Continuous Probability Distributions

Dr Richard Kenderdine

www.kenderdinemathstutoring.com.au

The probability of each possible outcome occurring in a discrete probability distribution can be calculated using a specific formula. In the continuous case, where there are an uncountable number of outcomes, it is impossible to assign a probability to each possible outcome such that the sum of the probabilities equals 1. We avoid this difficulty by calculating the probability of an outcome of the random variable y lying between two values, $P(a \leq y \leq b)$. In this case we are considering an area and therefore our thoughts turn to integration to find the area under a curve.

Probability Density Function (PDF)

A Probability Density Function, $f(y)$, is a non-negative function (probabilities cannot be negative) that integrates to 1. The function is non-zero in an interval [a, b] and zero elsewhere and hence we have

$$\int_a^b f(y)\,dy = 1$$

Cumulative Distribution Function (CDF)

The Cumulative Distribution Function, $F(y)$, is defined in terms of an integral

$$F(y) = \int_a^y f(t)\,dt$$

and therefore calculates the area under the PDF from the lower limit a to y. Note that $F(y)$ is a function of y and is used to find probabilities. For example,

$$P(c \leq y \leq d) = F(d) - F(c)$$

The CDF is often just referred to as the *distribution function*.

Two properties that $F(y)$ must satisfy are $F(a) = 0$ and $F(b) = 1$ when $f(y)$ is non-zero in $[a, b]$ and zero elsewhere.

Statistical Properties

The mode: if the PDF is an increasing function over $[a, b]$ then the mode is b while if it is a decreasing function the mode is a. Otherwise, it is more common for a PDF to have a maximum turning point and therefore we need to find the value of y for which $f'(y) = 0$ and $f''(y) < 0$.

The median: the k^{th} –percentile of a distribution is found from $\frac{k}{100} = F(y)$ and hence $y = F^{-1}\left(\frac{k}{100}\right)$ where F^{-1} is the inverse of the CDF. The median is therefore found from $y = F^{-1}(0.5)$. Sometimes the inverse function an be found algebraically, in other situations the solution must be found numerically.

The mean (Expected Value): replacing summation in the discrete case is the integral

$$E(Y) = \int_a^b y f(y)\, dy$$

The variance: using the relationship $Var(Y) = E(Y^2) - [E(Y)]^2$ we need

$$E(Y^2) = \int_a^b y^2 f(y)\, dy$$

The Uniform Distribution

The simplest continuous distribution is the Uniform Distribution where the probability is constant for all values of the random variable in the domain, that is

$$f(y) = c \quad \text{for} \quad a \le y \le b$$

Since the area under the curve is a rectangle with width $b - a$ and height c we have $c(b - a) = 1$ and hence $c = \frac{1}{b-a}$. Thus the PDF is $f(y) = \frac{1}{b-a}$.

The mean is then $\int_a^b y\left(\frac{1}{b-a}\right) dy = \frac{1}{b-a}\int_a^b y\, dy = \frac{a+b}{2}$

This is logical as it is the average of the endpoints.

The variance is $\frac{1}{b-a}\int_a^b y^2\, dy - \left(\frac{a+b}{2}\right)^2 = \left(\frac{b-a}{12}\right)^2$

The moment generating function is $\frac{1}{b-a}\int_a^b e^{ty}\, dy = \frac{e^{tb}-e^{ta}}{t(b-a)}$

The Normal Distribution

This most commonly used continuous distribution has PDF

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right) \quad \text{for} \quad -\infty < y < \infty$$

with mean μ, variance σ^2 and moment generating function $\exp\left(\mu t + \frac{t^2\sigma^2}{2}\right)$.

