NOTEBOOK

The ties that bind

In this special instalment of the Notebook series, **Lawrence M. Leemis** and **Raghu Pasupathy** explore the hidden links between probability distributions

Probability distributions and, more generally, probability models have become indispensable to modelling and predicting the functioning of physical phenomena. From the movement of asset price fluctuations or travel demand to the spread of an epidemic in a population, probability models are the principal tool for analysis, inference, forecasting and decision-making under uncertainty.

Given such importance, it seems useful to have a full understanding of which distributions may be appropriate for a given context, and how various distributions may be related to each other. A shining example of where such an understanding pays off is the case of the normal distribution.

The earliest record of the use of the normal distribution appears to be as a convenient approximation to the binomial distribution.[1] This early work, by de Moivre, led to later work by eminent researchers, including Laplace and Gauss, on approximating other distributions using the normal distribution, and ultimately to the construction of the now indispensable central limit theorem (CLT). Just as the CLT relates the normal to virtually any other distribution, numerous other useful relationships between distributions have been discovered over many decades.

Visualising the relationships

The boxes in Figure 1 attempt to succinctly represent many of the known relationships between 12 discrete distributions (with square corners) and 16 continuous distributions (with rounded corners) arising frequently in applications of probability and statistics.

The arrows connecting these probability distributions describe the relationships. Solid arrows containing an X in their label are transformations; solid arrows without an X in their label are special cases. Dashed arrows depict limiting distributions; dotted arrows depict parameters that are themselves random variables. Some of the distributions have properties given in the second line of their box, and these are described in the legend in the lower left-hand corner.

These relationships can be helpful in determining appropriate models when developing a probability model and for deriving the distribution of certain quantities which are useful in statistical inference.

Putting the relationships to work

One simple application of the relationship between distributions is that of the standard normal – $N(0, 1)$ in the figure – and the Cauchy distribution. If X_1 and X_2 are independent random variables having the standard normal distribution, then their ratio X_1/X_2 has the standard Cauchy distribution. Knowledge of this relationship means that in settings where a ratio of two random quantities that can be reasonably approximated by standard normal random variables is of interest, the ratio could be directly modelled by the Cauchy distribution.

Another example of the ratio distribution is that between Fisher's F and chi-squared – $\chi^2(n)$ – random variables. Specifically, if $\chi^2(n_1)$ and $\chi^2(n_2)$ are independent chi-squared random variables having n_1 and n_2 degrees of freedom respectively, then $(\chi^2(n_1)/n_1)/(\chi^2(n_2)/n_2)$ has Fisher's $F(n_1, n_2)$ distribution.

For a more specific contextual example, consider an unfinished part entering an assembly line in a factory. Suppose the part is to work its way through k stations in sequence before exiting the line as a finished good. Suppose further that an analyst is interested in modelling the time it takes from when the part first enters the line to when it exits the line. Reasoning that the service time at each station can be modelled as coming from an exponential distribution, and that the service times at the stations are mutually independent and identically distributed (i.i.d.), the analyst might use Figure 1 and follow the arrow connecting the exponential and the Erlang distributions to conclude that an Erlang distribution might be an appropriate probability model for the total time taken by the part to traverse the assembly line. (Such reasoning assumes little to no waiting time at the stations.) The decision to use an Erlang distribution comes from noticing from Figure 1 that the sum of i.i.d. exponential random variables has the Erlang distribution.

Lawrence M. Leemis is a professor in the Department of Mathematics at the College of William & Mary in Williamsburg, Virginia.

Raghu Pasupathy is an associate professor in the Department of Statistics at Purdue University in West Lafayette, Indiana.

Do some distributions arise more often than others?

Certain distributions in Figure 1 appear to be more central, while others appear to be at the periphery. One criterion for this centrality is the number of incoming and outgoing arrows. Here are three examples.

- The normal distribution – along with its special case, the standard normal distribution – is central to classical statistics. These two distributions are connected to several critical distributions that arise in classical statistics, such as the Student's t and chi-square distributions. The CLT provides the basis for some of the limiting relationships that are connected to the normal distribution.
- The exponential distribution is central in the study of stochastic processes. It

is connected to many popular survival distributions with positive support, such as the Weibull distribution.

- The uniform – $U(0, 1)$ – distribution is also heavily connected with other distributions. The probability integral transformation indicates that the $U(0, 1)$ distribution should be connected to all of the other distributions in the chart. Keeping the chart flat, rather than three-dimensional, is the only reason not to make these connections. To cite one specific example, if $X \sim U(0, 1)$ then the arrow connecting the $U(0, 1)$ distribution to the exponential(λ) distribution tells us that $-\frac{1}{\lambda} \ln X \sim$ exponential(λ). Distributions with the "V" property (for variate generation) have a simple, algebraic conversion from a random number to a random variate.

Keep in mind...

An interactive version of Figure 1, with more probability distributions and many of the proofs of the relationships and properties, is online at bit.ly/2WVblrl. A static version is also available.[2] A moment-ratio diagram, which allows a modeller to see several probability distributions at once, can be explored at significancemagazine.com/628.

