2012

Univariate Probability Distributions

Lawrence Leemis

Daniel J. Luckett

Austin Powell

Peter J. Vermeer

Univariate Probability Distributions

Lawrence M. Leemis
Daniel J. Luckett
Austin G. Powell
Peter E. Vermeer
The College of William & Mary

Journal of Statistics Education Volume 20, Number 3 (2012),
http://www.amstat.org/publications/jse/v20n3/leemis.pdf

Key Words: Continuous distributions; Discrete distributions; Distribution properties; Limiting distributions; Special Cases; Transformations; Univariate distributions.

Abstract

We describe a web-based interactive graphic that can be used as a resource in introductory classes in mathematical statistics. This interactive graphic presents 76 common univariate distributions and gives details on (a) various features of the distribution such as the functional form of the probability density function and cumulative distribution function, graphs of the probability density function for various parameter settings, and values of population moments; (b) properties that the distribution possesses, for example, linear combinations of independent random variables from a particular distribution family also belong to the same distribution family; and (c) relationships between the various distributions, including special cases, transformations, limiting distributions, and Bayesian relationships. The interactive graphic went online on 11/30/12 at the URL www.math.wm.edu/ leemis/chart/UDR/UDR.html.

1. Introduction

Introductory textbooks in probability and statistics often introduce univariate probability distributions in separate sections, which obscures both an understanding of the relationships between distributions and the properties that many of them have in common. Leemis and McQueston (2008) designed a figure to highlight these relationships and properties. Song and Chen (2011) redesigned the figure in a matrix format, making it easier to locate distributions. This article contains a description of a website that contains an interactive graphic of the figure at www.math.wm.edu/∼leemis/chart/UDR/UDR.html. The user can also find proofs of the relationships and properties at this website.

The figure in Leemis and McQueston (2008) contains 76 probability distributions. Of these, 19 are discrete and 57 are continuous. Figure 1 contains a screenshot of the upper-left-hand corner of the interactive graphic. An alphabetical list of the discrete and continuous probability distributions is displayed on the left-hand side of the screen, along with a slider bar to scroll through the list. Just above this list are four buttons labeled CHART, ABOUT, LINKS, and CONTACT. The CHART button displays the chart that contains the distributions. The ABOUT button gives information concerning the chart. The LINKS button contains links that might be of interest to a visitor. One such link connects the user to plots of the coefficient of variation vs. the skewness and the skewness vs. the kurtosis for the various distributions included in the website. The CONTACT button gives contact information for the developers of the interactive graphic. There are two ways to highlight a probability distribution. The first is to hover over the name of the distribution on the list at the far left side of the screen, which brings the probability distribution into view

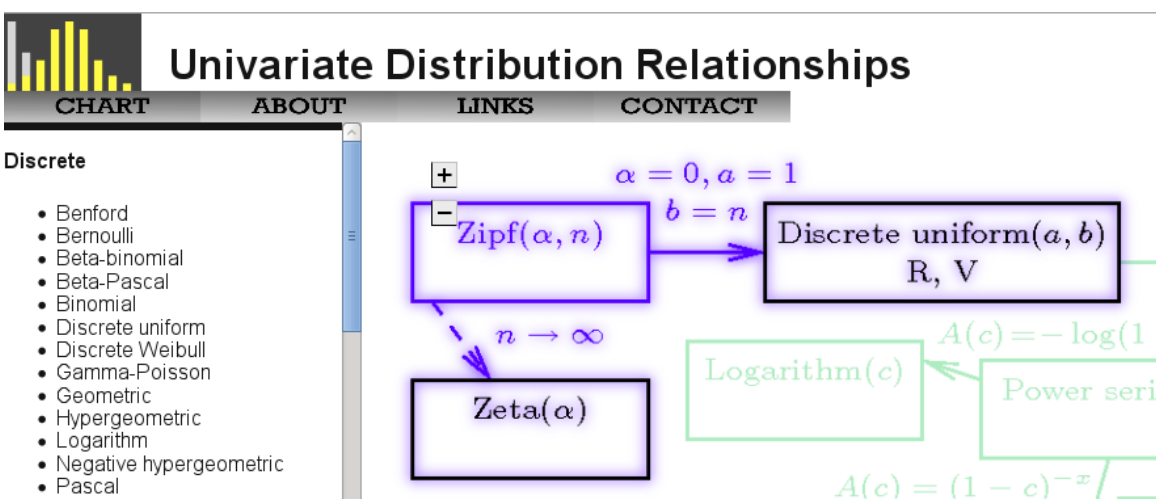

Figure 1. Hovering over the Zipf distribution.

on the chart. The second is to hover over the name of the distribution in the body of the chart. In Figure 1, the cursor is hovering over the Zipf distribution. When hovering over the Zipf distribution, in either of the two manners just described, five changes occur on the interactive graphic. First, the Zipf distribution is highlighted in blue (colors may vary depending on the browser). Second, the boxes associated with distributions connected to the Zipf distribution are highlighted in black. Third, the arrows and labels associated with outgoing relationships from the Zipf distribution are highlighted in blue. Fourth, the arrows and labels associated with incoming relationships to the Zipf distribution (there are no such relationships for the Zipf distribution) are highlighted in black. Fifth, the other distributions that are not connected to the Zipf distribution are lightened in intensity. The number of distributions that are highlighted is helpful to indicate which distributions are at the center of probability theory (e.g., the normal, binomial, exponential, uniform, and chi-square distributions) and which are at the periphery (e.g., the Pólya, generalized Pareto, and doubly noncentral F distributions). Finally, just to the northwest of the box for the Zipf distribution are a $+$ and $-$ button for zooming.

Each box contains the distribution's name, its parameters, and a list of its properties using a single letter as a code (e.g., C for the convolution property). For example, Figure 2 shows the interactive graphic when the cursor is hovering over the geometric distribution. The geometric distribution is highlighted in blue, while the Pascal (negative binomial) and discrete Weibull distributions are highlighted in black because they are connected to the geometric distribution. Since all three of these distributions are discrete, they are placed in rectangular boxes (the boxes for continuous distributions have rounded corners). The geometric distribution has one parameter, p, the probability of success. The Pascal distribution has two parameters, n, the number of successes, and p, the probability of success. The discrete Weibull distribution has two parameters, p and β. The geometric distribution has properties F (forgetfulness), M (minimum), and V (variate generation). Clicking the ABOUT tab reveals more detail about these properties. A legend in the lower-left-hand corner of the chart portion of the interactive graphic also contains the definitions of these

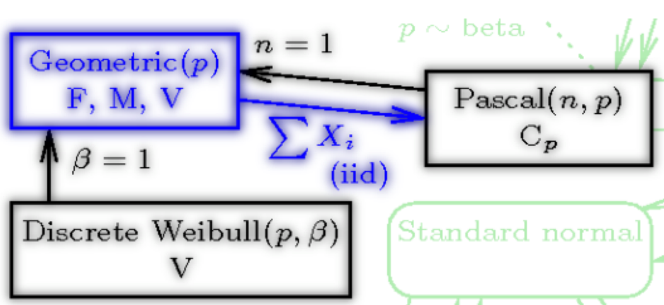

Figure 2. Hovering over the geometric distribution.

properties. The Pascal distribution has the C (convolution) property, but only when the p parameter is fixed. The discrete Weibull distribution has the V (variate generation) property. More detail about the properties is given in Section 2. The interactive graphic also indicates how the two distributions are related by the arrows connecting the distributions. The cursor is hovering over the geometric distribution, so outgoing arrows are blue and the incoming arrows are black. The blue outgoing arrow indicates that the sum of mutually independent and identically distributed (iid) geometric(p) random variables has the Pascal distribution. This is known as a transformation. The incoming black arrows indicates that the geometric distribution is a special case of the Pascal distribution when $n = 1$ and a special case of the discrete Weibull distribution when $\beta = 1$. These are known as special cases. More detail about the relationships between distributions is given in Section 3.

