Brazilian Journal of Probability and Statistics
2015, Vol. 29, No. 4, 813–832
DOI: 10.1214/14-BJPS248

On the Harris-G class of distributions: General results and application

Luis Gustavo Bastos Pinho[a], **Gauss Moutinho Cordeiro**[a] **and Juvêncio Santos Nobre**[b]

[a]*Federal University of Pernambuco*
[b]*Federal University of Ceará*

Abstract. We investigate some properties of the Harris-G class of distributions (*Sankhya B* **73** (2011) 70–80). We demonstrate that the density function of the Harris-G class can be expressed as a linear combination of density functions of the exponentiated baseline distribution. We provide general formulas for moments (raw, centered, incomplete and factorial), quantile function, generating functions and entropies. Two numerical examples are presented to demonstrate the potentiality of the models in this class. The first one applies the Harris–Burr XII distribution to model bimodal data. The second example uses the Harris-exponential distribution to model SAR image data. The results of the fitted models look promising.

1 Introduction

In this paper, we present some results on the Harris extended class of distributions, which we refer to as the *Harris-G class*. The distributions in this class are obtained from a general construction method explained in Aly and Benkherouf (2011) and the distribution proposed in Harris (1948).

Consider a sequence of i.i.d. random variables $Z_1, Z_2, \dots$ with cumulative distribution function (c.d.f.) $G(x)$, probability density function (p.d.f.) $g(x)$ and hazard rate function (h.r.f.) $h(x)$. Let $X = \min\{Z_1, Z_2, \dots, Z_N\}$, where N is a discrete random variable with probability generating function (p.g.f.) $\varphi(\cdot, \theta)$, $\theta > 0$ and support in $\mathbb{N}$. It is known that the survival function $\bar{F}(x)$ of X is given by

$$\bar{F}(x) = \varphi[\bar{G}(x), \theta], \tag{1.1}$$

where $\bar{G}(x) = 1 - G(x)$. Equation (1.1) provides a strong physical motivation for explaining the failure time of a system made of N serial subsystems, where the subsystems are identical although the number N of subsystems is unknown. Many properties of (1.1) were presented in Aly and Benkherouf (2011). Unless otherwise stated, all of the results presented in the paper are new and original. Their method introduces new parameters in an existing distribution and had already been used in other papers. For example, the case $\bar{G}(x) = \exp(-\lambda x)$, $\lambda > 0$

Key words and phrases. Burr XII distribution, entropies, extended exponential distributions, Harris-G class, power series expansion.

Received February 2013; accepted May 2014.

and $\varphi(s,\theta)$ taken as the zero truncated Poisson distribution, namely $\varphi(s,\theta) = \exp[\theta(s-1)][1-\exp(-\theta)]^{-1}$, leads to the *exponential Poisson* (EP) distribution (Kuş, 2006). When $\bar{G}(x) = \exp(-\lambda x)$ and $\bar{G}(x) = \exp[-(\lambda x)^{\alpha}]$ and $\varphi(s,\theta)$ is the generic power series distribution, we obtain the models developed by Chahkandi and Ganjali (2009) and Morais and Barreto-Souza (2011), respectively. Other proposals defined by this kind of mixture can be found in Barreto-Souza and Bakouch (2013), Lu and Shi (2012), Ristić and Nadarajah (2012) and references therein.

One of the most popular versions of (1.1) was proposed in Marshall and Olkin (1997) by replacing $\varphi(\cdot,\theta)$ with the geometric p.g.f. $\varphi(s,\theta) = s\theta(1-\bar{\theta}s)^{-1}$, where $\bar{\theta} = 1-\theta$. In the same paper, the method was applied to the exponential and Weibull distributions, resulting in the Marshall–Olkin extended exponential (MOEE) and Marshall–Olkin extended Weibull (MOEW) models, respectively. Aly and Benkherouf (2011) generalized the Marshall–Olkin class considering the p.g.f. of the Harris distribution (Harris, 1948) for obtaining new distributions. This p.g.f. is given by

$$\varphi(s,k,\theta) = \left(\frac{\theta s^k}{1-\bar{\theta}s^k}\right)^{1/k}, \qquad k > 0. \tag{1.2}$$

Aly and Benkherouf (2011) considered $\theta > 0$ although Harris (1948) restricted θ to the interval $(0,1)$. This restriction comes from the fact that the Harris distribution arises from a branching process for which each node may originate k new nodes with probability proportional to $-\log(\theta)$. Inserting (1.2) in (1.1) leads to

$$\bar{F}(x) = \left[\frac{\theta\bar{G}(x)^k}{1-\bar{\theta}\bar{G}(x)^k}\right]^{1/k}. \tag{1.3}$$

The p.d.f. corresponding to (1.3) is given by

$$f(x) = \frac{\theta^{1/k}g(x)}{[1-\bar{\theta}\bar{G}(x)^k]^{1+1/k}}. \tag{1.4}$$

Equation (1.4) reduces to that one proposed by Marshall and Olkin when $k = 1$. As explained below, equation (1.4) provides better fits than several of the known lifetime distributions, which is a remarkable feature. Particularly, the Harris-exponential distribution obtained by replacing $G(x)$ by $\lambda\exp(-\lambda x)$ seems to provide a simple alternative for SAR image modeling, as illustrated in Section 7. This is very interesting, since most distributions used in the modeling of SAR images are considerably complicated.

This rest of the paper is divided as follows. In Section 2, we demonstrate that the Harris-G p.d.f. can be expressed as a linear combination of *exponentiated-G* ("exp-G") densities. This result allows us to obtain some structural properties of the Harris-G class. In Section 3, we present the ordinary and incomplete moments and the moment generating function (m.g.f.) for this class. The Shannon and Rényi entropies are derived in Section 4. In Section 5, we present a double power series for the quantile function of the Harris-G class. Estimation by maximum likelihood is briefly discussed in Section 6. Applications to two real datasets are explored in Section 7. Finally, concluding remarks are addressed in the last section. In the Appendix, we present the proofs of the theorems given in the text.

2 Useful expansions

We demonstrate that the Harris-G density function is a linear combination of exp-G densities. Let $g_\alpha(x) = \alpha g(x) G(x)^{\alpha-1}$ be the exp-G density function with power parameter $\alpha > 0$.

