$

Now suppose a random variable T has support [0, ∞] with PDF r_T and CDF R_T , and the random variable X has CDF F_X and PDF f_X , we define the CDF of the r_T - X family induced by V* with the following integral

$$K(x;\alpha;\xi,\beta)=\int_{0}^{\frac{\log\left(\frac{e^{\alpha+F_X(x;\xi)}}{e^{\alpha}+e^{F_X(x;\xi)}-1}\right)}{\alpha}} r_T(t,\beta)\,dt$$

where $x, \alpha \in$ R and $\alpha \neq 0$, ξ is a vector of parameters in the distribution of X, and β is a vector of parameters in the distribution of T. The parameter space for ξ and β, depends on the chosen distribution of the random variables X and T, respectively. By evaluating the above integral, we get the following

Proposition 4.1: The *CDF* of the γ_T-X family induced by V* is given by

$$K(x;\alpha,\xi,\beta)=r_T\left(\frac{\log\left(\frac{e^{\alpha+F_X(x;\xi)}}{e^{\alpha}+e^{F_X(x;\xi)}-1}\right)}{\alpha};\beta\right)$$

where T has support [0, ∞) with PDF γ_T and CDF R_T, and the random variable X has *CDF* F_X and *PDF* f_X, with x, α ∈ R and α ≠ 0 , ξ is a vector of parameters in the distribution of X, and β is a vector of parameters in the distribution of T. The parameter space for ξ and β, depends on the chosen distribution of the random variables X and T, respectively.

Remark 4.2. The *PDF* can be obtained by differentiating the *CDF*

Power Series Application of the New Family

Using Proposition 4.1 in Definition 3.4, we have the following

Proposition 5.1: *The* r_T – $X(V^{*})$ minimal family of power series distributions has *CDF*

$$1-\frac{C\left(\lambda\left(1-R_T\left(\frac{\log\left(\frac{e^{\alpha+F_X(x;\xi)}}{e^{\alpha}+e^{F_X(x;\xi)}-1}\right)}{\alpha};\beta\right)\right)\right)}{C(\lambda)}$$

where T has support $[0, \infty)$ with *PDF* γ_T and *CDF* R_T, and the random variable X has *CDF* F_X and *PDF* f_X, with $x, \alpha \in \mathrm{R}$ and $\alpha \neq 0$, ξ is a vector of parameters in the distribution of X, and β is a vector of parameters in the distribution of T, and C is some useful quantity for power series distributions.

Remark 5.2: The *PDF* can be obtained by differentiating the *CDF*

Practical Illustrations

In this section we show applicability of a so-called Exponential-Normal(V^*) family of distributions, and a so-called Exponential-Normal(V^*) Poisson family of distributions in fitting the breaking stress of carbon fibers data contained in Table 2 of Alzaatreh A, Lee C, Famoye F [8]. We make the following assumptions throughout this section. First, we assume T is an Exponential random variable with CDF

$R_T(t; b) = 1 - e^{-bt}$

where $x, b > 0$, and X is a Normal random variable with CDF

$$F_X(x;c,d)=\frac{1}{2}\,erfc\left(\frac{c-x}{\sqrt{2}d}\right)$$

where $\mathrm{erfc}(z) = 1 - \mathrm{erf}(z) = 1-\frac{2}{\sqrt{\pi}}\int_0^z e^{-t^2}dt$, $d > 0$, $x, c \in \mathrm{R}$

Using Table 1 above, we assume C is Poisson, that is, $C(\lambda) = e^{\lambda} - 1$, where $\lambda > 0$

Exponential-Normal (V*) Family of Distributions

Now we have the following from Proposition 4.1

Corollary 6.1. The CDF of the Exponential-Normal (V*) family of distributions is given by

$$K(x;b,c,d,\alpha)=1-\left(\frac{e^{\alpha}}{(e^{\alpha}-1)e^{-\frac{1}{2}erfc\left(\frac{c-x}{\sqrt{2}d}\right)}+1}\right)^{-\frac{b}{\alpha}}$$

where $b,d > 0$, and $x,c \in \mathrm{R}$

Remark 6.2. We write $S \sim EN\,(b, c, d, \alpha)$, if S is an Exponential-Normal (V*) random variable

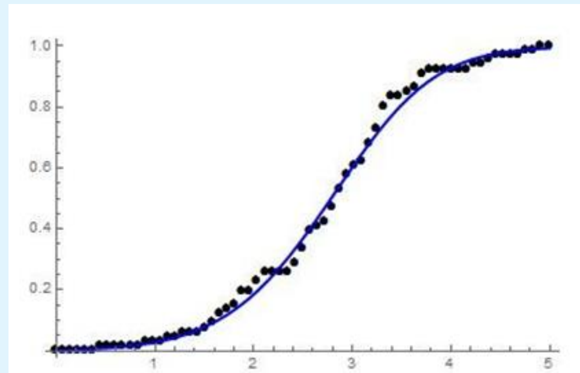

Figure 1: The CDF of *EN* (15.1051, 4.92022, 1.40262, 0.663882) fitted to the empirical distribution of the breaking stress of carbons fiber data, Table 2 [8].

Remark 6.3. The *PDF* of the Exponential-Normal (V*) family of distributions can be obtained by differentiating the *CDF*.

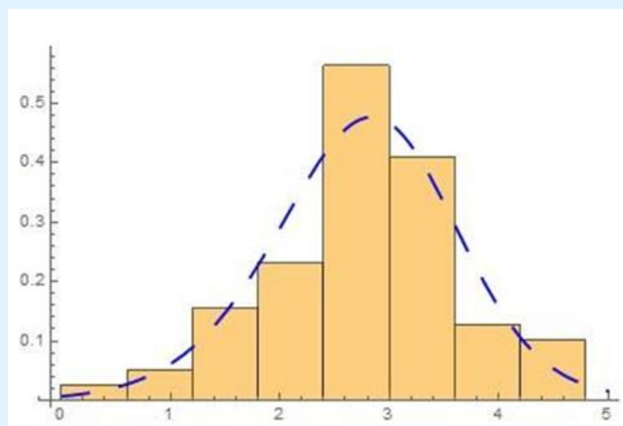

Figure 2: The PDF of *EN* (15.1051, 4.92022, 1.40262, 0.663882) fitted to the histogram of the breaking stress of carbons fiber data, Table 2 [8].

Exponential-Normal (V*) Poisson Family of Distributions

Now we have the following from Proposition 5.1

Corollary 6.4. The *CDF* the Exponential-Normal (V*) Poisson family of distributions is given by

$$Q\left(x;b,c,d,\alpha,\lambda\right)=\frac{e^{\lambda}-\exp\left(\lambda\left(\frac{e^{\alpha}}{\left(e^{\alpha}-1\right)e^{-\frac{1}{2}erfc\left(\frac{c-x}{\sqrt{2}d}\right)}+1}\right)^{-\frac{b}{\alpha}}\right)}{e^{\lambda}-1}$$

where $b,d,\lambda>0$, and $x,c\in\mathbb{R}$

Remark 6.5. We write $W \sim ENP\,(b, c, d, \alpha, \lambda)$, if W is an Exponential-Normal (V*) Poisson random variable. When $\lambda = 1$ is fixed, we simply write $ENP\ (b, c, d, \alpha)$

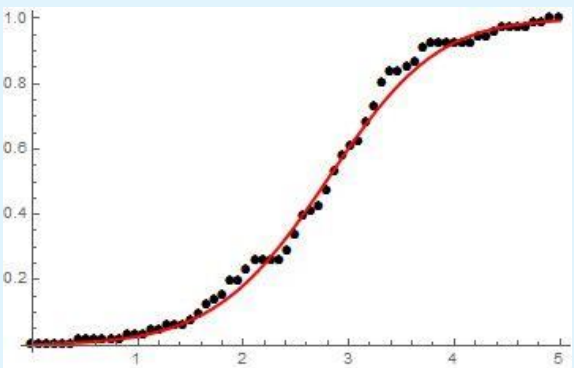

Figure 3: The CDF of *ENP* (12.1332, 6.03624, 1.57525, – 1.34131) fitted to the empirical distribution of the breaking stress of carbons fibers data, Table 2 [8].

Remark 6.6: The PDF of the Exponential-Normal (V*) Poisson family of distributions can be obtained by differentiating the *CDF*

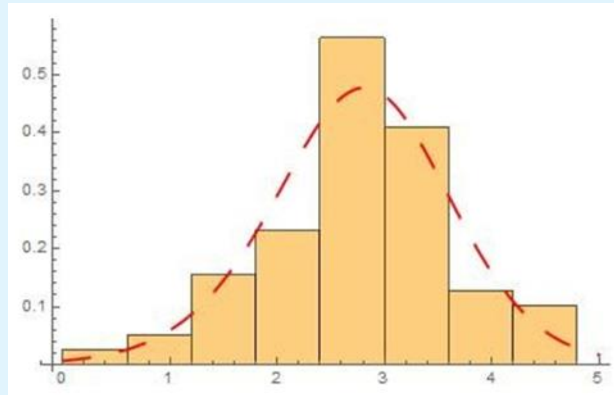

Figure 4: The PDF of ENP (12.1332, 6.03624,1.57525, –1.34131) fitted to the histogram of the breaking stress of carbons fibers data, Table 2 [8].

Numerical Comparisons

In this section we compare the Exponential-Normal (V^*), and the Exponential-Normal (V^*) Poisson, family of distributions, respectively, in fitting the breaking stress of carbon fibers data contained in Table 2 of Alzaatreh A, Lee C, Famoye F [8].

Model	Parameter Estimate	Standard Error
EN (b, c, d, α)	(15.1051,4.92022,1.40262,0.663882)	(103.117, 6.40881, 1.26158, 30.1413)
ENP (b, c, d, α)	(12.1332, 6.03624,1.57525,-1.34131)	(12.7002, 5.73428, 0.997658, 10.9083)

Table 2: Estimated Parameters for the breaking stress of carbon fibers data.

Model	-2(Log-likelihood)	AIC	AICC	BIC
EN (b, c, d, α)	170.636	178.636	179.292	187.395
ENP (b, c, d, α)	170.648	178.648	179.304	187.407

Table 3: Criteria for Comparison.

In order to compare the two distribution models, we used the following criteria: -2(Log- likelihood) and AIC(Akaike information criterion), AICC (corrected Akaike information criterion), and BIC (Bayesian information criterion) for the data set. The better distribution corresponds to the smaller -2(Log-likelihood) AIC, AICC, and BIC values:

$$AIC = 2k - 2l$$

$$AICC = AIC + \frac{2k\left(k+1\right)}{n-k-1}$$

$$BIC = k\log\left(n\right) - 2l$$

where k is the number of parameters in the statistical model, n is the sample size, and l is the maximized value of the log-likelihood function under the considered model. From Table 3 above, it is clear that the Exponential-Normal (V^*) distribution has the smallest values across all criteria considered, hence we see the Exponential-Normal (V *) distribution is a better fit than the Exponential-Normal (V*) Poisson distribution to the breaking stress of carbons fibers data.

Concluding Remarks and Further Recommendations

In this paper we have introduced two new broad classes of statistical distributions, and special members of these two classes of distributions are shown to be a good fit to real life data.

Now supposing $X(n) = max\{X_1, \cdots, X_N\}$, then the conditional CDF of $X_{(n)}|N = n$ is given by

$$F(x)^n$$

The CDF of the power series class of distributions associated with some parent distribution with PDF $f(x)$ and CDF $F(x)$ is the marginal CDF of $X_{(n)}$ which is given by

$$\frac{C(\lambda F(x))}{C(\lambda)}$$

Now we introduce the following

Definition 8.1. A power series class of distributions associated with some parent distribution whose *PDF* is *f(x)* and whose *CDF* is *F(x)*, and defined by the marginal *CDF* of $X_{(n)}$ which is given by

$$\frac{C(\lambda F(x))}{C(\lambda)}$$

will be called a Maximal Family of Power Series Distributions

Using Proposition 4.1 in Definition 8.1, we introduce the following

Proposition 8.2. The $r_T - X$ *(V*)* maximal family of power series distributions has *CDF*

$$\frac{C\left(\lambda\left(R_T\left(\frac{\log\left(\frac{e^{\alpha+F_X(x;\xi)}}{e^{\alpha}+e^{F_X(x;\xi)}-1}\right)}{\alpha};\beta\right)\right)\right)}{C(\lambda)}$$

where T has support $[0, \infty)$ with *PDF* r_T and *CDF* R_T, and the random variable X has *CDF* F_X and *PDF* f_X, with $x, \alpha \in R$ and $\alpha \neq 0$, ξ is a vector of parameters in the distribution of X, and β is a vector of parameters in the distribution of T , and C is some useful quantity for power series distributions.

Remark 8.3. The PDF of the $r_T - X(V^*)$ maximal family of power series distributions can be obtained by differentiating the *CDF*

The future interesting problem is to obtain some properties and applications of the $r_T - X$ (V*) maximal family of power series distributions

References

1. Ampadu CB (2019) The $r_T - X$ Family of Distributions induced by V: A New Method for Generating Continuous Distributions with Illustration to Cancer Patients Data. Annal Biostat & Biomed Appli 2(2): 1-3.

2. Moolath GB, Jayakumar K (2017) T-Transmuted X Family of Distributions. STATISTICA 77(3): 252-276.

3. Noack A (1950) A class of random variables with discrete distributions. The Annals of Mathematical Statistics 21(1): 127-132.

4. Johnson NL, Kemp AW, Kotz S (2005) Univariate discrete distributions. 3rd (Edn.), John Wiley and Sons.

5. Ampadu CB (2019) The Ampadu APT qT–X Family of Distributions Induced by V with an Illustration to Data in the Health Sciences. Annal Biostat & Biomed Appli 2(1): 1-5.

6. Anafo AY (2019) The New Alpha Power Transform: Properties and Applications. Master of Science in Mathematical Sciences Essay, African Institute for Mathematical Sciences (Ghana).

7. Ampadu CB (2019) A New T-X Family of Distributions. Annal Biostat & Biomed Appli 3(1).

8. Alzaatreh A, Lee C, Famoye F (2014) T-normal family of distributions: a new approach to generalize the normal distribution. Journal of Statistical Distributions and Applications 1: 16.