Figure 3 shows the screen when the cursor hovers over the standard normal distribution. Solid arrows denote transformations and special cases. Dashed arrows denote asymptotic relationships, and dotted arrows denote Bayesian relationships. Some of the facts concerning the standard normal distribution from Figure 3 are listed below.

- The ratio of two independent standard normal random variables has the standard Cauchy distribution.
- The t distribution approaches the standard normal distribution as its degrees of freedom $n \to \infty$.
- The absolute value of a standard normal random variable has the chi distribution.
- The sum of squares of mutually independent standard normal random variables has the chi-square distribution.
- The standard normal distribution is a special case of the normal distribution in which $\mu = 0$ and $\sigma = 1$.
- There is a one-to-one transformation between the normal distribution and standard normal distribution. Standardizing a normal random variable results in a standard normal random variable, which is useful for probability calculations. Multiplying a standard normal random variable by σ and adding μ results in a normal random variable, which is useful for random variate generation.

When the name of a distribution is clicked (either from the list at the far left or on the chart itself), a window appears or a download commences with a short description of the distribution. This description typically contains the probability density function $f(x)$, the cumulative distribution function $F(x)$, a graph of the probability density function for various parameter values, the moment generating function, and the first four moments (the

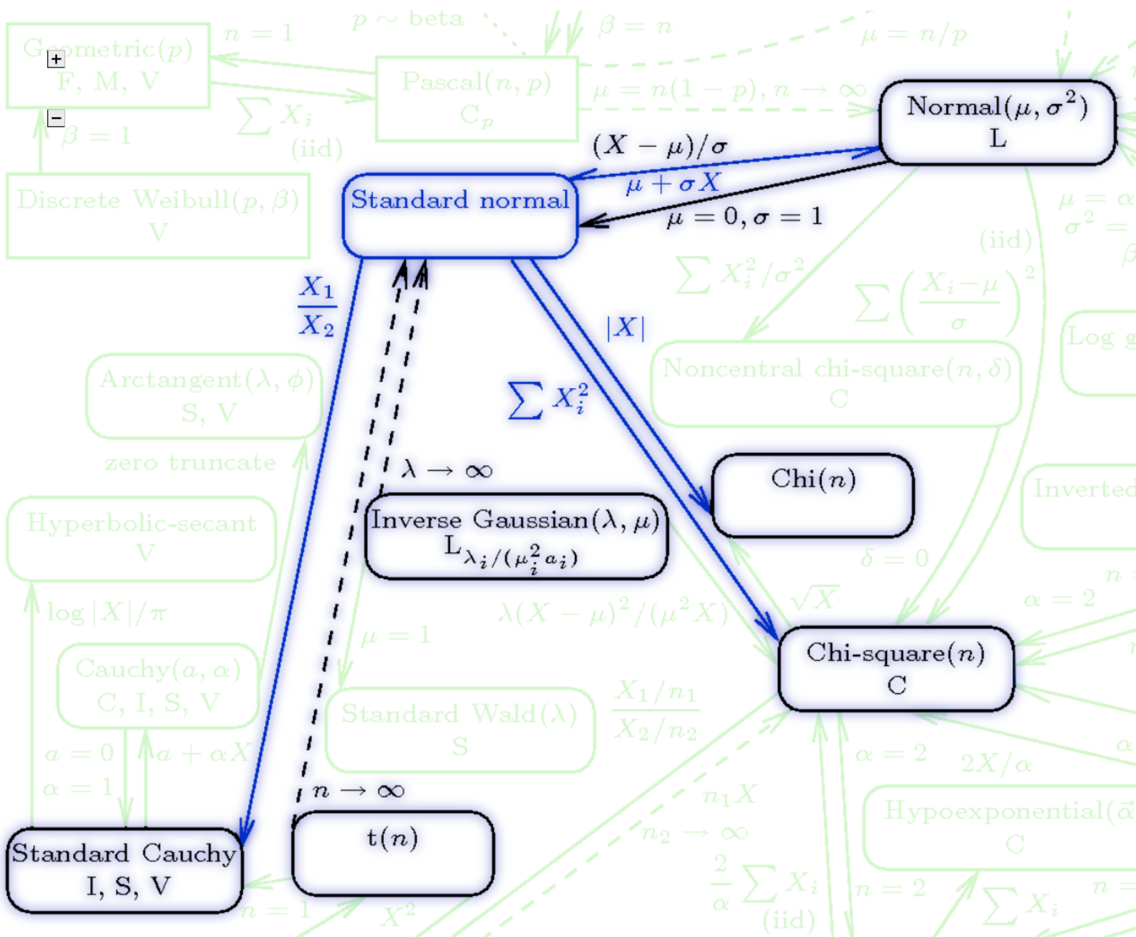

Figure 3. Hovering over the standard normal distribution.

population mean, variance, skewness, and kurtosis). When these values are not mathematically tractable, they are excluded. In addition, when the cursor hovers over a property of a distribution, the associated letter turns red. Furthermore, when the user clicks on the letter associated with a property of a distribution, a window opens with a proof of that particular property if a proof exists. For example, the URL for the minimum property of the geometric distribution is http://www.math.wm.edu~leemis/chart/UDR/PDFs/GeometricM.pdf. Similar naming conventions are used for other distributions and their properties. An arrow turns red when the cursor is placed over the arrow. When an arrow is clicked, a window opens with a proof of the associated relationship if a proof exists. In some browsers, clicking on the backarrow button will return the user to the chart. In other browsers, simply closing the window returns the user to the chart. The authors were not able to prove the relationship between the inverse Gaussian distribution and the standard normal distribution. So when the associated arrow is clicked, a web page opens that cites the result and gives an unfinished proof. An invitation to the statistics community to complete any unfinished proofs and share the proofs with the developers for posting is hereby extended.

