The GLE Distributions Family

Aleixo, S. M., Diamantino, M. F. and Pestana, D. D.

Abstract Structural properties of a Gaussian-Laplace Extended family of symmetrical distributions, which contains as special cases the Gaussian and the Laplace distributions, are investigated, and this family is further extended by folding. The main purpose is to compare pseudo-random numbers generators for members of this family, useful in investigating problems that may be influenced by kurtosis, and namely the investigation of robustness issues.

1 Introduction

In their investigation on the implications of Rényi's rarefaction, [5], in new ways of dealing with non-response in sample surveys, [1] used the family of positive random variables $W_{\beta,\lambda,\delta} = \lambda + \delta W_\beta$, where the standard $W_\beta = W_{\beta,0,1}$ has probability density functions

$$f_{W_\beta}(x) = \frac{\exp\left(-x^\beta\right)}{\Gamma\left(1+\frac{1}{\beta}\right)} \quad \mathrm{I}_{(0,\infty)}(x),\ \beta > 0,$$

and the family of symmetric random variables $X_{\beta,\lambda,\delta}$, which includes the Gaussian and the Laplace distributions.

Aleixo, S. M
Instituto Politécnico de Lisboa, ISEL, and CEAUL, Centro de Estatística e Aplicações da Universidade de Lisboa, e-mail: sandra.aleixo@dec.isel.pt

Diamantino, M. F.
Faculdade de Ciências, Universidade de Lisboa, and CEAUL, Centro de Estatística e Aplicações da Universidade de Lisboa, e-mail: mfdiamantino@fc.ul.pt

Pestana, D. D.
Faculdade de Ciências, Universidade de Lisboa, and CEAUL, Centro de Estatística e Aplicações da Universidade de Lisboa, e-mail: dinis.pestana@fc.ul.pt

We define standard symmetric random variables $X_{\beta,0,1}$ using the usual symmetrization procedure $X_\beta = BW_\beta$, where $B = \begin{cases} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{cases}$, with B and W_β independent. For convenience, we use a reparametrization $X_{\beta^*} = X_{\frac{2-\beta}{\beta}}$, thus using as shape parameter $\beta^* = \frac{2-\beta}{\beta} > -1$ for the symmetric random variables, and use location and shape parameters $\lambda \in \mathbb{R}$ and $\delta > 0$, respectively:

$$f_{X_{\beta^*,\lambda,\delta}}(x) = \frac{1}{2^{\frac{\beta^*+3}{2}} \Gamma\left(\frac{\beta^*+3}{2}\right) \delta} \exp\left\{ -\frac{1}{2} \left| \frac{x-\lambda}{\delta} \right|^{\frac{2}{1+\beta^*}} \right\} \mathrm{I}_{\mathbb{R}}(x)$$

$\beta^* > -1, \lambda \in \mathbb{R}, \delta > 0.$

$\{X_{\beta^*,\lambda,\delta}\}_{\beta^*>-1}$ is the symmetric Gauss-Laplace Extended (GLE) family, and $\{W_{\beta,\lambda,\delta}\}_{\beta>0}$ is the folded GLE family.

2 The folded GLE family

Let $W_{\beta,\lambda,\delta} = \lambda + \delta W_\beta, \ \ \beta > 0$, with

$$f_{W_\beta}(x) = \frac{\exp\left(-x^\beta\right)}{\Gamma\left(1+\frac{1}{\beta}\right)} \quad \mathrm{I}_{(0,\infty)}(x).$$

(W_1 is an Exponencial(1) random variable (or folded Laplace) and $\sqrt{2}W_2$ is the folded Gaussian random variable.)

The mean of $W_{\beta,\lambda,\delta}$ is

$$\mu = \lambda + \delta \frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} = \lambda + \frac{2^{\frac{2}{\beta}-1}\delta}{\sqrt{\pi}} \Gamma\left(\frac{2+\beta}{2\beta}\right),$$

and the variance of $W_{\beta,\lambda,\delta}$ is

$$\sigma^2 = \left[\frac{\Gamma\left(\frac{3}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} - \left(\frac{\Gamma\left(\frac{2}{\beta}\right)}{\Gamma\left(\frac{1}{\beta}\right)} \right)^2 \right] \delta^2.$$

In Fig. 1 the influence of the shape parameter β on the tailweight of W_β is evident; Fig. 2 shows that the kurtosis of W_β is indeed very large for very small values of $beta$, decreasing steadily for values of β smaller than 1.5, and becoming negative for $\beta > 1.5$.

