IIE Transactions (2005) **37**, 651–656

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170590948512

Relationships among some univariate distributions

WHEYMING TINA SONG

Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Hsinchu, Taiwan, Republic of China
E-mail: wheyming@ie.nthu.edu.tw, wheyming_song@yahoo.com

Received March 2004 and accepted August 2004

The purpose of this paper is to graphically illustrate the parametric relationships between pairs of 35 univariate distribution families. The families are organized into a seven-by-five matrix and the relationships are illustrated by connecting related families with arrows. A simplified matrix, showing only 25 families, is designed for student use. These relationships provide rapid access to information that must otherwise be found from a time-consuming search of a large number of sources. Students, teachers, and practitioners who model random processes will find the relationships in this article useful and insightful.

1. Introduction

An understanding of probability concepts is necessary if one is to gain insights into systems that can be modeled as random processes. From an applications point of view, univariate probability distributions provide an important foundation in probability theory since they are the underpinnings of the most-used models in practice.

Univariate distributions are taught in most probability and statistics courses in schools of business, engineering, and science. A figure illustrating the relationships among univariate distributions is useful to indicate how distributions correspond to one another. Nakagawa and Yoda (1977), Leemis (1986), and Kotz and van Dorp (2004) offer diagrams of relations among univariate distributions, but their diagrams are not formatted in a matrix form, so it is difficult for the instructor to refer to any particular distribution from the diagram, and it might be difficult for users to find the required distribution quickly. Two versions of the figure presented in this article overcome this shortcoming. Figure 1 is the complete figure with 35 univariate distributions, whereas Fig. 2 is a simpler version with 25 distributions designed for more elementary needs.

Some useful references for studying univariate distributions can be found in Patil, Boswell, Joshi, and Ratnaparkhi (1985), Patil, Boswell, and Ratnaparkhi (1985), Stigler (1986), Hald (1990), Johnson *et al.* (1993, 1994, 1995), Stuart and Ord (1994), Wimmer and Altmann (1999), Evans, Hastings and Peacock (2000), and Balakrishnan and Nevzorov (2003). These books individually do not provide all of the simple relationships shown in this article.

2. A seven-by-five matrix

Figure 1 illustrates 35 univariate distributions in 35 rectangle-like entries. The row and column numbers are labeled on the left and top of Fig. 1, respectively. There are 10 discrete distributions, shown in the first two rows, and 25 continuous distributions. Five commonly used sampling distributions are listed in the third row. For example, the Normal distribution is indexed as R3C1, which indicates Row 3 and Column 1. The range, probability mass (density) function, mean, and variance for discrete and continuous distributions are summarized in Tables 1 and 2, respectively. Distributions are ordered alphabetically in Tables 1 and 2. According to the index number shown in the second column of Tables 1 and 2, users can easily find the distribution required in Fig. 1.

In Fig. 1, the distribution name, parameters and range are shown in each entry. The parameters adopted satisfy the following conventions: n and k are integers; $0 \leq p \leq 1$; a is the minimum; b is the maximum; μ is the expected value; m is either the median or mode; σ is the standard deviation; θ is the location parameter; β and γ are scale parameters; $\alpha, \alpha_1, \alpha_2$ are shape parameters; and $\upsilon, \upsilon_1, \upsilon_2$ are degrees of freedom. We do not always use α to denote the shape parameter. For example, the parameter m (denoted as the mode) in the Triangular (a, m, b); the parameter k (denoted as an integer) in the Erlang (k, β); υ (denoted as the degrees of freedom) in the Chi-Squared (υ); and υ_1, υ_2 (both denoted as the degrees of freedom) in the F(υ_1, υ_2) are shape parameters.

Relations between entries are indicated with a dashed arrow or solid arrow. A dashed arrow shows asymptotic relations, and a solid arrow shows transformations or special cases. The random variable X is used for all distributions.

