187 / 2-7-78 DR 1822

LA-7095-MS
Informal Report

UC-32
Issued: January 1978

MASTER

A New Probability Distribution with Applications in Monte Carlo Studies

Mark E. Johnson
Myrle M. Johnson

los alamos scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87545

UNITED STATES
DEPARTMENT OF ENERGY
CONTRACT W-7405-ENG. 36

A NEW PROBABILITY DISTRIBUTION WITH APPLICATIONS IN MONTE CARLO STUDIES

by

Mark E. Johnson and Myrle M. Johnson

ABSTRACT

A new symmetric univariate probability distribution is proposed. Several properties are derived, and the distribution's applicability to Monte Carlo studies is discussed.

I. INTRODUCTION

A frequent goal in Monte Carlo studies is to investigate the performance of a statistical procedure against departures from normality. The departures generally studied involve a collection of univariate nonnormal distributions. The resulting analysis is consequently qualitative in the sense of clustering the distributions that yield similar performances. In this report we propose a symmetric univariate probability distribution, which lends itself to a more quantitative analysis of the departure from normality. This distribution includes the uniform and normal distributions as special cases, and moreover, brackets the normal distribution in the sense that its kurtosis can range from 1.8 to 5.4. The kurtosis of the normal distribution equals 3.0. The proposed distribution and some of its properties are derived in Sec. II. Applications to Monte Carlo studies are discussed in Sec. III.

II. THE PROPOSED DISTRIBUTION

The density of the proposed distribution is

$$f(x) = \frac{\sqrt{\alpha/3}\ \Gamma(\alpha-\frac{1}{2})}{2\Gamma(\alpha)} \left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right], \qquad -\infty < x < \infty,\ \alpha > \tfrac{1}{2}, \tag{2.1}$$

where G is the distribution function of a gamma random variable with shape parameter $\alpha-\frac{1}{2}$ and scale parameter 2. This function is indeed a density since it is obviously nonnegative, and we prove in the Appendix that f integrates to 1.

Using techniques similar to those used in the Appendix, the distribution function of the proposed distribution can be derived to be

$$F(x) = \begin{cases} \dfrac{\Gamma(\alpha-\frac{1}{2})\sqrt{\alpha}\,x}{2\sqrt{3}\;\Gamma(\alpha)}\left[1 - G_1\left(\dfrac{2\alpha x^2}{3}\right)\right] + \dfrac{1}{2}\left[1 - G_2\left(\dfrac{2\alpha x^2}{3}\right)\right] & x < 0 \\ \dfrac{1}{2} + \dfrac{\Gamma(\alpha-\frac{1}{2})\sqrt{\alpha}\,x}{2\sqrt{3}\;\Gamma(\alpha)}\left[1 - G_1\left(\dfrac{2\alpha x^2}{3}\right)\right] + \dfrac{1}{2}G_2\left(\dfrac{2\alpha x^2}{3}\right) & x > 0, \end{cases}$$

where G_1 and G_2 are distribution functions corresponding to $\Gamma(\alpha-\frac{1}{2},2)$ and $\Gamma(\alpha,2)$ random variables, respectively.

For $\alpha = 3/2$, f simplifies, as follows:

$$f(x) = \frac{\sqrt{1/2}\,\Gamma(1)}{2\Gamma(3/2)}\left[1 - \left(1 - e^{-x^2/2}\right)\right]$$

$$= \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2},$$

which is the standard normal density.

For all values of α, the distribution is symmetric and has the real line as its support. For the density form given in Eq. (2.1), the mean is 0, the variance is 1, and the kurtosis (denoted β_2) is

$$\beta_2 = \frac{9(\alpha+1)}{5\alpha}, \tag{2.2}$$

which is derived in the Appendix. Note that for $\alpha = 1.5$, $\beta_2 = 3$ as expected. When α equals 0.5, β_2 equals 5.4; as α tends to infinity, β_2 approaches 1.8, which is the kurtosis of a uniform distribution. Moreover, the limiting form of f as α tends to infinity is the uniform density on the interval $(-\sqrt{3}, \sqrt{3})$.