Theorem 1. *For* $0 < \theta < 1$, *we can write* $f(x)$ *in* (1.4) *as*

$$f(x) = \sum_{i=1}^{\infty} w_i g_i(x),$$

where

$$w_i = (-1)^{i+1} \sum_{j=i}^{\infty} \binom{j}{i} \frac{1}{j!} \frac{\partial^j \varphi(t, s, \theta)}{\partial t^j}\bigg|_{t=0}.$$

The second way of writing $f(x)$ as a linear combination of exp-G densities is obtained after some algebra. In the following results, we will make extensive use of the generalized binomial coefficient given by

$$\binom{a}{b} = \frac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)},$$

where $\Gamma(a) = \int_0^\infty t^{a-1} \mathrm{e}^{-t}\, dt$ is the gamma function.

Theorem 2. *For* $\theta \in (0, \infty)$, *we can write* $f(x)$ *in* (1.4) *as*

$$f(x) = \sum_{i=0}^{\infty} w_i g_{i+1}(x), \tag{2.1}$$

where $g_{i+1}(x) = (i+1) g(x) G(x)^i$, $\tau = \theta^{-1}$, $\bar{\tau} = 1 - \tau$ and

$$w_i = \begin{cases} \dfrac{(-1)^i \theta^{1/k}}{i+1} \left[\displaystyle\sum_{j=0}^{\infty} \bar{\theta}^j \binom{j + k^{-1}}{j} \binom{kj}{i} \right], & \text{if } 0 < \theta < 1, \\ \dfrac{(-1)^i \tau}{i+1} \left\{ \displaystyle\sum_{j=0}^{\infty} \sum_{l=0}^{j} (-1)^l \bar{\tau}^j \binom{j + k^{-1}}{j} \binom{j}{l} \binom{lk}{i} \right\}, & \text{if } \theta > 1. \end{cases}$$

Henceforth, we use w_i for the coefficients in Theorem 2. Writing the p.d.f. of the Harris-G distribution as a linear combination of exp-G densities allows us to obtain several structural properties for the Harris-G class, since we can use the results on exponentiated distributions.

For $0 < \theta < 1$, we can choose between Theorems 1 and 2. The lack of closed-form expression for the derivatives of the Harris p.g.f. suggests that the Theorem 2 may be a better choice for computation. However, it is possible to obtain the derivatives in Theorem 1 with an error as small as we want using finite difference methods, which do not require great computational effort. Very often, we can write these densities as linear combination of simpler densities than those from the exponentiated distributions. For example, the Harris-exponential p.d.f. may be written as a linear combination of exponential densities. The same holds for the Harris–Burr XII distribution. The proof of these results mimics closely the proof of Theorem 2. Equation (2.1) is the main result of this section.

Finally, the following result may be used for the parametric characterization of the Harris-G density function.

Theorem 3. *If $g(x)$ is monotonically decreasing in an open interval D and $0 < \theta < 1$, then $f(x)$ is also monotonically decreasing in D. Besides, if $g(x)$ is monotonically increasing in D and $\theta > 1$, then $f(x)$ is also monotonically increasing in D.*

3 Moments and generating function

Henceforth, let X be a random variable distributed according to (1.4) and Y_i be a random variable having the exp-G distribution with power parameter $i + 1$. Based on Theorem 2, we can write the m.g.f. of X as a linear combination of m.g.f.'s of Y_i (for $i = 0, 1, \dots$), namely

$$M_X(t) = \sum_{i=0}^{\infty} w_i M_{Y_i}(t).$$

The raw jth moment of X becomes

$$\mu'_j = \sum_{i=0}^{\infty} w_i \mu'^{(i)}_j, \tag{3.1}$$

where $\mu'^{(i)}_j$ represents the jth raw moment of Y_i. Central moments (μ_j), cumulants (κ_j) and factorial moments ($\mu_{(j)}$) can be obtained from the following relations:

$$\mu_j = \sum_{i=0}^{j} \binom{j}{i} (-1)^{j-i} \mu'_j \mu'^{\,j-i}_1, \qquad \kappa_j = \mu'_j - \sum_{i=1}^{j-1} \binom{i-1}{k-1} \kappa_i \mu'_{j-i},$$

and

$$\mu_{(j)} = \mathbb{E}\big[X(X-1) \times \cdots \times (X-j+1)\big] = \sum_{i=0}^{j} s(j,i)\mu'_i,$$

where $s(r,i) = (i!)^{-1}[d^i x^{(r)}/dx^i]_{x=0}$ represents the Stirling number of the first kind, and $x^{(n)} = x(x+1)\cdots(x+n-1)$ denotes the ascending factorial.

Incomplete moments are useful in several areas such as insurance and credibility, and allow us to obtain the mean deviations around the mean and median. The jth incomplete moment of X, denoted by $J_j(z) = \int_0^z x^j f(x)\,dx$, can also be obtained from Theorem 2 as a linear combination of the incomplete moments of Y_i. Denoting the jth incomplete moment of Y_i as $J_j^{(i)}(z) = \int_0^z x^j dG(x)^{i+1}$, we have

$$J_j(z) = \sum_{i=0}^{\infty} w_i J_j^{(i)}(z). \tag{3.2}$$

Equations (3.1) and (3.2) are the main results of this section.

4 Entropies

Entropy refers to the amount of uncertainty (or surprisal) associated to a random variable. It is an important concept in many areas of knowledge, specially theory of communication, physics and probability. Possibly the most famous measure of entropy was introduced in the seminal work of Shannon (Shannon, 1948). For a continuous distribution $F(x)$ with density $f(x)$, the Shannon entropy is given by

$$\mathcal{H}_{\mathrm{Sh}}[f(x)] = \mathbb{E}_X\{-\log[f(X)]\} = -\int_{-\infty}^{+\infty} \{\log[f(x)]\} f(x)\,dx.$$

It is possible to relate $\mathcal{H}_{\mathrm{Sh}}[f(x)]$ with $\mathcal{H}_{\mathrm{Sh}}^{i+1}[g_{i+1}(y)]$, $i \in \mathbb{N}$, where $\mathcal{H}_{\mathrm{Sh}}^{i+1}$ represents the measure of entropy based in the distribution of Y_i. For doing that, the following lemma will be useful.

Lemma 1. *For the random variable Y_i having the exp-G distribution with power parameter $i+1$, we have*

$$\mathbb{E}_{Y_i}\{\log[g(Y_i)]\} = \mathcal{H}_{\mathrm{Sh}}^{i+1}[g_{i+1}(y)] - \log(i+1) + \frac{1}{i+1}.$$

The proof of Lemma 1 is given in the Appendix. We now give the following result.