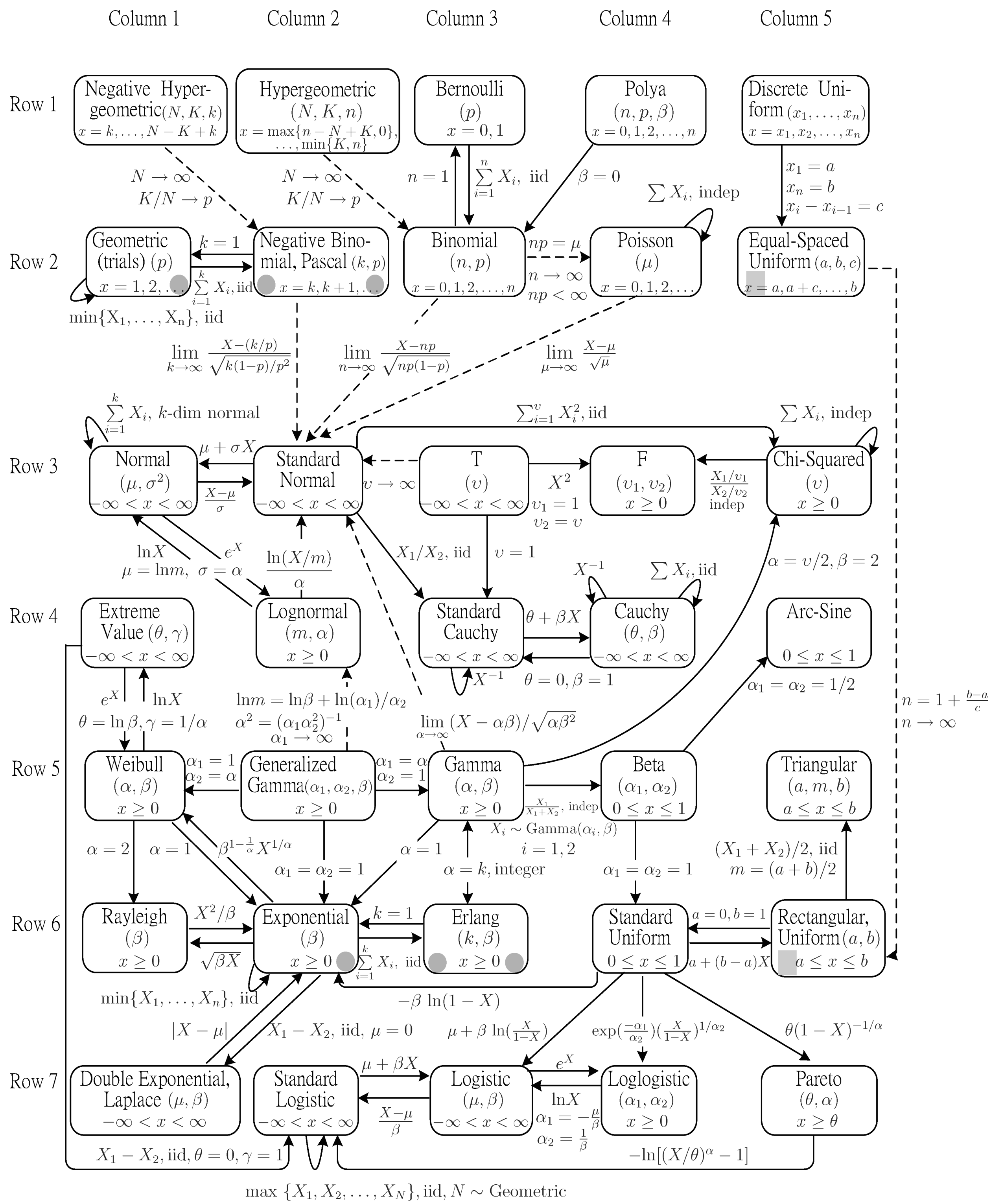

Fig. 1. Relationships among 35 distributions. (An arrow beginning and ending at the same rectangle indicates that it remains in the same distribution family, but the parameter values might change.)