Any value for β_2 in the interval (1.8, 5.4) is attainable, since β_2 is a monotone, continuous function of α. In particular, for specified β_2, the corresponding α value is $1.8/(\beta_2 - 1.8)$. Some density plots for several values of α are given in Fig. 1. Percentiles of the distribution for selected α values and cumulative probabilities p appear in Table I. These computations used the routines developed by Amos and Daniel [1972] for computing gamma distribution function values.

III. APPLICATIONS TO MONTE CARLO STUDIES

In this section, an algorithm for generating variates having the proposed distribution is derived. We then describe how this distribution can be employed in Monte Carlo studies.

Consider the following algorithm for generating variates having the density in Eq. (2.1).

1. Generate X having a gamma distribution with shape parameter α and scale parameter 2.
2. Generate Y having a uniform distribution on the interval $(-\sqrt{X}, \sqrt{X})$.
3. The desired variate is $(3/2\alpha)^{\frac{1}{2}}Y$.

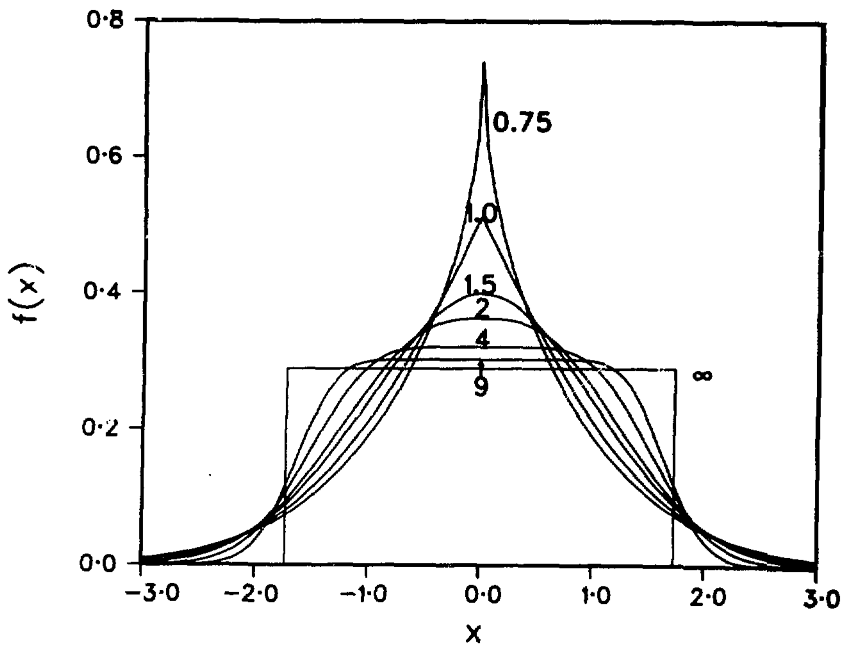

Fig. 1. Densities for Selected α Values.

TABLE I

PERCENTILES OF THE PROPOSED DISTRIBUTION

α\p	0.75	0.90	0.95	0.975	0.99
0.55	0.4744	1.1804	1.6868	2.1604	2.7392
0.60	0.4958	1.1939	1.6859	2.1428	2.6987
0.65	0.5148	1.2052	1.6843	2.1264	2.6621
0.70	0.5318	1.2149	1.6823	2.1112	2.6288
0.75	0.5471	1.2233	1.6800	2.0969	2.5984
0.80	0.5610	1.2306	1.6776	2.0836	2.5704
0.85	0.5737	1.2370	1.6751	2.0712	2.5447
0.90	0.5853	1.2428	1.6726	2.0595	2.5208
0.95	0.5959	1.2479	1.6700	2.0485	2.4985
1.00	0.6057	1.2524	1.6674	2.0382	2.4778
1.10	0.6233	1.2604	1.6624	2.0191	2.4402
1.20	0.6386	1.2670	1.6577	2.0021	2.4069
1.30	0.6520	1.2726	1.6531	1.9867	2.3773
1.40	0.6639	1.2774	1.6489	1.9727	2.3505
1.50	0.6745	1.2816	1.6449	1.9600	2.3263
1.60	0.6840	1.2852	1.6411	1.9483	2.3043
1.70	0.6927	1.2885	1.6375	1.9375	2.2841
1.80	0.7005	1.2915	1.6342	1.9275	2.2655
1.90	0.7077	1.2941	1.6311	1.9183	2.2484
2.00	0.7143	1.2965	1.6281	1.9097	2.2324
2.20	0.7260	1.3008	1.6227	1.8941	2.2038
2.40	0.7361	1.3044	1.6179	1.8804	2.1786
2.60	0.7449	1.3076	1.6136	1.8683	2.1564
2.80	0.7526	1.3104	1.6097	1.8574	2.1365
3.00	0.7594	1.3129	1.6061	1.8475	2.1186
3.50	0.7734	1.3182	1.5986	1.8267	2.0807
4.00	0.7843	1.3224	1.5926	1.8099	2.0501
4.50	0.7930	1.3260	1.5876	1.7960	2.0247
5.00	0.8001	1.3291	1.5835	1.7842	2.0033
6.00	0.8110	1.3342	1.5770	1.7655	1.9690
7.00	0.8188	1.3383	1.5722	1.7511	1.9425
8.00	0.8248	1.3417	1.5685	1.7397	1.9213
9.00	0.8294	1.3446	1.5656	1.7304	1.9038
10.00	0.8331	1.3472	1.5633	1.7227	1.8892
11.00	0.8361	1.3494	1.5614	1.7161	1.8767
12.00	0.8386	1.3514	1.5599	1.7105	1.8659
13.00	0.8408	1.3532	1.5586	1.7056	1.8564
14.00	0.8426	1.3548	1.5575	1.7013	1.8480
15.00	0.8442	1.3563	1.5566	1.6976	1.8406
20.00	0.8497	1.3620	1.5537	1.6838	1.8127
25.00	0.8530	1.3660	1.5523	1.6750	1.7942
30.00	0.8551	1.3689	1.5517	1.6690	1.7810
35.00	0.8567	1.3711	1.5514	1.6646	1.7710
40.00	0.8579	1.3728	1.5514	1.6613	1.7632
45.00	0.8588	1.3742	1.5515	1.6587	1.7569
50.00	0.8595	1.3753	1.5516	1.6566	1.7516
60.00	0.8606	1.3770	1.5520	1.6536	1.7434
70.00	0.8614	1.3782	1.5524	1.6514	1.7373
80.00	0.8620	1.3791	1.5529	1.6499	1.7325
90.00	0.8624	1.3799	1.5533	1.6487	1.7287
100.00	0.8628	1.3804	1.5537	1.6478	1.7256
∞	0.8660	1.3856	1.5588	1.6454	1.6974

This algorithm is proved directly. The joint distribution of X and Y is

$$f(x,y) = \frac{1}{2\sqrt{x}} \frac{x^{\alpha-1} e^{-x/2}}{\Gamma(\alpha)\ 2^{\alpha}}, \qquad x > y^2.$$

The marginal distribution of Y is

$$f(y) = \int_{y^2}^{\infty} f(x,y)\ dx$$

$$= \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{2}\ \Gamma(\alpha)} \int_{y^2}^{\infty} \frac{x^{\alpha-\frac{1}{2}-1} e^{-x/2}}{\Gamma(\alpha-\frac{1}{2})\ 2^{\alpha-\frac{1}{2}}}\ dx$$

$$= \frac{\Gamma(\alpha-\frac{1}{2})}{2\sqrt{2}\ \Gamma(\alpha)} [1 - G(y^2)],$$

where G is the distribution function of a $\Gamma(\alpha-\frac{1}{2},2)$ random variable. Changing the scale via the transformation $(3/2\alpha)^{\frac{1}{2}}Y$ yields the density in Eq. (2.1).

Steps 2 and 3 of the algorithm are simple to perform. Step 1 can be accomplished by the following methods recommended by Atkinson and Pearce [1976] and by Cheng [1977].

<u>Case 1</u>: $\alpha < 1$

1. Generate a beta variate Y with parameters α and $1-\alpha$, as follows. Take $W_1 = U_1^{1/\alpha}$ and $W_2 = U_2^{1/(1-\alpha)}$, where U_1 and U_2 are generated as iid uniform 0-1 variates. If $W_1 + W_2 \leq 1$, take $Y = W_1/(W_1 + W_2)$. Otherwise, generate a new pair U_1 and U_2 and repeat this procedure.

2. Generate an exponential variate Z by the transformation $Z = -\ell n\ U$, where U is uniform 0-1.

3. The variate $X = 2 \cdot Y \cdot Z$ has a $\Gamma(\alpha,2)$ distribution.

<u>Case 2</u>: $1 < \alpha < 2$

1. Generate $Y_1 = -\ell n\ U_1$, $Y_2 = -\ell n\ U_2$, where the U_i's are iid uniform 0-1.

2. If $Y_2 \geq (\alpha-1) \cdot (Y_1 - \ell n\ Y_1 - 1)$, take $X = 2\alpha Y_1$. Otherwise, return to Step 1.

Case 3: $\alpha > 2$, α nonintegral

1. Compute $\lambda = \sqrt{2\alpha-1}$, $\mu = \alpha^{\lambda}$, $a = \lambda^{-1}$, $b = \alpha - \ln 4$, $c = \alpha + \lambda$.
2. Generate U_1, U_2, which are iid uniform 0-1 variates.
3. Set $Y = a \ln [U_1/(1-U_1)]$, $X = \alpha e^{Y}$.
4. If $b + cY - X \geq \ln(U_1{}^2U_2)$, accept X. Otherwise, return to Step (2).

Case 4: α integer, $\alpha \in \{1,2,3,4,5\}$

1. $X = -\ln \prod_{i=1}^{\alpha} U_i$, where the U_i's are iid uniform 0-1, has a $\Gamma(\alpha,1)$ distribution.

The proposed distribution can be used in Monte Carlo simulation studies to measure departures from univariate normality. For example, in assessing the performance of a univariate statistical procedure, one could test the procedure in a Monte Carlo computer program with the proposed distribution for a spectrum of α values (including $\alpha = 1.5$, the normal). The performance measure could then be plotted against the kurtosis $[1.8(\alpha+1)/\alpha]$ to observe this effect of nonnormality.

The subsequent example illustrates the method. In a particular nuclear safeguards application, we are interested in the distribution of the random variable Z defined by

$$Z = Y_1Y_2 - Y_3Y_4 - Y_5Y_6 - Y_7,$$

where the Y_i's are independent with means and standard deviations given as follows:

	MEAN	ST. DEV.
Y_1	40	0.05
Y_2	250	0.0501
Y_3	10 657	2.0
Y_4	0.844514	0.002351
Y_5	40	0.5
Y_6	22.5	0.0501
Y_7	100	5.0

The expected value of Z is 0. We wish to estimate 98% confidence intervals for Z as a function of the distributions of the Y_i's. We generate random variates from Z according to the following scheme.

1. Consider the set of nine parameter-value-kurtosis combinations: $\{(\alpha,\beta_2)$: (0.5, 5.4), (0.6, 4.8), (0.75, 4.2), (1.0, 3.6), (1.5, 3.0), (2.0, 2.7), (4.0, 2.25), (9.0, 2.0), $(\infty, 1.8)\}$.

2. For each of the nine values of α, let the Y_i's, i = 1,2,...,9, have the proposed distribution with shape parameter α and mean and standard deviation as above.

3. For each α value, generate 10 sets of 1000 random variates, Z. Sort each set of 1000 variates. An estimate of a 98% confidence interval on the distribution of Z is given by $(X_{(10)}, X_{(990)})$, where $X_{(i)}$ is the i^{th} order statistic.

The values of $X_{(10)}$ and $X_{(990)}$ generated from the above scheme are plotted in Fig. 2. As the kurtosis increases, the estimated 98% confidence interval becomes wider, and the variability in the order statistics increases. The plot gives a visual representation of the effect on Z with departures from normality.

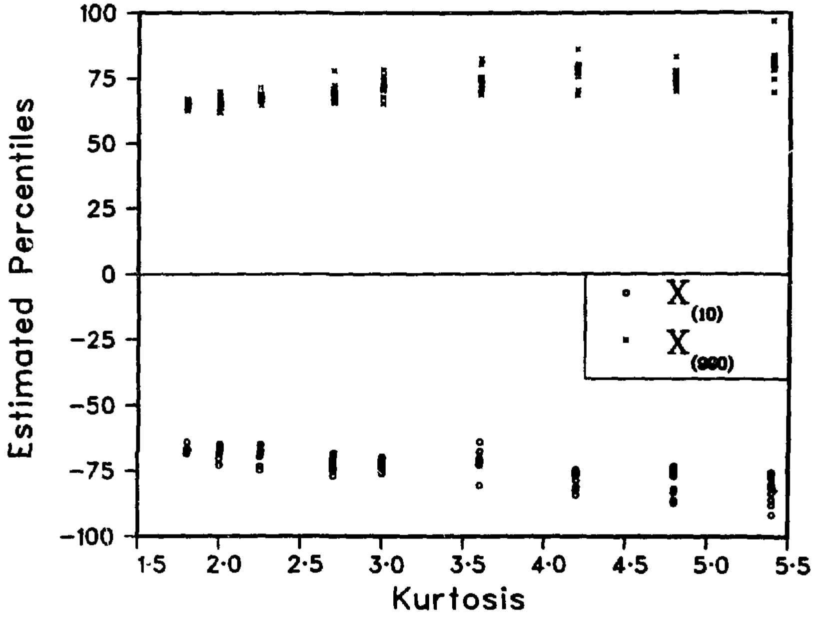

Fig. 2. Safeguards Application.

IV. SUMMARY

A new symmetric probability distribution has been proposed which has desirable properties for use in Monte Carlo simulation studies. The proposed distribution includes both the normal and uniform distributions as special cases and can have a kurtosis in the range 1.8 to 5.4. An algorithm for generating variates having this distribution was derived in Sec. III. An outline of potential applications and an example were also described.

REFERENCES

Amos, D. E. and Daniel, S. L., "Significant Digit Incomplete Gamma Ratios," (1972), Sandia Laboratories Development Report SC-DR-72 0303.

Atkinson, A. C. and Pearce, M. C., "The Computer Generation of Beta, Gamma, and Normal Random Variables." Journal of the Royal Statistical Society A, (1976), 139, Part 4, pp. 431-448.

Cheng, R. C. H., "The Generation of Gamma Variables with Non-integral Shape Parameter," Applied Statistics (1977), 26, No. 1, pp. 71-75.

APPENDIX

MATHEMATICAL DERIVATIONS

We first demonstrate that the density in Eq. (2.1) actually integrates to 1, as follows:

$$\int_{-\infty}^{\infty} \left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right] dx$$

$$= \int_{0}^{\infty} \left(\frac{3}{2\alpha}\right)^{\frac{1}{2}} y^{-\frac{1}{2}} \, [1 - G(y)] \; dy$$

$$= \left(\frac{3}{2\alpha}\right)^{\frac{1}{2}} \left\{ 2y^{\frac{1}{2}}[1 - G(y)] \Big|_0^{\infty} + \int_0^{\infty} 2y^{\frac{1}{2}} g(y) \; dy \right\}$$

$$= \left(\frac{3}{2\alpha}\right)^{\frac{1}{2}} \left\{ \lim_{y\to\infty} \frac{2[1 - G(y)]}{y^{-\frac{1}{2}}} + \int_0^{\infty} 2y^{\frac{1}{2}} \frac{y^{\alpha-3/2} e^{-y/2}}{\Gamma(\alpha-\frac{1}{2})\, 2^{\alpha-\frac{1}{2}}} \, dy \right\}$$

$$= \left(\frac{3}{2\alpha}\right)^{\frac{1}{2}} \left\{0 + \frac{2^{3/2}\Gamma(\alpha)}{\Gamma(\alpha-\frac{1}{2})}\right\}$$

$$= \frac{2\Gamma(\alpha)}{\Gamma(\alpha-\frac{1}{2})(\alpha/3)^{\frac{1}{2}}} .$$

By a symmetry argument, the mean and skewness of the proposed distribution can be shown to be 0. For example, for the mean,

$$E(X) = \int_{-\infty}^{\infty} \frac{(\alpha/3)^{\frac{1}{2}}\ \Gamma(\alpha-\frac{1}{2})}{2\Gamma(\alpha)}\ x\left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right] dx$$

$$= \frac{(\alpha/3)^{\frac{1}{2}}\Gamma(\alpha-\frac{1}{2})}{2\Gamma(\alpha)} \left\{\int_0^{\infty} x\left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right] dx + \int_0^{\infty} (-x)\left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right] dx\right\}$$

$$= 0.$$

The variance of the proposed distribution is

$$\mathrm{Var}(X) = \frac{(\alpha/3)^{\frac{1}{2}}\Gamma(\alpha-\frac{1}{2})}{2\Gamma(\alpha)} \int_{-\infty}^{\infty} x^2 \left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right] dx$$

$$= \frac{3\,\Gamma(\alpha-\frac{1}{2})}{2^{5/2}\alpha\Gamma(\alpha)} \int_0^{\infty} y^{\frac{1}{2}} \left[1 - G(y)\right] dy$$

$$= \frac{3\Gamma(\alpha-\frac{1}{2})}{2^{5/2}\ \alpha\Gamma(\alpha)} \int_0^{\infty} \frac{2}{3}\ y^{3/2} \cdot \frac{y^{\alpha-\frac{1}{2}-1}e^{-y/2}}{\Gamma(\alpha-\frac{1}{2})\ 2^{\alpha-\frac{1}{2}}}\, dy$$

$$= 1.$$

A similar derivation yields the result

$$\beta_2 = \frac{9(\alpha+1)}{5\alpha}.$$

We now consider the behavior of the proposed distribution as the parameter α tends to infinity. We might expect that the proposed density converges to a uniform density since the kurtosis converges to 1.8 as α goes to infinity. This suspicion is confirmed as follows:

$$\lim_{\alpha\to\infty} f(x) = \lim_{\alpha\to\infty} \frac{(\alpha/3)^{\frac{1}{2}}\Gamma(\alpha-\frac{1}{2})}{2\Gamma(\alpha)} \left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right]$$

$$= \lim_{\alpha\to\infty} \frac{(\alpha/3)^{\frac{1}{2}}}{2}\left[\frac{(2\pi)^{\frac{1}{2}}e^{-\alpha}\alpha^{\alpha-1}}{(2\pi)^{\frac{1}{2}}e^{-\alpha}\alpha^{\alpha-\frac{1}{2}}}\right]\cdot\left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right]$$

$$= \frac{1}{2\sqrt{3}} \lim_{\alpha\to\infty}\left[1 - G\left(\frac{2\alpha x^2}{3}\right)\right]$$

$$= \frac{1}{2\sqrt{3}} \lim_{\alpha\to\infty}\left[1 - H\left(x^2/3\right)\right],$$

where H is the distribution function of a Γ $(\alpha-\frac{1}{2},\ 1/\alpha)$ random variable. The distribution corresponding to H has mean $(\alpha-\frac{1}{2})/\alpha$ and variance $(\alpha-\frac{1}{2})/\alpha^2$. As α tends to infinity, the mean goes to 1 and the variance goes to 0. Thus, for $|x| < \sqrt{3}$, $\lim_{\alpha\to\infty}\left[1 - H\left(x^2/3\right)\right] = 1$ and for $|x| \geq \sqrt{3}$, this limit is 0. Hence, the density f coverges to a uniform density as α tends to infinity.