The Gamma Distribution

Some random variables are always non-negative and right skewed. For example, time between failure for engines and other manufactured items, time between arrivals at a checkout and repair time for breakdowns. In such cases the distribution of the data can be modelled by gamma density function

$$f(y) = \frac{y^{\alpha-1}e^{-\frac{y}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} \quad \text{for} \quad 0 \leq y < \infty$$

The parameter α determines the shape of the distribution while β is a scale parameter. The gamma function, $\Gamma(\alpha)$, is defined as $\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\,dy$ and is the continuous equivalent of the factorial function. Two useful properties are $\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$ and $\Gamma(n) = (n-1)!$ for integer n. The first property comes from integration by parts:

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\,dy = y^{\alpha-1}(-e^{-y})|_0^{\infty} + \int_o^{\infty} (\alpha-1)y^{\alpha-2}e^{-y}\,dy$$

$$= 0 + (\alpha-1)\int_0^{\infty} y^{(\alpha-1)-1}e^{-y}\,dy = (\alpha-1)\Gamma(\alpha-1)$$

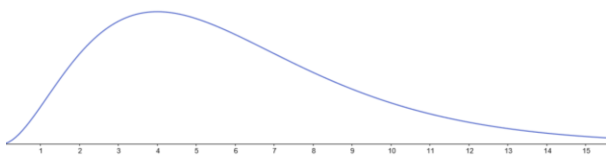
From the definition, $\Gamma(1) = \int_0^{\infty} e^{-y}\,dy = -e^{-y}|_0^{\infty} = 1 = 0!$

And $\Gamma(2) = \int_0^{\infty} ye^{-y}\,dy = -ye^{-y}|_0^{\infty} + \int_0^{\infty} e^{-y}\,dy = 1 = 1!$

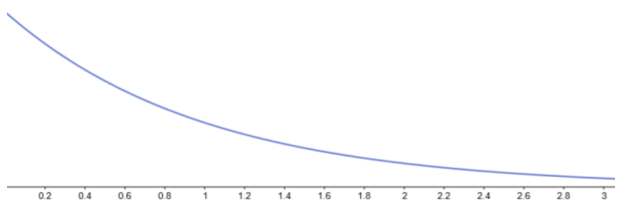
Then $\Gamma(3) = 2\Gamma(2) = 2 \times 1 = 2!$ etc.

The Exponential, Chi-Square and Weibull Distributions are related to the Gamma Distribution.

Here is a gamma density function with $\alpha = 3$ and $\beta = 2$



When both α and β equal 1 the density function is



To find the Expected Value we use the fact that $\int_0^\infty \frac{y^{\alpha-1}e^{-\frac{y}{\beta}}}{\beta^\alpha \Gamma(\alpha)} dy = 1$ and therefore $\int_0^\infty y^{\alpha-1}e^{-\frac{y}{\beta}} dy = \beta^\alpha \Gamma(\alpha)$. Then

$$E(Y) = \int_0^\infty y \frac{y^{\alpha-1}e^{-\frac{y}{\beta}}}{\beta^\alpha \Gamma(\alpha)} dy = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty y^\alpha e^{-\frac{y}{\beta}} dy$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \beta^{\alpha+1}\Gamma(\alpha+1) = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha\beta$$

Here use is made of the relationship $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

A similar approach yields

$$E(Y^2) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \left(\beta^{\alpha+2}\Gamma(\alpha+2)\right)$$

$$= \frac{\beta^2(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \alpha(\alpha+1)\beta^2$$

Then $\text{Var}(Y) = \alpha(\alpha+1)\beta^2 - (\alpha\beta)^2 = \alpha\beta^2$

Hence the Expected Value and Variance of the Gamma Distribution are $\alpha\beta$ and $\alpha\beta^2$ respectively.

The moment generating function, $m(t)$, is $(1-\beta t)^{-\alpha}$, derived in the following manner.

$$m(t) = E(e^{ty}) = \int_0^\infty e^{ty} \frac{y^{\alpha-1}e^{-\frac{y}{\beta}}}{\beta^\alpha \Gamma(\alpha)} dy = \int_0^\infty \frac{y^{\alpha-1}e^{ty-\frac{y}{\beta}}}{\beta^\alpha \Gamma(\alpha)} dy$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty y^{\alpha-1} e^{-\left(\frac{1-\beta t}{\beta}\right)y} = \frac{\left(\frac{\beta}{1-\beta t}\right)^\alpha \Gamma(\alpha)}{\beta^n \Gamma(\alpha)} = (1-\beta t)^{-\alpha}$$

Note that in the standard form the power of e is $-\frac{y}{\beta}$ hence whatever divides $-y$ is β. In $m(t)$ the coefficient of $-y$ is $\frac{1-\beta t}{\beta}$ which means division by $\frac{\beta}{1-\beta t}$.