For example, the arrow from the Normal (R3C1) to the Standard Normal (R3C2) indicates that subtracting the mean from a Normal and dividing by its standard deviation yields a Standard Normal distribution.

In Fig. 1, if more than one random variable is involved to form a transformation, the relationship between these random variables is denoted by "iid" (independent and identically distributed), "indep" (independent), or "k-dim

Table 1. Discrete distributions

Distribution name	*Index no.*	*Range*	*Probability mass function*	*Expected value*	*Variance*
Bernoulli (p)	R1C3	$x = 0, 1$	$p^x(1-p)^{1-x}$	p	$p(1-p)$
Binomial (n, p)	R2C3	$x = 0, 1, \ldots, n$	$C_x^n p^x (1-p)^{n-x}$	np	$np(1-p)$
Discrete Uniform $(x_1, \ldots, x_n)$	R1C5	$x = x_1, x_2, \ldots, x_n$	$1/n$	$\sum_{i=1}^n x_i/n$	$\sum_{i=1}^n x_i^2/n - (\sum_{i=1}^n x_i/n)^2$
Equal-Spaced Uniform (a, b, c)	R2C5	$x = a, a+c, a+2c, \ldots, b$	$1/n$ where $n = 1 + \frac{b-a}{c}$	$(a+b)/2$	$c^2(n^2-1)/12$
Geometric (trials) (p)	R2C1	$x = 1, 2, \ldots$	$p(1-p)^{x-1}$	$1/p$	$(1-p)/p^2$
Hyper-Geometric (N, K, n)	R1C2	$x = \max\{n-N+K, 0\}, \ldots, \min\{K, n\}$	$C_x^K C_{n-x}^{N-K} / C_n^N$	np	$np(1-p)(\frac{N-n}{N-1})$ where $p = K/N$
Negative Binomial (k, p)	R2C2	$x = k, k+1, \ldots$	$C_{k-1}^{x-1} p^k (1-p)^{x-k}$	k/p	$k(1-p)/p^2$
Negative Hyper-Geometric (N, K, k)	R1C1	$x = k, \ldots, N-K+k$	$\frac{C_{k-1}^K C_{x-k}^{N-K}}{C_{x-1}^N} \frac{K-k+1}{N-x+1}$	$\frac{k(N+1)}{K+1}$	$\frac{k(N+1)(N-K)}{(K+1)^2(K+2)}(K+1-k)$
Poisson (μ)	R2C4	$x = 0, 1, 2, \ldots$	$\mu^x \exp(-\mu)/x!$	μ	μ
Polya (n, p, β)	R1C4	$x = 0, 1, 2, \ldots, n$	$C_x^n \prod_{j=0}^{x-1}(p + j\beta) \cdot \prod_{k=0}^{n-x-1}(1-p+k\beta) \cdot [\prod_{i=0}^{n-1}(1+i\beta)]^{-1}$	np	$\frac{np(1-p)(1+n\beta)}{1+\beta}$

Note: $C_m^n = n!/[m!(n-m)!]$.