Probability distributions arranged in a matrix fashion are given by Song.[3] Finally, most probability distributions have probabilistic interpretations (e.g., the binomial distribution models the number of successes in n independent Bernoulli trials), and these are discussed in the encyclopaedic series by Johnson, Kotz and Balakrishnan.[4] ■

Understanding the relationships between distributions can be helpful in determining appropriate models when developing a probability model

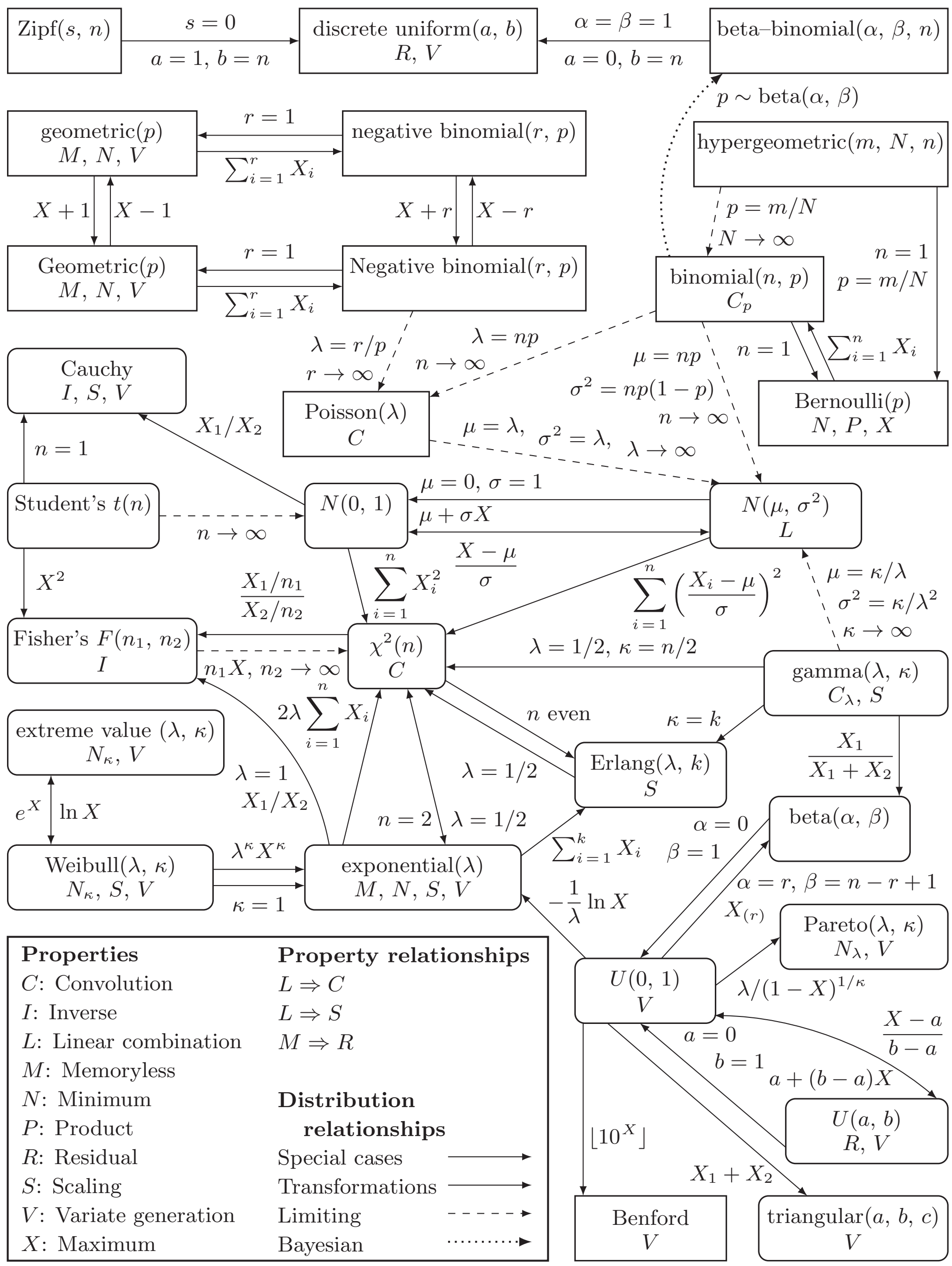

FIGURE 1 Univariate distributions: parameters, properties, and relationships.

References

1. Pearson, K., de Moivre, A. and Archibald, R. C. (1926) A rare pamphlet of Moivre and some of his discoveries. *Isis*, **8**(4), 671–683.

2. Leemis, L. M. and McQueston, J. T. (2008) Univariate distribution relationships. *American Statistician*, **62**(1), 45–53.

3. Song, W. T. (2005) Relationships among some univariate distributions. *IIE Transactions*, **37**, 651–656.

4. Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994–1995) *Continuous Univariate Distributions*, Volumes 1 and 2 (2nd edition). New York: Wiley.