The expressions for $E(Y)$ and $E(Y^2)$ and then found using $m'(0)$ and $m''(0)$ respectively, remembering that differentiation is in respect of the variable t.

The distribution function, $F(y)$, has to be found using integration by parts. There is a connection with the Poisson Distribution when the shape parameter, α, is an integer.

Example

Consider the Gamma Distribution with $\alpha = 3$ and $\beta = 2$. The density function is

$$f(y) = \frac{y^{\alpha-1}e^{-\frac{y}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)} = \frac{y^2 e^{-\frac{1}{2}y}}{8 \times 2} = \frac{1}{16}y^2 e^{-\frac{1}{2}y}$$

Then the distribution function is

$$F(y) = \frac{1}{16}\int_0^y t^2 e^{-\frac{1}{2}t}\,dt$$

Using integration by parts, this becomes when integrating the exponential and differentiating the power,

$$F(y) = \frac{1}{16}\left[-2t^2 e^{-\frac{1}{2}t} + 2\int 2te^{-\frac{1}{2}t}\,dt\right]$$

Now the integral inside the brackets becomes

$$\int te^{-\frac{1}{2}t}\,dt = -4te^{-\frac{1}{2}t} + 4\int e^{-\frac{1}{2}t}\,dt$$

with $\int e^{-\frac{l}{2}t}\,dt = -2e^{-\frac{1}{2}t}$. Putting all this together, we have

$$F(y) = -\frac{1}{8}e^{-\frac{1}{2}t}[t^2 + 4t + 8]_0^y = 1 - \frac{1}{8}e^{-\frac{1}{2}y}(y^2 + 4y + 8)$$

Then, for example, $F(8) = 1 - 13e^{-4} \approx 0.7619$

Since $\alpha(3)$ is an integer we can use the Poisson approach. The distribution function is

$$F(y) = 1 - \sum_{k=0}^{\alpha-1} e^{-\frac{y}{\beta}}\frac{y^k}{\beta^k k!}$$

Then, using $y = 8, \alpha = 3, \beta = 2$ we have, as previously,

$$F(8) = 1 - e^{-4}\left(1 + \frac{8}{2 \times 1} + \frac{8^2}{2^2 \times 2}\right) = 1 - 13e^{-4}$$

Related distributions

1) Exponential, when $\alpha = 1$
2) Chi-square, when $\alpha = \frac{v}{2}, \beta = 2$
3) Weibull, $f(y) = \frac{m}{\alpha}y^{m-1}e^{-\frac{y^m}{\alpha}}$ for $y > 0$

The Exponential Distribution

As stated, the Exponential Distribution is a particular case of the Gamma Distribution when $\alpha = 1$. Hence the PDF is $f(y) = \frac{1}{\beta}e^{-\frac{y}{\beta}}$ for $y > 0$, the Expected Value is β, the Variance is β^2 and the MGF is $(1-\beta t)^{-1}$.

The Cumulative Distribution Function is

$$F(y) = \frac{1}{\beta}\int_0^y e^{-\frac{t}{\beta}}\,dt = \left[-e^{-\frac{t}{\beta}}\right]_0^y = 1 - e^{-\frac{y}{\beta}}$$

Hence $P(Y > y) = e^{-\frac{y}{\beta}}$ and, consequently,

$$P(Y > a + y | Y > a) = \frac{P(Y > a + y)}{P(Y > a)} = \frac{e^{\frac{-(a+y)}{\beta}}}{e^{-\frac{a}{\beta}}} = e^{-\frac{y}{\beta}}$$

This is the memoryless property of the Exponential Distribution. For example, if we are modelling a waiting line, the probability of waiting one further minute, having been in the queue for five minutes, is the same as joining the queue and waiting one minute.