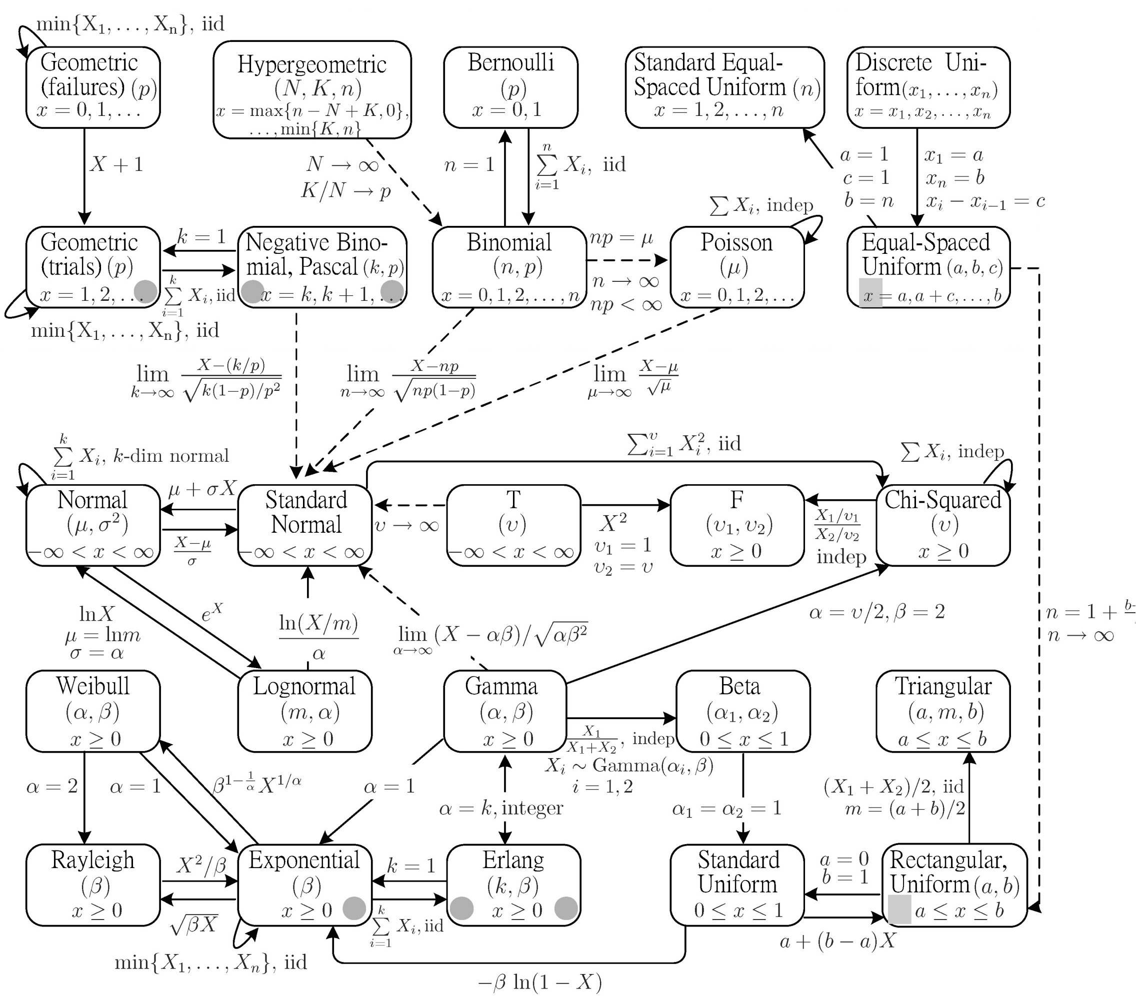

Fig. 2. Relationships among 25 distributions.