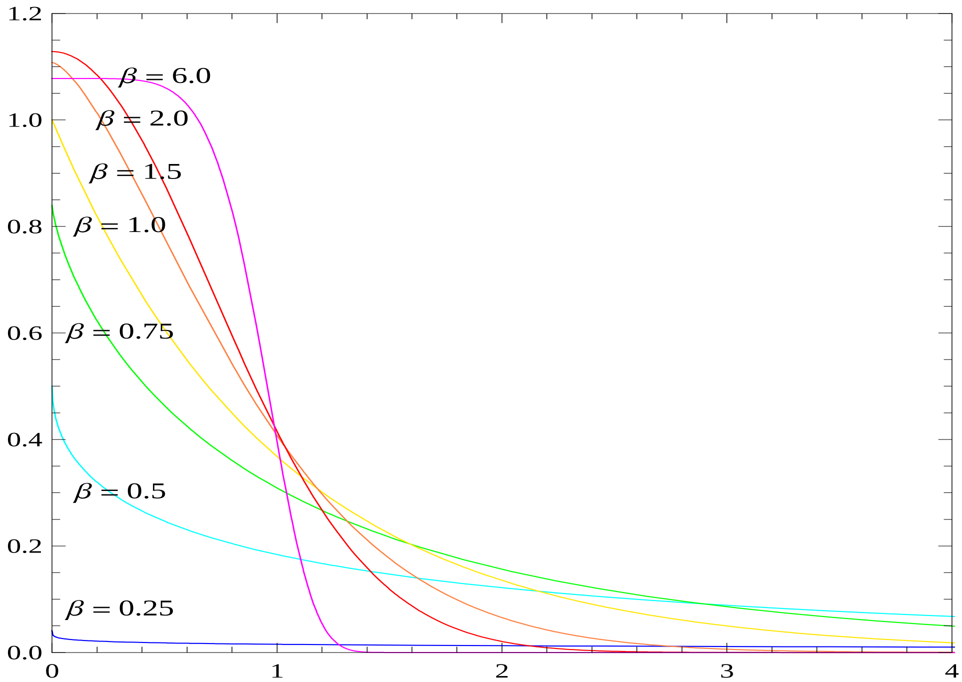

Fig. 1 Probability density functions of the random variables W_β.

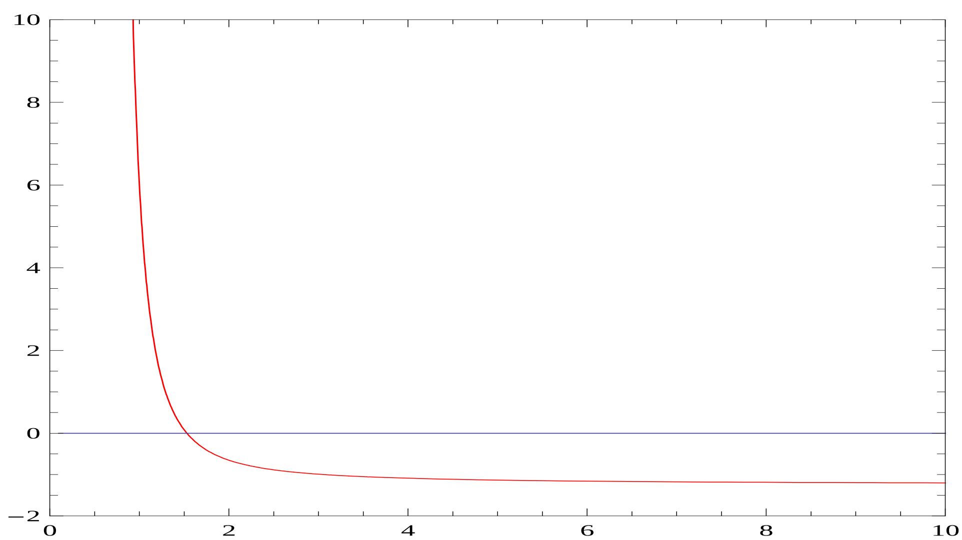

Fig. 2 Kurtosis of W_β.

For the generation of pseudo-random numbers from W_β, we may observe that if $Y_\alpha \frown Gama(\alpha, 1)$, then the probability density function of $V_\alpha = Y_\alpha^\alpha$ is

$$f_{V_\alpha(x)} = \frac{\mathrm{e}^{-x^{\frac{1}{\alpha}}}}{\Gamma(\alpha+1)}\ \mathrm{I}_{(0,\infty)}(x),$$

and therefore $W_\beta = V_{\frac{1}{\beta}} = Y_{\frac{1}{\beta}}^{\frac{1}{\beta}}$.