Further properties of the Exponential Distribution are

i) The sum of n independent and identically distributed exponential random variables with mean β has a Gamma distribution with parameters n and β. Proof of this requires knowledge of joint distributions, not yet covered.

ii) Given two independent exponential random variables Y_1 and Y_2 with parameters β_1 and β_2 respectively, we can determine the probability that Y_1 is less than Y_2.

$$\int_0^\infty \int_0^{x_2} \frac{1}{\beta_1} e^{-\frac{x_1}{\beta_1}}\,dx_1 \frac{1}{\beta_2} e^{-\frac{x_2}{\beta_2}}\,dx_2$$

$$= \int_0^\infty \left[-e^{-\frac{x_1}{\beta_1}}\right]_0^{x_2} \frac{1}{\beta_2} e^{-\frac{x_2}{\beta_2}}\,dx_2 = \int_0^\infty \left(1 - e^{-\frac{x_2}{\beta_1}}\right)\frac{1}{\beta_2} e^{-\frac{x_2}{\beta_2}}\,dx_2$$

$$= \frac{1}{\beta_2}\int_0^\infty \left(e^{-\frac{x_2}{\beta_2}} - e^{-\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)x_2}\right)dx_2$$

$$= \frac{1}{\beta_2}\left[-\beta_2 e^{-\frac{x_2}{\beta_2}} + \frac{1}{\frac{1}{\beta_1}+\frac{1}{\beta_2}} e^{-\left(\frac{1}{\beta_1}+\frac{1}{\beta_2}\right)x_2}\right]_0^\infty$$

$$= \frac{1}{\beta_2}\left[-\beta_2 + \frac{\beta_1\beta_2}{\beta_1+\beta_2}\right] = -1 + \frac{\beta_1}{\beta_1+\beta_2} = \frac{\beta_2}{\beta_1+\beta_2}$$

For example, suppose a machine has two components, *A* and *B*, with expected lifetimes of 1000 hours and 500 hours respectively. The probability that component *A* fails before component *B* is $\frac{500}{1000+500} = \frac{1}{3}$.

iii) Hazard rate (failure rate) function is defined as

$$r(t) = \frac{f(t)}{1 - F(t)}$$

For the Exponential Distribution we have

$$r(t) = \frac{\frac{1}{\beta}e^{-\frac{u}{\beta}}}{1-\left(1-e^{-\frac{y}{\beta}}\right)} = \frac{1}{\beta}$$

Hence the failure rate is constant and the reciprocal of the mean.

If the arrivals to a queue is modelled by a Poisson distribution then the time between arrivals is Exponential.

The Beta Distribution

The Beta Distribution is based on the Beta function $B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy$ that equals $\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Since a valid PDF equals 1, we have the Beta Distribution PDF

$$f(y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\,dy$$

Note that the domain is [0,1] (any other domain will need to be transformed before the Beta Distribution can be used), contrasting with Gamma-based distributions with infinite domains, and therefore suitable for use with proportions. As the integrand is similar to the Binomial Distribution the Beta Distribution is often used as a prior in Bayesian statistics.

The Expected Value is easily found.

Since $\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ then

$$E(Y) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 y \cdot y^{\alpha-1}(1-y)^{\beta-1}\,dy$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1 y^{\alpha}(1-y)^{\beta-1}\,dy$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)}$$

$$= \frac{\Gamma(\alpha+1)}{\Gamma(\alpha)} \cdot \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+1)} = \frac{\alpha}{\alpha+\beta} \quad \text{(using } \Gamma(\alpha+1) = \alpha\Gamma(\alpha)\text{)}$$

Similarly,

$$E(Y^2) = \frac{\Gamma(\alpha+\beta)}{\Gamma|\alpha|\Gamma(\beta)} \int_0^1 y^{\alpha+1}(1-y)^{\beta-1}\,dy$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma|\alpha|\Gamma(\beta)} \; \frac{\Gamma(\alpha+2)\Gamma(\beta)}{\Gamma(\alpha+\beta+2)}$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma|\alpha|\Gamma(\beta)} \; \frac{\alpha(\alpha+1)\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)(\alpha+\beta+1)\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

Then

$$\text{Var}(Y) = E(Y^2) - [E(Y)]^2$$

$$= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2$$

$$= \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

A closed form for the CDF does not exist. When α and β are integers the CDF integral can be evaluated either by integration by parts or expanding and integrating each term. For non-integral α, β numerical integration can be used. Most users will have access to software that can provide the solution.