Table 2. Continuous distributions

Distribution name	*Index no.*	*Range*	*Probability density function*	*Expected value*	*Variance*
Arc-Sine	R4C5	[0,1]	$[\pi\sqrt{x(1-x)}]^{-1}$	1/2	1/8
Beta (α_1, α_2)	R5C4	[0,1]	$\frac{\Gamma(\alpha_1+\alpha_2)x^{\alpha_1-1}(1-x)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}$	$\frac{\alpha_1}{\alpha_1+\alpha_2}$	$\frac{\alpha_1\alpha_2}{(\alpha_1+\alpha_2)^2(\alpha_1+\alpha_2+1)}$
Cauchy (θ, β)	R4C4	$\mathbb{R}$	$\{\beta\pi[1+(\frac{x-\theta}{\beta})^2]\}^{-1}$	NA	NA
Chi-Squared (υ)	R3C5	$\mathbb{R}^+$	$\frac{x^{\frac{\upsilon}{2}-1}e^{-x/2}}{\Gamma(\upsilon/2)2^{\upsilon/2}}$	υ	2υ
Erlang (k, β)	R6C3	$\mathbb{R}^+$	$\frac{x^{k-1}\exp(-x/\beta)}{\beta^k(k-1)!}$	$k\beta$	$k\beta^2$
Exponential (β)	R6C2	$\mathbb{R}^+$	$\beta^{-1}\exp(-x/\beta)$	β	β^2
Extreme Value (θ, γ)	R4C1	$\mathbb{R}$	$\gamma^{-1}\exp[-(x-\theta)/\gamma]$ $\cdot\exp\{-\exp[-(x-\theta)/\gamma]\}$	$\theta+0.57722\gamma$	$\gamma^2\pi^2/6$
$\mathrm{F}(\upsilon_1, \upsilon_2)$	R3C4	$\mathbb{R}^+$	$\frac{\Gamma(\frac{\upsilon_1+\upsilon_2}{2})(\upsilon_1/\upsilon_2)^{\upsilon_1/2}}{\Gamma(\upsilon_1/2)\Gamma(\upsilon_2/2)}\frac{x^{(\upsilon_1/2)-1}}{(1+\upsilon_1\upsilon_2^{-1}x)^{(\upsilon_1+\upsilon_2)/2}}$	$\frac{\upsilon_2}{\upsilon_2-2}, \upsilon_2>2$	$\frac{2\upsilon_2^2(\upsilon_1+\upsilon_2-2)}{\upsilon_1(\upsilon_2-2)^2(\upsilon_2-4)}, \upsilon_2>4$
Gamma (α, β)	R5C3	$\mathbb{R}^+$	$\frac{(x/\beta)^{\alpha-1}\exp(-x/\beta)}{\beta\Gamma(\alpha)}$	$\alpha\beta$	$\alpha\beta^2$
Generalized Gamma $(\alpha_1, \alpha_2, \beta)$	R5C2	$\mathbb{R}^+$	$\frac{\alpha_2}{\Gamma(\alpha_1)\beta}(\frac{x}{\beta})^{\alpha_1\alpha_2-1}\exp[-(\frac{x}{\beta})^{\alpha_2}]$	$\frac{\beta\Gamma(\frac{1}{\alpha_2}+\alpha_1)}{\Gamma(\alpha_1)}$	$\beta^2\{\frac{\Gamma(\frac{2}{\alpha_2}+\alpha_1)}{\Gamma(\alpha_1)}-\frac{\Gamma^2(\frac{1}{\alpha_2}+\alpha_1)}{\Gamma^2(\alpha_1)}\}$
Laplace (μ, β)	R7C1	$\mathbb{R}$	$(2\beta)^{-1}\exp(-\|x-\mu\|/\beta)$	μ	$2\beta^2$
Logistic (μ, β)	R7C2	$\mathbb{R}$	$\frac{\exp[-(x-\mu)/\beta]}{\beta\{1+\exp[-(x-\mu)/\beta]\}^2}$	μ	$\pi^2\beta^2/3$
Loglogistic (α_1, α_2)	R7C4	$\mathbb{R}^+$	$\frac{\alpha_2\exp(-\alpha_1)x^{-\alpha_2-1}}{[1+\exp(-\alpha_1)x^{-\alpha_2}]^2}$	$\delta\exp(-\eta)\csc(\delta)$	$\delta[\exp(-2\eta)]$ $\cdot[\tan(\delta)-\delta]\csc^2(\delta)$ where $\delta=\frac{\pi}{\alpha_2}, \eta=\frac{\alpha_1}{\alpha_2}$
Lognormal (m, α)	R4C2	$\mathbb{R}^+$	$(x\alpha\sqrt{2\pi})^{-1}\exp[-\frac{1}{2}(\frac{\ln(x/m)}{\alpha})^2]$	$m\exp(\alpha^2/2)$	$m^2\exp(\alpha^2)$ $\cdot[\exp(\alpha^2)-1]$
Normal (μ, σ^2)	R3C1	$\mathbb{R}$	$\frac{\exp[(-1/2)(\frac{x-\mu}{\sigma})^2]}{\sqrt{2\pi}\sigma}$	μ	σ^2
Pareto (θ, α)	R7C5	$[\theta, \infty)$	$\alpha\theta^\alpha/x^{\alpha+1}$	$\frac{\alpha\theta}{\alpha-1}, \alpha>1$	$\frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, \alpha>2$
Rayleigh (β)	R6C1	$\mathbb{R}^+$	$(2x/\beta^2)\exp[-(x/\beta)^2]$	$\beta\sqrt{\pi}/2$	$\beta^2-(\frac{\beta\sqrt{\pi}}{2})^2$
Rectangular (a, b)	R6C5	$[a, b]$	$(b-a)^{-1}$	$(a+b)/2$	$(b-a)^2/12$
Standard Cauchy $(\theta=0, \beta=1)$	R4C3	$\mathbb{R}$	$[\pi(1+x^2)]^{-1}$	NA	NA
Standard Logistic $(\mu=0, \beta=1)$	R7C3	$\mathbb{R}$	$\frac{\exp(-x)}{[1+\exp(-x)]^2}$	0	$\pi^2/3$
Standard Normal $(\mu=0, \sigma=1)$	R3C2	$\mathbb{R}$	$\frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$	0	1
Standard Uniform $(a=0, b=1)$	R6C4	[0,1]	1	1/2	1/12
$\mathrm{T}(\upsilon)$	R3C3	$\mathbb{R}$	$\frac{\Gamma[(\upsilon+1)/2]}{\Gamma(\upsilon/2)\sqrt{\pi\upsilon}}(1+\frac{x^2}{\upsilon})^{-(\upsilon+1)/2}$	$0, \upsilon>1$	$\frac{\upsilon}{\upsilon-2}, \upsilon>2$
Triangular (a, m, b)	R5C5	$[a, b]$	$\begin{cases}\frac{2(x-a)}{(m-a)(b-a)}, a\le x\le m\\ \frac{2(b-x)}{(b-a)(b-m)}, m<x\le b\end{cases}$	$\frac{1}{3}(a+m+b)$	$\frac{1}{18}\{a^2+m^2+b^2$ $-(am+ab+mb)\}$
Weibull (α, β)	R5C1	$\mathbb{R}^+$	$\alpha\beta^{-\alpha}x^{\alpha-1}\exp[-(x/\beta)^\alpha]$	$\beta\Gamma(1+\frac{1}{\alpha})$	$\beta^2\Gamma(1+\frac{2}{\alpha})$ $-[\beta\Gamma(1+\frac{1}{\alpha})]^2$