The interactive graphic is designed for practitioners who are interested in specific information about a distribution. It is also designed for instructors who are teaching courses in a mathematical statistics sequence. Listed here are several uses of classroom applications for the interactive graphic. First, in the initial probability class, typically fewer than a dozen discrete and continuous probability distributions are introduced. Exposing the students to the interactive graphic, however briefly, will let them know that there are a large number of more obscure univariate distributions that might arise in a modeling situation, many with well-developed theory. Second, the interactive graphic is an important reminder to the students that the univariate distributions are oftentimes related to one another. Third, when a proof of one of the properties or relationships is assigned as a homework exercise, the website can be used by students to see some of the techniques that are used to prove the relationships in the interactive graphic. Certain relationships between distributions were not included in the interactive graphic to keep the network of distributions planar. These "missing" proofs remain as viable homework exercises for students using the interactive graphic. The relationships can be between two distributions (e.g., the floor of an exponential random variable has the geometric distribution), or combining two distributions to arrive at a third distribution (e.g., the ratio of a standard normal random variable to the square root of an independent chi-square random variable with n degrees of freedom divided by n is a t random variable with n degrees of freedom).

In Section 2 of this article we discuss the distribution properties that the interactive graphic includes. In Section 3, we discuss the relationships between distributions that the interactive graphic includes. Section 4 contains conclusions.

2. Discussion of Properties

In this section we discuss a number of distribution properties. These properties apply to a single distribution, and are denoted with a capital letter in the online interactive graphic. Clicking on a property on the chart opens up a file containing a proof of the associated property that the distribution satisfies. In our discussion, the abbreviations used in the interactive graphic are listed in parentheses after the name of the property.

- The *Convolution* property (C) means that the sum of mutually independent random variables following this distribution belongs to the same family of distributions. In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent random variables from a given distribution family with this property, then $\sum_{i=1}^{n} X_i$ belongs to the same distribution family.

- The *Forgetfulness* property (F), also known as the *memoryless* property, means that the conditional distribution of the random variable following this distribution is identical to the unconditional distribution. In other words, a random variable X that follows a distribution with the forgetfulness property satisfies $P(X > t + s | X > t) = P(X > s)$ for all values of s and t. The term forgetfulness is used in place of the more common "memoryless" designation because the letter M was taken by the minimum property.