Theorem 4. *The Shannon entropy for the Harris-G distribution can be expressed as a linear combination of entropies of exp-G distributions, namely*

$$\mathcal{H}_{\mathrm{Sh}}[f(x)] = -\frac{1}{k}\log(\theta) + \left(1 + \frac{1}{k}\right) I(k,\theta) + \sum_{i=0}^{\infty} w_i \left\{\mathcal{H}_{\mathrm{Sh}}^{i+1}[g_{i+1}(x)] - \log(i+1) + \frac{1}{i+1}\right\},$$

where

$$I(k,\theta)=\theta^{1/k}\int_0^1 \frac{\log(1-\bar{\theta}t^k)}{(1-\bar{\theta}t^k)^{1+1/k}}\,dt.$$

Notice that $I(k,\theta)$ does not depend on the choice of $g(x)$. The proof of Theorem 4 is given in the Appendix.

The entropy of Shannon may also be used to identify probability models. Such an approach can be seen in Jaynes (1957). Consider a class of distributions defined by a set of constraints such as

$$\mathcal{F}=\{f(x)|\mathbb{E}_X[T_i(X)]=a_i, i=1,2,\dots,m\},$$

where $a_i\in\mathbb{R}$, $\forall i$. We can choose a member of $\mathcal{F}$ as the p.d.f. of a random variable X if it maximizes the Shannon entropy under these constraints. The chosen p.d.f. is called maximum entropy distribution. This approach ensures that no other assumptions except those from the constraints are made. For instance, we can show that if the first and second moments are constrained, the maximum entropy distribution is the normal distribution. The statistical approach for the theory of information is a vast and interesting field. The following result gives the maximum entropy characterization for the Harris-G class.

Theorem 5. *Let $h(x)$ be a p.d.f. under the following constraints*:

- $\mathbb{E}_X[1-\bar{\theta}\bar{G}(X)^k]=I(k,\theta)$,
- $\mathbb{E}_X\{\log[g(X)]\}=\mathbb{E}_W[\log\{g[G^{-1}(W)]\}]$,

where W has p.d.f. given by $m(w)=\theta^{k^{-1}}(1-\bar{\theta}w^k)^{1+k^{-1}}$. Then, $f(x)$ defined in (1.4) *is the only solution for the optimization problem*

$$f(x)=\arg\max_h \mathcal{H}_{\mathrm{Sh}}[h(x)].$$

The Rényi entropy is another very popular measure. It is defined for continuous distributions by

$$\mathcal{H}_R^{\alpha}[f(x)]=\frac{1}{1-\alpha}\log\left[\int_{-\infty}^{+\infty} f(x)^{\alpha}\,dx\right],\qquad \alpha>0.$$

The Shannon entropy follows as a special case when $\alpha\to 1$. We derive the following result.

Theorem 6. *The Rényi entropy of the Harris-G distribution is given by*

$$\mathcal{H}_R^{\alpha}[f(x)]=\frac{1}{1-\alpha}\log\left[\sum_{i=0}^{\infty}a_i\int_{-\infty}^{+\infty} g(x)^{\alpha}G(x)^i\,dx\right],$$

where

$$a_i = \begin{cases} (-1)^i \theta^{\alpha/k} \left[\sum_{j=0}^{\infty} \bar{\theta}^j \binom{kj}{i} \binom{j+\alpha+\alpha k^{-1}-1}{j} \right], & \text{if } 0 < \theta < 1, \\ (-1)^i \tau^{\alpha} \left[\sum_{j=0}^{\infty} \sum_{l=0}^{j} (-1)^l \bar{\tau}^j \binom{j}{l} \binom{lk}{i} \binom{j+\alpha+\alpha k^{-1}-1}{j} \right], & \text{if } \theta > 1 \end{cases}$$

and $\tau = \theta^{-1}$ *as before.*

The above integral can be computed, at least numerically, for most baseline distributions. The proof of this result is much similar to the that of Theorem 2, and thus it is not included here.

5 Quantile function

The quantile function of X is easily obtained by inverting (1.3). We have

$$x = Q_F(u) = Q_G\left(1 - \frac{(1-u)}{[\theta + \bar{\theta}(1-u)^k]^{1/k}}\right), \tag{5.1}$$

where $Q_G(u) = G^{-1}(u)$ is the quantile function of G.

The use of power series methods is at the heart of many aspects of applied mathematics and statistics. If the function $Q_G(u)$ does not have a closed-form expression, it can usually be expressed as a power series of the form

$$Q_G(u) = \sum_{i=0}^{\infty} e_i u^i, \tag{5.2}$$

where the coefficients e_i are suitably chosen real numbers. For several important distributions, such as the normal, Student t, gamma and beta distributions, $Q_G(u)$ does not have closed-form but it can be expressed as (5.2). For example, for the standard normal distribution, the coefficients e_i are given by

$$e_i = (2\pi)^{i/2} \sum_{m=i}^{\infty} \left(\frac{-1}{2}\right)^{m-j} \binom{m}{i} p_i,$$

where the quantities p_i are defined by $p_i = 0$ for $i = 0, 2, 4, \ldots$ and $p_i = q_{(i-1)/2}$ for $i = 1, 3, 5, \ldots$, and the q_k's are calculated recursively from

$$q_{k+1} = \frac{1}{2(2k+3)} \sum_{r=0}^{k} \frac{(2r+1)(2k-2r+1) q_r q_{k-r}}{(r+1)(2r+1)}.$$

Here, $q_0 = 1$, $q_1 = 1/6$, $q_2 = 7/120$, $q_3 = 127/7560, \ldots$.

For $\theta > 1/2$, we can obtain a power series for the denominator of (5.1) using the expansion given in the Appendix (Theorem 2). We have

$$[\theta + \bar{\theta}(1-u)^k]^{-1/k} = \theta^{-1/k}\left[1 + \frac{\bar{\theta}}{\theta}(1-u)^k\right]^{-1/k} = \sum_{j=0}^{\infty} m_j (1-u)^{kj},$$

where

$$m_j = \theta^{-1/k}\binom{j-k^{-1}-1}{j}\left(-\frac{\bar{\theta}}{\theta}\right)^j.$$

Then the argument of $Q_G(\cdot)$ in (5.1) becomes

$$1 - \frac{(1-u)}{[\theta + \bar{\theta}(1-u)^k]^{1/k}} = 1 - \sum_{j=0}^{\infty} m_j (1-u)^{kj+1}.$$

Using (5.2), $Q_F(u)$ reduces to

$$\begin{aligned} Q_F(u) &= \sum_{i=0}^{\infty} e_i \left(1 - \sum_{j=0}^{\infty} m_j (1-u)^{kj+1}\right)^i \\ &= \sum_{i=0}^{\infty} e_i \sum_{p=0}^{i} (-1)^p \binom{i}{p} \left(\sum_{j=0}^{\infty} m_j (1-u)^{kj+1}\right)^p. \end{aligned}$$