Note: $\mathbb{R}^+$ denotes $[0, \infty)$ and $\mathbb{R}$ denotes $(-\infty, \infty)$. NA denotes "not applicable."

normal" (k-dimensional Normal Distribution). For example, the relation from the Geometric (R2C1) to the Negative Binomial (R2C2) is marked "$\sum_{i=1}^k X_i$, iid," which indicates that the sum of k iid Geometric random variables yields a Negative Binomial distribution. The relation from Chi-Squared (R3C5) back to the Chi-Squared is marked "$\sum X_i$, indep," which indicates that the sum of independent Chi-Squared random variables yields a Chi-Squared distribution. The relation from the Normal (R2C1) to the Normal is marked "$\sum_{i=1}^k X_i$, k-dim normal," which indicates that the sum of k random variables has a Normal distribution if they are components of a k-dimensional Normal distribution.

The multivariate Normal is discussed in many places, including Johnson *et al.* (1997), Johnson and Wichern (1998), and Kotz et al. (2000).

The transformation relationships in Fig. 1 can be combined to form other relationships. For instance, a path from the Standard Normal (R3C2) to the Chi-Squared (R3C5) to the Gamma (R5C3) to the Exponential (R6C2) indicates that the random variable $X_1^2 + X_2^2$ has the Exponential ($\beta = 2$) distribution if X_1 and X_2 are independent Standard Normal random variables.