- The *Inverse* property (I) means that the inverse (reciprocal) of a random variable following this distribution family belongs to the same distribution family. In other words, if X belongs to a given distribution family with this property, then $1/X$ belongs to the same distribution family.

- The *Linear Combination* property (L) means that a linear combination of mutually independent random variables following this distribution belongs to the same family of distributions. In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent random variables from a given distribution family with this property, and $a_1, a_2, \ldots, a_n$ are real-valued constants, then $\sum_{i=1}^{n} a_i X_i$ belongs to the same distribution family.

- The *Minimum* property (M) means that the minimum of n mutually independent random variables following this distribution belongs to the same family of distributions. In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent random variables following a given distribution family with this property, then $X_{(1)}$ (the first order statistic) belongs to the same distribution family.

- The *Product* property (P) means that the product of mutually independent random variables from this distribution belongs to the same family of distributions. In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent random variables from a given distribution family with this property, then $\prod_{i=1}^{n} X_i$ belongs to the same distribution family.

- The *Residual* property (R) means that the conditional distribution of a random variable from this distribution family that is left-truncated at a value in its support belongs to the same distribution family.

- The *Scaling* property (S) means that a random variable from this distribution family that is multiplied by a positive, real-valued constant belongs to the same distribution family. In other words, if X is a random variable following this distribution family, and a is a positive, real-valued constant, then aX belongs to the same distribution family.

- The *Variate Generation* property (V) means that the inverse of the cumulative distribution function for this distribution family can be obtained in closed form. The inverse cumulative distribution function can be used in a simple and fast algorithm to generate random variates from this distribution family.

- The *Maximum* property (X) means that the maximum of mutually independent random variables from this distribution family belongs to the same distribution family. In other words, if $X_1, X_2, \ldots, X_n$ are mutually independent random variables from a given distribution family with this property, then $X_{(n)}$ (the nth order statistic) belongs to the same distribution family.

Notice that property L implies properties C and S, and property F implies property R. These implications are listed in the legend in the lower-left-hand corner of the interactive graphic. We also note that certain properties only hold under restricted conditions. These properties are denoted on the chart with subscripts. For example, the binomial distribution is marked with C_p to show that the binomial distribution satisfies the *Convolution* property only if p is fixed.

3. Discussion of Relationships

There are many relationships between distributions. Some of the distributions that are included in the interactive graphic are generalizations of other distributions. For example, the exponential and chi-square distributions are special cases of the gamma distribution. The Erlang distribution is the sum of mutually independent and identically distributed exponential random variables. The chi-square distribution is the sum of squares of mutually independent standard normal random variables. The gamma distribution approaches the normal distribution as its shape parameter goes to infinity. We will now discuss each different type of relationship in more detail.

3.1 Special Cases

Many distributions are simply special cases of others. One well-known example is the standard normal distribution, which is a special case of the normal distribution where $\mu = 0$ and $\sigma^2 = 1$. In this example, the values of the parameters of the distribution are fully specified. There are other examples where only some of the parameters need to be specified. The exponential(α) distribution is a special case of the Weibull(α, β) distribution in which $\beta = 1$. In this case, α can remain unspecified.

3.2 Transformations

Other distributions are created through transformations. One way of distinguishing a transformation from a special case (since both are depicted by a solid arrow) is to recognize that a capital X will appear in the label by the arrow for a transformation. One of the most common transformations is to sum random variables. The binomial distribution with parameters n and p, for example, is the sum of n mutually independent Bernoulli(p) random variables. The Pascal distribution with parameters n and p is the sum of n mutually independent geometric(p) random variables. In situations in which a sum is involved, the relationship between two distributions can be adequately described either with a transformation or a special case. For example, an Erlang(α, n) random variable is the sum of n mutually independent exponential(α) random variables. The exponential(α) distribution is a special case of the Erlang(α, n) distribution in which $n = 1$.

There are other transformations that are used to define random variables that do not involve a sum. For example, a chi random variable is the square root of a chi-square random variable. A Rayleigh random variable is defined as the square of an exponential random variable. Transformations like these, which are expressed as one-to-one functions, can be inverted.

3.3 Limiting Distributions

A third type of relationship is a limiting distribution. This occurs when one distribution converges asymptotically to another distribution as one or more of the parameters approaches a limiting value. One important example is the gamma distribution, which converges to the normal distribution as the shape parameter β goes to infinity. This means that the gamma distribution can be used as a model when a normal distribution does not quite fit, or when an approximately normal distribution is needed, but only on positive support. Another important example is the t distribution, which approaches the standard normal distribution as the degrees of freedom, n, goes to infinity, which is widely used in statistical hypothesis tests.

Many noncentral models have a limiting distribution as the noncentrality parameter approaches 0. For example, the noncentral F distribution converges to the F distribution as δ approaches 0. Finally, limiting distributions can give us insight into the behavior of the sum of random variables as the sample size becomes very large. The binomial, as the sum of mutually independent Bernoulli random variables, and the Erlang, as the sum of mutually independent exponential random variables, are examples of this. These both approach a normal distribution as $n \to \infty$ by the central limit theorem.

3.4 Bayesian Models

A fourth type of relationship that is included on the chart is a stochastic parameter model. The familiar binomial model has parameters n and p, where n is a positive integer and $0 < p < 1$. Both parameters are assumed to be constants. But what of the case when one or both of the parameters are random variables? The dotted arrow that connects the binomial distribution to the beta–binomial distribution illustrates the case of allowing the parameter p to be a random variable having the beta distribution. The beta–Pascal, gamma–Poisson, and beta–binomial illustrate this particular type of relationship.

4. Conclusions

The figure presented by Leemis and McQueston (2008) is a helpful tool for students and instructors in the study of univariate probability distributions. It presents distributions simultaneously, as opposed to one at a time. It highlights how the distributions are related to each other, how distributions share many important properties, and how distributions are formed from other distributions. This figure is now available online at the URL www.math.wm.edu∼leemis/chart/UDR/UDR.html as an interactive graphic. The on-line chart is more useful than the figure because it allows a user to click on distributions, properties, and relationships to display additional information, such as a proof, a graph, or various moments associated with a distribution.

Acknowledgments

The authors thank the students in the Computational Probability class at The College of William & Mary, namely, Ryan Carpenter, Erin Catlett, Danny Cogut, Chad Conrad, Laura Decena, Evan Elsaesser, Luke Godwin–Jones, Matt Goldman, Matt Hanson, Yan Hao, Billy Kaczynski, James McCormack, Lee McDaniel, Lauren Merrill, Erik Olsen, Scott Percic, Evan Saltzman, Chris Thoma, Erik Vargo, Jeff Yang, Vincent Yannello, and Hang Yu for their help with the materials associated with the on-line version of the univariate distribution relationship chart. The authors thank Barry Nelson and Shane Henderson for helpful comments concerning the website. The authors also thank the editor, John Gabrosek, and two anonymous referees for their careful reading of the paper and their helpful suggestions on both the paper and the website. This research is partially supported by an NSF CSUMS grant DMS–0703532 at The College of William & Mary.

References

Leemis, L., and McQueston, J. (2008), "Univariate Distribution Relationships," *The American Statistician*, 62, 45–53.

Song, W.T., and Chen, Y. (2011), "Eighty Univariate Distributions and their Relationships Displayed in a Matrix Format," *IEEE Transactions on Automatic Control*, 56, 1979–1984.

Lawrence M. Leemis
The College of William & Mary
Department of Mathematics
P.O. Box 8795
Williamsburg, VA 23187-8795
leemis@math.wm.edu

Daniel J. Luckett
The College of William & Mary
Department of Mathematics
P.O. Box 8795
Williamsburg, VA 23187-8795

Austin G. Powell
The College of William & Mary
Department of Mathematics
P.O. Box 8795
Williamsburg, VA 23187-8795

Peter E. Vermeer
The College of William & Mary
Department of Mathematics
P.O. Box 8795
Williamsburg, VA 23187-8795