There is no closed form MGF for the Beta Distribution.

The Lognormal Distribution

A random variable Y that can only take positive values has a Lognormal distribution if $X = Log\ (Y)$ is normal. Equivalently, $Y = e^X$ for a normally distributed variable X. The distribution is used in both biological and physical sciences to model right-skewed data.

To obtain the PDF of Y we can use the Method of Transformations. If we have a random variable X with density $f_X(x)$ and an increasing function $h(x)$ then we can find the density of $Y = h(X)$.

We have

$$P(Y \leq y) = P(h(X) \leq y) = P\left(h^{-1}\big(h(X)\big) \leq h^{-1}(y)\right) = P\big(X \leq h^{-l}(y)\big)$$

This means, in terms of the CDFs,

$$F_Y(y) = F_X\big(h^{-1}(y)\big)$$

To obtain the density function for Y we differentiate

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\big(h^{-1}(y)\big)$$

$$= \frac{d}{d\big(h^{-1}(y)\big)} F_X\big(h^{-1}(y)\big) \frac{d\big(h^{-1}(y)\big)}{dy}$$

$$= f_X\big(h^{-1}(y)\big) \frac{d\big(h^{-1}(y)\big)}{dy}$$

Thus, to obtain the density of Y we substitute the inverse function into the density of X and multiply by the derivative of the inverse function.

For the Lognormal example we have $y = h(x) = e^x$ and $x = h^{-1}(y) = \log(y)$. Note that this function satisfies the conditions that $y > 0$ and the function is always increasing.

The Normal PDF is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \text{for } -\infty < x < \infty$$

Hence the Lognormal PDF is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right)\left(\frac{1}{y}\right) \quad = \quad \frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\log(y)-\mu}{\sigma}\right)^2\right) \quad \text{for } y > 0$$

We can obtain the mean and variance using integration but it is more efficient to use the Method of Moments. The MGF for the Normal Distribution is

$$E(e^{tx}) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \quad \text{where} \quad x \sim N(\mu, \sigma^2)$$

Now for Lognormal $Y = e^X$, hence we can find the expected values of the powers of Y:

$$E(Y) = E(e^X) = e^{\mu + \frac{1}{2}\sigma^2}$$

$$E(Y^2) = E(e^{2X}) = e^{2(\mu + \sigma^2)}$$

$$\text{var}(Y) = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}\big(e^{\sigma^2} - 1\big)$$

The Standard Normal Distribution can be used to obtain quantiles for the Lognormal Distribution, $F(y)$:

$$F(y) = \Phi\left(\frac{\log(y) - \mu}{\sigma}\right)$$

Thus

$$p = \Phi\left(\frac{\log(y) - \mu}{\sigma}\right)$$

$$\emptyset^{-1}(p) = \frac{\log(y) - \mu}{\sigma}$$

$$y = e^{\mu + \sigma\, \Phi^{-1}(p)}$$

As an example, consider the case where $\mu = 3,\ \sigma = 0.5$ and $p = 0.25$ (the lower quartile). The Standard Normal lower quartile is z= -0.67449, and therefore the lower quartile for this Lognormal distribution is

$$y = e^{3+0.5(-0.67449)} = 14.3357$$

The mode is found by setting the derivative of the density to 0. Using the product rule, the derivative is, upon simplification,

$$f'(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log y - \mu}{\sigma}\right)^2} \left[-\frac{1}{y^2}\left(1 + \frac{\log y - \mu}{\sigma^2}\right)\right]$$

Then

$$1 + \frac{\log y - \mu}{\sigma^2} = 0 \;\Rightarrow\; y = e^{\mu - \sigma^2}$$

Thus the mode occurs at $y = e^{\mu-\sigma^2}$ with corresponding function value $\frac{1}{\sqrt{2\pi}\sigma} e^{\frac{1}{2}\sigma^2 - \mu}$

Using the example above, the mode of the distribution occurs at $y = e^{3-0.5^2} = 15.642$ with function value 0.045.

This plot shows the lower quartile at 14.3357: