The Beta Distribution

1 Introduction

One particularly useful random variable is the Beta Distribution, which model proportions relatively well, as it only takes values between 0 and 1, and which also retains the uniform distribution as a special case.

2 Density Function

A random variable Y has a *beta probability distribution* if and only if it has density function

$$f_Y(y) = \begin{cases} \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}, & 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases} \qquad \alpha, \beta > 0 \tag{1}$$

where the function B is defined

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

Varying α and β can lead to a vast array of different shapes.

3 Key Statistics

By a few easy manipulations, it can be shown that the beta distribution has mean and variance

$$\mu = \frac{\alpha}{\alpha+\beta}, \qquad \sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{2}$$

4 The Beta Function

The Beta Function can be defined as

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

This happens to have several integral representations, two of which we list:

$$B(\alpha, \beta) = \int_0^\infty t^{\alpha-1}(1+t)^{-(\alpha+\beta)}\, dt \tag{3}$$

$$B(\alpha, \beta) = \int_0^1 t^{x-1}(1-t)^{y-1} dt \tag{4}$$

The proof[1] that Equation 3 does indeed equal that of 3 requires a bit of clever manipulation, while the Equation 4 uses Equation 3 and makes the substitution

$$s = \frac{t}{1+t} = 1 - \frac{1}{1+t}, \qquad \Rightarrow t = \frac{s}{1-s}$$

[1] All proofs and further information can be found on the Statlect.com website: `http://www.statlect.com/subon2/betfun1.htm`.