There is a special relationship between the Standard Uniform (R6C4) and any continuous distribution. That is, $F_X(X) \sim \text{Uniform}(0, 1)$, where F_X is the cumulative distribution function (cdf) of any continuous random variable X. Therefore, an arrow with the relation $F_X(X)$ could be drawn from any continuous distribution to the Standard Uniform, and an arrow with the relation $F_X^{-1}(U)$ (where $U \sim \text{Uniform}\,(0,1)$) could be drawn from the Standard Uniform to any continuous distribution where the random variable $F_X^{-1}(U)$ is known as the transformation of inverse-cdf, which is a popular random variate generation technique. Random variate generation is discussed in many places, including Schmeiser (1980), Devroye (1986) and Gentle (2003).

Let us continue the discussion of the relationship between the Standard Uniform and any continuous distribution. We do not draw arrows connecting the Standard Uniform U with all continuous distributions X in Fig. 1 because that would render the diagram unnecessarily difficult to read. In Fig. 1, we simply draw arrows from the Standard Uniform to certain continuous distributions. For example, we show the arrows from the Standard Uniform to the Exponential distribution (R6C2), Rectangle (R6C5), Loglogistic (R7C4), Logistic (R7C2), and Pareto (R7C5). The relations from the Standard Uniform to some other continuous distributions, including the Triangular (R5C5), Standard Logistic (R7C4), and Rayleigh (R6C1), such that the closed-form inverse cdf exists, are not shown in Fig. 1 because of space limitations. These relations can be easily derived via the combined paths, as discussed earlier.

We mark some entries to show the relationship between the corresponding discrete and continuous distributions. To indicate that the Geometric distribution (R2C1) corresponds to the Exponential distribution (R6C2), we mark a circle in the bottom of these two entries. To indicate that the Negative Binomial distribution (R2C2) corresponds to the Erlang (R6C3), we mark two circles in the bottom of these two entries. To indicate that the Equal-Spaced Uniform (R2C5) corresponds to the Continuous Uniform (R6C5), we mark a rectangle in the bottom of these two entries.

One final comment is that the Logistic, Standard Logistic, and Loglogistic are analogous to the Normal, Standard Normal, and Lognormal, respectively. See Balakrishnan (1992, p. 190) for a detailed explanation.

3. A simplified matrix

Figure 2 is a simpler version of Fig. 1, not including the Negative Hypergeometric (R1C1), Polya (R1C4), Extreme Value (R4C1), Standard Cauchy (R4C3), Cauchy (R4C4), Arc-Sine (R4C5), Generalized Gamma (R5C2), and the five distributions in Row 7. In a probability introductory course, we have found that students are motivated to learn the transformation of random variables and uses of the moment generating functions, so that they can derive all 44 relations in Fig. 2. Discussions on the transformation of random variables and moment generating functions are typically optional in standard probability and statistics texts, such as Walpole *et al.* (2002). To be able to prove all relations in Fig. 1, some additional mathematical background is required.

We add two distributions in the first row of Fig. 2: the Geometric (failures) distributions and the Standard Discrete Uniform (with range $1, 2, \ldots, n$), which is a special case of the Equal-Spaced Uniform. The Geometric (failures) random variable is the total number of failures before the first success occurs. The Geometric (R2C1) in Fig. 1 now becomes Geometric (trials) in which the random variable indicates the number of trials needed until the first success occurs. To avoid confusion with index numbers used in Tables 1 and 2, we do not list row and column numbers in Fig. 2.

Acknowledgements

This paper is based upon work supported by the National Science Council under grant NSC-90-2218-E-007-018. Many people contributed information that was of value in the preparation of this paper and for that I am most appreciative. In particular, the author sincerely thanks professors Bruce Schmeiser and David Goldsman for their perceptive comments. Further, the author deeply thanks Ph.D. students Mingchang Chih and Sn-Long Liu for their very helpful proofreading.

References

Balakrishnan, N. (1992) *The Handbook of the Logistic Distribution*, Marcel Dekker, New York, NY.

Balakrishnan, N. and Nevzorov V.B. (2003) *A Primer on Statistical Distributions*, Wiley, New York, NY.

Devroye, L. (1986) *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, NY.

Evans, M., Hastings, N. and Peacock, B. (2000) *Statistical Distributions*, Wiley, New York, NY.

Gentle, J.E. (2003) *Random Number Generation and Monte Carlo Methods*, Springer-Verlag, New York, NY.

Hald, A. (1990) *A History of Probability and Statistics and Their Applications Before 1750*, Wiley, New York, NY.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) *Continuous Univariate Distributions—1*, 2nd eds., Wiley, New York, NY.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995) *Continuous Univariate Distributions—2*, 2nd eds., Wiley, New York, NY.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1997) *Discrete Multivariate Distributions*, Wiley, New York, NY.

Johnson, N.L., Kotz, S. and Kemp, W.W. (1993) *Univariate Discrete Distributions*, 2nd eds., Wiley, New York, NY.

Johnson and Wichern (1998) *Applied Multivariate Statistical Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000). *Continuous Multivariate Distributions*, 2nd eds., Wiley, New York, NY.

Kotz, S. and Jogn René van Dorp (2004) *Beyond Beta—Other Continuous families of Distributions with Bounded Support and Applications*, World Scientific, NJ, p. 251.

Leemis, M.L. (1986) Relationships among common univariate distributions. *The American Statistician*, **40**(2), 143–145.

Nakagawa, T. and Yoda, H. (1977) Relationships among distributions. *IEEE Transactions on Reliability*, **26**(5), 352–353.

Patil, G.P., Boswell, M.T., Joshi, S.W. and Ratnaparkhi, M.V. (1985) *Discrete Continuous Models*, International Co-operative Publishing House, Burtonsville, MD.

Patil, G.P., Boswell, M.T. and Ratnaparkhi, M.V. (1985) *Univariate Continuous Models*, International Co-operative Publishing House, Burtonsville, MD.

Schmeiser, B.W. (1980) Random variate generation: a survey, in *Simulation with Discrete Models: A State-of-the-Art View*, Oren, T.I., Shub, C.M. and Roth, P.F. (eds). Institute of Electrical and Electronic Engineers, New York, NY, pp. 79–104.

Stigler, S.M. (1986) *The History of Statistics*, Cambridge, MA: The Balknap Press.

Walpole, R.E., Myers, R.H., Myers, S.L. and Ye, K. (2002) *Probability & Statistics for Engineering and Scientists*, 7th ed., Prentice Hall, Upper Saddle River, NJ.

Wimmer, G. and Altmann, G. (1999) *Thesaurus of Univariate Discrete Distribution Probability Distributions*, STAMM Verlag, Essen, Germany.

Biography

Wheyming Tina Song is a Professor in the Department of Industrial Engineering at the National Tsing Hua University in Taiwan. She received her undergraduate degree in Statistics and Master's degree in Industrial Engineering at Cheng-Kung University in Taiwan in 1979. She then received Master's degrees in Applied Mathematics in 1983 and Industrial Engineering in 1984, both from the University of Pittsburgh. She received her Ph.D. from the School of Industrial Engineering at Purdue University in 1989. She joined Tsing Hua in 1990 after spending a year as a Visiting Assistant Professor at Purdue. She received the outstanding university-wide teaching award from the National Tsing Hua University in 1993. Her research honors include the 1988 Omega Rho student paper award, the 1990 IIE Doctoral Dissertation award, and the 1996 distinguished research award from the National Science Council of the Republic of China. Her research interests are in applied operations research; probability and statistics; and the statistical aspects of stochastic simulation, including input modeling, output analysis, variance reduction, and ranking and selection.

Contributed by the Simulation Department