Hence, the well-developed theory of gamma pseudo-random numbers generation can be used to generate pseudo-random numbers of the folded GLE populations.

3 The symmetric GLE family

As, for $\beta^* > -1$, we have

$$\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}|x|^{\frac{2}{1+\beta^*}}\right\} dx = 2^{\frac{\beta^*+1}{2}}(\beta^*+1)\,\Gamma\left(\frac{\beta^*+1}{2}\right) = 2^{\frac{\beta^*+3}{2}}\Gamma\left(\frac{\beta^*+3}{2}\right),$$

the function

$$f(x \mid \beta^*, \lambda, \delta) = \frac{1}{2^{\frac{\beta^*+3}{2}}\Gamma\left(\frac{\beta^*+3}{2}\right)\delta}\ \exp\left\{-\frac{1}{2}\left|\frac{x-\lambda}{\delta}\right|^{\frac{2}{\beta^*+1}}\right\}\ \mathrm{I}_{\mathbb{R}}(x)$$

is the probability density function of a random variable $X_{\beta^*,\lambda,\delta}$ for any $\beta^* > -1$, $\lambda \in \mathbb{R}$ and $\delta > 0$.

Hence we have a parametrized family which contains in particular the Gaussian ($\beta^* = 0$) and Laplace ($\beta^* = 1$) random variables. In Fig. 3 we see that for $\beta^* > 0$ the density of X_{β^*} is peaked, while for $\beta^* < 0$ the density of X_{β^*} has a "plateau" in a neighborhood of 0, that approaches $[-1, 1]$ when $\beta^* \downarrow -1$.

As

$$E\left(X_{\beta^*}^{2k}\right) = 2^{k(\beta^*+1)}\,\frac{\Gamma\left(k(\beta^*+1)+\frac{\beta^*+1}{2}\right)}{\Gamma\left(\frac{\beta^*+1}{2}\right)},$$

the mean μ of $X_{\beta^*,\lambda,\delta}$ is $\mu = \lambda$, the variance σ^2 of $X_{\beta^*,\lambda,\delta}$ is

$$\sigma^2 = \frac{2^{\beta^*+1}\Gamma\left(3\,\frac{\beta^*+1}{2}\right)}{\Gamma\left(\frac{\beta^*+1}{2}\right)}\ \delta^2,$$

and the kurtosis γ_2 of $X_{\beta^*,\lambda,\delta}$ is

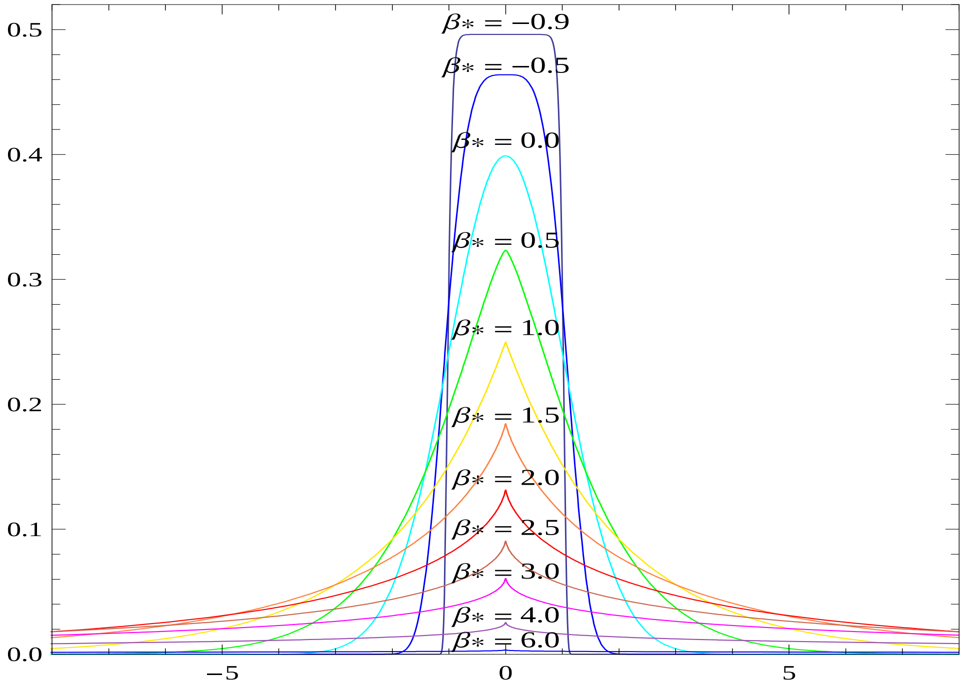

Fig. 3 Probability density functions of the random variables X_{β^*}.