Now, we consider a result of Gradshteyn and Ryzhik (2000, Section 0.314) for a power series raised to a positive integer p

$$\left(\sum_{j=0}^{\infty} m_j z^j\right)^p = \sum_{j=0}^{\infty} c_{p,j} z^j, \tag{5.3}$$

where the coefficients $c_{p,j}$ (for $j = 1, 2, \ldots$) are easily obtained from the recurrence equation

$$c_{p,j} = (j m_0)^{-1} \sum_{\ell=1}^{i} [\ell(p+1) - j] m_\ell c_{p,i-\ell}, \tag{5.4}$$

and $c_{p,0} = m_0^n$. The coefficient $c_{p,m}$ can be determined from $c_{p,0}, \ldots, c_{p,j-1}$, and hence from the quantities $m_0, \ldots, m_j$. We can write

$$\left(\sum_{j=0}^{\infty} m_j (1-u)^{kj+1}\right)^p = (1-u)^{pj} \left(\sum_{j=0}^{\infty} m_j (1-u)^{kj}\right)^p = \sum_{j=0}^{\infty} c_{p,j} (1-u)^{(p+k)j},$$

where the coefficients $c_{p,j}$ are determined by (5.4). Finally, changing $\sum_{i=0}^{\infty}\sum_{p=0}^{i}$ by $\sum_{p=0}^{\infty}\sum_{i=p}^{\infty}$, we can write $Q_F(u)$ as a double power series given by

$$Q_F(u) = \sum_{p,j=0}^{\infty} d_{p,j} (1-u)^{(p+k)j}, \tag{5.5}$$

where $d_{p,j} = (-1)^p c_{p,j} \sum_{i=p}^{\infty} e_i \binom{i}{p}$. Combining equations (5.3) and (5.5) gives a useful procedure to obtain alternative expressions for the ordinary and incomplete moments, generating function, mean deviations of any distribution in the Harris-G class.

6 Estimation

Estimation of the model parameters of the Harris-G distribution can be accomplished by the maximum likelihood method. Let $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_j)^\top$ be the vector of parameters of $G(x)$. Based on a sample $x_1, x_2, \ldots, x_n$, the logarithm of the likelihood function for the parameters in (1.4) is given by

$$\ell = \ell(\theta, k, \boldsymbol{\phi}^\top) = \frac{n}{k}\log(\theta) + \sum_{i=1}^{n}\log[g(x_i)] - \left(1 + \frac{1}{k}\right)\sum_{i=1}^{n}\log[1 - \bar{G}(x_i)^k].$$

Then we can write

$$\frac{\partial \ell}{\partial k} = -\frac{n}{k^2} + \frac{1}{k^2}\sum_{i=1}^{n}\log[1 - \bar{\theta}\bar{G}(x_i)^k] + \bar{\theta}\left(1 + \frac{1}{k}\right)\frac{\bar{G}(x_i)^k \log[\bar{G}(x_i)^k]}{1 - \bar{\theta}\bar{G}(x_i)^k},$$

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{k\theta} - \left(1 + \frac{1}{k}\right)\sum_{i=1}^{n}\frac{\bar{G}(x_i)^k}{1 - \bar{\theta}\bar{G}(x_i)^k}$$

and

$$\frac{\partial \ell}{\partial \phi_j} = \sum_{i=1}^{n}\frac{\partial \log[g(x_i)]}{\partial \phi_j} + \bar{\theta}\left(1 + \frac{1}{k}\right)\sum_{i=1}^{n}\frac{1}{1 - \bar{\theta}\bar{G}(x)^k}\frac{\partial \bar{G}(x_i)^k}{\partial \phi_j}.$$

Setting these derivatives to zero yields the maximum likelihood estimators (MLEs) of the Harris-G parameters. Unfortunately, the $j + 2$ equations cannot seem to be simplified any further for a generic distribution G and we require any iterative numerical method such as the Newton–Raphson or quasi-Newton procedures, even in the simple cases. Under general regularity conditions, the asymptotic distribution of $(\hat{\theta}, \hat{k}, \hat{\boldsymbol{\phi}}^\top)$ is $N_{j+2}(\mathbf{0}, \mathbf{K}^{-1})$, where $\mathbf{K} = \mathbf{K}(\theta, k, \boldsymbol{\phi}^\top)$ is the expected information matrix. The matrix $\mathbf{K}$ can be replaced by the observed information matrix for constructing approximate confidence intervals for the model parameters.

Care is advised when extracting a numerical approximation for the matrix $\mathbf{K}$ from the iterative methods used to obtain the point estimates of the parameters. For some methods, such as the BFGS, an approximation of the Hessian matrix is used in the calculations at each iteration. This approximation may not be reliable if the convergence of the methods happens too fast. If the number of iterations is small (e.g., five iterations), using the BFGS method, the output it gives for the approximate Hessian matrix may be unreliable, even when the point estimates are very accurate. Bootstrap confidence intervals are a reliable alternative in these cases.

The convergence of the estimation procedures usually depends on the choice of the starting values of the parameters. We advise first using a nondeterministic optimization routine, such as simulated annealing, to obtain the initial guesses of the parameters and then using the Newton or quasi-Newton methods. Although it adds to the total computational time, this is usually reliable, especially when dealing with simulation and bootstrap.

Other estimation methods such as the method of moments (see Cramér, 1946, Section 33) or the generalized method of moments (Hansen, 1982) may be used. Particularly, the generalized method of moments may be used in conjunction with the maximum entropy characterization given in this paper to produce estimates of the parameters. This is, however, a discussion which may be long, and thus fit to be presented in a separate work.

7 Applications

To motivate the use of the distributions in the Harris-G class, we present two applications for the Harris-exponential and Harris–Burr XII distributions.

The Harris–Burr XII distribution is obtained by inserting the Burr XII survival function in (1.3). The Burr XII survival function is given by

$$\bar{G}(x) = \left(1 + x^c\right)^{-\alpha}, \qquad x, c, \alpha > 0.$$

Inserting $\bar{G}(x)$ in (1.3) yields

$$\bar{F}(x) = \left[\frac{\theta(1 + x^c)^{-k\alpha}}{1 - \bar{\theta}(1 + x^c)^{-k\alpha}}\right]^{1/k}, \qquad k, \theta > 0.$$

If we add location and scale parameters to the Harris–Burr XII distribution, it becomes a very flexible model. For doing this, we just change x by $\sigma^{-1}(x - \mu)$, with $\sigma > 0$ and $\mu \in \mathbb{R}$ in the equation for $\bar{F}(x)$ as usual. In this case, we have $x > \mu$.

Consider the `faithful` dataset found in the `datasets` package of the default installation of `R`. This dataset contains 272 observations of two random variables which are traditionally modeled using a mixture of two normal distributions. The second column of this dataset corresponds to the waiting times between consecutive eruptions of the Old Faithful, a geyser located in the Yellowstone area in the United States. Figure 1 displays the histogram for the dataset. It is fairly evident that this dataset presents two modes.

The six parameter Harris–Burr XII distribution was fitted to these data. The parameters were estimated using the `fitdistr()` function from the `MASS` package, with initial values provided by simulated annealing using the `GenSA()` func-

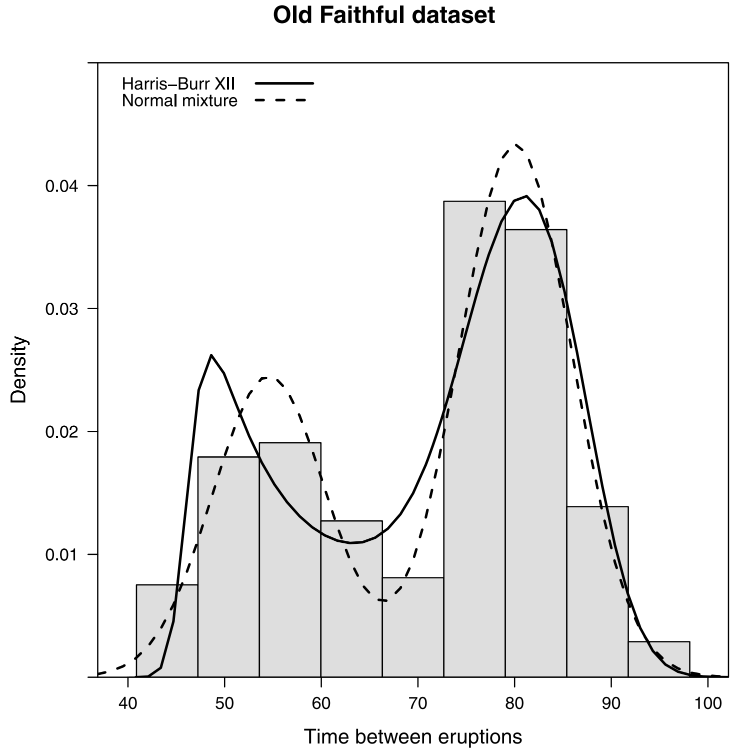

Figure 1 *The histogram and two fitted densities to this dataset.*

tion from the `GenSA` package in `R`. Also, a mixture of normal densities given by

$$h(x) = \lambda\sigma_1^{-1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right) + (1-\lambda)\sigma_2^{-1}\phi\left(\frac{x-\mu_2}{\sigma_2}\right),$$

with $0 < \lambda < 1$, $\sigma_1, \sigma_2 > 0$ and $\phi(\cdot)$ representing the standard Gaussian function was fitted using the `nomalmixEM()` function from the `mixtools` in `R`. The two fitted distributions are given in Table 1. The logarithm of the maximized likelihood function for both models were close: -1034.03 for the normal mixture and -1038.44 for the Harris–Burr XII model. The AIC and BIC were -1024.04 and -1021.17 for the normal mixture and -1026.42 and -1023.35 for the Harris–Burr XII model, respectively. Both estimated densities are superposed to the data histogram in Figure 1.

The normal mixture model has as advantage the good mathematical properties of the normal distribution. However, when used to model strictly positive data, with prediction being the main focus, this model may assign positive probability to negative values. Such a "trap" is reasonably more dangerous in the context of regression models, since points in the remote regions of the design matrices column space may lead to unreasonable predictions. From this point of view, the Harris–

Table 1 *Fitted distributions to the old faithful dataset*

Model	Parameter	MLE	Standard error	AIC	BIC
Harris–Burr XII	k	22.3062	2.0982×10^{-2}	-1026.42	-1023.35
	θ	0.0001	8.5370×10^{-5}		
	c	48.7521	7.3411×10^{-2}		
	α	6.6527	4.5274×10^{-4}		
	σ	75.2536	1.9180×10^{-4}		
	μ	40.0204	2.7802×10^{-1}		
Normal mixture	λ	0.3609	0.0322	-1024.04	-1021.17
	μ_1	54.6149	0.7047		
	σ_1	5.8712	0.5448		
	μ_2	80.0911	0.4980		
	σ_2	5.8677	0.3909		

Burr XII model is preferred for modeling this dataset although at the expense of an extra parameter.

In the second application, we present the Harris-exponential (HEE) distribution. This model is obtained by inserting the exponential distribution survival function in (1.3). The resulting survival function is given by (for $x, \lambda, k, \theta > 0$)

$$\bar{F}(x) = \left[\frac{\theta\exp(-\lambda k x)}{1-\bar{\theta}\exp(-\lambda k x)}\right]^{1/k}. \tag{7.1}$$

Many properties of this distribution are given in Pinho et al. (2015). This rather simple distribution finds application in the field of SAR [synthetic aperture radar] image statistical modeling. SAR images are used in remote sensing and mapping of surfaces, even outside our planet. The statistical treatment of SAR images allows for target detection, boundary identification and noise reduction, for example. This application is indeed the main reason why this paper was written.

We sampled an homogeneous area of a SAR image from Foulum (Denmark) and extracted the values of the intensity of the pixels in the sampled area. The details and actual dataset, which consists of 5588 points, are available in plain text format upon request. Two distributions are very popular alternatives for modeling this kind of dataset: the K distribution and the G^0 distribution. For more details on these two distributions, see Frery et al. (1997). The five parameter beta-generalized normal (BGN) is presented in Cintra et al. (2014) as an alternative to the K and G^0 distributions. The gamma distribution is sometimes used in this context. We compared the Harris-exponential fit to the ones of the K, G^0, BGN, gamma and exponentiated-exponential (EE) distributions. The distributions were fitted using the `fitdistr()` routine in `R`. Table 2 gives a summary of the three fitted distributions. We consider the logarithm of the maximized likelihood, Akaike and Bayesian information criteria (AIC and BIC) and Cramér–von-Mises (CVM)

Table 2 *Fitted distributions to the foulum dataset*

Model	Parameter	MSE	Standard error	AIC	BIC	CVM
BGN	a	7.20×10	6.64×10^{-1}	−60058.15	−60025.01	28.51
	b	6.11×10	5.93×10^{-1}			
	μ	1.34×10^{-3}	2.43×10^{-5}			
	σ	5.99×10^{-3}	1.63×10^{-4}			
	s	5.99×10^{-1}	1.77×10^{-2}			
HEE	k	1.26×10	2.30×10^{-1}	−62615.02	−62595.16	2.23
	θ	3.09×10^{5}	1.05×10^{3}			
	λ	8.31×10^{2}	1.22×10			
K	μ	2.37×10^{-3}	1.47×10^{-5}	−61218.88	−61199.00	19.65
	ν	5.76	5.53×10^{-1}			
	L	3.40×10	4.44×10^{-1}			
G^0	n	7.82×10	1.28	−62413.62	−62393.74	962.31
	α	−6.23	1.92×10^{-1}			
	γ	2.47×10^{-2}	8.61×10^{-4}			
EE	α	9.04	0.28	−61650.39	−61637.14	14.66
	λ	1208.12	15.66			
Gama	α	4.53	0.08	−60897.36	−60884.10	24.06
	β	1893.89	36.55			

statistics as measures of the quality of the fits. For information on the Cramér–von-Mises, see Chen and Balakrishnan (1995). Generally speaking, we may favor models with lower values of the CVM statistic. The Harris-exponential distribution outperforms the K, G^0 and the other distributions for the current dataset based on the chosen criteria. Figure 2 displays the HEE, K and G^0 fitted densities overlapping the data histogram. It suggests that the three distributions are well suited to the current dataset.

However, the Harris-exponential distribution seems not to be adequate for heterogeneous or highly heterogeneous areas, such as cities for example. Although the K and G^0 distributions fitted to the data almost as well as the Harris-exponential distribution, it is computationally much easier to deal with the latter distribution. The mathematical properties of the Harris-exponential distribution are also much simpler than those from the K and G^0 distribution.

8 Conclusions

We present several new structural properties of the Harris-G class of distributions, including moments, generating and quantile functions, mean deviations, entropies

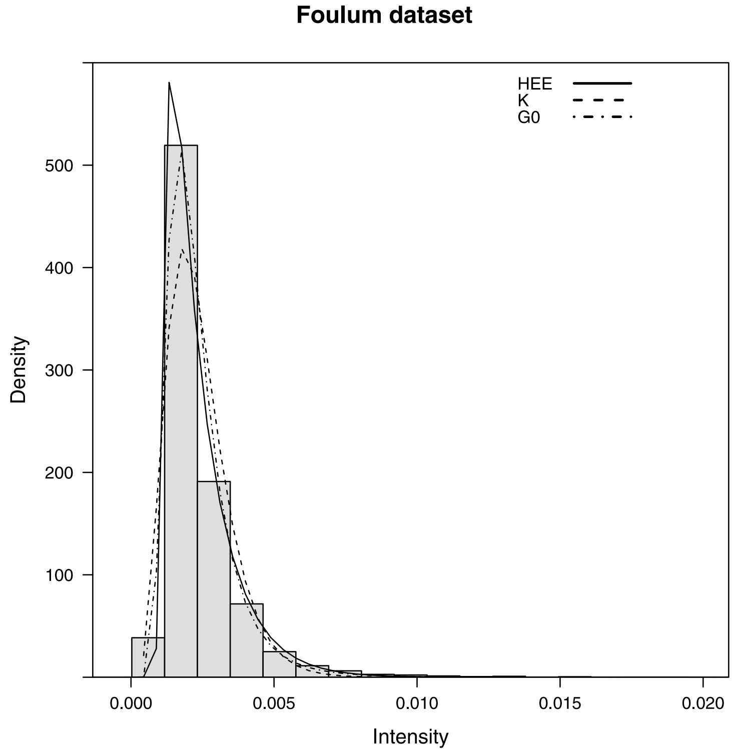

Figure 2 *The histogram and three fitted densities to this dataset.*

and order statistics. These results are different from those obtained by Aly and Benkherouf (2011). Also, we provide two applications of special models of the Harris-G class to real univariate datasets to promote the use of this class. The applications indicate that some distributions in this class are very promising. Particularly, we expect this fact will attract more attention to the Harris-G class in reliability, engineering and in other areas of research. Properties of equation (1.4) not considered in this paper are: cumulative residual entropy, Song's measure, goodness-of-fit tests, Bayesian estimation, estimation using bootstrap and estimation using L-moments. We hope to address some of these in a future paper.

The infinite sums presented in this paper, for practical purposes, can be limited to 30 or 40 terms and the error is already very small. Of course, one can always use numerical integration where those sums are suggested. However, numerical integration often misbehaves or fails completely. When this happens inside some software routines, the results may be very different from the ones intended. These sums are a useful alternative in some cases.

Appendix: Proofs of the results

Proof of Theorem 1. For $0 < \theta < 1$, if $\varphi(s, k, \theta)$ is a p.g.f., we can write

$$\begin{aligned}
\bar{F}(x) &= \varphi\big(\bar{G}(x), k, \theta\big) \\
&= \sum_{j=1}^{\infty} P(N = j)\bar{G}(x)^j = \sum_{j=1}^{\infty} \frac{1}{j!} \frac{\partial^j \varphi(t, s, \theta)}{\partial t^j}\bigg|_{t=0} \bar{G}(x)^j \\
&= \sum_{j=1}^{\infty} \frac{1}{j!} \frac{\partial^j \varphi(t, s, \theta)}{\partial t^j}\bigg|_{t=0} \left[\sum_{i=0}^{j} \binom{j}{i} (-1)^i G(x)^i\right] \\
&= \sum_{i=1}^{\infty} (-1)^i \left[\sum_{j=i}^{\infty} \binom{j}{i} \frac{1}{j!} \frac{\partial^j \varphi(t, s, \theta)}{\partial t^j}\bigg|_{t=0}\right] G(x)^i.
\end{aligned}$$

By differentiating both sides, we have

$$f(x) = \sum_{i=1}^{\infty} (-1)^{i+1} \left[\sum_{j=i}^{\infty} \binom{j}{i} \frac{1}{j!} \frac{\partial^j \varphi(t, s, \theta)}{\partial t^j}\bigg|_{t=0}\right] i g(x) G(x)^{i-1},$$

and then

$$f(x) = \sum_{i=1}^{\infty} w_i g_i(x).$$

$\square$

Proof of Theorem 2. For $0 < \theta < 1$, we consider the negative binomial series,

$$(1 - y)^{-r} = \sum_{i=0}^{\infty} \binom{i + r - 1}{i} y^i,$$

which holds for $|y| < 1$ and any real number $r > 0$. Using this expansion in the denominator of (1.4), we have

$$\begin{aligned}
f(x) &= \theta^{1/k} g(x) \sum_{j=0}^{\infty} \binom{j + k^{-1}}{j} \bar{\theta}^j \bar{G}(x)^{kj} \\
&= \theta^{1/k} g(x) \sum_{j=0}^{\infty} \binom{j + k^{-1}}{j} \bar{\theta}^j \left[\sum_{i=0}^{\infty} (-1)^i \binom{kj}{i} G(x)^i\right] \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{\theta^{1/k} \bar{\theta}^j (-1)^j}{i + 1} \binom{j + k^{-1}}{j} \binom{kj}{i} (i + 1) g(x) G(x)^i \\
&= \sum_{i=0}^{\infty} w_i g_{i+1}(x),
\end{aligned}$$

where

$$w_i = \frac{\theta^{1/k}(-1)^i}{i+1}\left[\sum_{j=0}^{\infty}\bar{\theta}^j\binom{j+k^{-1}}{j}\binom{kj}{i}\right].$$

For $\theta > 1$, we replace $\tau = \theta^{-1}$. Then we can rewrite (1.4) as

$$\begin{aligned} f(x) &= \frac{\tau g(x)}{[\tau - (\tau-1)\bar{G}(x)^k]^{1+1/k}} = \frac{\tau g(x)}{1-(1-\tau)[1-\bar{G}(x)^k]^{1+1/k}} \\ &= \frac{\tau g(x)}{\{1-\bar{\tau}[1-\bar{G}(x)^k]\}^{1+1/k}}, \end{aligned}$$

where $\bar{\tau} = 1-\tau$. Now, we can use the negative binomial expansion in the denominator of the last expression to obtain

$$\begin{aligned} f(x) &= \tau g(x)\sum_{j=0}^{\infty}\binom{j+k^{-1}}{j}\bar{\tau}^j\sum_{l=0}^{j}(-1)^l\binom{j}{l}\bar{G}(x)^{lk} \\ &= \tau g(x)\sum_{j=0}^{\infty}\binom{j+k^{-1}}{j}\bar{\tau}^j\sum_{l=0}^{j}(-1)^l\binom{j}{l}\sum_{i=0}^{\infty}(-1)^i\binom{lk}{i}G(x)^i \\ &= \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{j}\frac{\tau\bar{\tau}(-1)^{l+i}}{i+1}\binom{j+k^{-1}}{j}\binom{j}{l}\binom{lk}{i}(i+1)g(x)G(x)^i \\ &= \sum_{i=0}^{\infty}w_i g_{i+1}(x), \end{aligned}$$

where

$$w_i = \frac{(-1)^i\tau}{i+1}\left[\sum_{j=0}^{\infty}\sum_{l=0}^{j}(-1)^l\bar{\tau}^j\binom{j+k^{-1}}{j}\binom{j}{l}\binom{lk}{i}\right].$$

□

Proof of Theorem 3. The first derivative of $f(x)$ is given by

$$f'(x) = \frac{[1-\bar{\theta}\bar{G}(x)^k]^{1/k}\{g'(x)[1-\bar{\theta}\bar{G}(x)^k] - \bar{\theta}(1+k)g(x)^2\bar{G}(x)^{k-1}\}}{\{1-\bar{\theta}\bar{G}(x)^k\}^{2(1+1/k)}}.$$

Since $[1-\bar{\theta}\bar{G}(x)^k]^{1/k}$ and the denominator are always positive, we have

$$\operatorname{sgn}[f'(x)] = \operatorname{sgn}\{g'(x)[1-\bar{\theta}\bar{G}(x)^k] - \bar{\theta}(1+k)g(x)^2\bar{G}(x)^{k-1}\},$$

where $\operatorname{sgn}(x)$ is the signal function ($\operatorname{sgn}(x) = 1$ if x is positive, $\operatorname{sgn}(x) = 0$, if $x = 0$ and $\operatorname{sgn}(x) = -1$ otherwise). Then a simple analysis of this last expression concludes the proof. □

Proof of Lemma 1.

$$\begin{aligned}\mathcal{H}_{\mathrm{Sh}}^{i+1}[g_{i+1}(y)] &= \mathbb{E}_{Y_i}\{-\log[g_{i+1}(Y_i)]\}\\ &= -\log(i+1) - \mathbb{E}_{Y_i}\{\log[g(Y_i)]\} - i\mathbb{E}_{Y_i}\{\log[G(Y_i)]\},\end{aligned}$$

and then

$$\mathbb{E}_{Y_i}\{\log[g(Y_i)]\} = -\mathcal{H}_{\mathrm{Sh}}^{i+1}(g_{i+1}) - \log(i+1) - i\mathbb{E}_{Y_i}\{\log[G(Y_i)]\}. \tag{A.1}$$

But

$$\begin{aligned}\mathbb{E}_{Y_i}\{\log[G(Y_i)]\} &= \int_{-\infty}^{+\infty}\{\log[G(y)]\}g_{i+1}(y)\,dy\\ &= (i+1)\int_{-\infty}^{+\infty}\{\log[G(y)]\}g(y)G(y)^i\,dy\\ &= (i+1)\int_{-\infty}^{+\infty}\frac{\partial}{\partial a}g(y)G(y)^{i+a}\Big|_{a=0}\,dy\\ &= (i+1)\frac{\partial}{\partial a}\left[\int_{-\infty}^{+\infty}g(y)G(y)^{i+a}\,dy\right]\Big|_{a=0}\\ &= -\frac{1}{i+1}.\end{aligned}$$

Substituting $\mathbb{E}_{Y_i}\{\log[G(Y_i)]\}$ in (A.1), the result follows. □

Proof of Theorem 4. From the definition of $\mathcal{H}_{\mathrm{Sh}}[f(x)]$

$$\begin{aligned}\mathcal{H}_{\mathrm{Sh}}[f(x)] &= \mathbb{E}_X\{-\log[f(X)]\}\\ &= -\frac{1}{k}\log(\theta) - \mathbb{E}_X\{\log[g(X)]\}\\ &\quad + \left(1+\frac{1}{k}\right)\mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\}.\end{aligned} \tag{A.2}$$

We can obtain $\mathbb{E}_X\{\log[g(X)]\}$ from Theorem 2 and Lemma 1

$$\begin{aligned}\mathbb{E}_X\{\log[g(X)]\} &= \int_{-\infty}^{+\infty}\log[g(x)]f(x)\,dx\\ &= \sum_{i=0}^{\infty}w_i\int_{-\infty}^{+\infty}\log[g(x)]g_{i+1}(x)\,dx\\ &= \sum_{i=0}^{\infty}w_i\mathbb{E}_{Y_i}\{\log[g(X)]\}\\ &= \sum_{i=0}^{\infty}w_i\left\{\mathcal{H}_{\mathrm{Sh}}^{i+1}[g_{i+1}(y)] - \log(i+1) + \frac{1}{i+1}\right\}.\end{aligned} \tag{A.3}$$

To obtain $\mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\}$, let $t = G(x)$. So,

$$\begin{aligned}\mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\} &= \frac{\theta^{1/k}\int_{-\infty}^{+\infty}\log[1-\bar{\theta}\bar{G}(x)^k]g(x)}{[1-\bar{\theta}\bar{G}(x)^k]^{1+1/k}}\,dx \\ &= \theta^{1/k}\int_0^1 \frac{\log(1-\bar{\theta}t^k)}{(1-\bar{\theta}t^k)^{1+1/k}}\,dt \qquad (A.4) \\ &= I(k,\theta).\end{aligned}$$

Inserting (A.3) and (A.4) in equation (A.2) concludes the proof. □

Proof of Theorem 5. Consider the Kullback–Leibler divergence between the density functions $f(x)$ and $h(x)$ given by

$$D(h,f) = \int_{-\infty}^{+\infty} h(x)\log\left[\frac{h(x)}{f(x)}\right]dx.$$

The Gibbs' inequality implies $D(h,f) \geq 0$ where the equality is attained iff $h(x)$ and $f(x)$ are equal almost everywhere. Then

$$\begin{aligned} 0 &\leq \int_{-\infty}^{+\infty} h(x)\log\left[\frac{h(x)}{f(x)}\right]dx, \\ 0 &\leq \int_{-\infty}^{+\infty} h(x)\log[h(x)]\,dx - \int_{-\infty}^{+\infty} h(x)\log[f(x)]\,dx, \\ \mathcal{H}_{\mathrm{Sh}}[h(x)] &\leq -\int_{-\infty}^{+\infty} h(x)\log[f(x)]\,dx, \\ \mathcal{H}_{\mathrm{Sh}}[h(x)] &\leq -\frac{1}{k}\log(\theta) - \mathbb{E}_X\{\log[g(X)]\} + \mathbb{E}_X[1-\bar{\theta}\bar{G}(X)^k].\end{aligned}$$

Under the imposed constraints and the definition of W given before, we have

$$\begin{aligned}\mathcal{H}_{\mathrm{Sh}}[h(x)] &\leq -\frac{1}{k}\log(\theta) - \mathbb{E}_W\{\log\{g[G^{-1}(W)]\}\} + I(k,\theta), \\ \mathcal{H}_{\mathrm{Sh}}[h(x)] &\leq -\frac{1}{k}\log(\theta) - \int_{-\infty}^{+\infty}\log\{g[G^{-1}(w)]\}\frac{\theta^{1/k}}{(1-\bar{\theta}w^k)^{1+1/k}}\,dw + I(k,\theta).\end{aligned}$$

Setting $w = G(x)$ and $I(k,\theta) = \mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\}$ gives

$$\begin{aligned}\mathcal{H}_{\mathrm{Sh}}[h(x)] \leq{}& -\frac{1}{k}\log(\theta) - \int_{-\infty}^{+\infty}\log[g(x)]\frac{\theta^{1/k}g(x)}{[1-\bar{\theta}\bar{G}(x)^k]^{1+1/k}}\,dx \\ &+ \mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\}, \\ \mathcal{H}_{\mathrm{Sh}}[h(x)] \leq{}& -\frac{1}{k}\log(\theta) - \mathbb{E}_X\{\log[g(X)]\} + \mathbb{E}_X\{\log[1-\bar{\theta}\bar{G}(X)^k]\}, \\ \mathcal{H}_{\mathrm{Sh}}[h(x)] \leq{}& \mathcal{H}_{\mathrm{Sh}}[f(x)],\end{aligned}$$

the equality holding iff $f(x)$ equals $h(x)$ almost everywhere. □

Acknowledgments

The Foulum SAR image data were downloaded from the European Space Agency (ESA) in http://earth.eo.esa.int/polsarpro/datasets.html. This dataset was made available by Professor Henning Skiver and Dr. Jorgen Dall. The first author acknowledges receipt of a grant from CAPES. The second author is grateful to CNPq.

References

Aly, E. and Benkherouf, L. (2011). A new family of distributions based on probability generating functions. *Sankhya B—Applied and Interdisciplinary Statistics* **73**, 70–80. MR2826321

Barreto-Souza, W. and Bakouch, H. S. (2013). A new lifetime model with decreasing failure rate. *Statistics* **47**, 465–476. MR3043713

Chahkandi, M. and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis* **53**, 4433–4330. MR2744336

Chen, G. and Balakrishnan, N. (1995). A general purpose approximate goodness-of-fit test. *Journal of Quality Technology* **27**, 154–161.

Cintra, R. J., Rêgo, L. C., Cordeiro, G. M. and Nascimento, A. D. C. (2014). Beta generalized normal distribution with an application for SAR image processing. *Statistics* **48**, 279–294. MR3175770

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton: Princeton Univ. Press. MR0016588

Frery, A. C., Müller, H. J., Yanasse, C. C. F. and Sant'Anna, S. J. S. (1997). A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing* **35**, 648–659.

Gradshteyn, I. S. and Ryzhik, I. M. (2000). *Table of Integrals, Series, and Products*. San Diego: Academic Press. MR1773820

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* **50**, 1029–1054. MR0666123

Harris, T. E. (1948). Branching processes. *The Annals of Mathematical Statistics* **19**, 474–494. MR0027465

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review* **106**, 620–630. MR0087305

Kuş, C. (2006). A new lifetime distribution. *Computational Statistics & Data Analysis* **51**, 4497–4509. MR2364461

Lu, W. and Shi, D. (2012). A new compounding life distribution: The Weibull–Poisson distribution. *Journal of Applied Statistics* **39**, 21–38. MR2872325

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. MR1603936

Morais, A. L. and Barreto-Souza, W. (2011). A compound class of Weibull and power series distributions. *Computational Statistics & Data Analysis* **55**, 1410–1425. MR2741424

Pinho, L. G. B., Cordeiro, G. M. and Nobre, J. S. (2015). The Harris extended exponential distribution. *Communications in Statistics: Theory and Methods* **44**, 3486–3502. DOI:10.1080/03610926.2013.851221

Ristić, M. and Nadarajah, S. (2012). A new lifetime distribution. *Journal of Statistical Computation and Simulation* **82**, 1191–1206.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–432. MR0026286

L. G. B. Pinho
G. M. Cordeiro
Federal University of Pernambuco
UFPE
Brazil

J. S. Nobre
Federal University of Ceará
UFC
Brazil