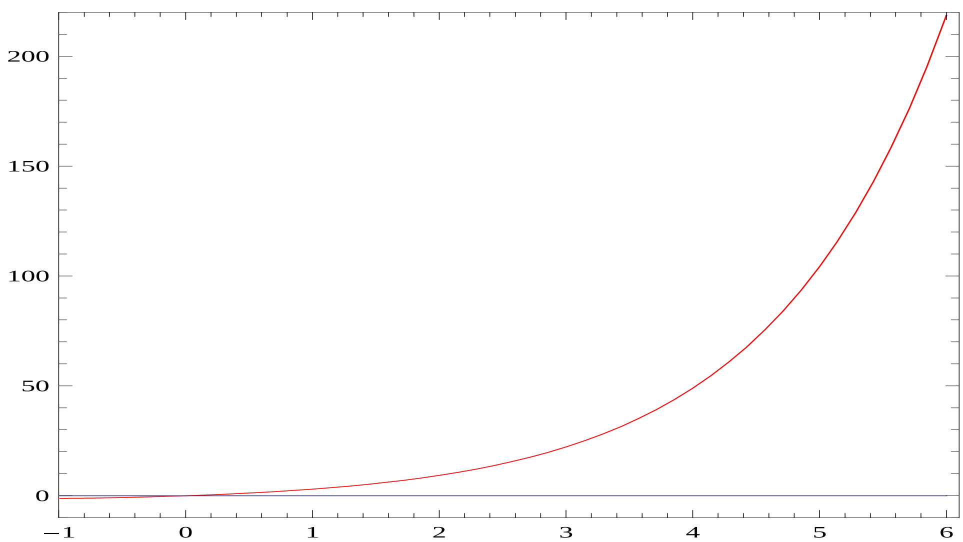

Fig. 4 Kurtosis of X_{β^*}.

$$\gamma_2 = \frac{\Gamma\left(\frac{\beta^*+1}{2}\right)\Gamma\left(5\,\frac{\beta^*+1}{2}\right)}{\left[\Gamma\left(3\,\frac{\beta^*+1}{2}\right)\right]^2} - 3$$

which increases to ∞ with β^*, see Fig. 4.

4 Generation methods

In order to generate directly sets of pseudo-random numbers from populations with distribution belonging to the GLE family, we compared the performance of several methods suggested in [2], [3], [4] and [6]. As the symmetrical family can easily be obtained from the folded GLE family, we state our findings only in terms of the folded family.

The parameter β has indeed an important bearing in this issue. The conclusions of our analysis may be summarized as follows:

- For $0 < \beta < 1$, rejection method using the Pareto(α) distribution, with optimal α.
- For $\beta = 1$, inversion method to generate unit exponentials.
- For $1 < \beta < 6$, rejection method using the Exponencial(1) distribution.
- For $\beta \geq 6$, rejection method using the Weibull$(0, \alpha, \gamma)$ distribution.

We generated sets of $\nu = 5000$ pseudo-random numbers x_k of X_β, for $\beta = -0.75(0.25)1.5(0.5)3$, and of W_β, for $\beta = 0.25(0.25)1.5(0.5)6$; in each case we calculated their mean $\mu^* = \frac{1}{\nu}\sum_{k=1}^{\nu} x_k$ and their variance $\sigma^{*2} = \frac{1}{\nu-1}\sum_{k=1}^{\nu}(x_k-\mu^*)^2$. These sets of pseudo-random numbers are available at www.ceaul.fc.ul.pt, for anyone wishing to use them.

5 Validity of the generated populations

The quality of the pseudo-random numbers generated as described has been assessed as follows:

- Comparing the empirical and populational moments of small order, μ_{W_β} with $\mu^*_{W_\beta}$ and $\sigma^2_{W_\beta}$ with $\sigma^{*2}_{W_\beta}$.
- Comparing a maximum likelihood estimate $\widehat{\beta}$ with the true parameter value β.
- Making several replicas of the Kolmogorov-Smirnov adjustment test.

In all cases we obtained satisfactory results, as reported in Table 1. There are no big differences between the empirical and populational values, either in terms of the mean values or in terms of standard deviations, and the maximum likelihood estimates $\widehat{\beta}$ are always close to the true β.

We have no explicit analytical expression for the maximum likelihood estimator for the parameter β of the GLE distribution with positive support, but using the software Wolfram Mathematica7 numerical evaluation of $\frac{d \ln L(\beta)}{d\beta} = 0$ is readily achieved, where the likelihood function is

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{e^{-x_i^{\beta}}}{\Gamma\left(1+\frac{1}{\beta}\right)} \right] = \frac{e^{-\sum_{i=1}^{n} x_i^{\beta}}}{\Gamma^n\left(1+\frac{1}{\beta}\right)},$$

and hence the log-likelihood is $\ln L(\beta) = -\sum_{i=1}^{n} x_i^{\beta} - n ln\left(\Gamma\left(1+\frac{1}{\beta}\right)\right)$.

Table 1 Validity of generated populations with GLE distribution with shape parameter β

β	$\hat{\beta}$	μ	σ	μ^*	σ^*	$K-S$
−0.75	−0.7319	0.0000	0.6117	−0.0019	0.6057	0.1278
−0.50	−0.4919	0.0000	0.6914	0.0017	0.6882	0.1255
−0.25	−0.2550	0.0000	0.8174	0.0159	0.8187	0.1045
0.00	0.0242	0.0000	1.0000	0.0002	1.0200	0.1116
0.25	0.2506	840.0000	2438.6900	811.7070	2375.5700	0.1144
0.50	0.4959	6.0000	9.1652	6.2544	9.8940	0.1034
0.75	0.7505	1.6849	1.9698	1.6840	1.9532	0.1111
1.00	1.0111	1.0000	1.0000	0.9924	0.9806	0.1084
1.25	1.2717	0.7675	0.6913	0.7572	0.6725	0.1161
1.50	1.5039	0.6595	.5510	0.6599	0.5491	0.1136
2.00	2.0089	0.5642	0.4263	0.5517	0.4249	0.1148
2.50	2.5095	0.5249	0.3721	0.5242	0.3721	0.1236
3.00	2.9590	0.5055	0.3432	0.5104	0.3485	0.1236
3.50	3.5080	0.4949	0.3259	0.4897	0.3241	0.0952
4.00	3.9659	0.4889	0.3146	0.4862	0.3143	0.1104
4.50	4.1715	0.4853	0.3070	0.4766	0.3133	0.1294
5.00	4.9462	0.4832	0.3015	0.4887	0.3031	0.1097
5.50	5.3100	0.4819	0.2976	0.4857	0.2927	0.1002
6.00	6.0160	0.4813	0.2946	0.4728	0.2927	0.1276

In Table 1 we reported only the worst p-value obtained when using the Kolmogorov–Smirnov test to investigate the goodness-of-fit of W_β to samples of size $n = 100$ randomly extracted from the sets of 5 000 pseudo-random numbers we had generated, for each β. Observe that at the significance level $\alpha = 5\%$ the rejection critical point is $\frac{1.36}{\sqrt{100}} = 0.136$. As all worst observed values of the Kolmogorov-Smirnov statistical test are smaller than 0.136, we maintain the null hypothesis that in all cases W_β fits well the generated random numbers.

The calculus of the empirical and populational moments, as well as the Kolmogorov-Smirnov adjustment test were also computationally implemented in the Wolfram Mathematica7 program.

Acknowledgements

This research has been partially supported by FCT/OE, CEAUL and ISEL.

References

1. Aleixo, S., Brilhante, F., Diamantino, F. Mendonça, Pestana, D.: Non–Response and Sample Size. Bulletin of the International Statistical Iinstitute , **LXII**, 4804–4807 (2007)
2. Devroye, L.: Non–Uniform Random Variate Generation. Springer–Verlag, New York (1986)
3. Kundu, D., Gupta, R. D.: A convenient way of generating gamma random variables using generalized exponential distribution. Science Direct–Comp. Stat. Data Analysis. **51**, 2796–2802 (2007)
4. Law, A. M., Kelton, W.D.: Simulation, Modeling and Analysis. Academic Press, New York (1982)
5. Rényi, A.: A characterization of the Poisson process. MTA Mat. Kut. Int. Közl. **1**, 519–527 (Reprinted in Selected Papers of Alfréd Rényi, P. Turán, ed, Akadémiai Kiadó, Budapest, 1976, vol. **I**, 622–628) (1956)
6. Ross, S. M.: Simulation, Statistical Modeling and Decision Science. 2nd edn,. McGraw-Hill Book Company, San Diego (1997)