Lesson 3.6

The Beta Distribution

Learning Outcomes

At the end of the lesson, students must be able to

1. Describe the key properties of a random variable having a beta distribution, such as the mean, variance, and moment generating function, and
2. Compute probabilities associated with random variables having a beta distribution.

Introduction

The beta distribution is useful for modeling random probabilities and proportions, particularly in the context of Bayesian analysis.

Definition:

A random variable Y is said to have a beta distribution with parameters $\alpha > 0$ and $\beta > 0$ if its density function is given by

$$f_Y(y) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}, \, 0 < y < 1 \\ 0, \text{ elsewhere} \end{cases}$$

We write $Y \sim Beta(\alpha, \beta)$.

Remarks:

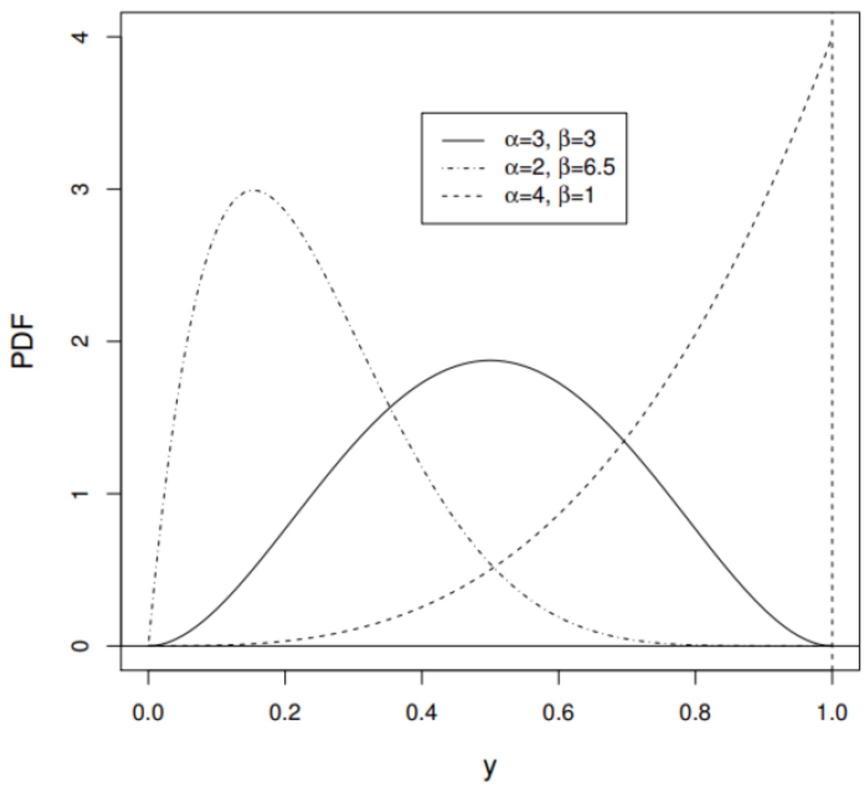
1. The nonzero part of the $Beta(\alpha, \beta)$ PDF consists of two parts:
   - the kernel: $y^{\alpha-1}(1-y)^{\beta-1}$

- the constant: $\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$

2. The kernel is the "guts" of the formula, while the constant out front is simply the "right quantity" that makes $f_Y(y)$ a valid PDF; i.e., the constant that makes $f_Y(y)$ integrate to 1.

3. The graphs of beta density functions assume widely differing shapes for various values of the two parameters α and β . In fact, the PDF of $Beta(\alpha, \beta)$ is very flexible, that is, the PDF can assume many shapes over the interval $(0, 1)$. For example,

   - If $\alpha = \beta$, then $f_Y(y)$ is symmetric at $y = \frac{1}{2}$. In fact, if $\alpha = \beta = 1$, then $Y \sim U(0, 1)$.
   - If $\alpha > \beta$, then $f_Y(y)$ is left skewed.
   - If $\alpha < \beta$, then $f_Y(y)$ is right skewed.



Example 1:

A gasoline wholesale distributor has bulk storage tanks that hold fixed supplies and are filled every Monday. Of interest to the wholesaler is the proportion of this supply that is sold during the week. Over many weeks of observation, the distributor found that this proportion could be modeled by a beta distribution with $\alpha = 4$ and $\beta = 2$. Find the probability that the wholesaler will sell at least 90% of her stock in a given week.

SOLUTION:

Let Y denote the proportion of the supply sold during the week, then $Y \sim Beta(4, 2)$ and

$$f_Y(y) = \begin{cases} 20y^3(1-y),\ 0 < y < 1 \\ 0,\ \text{elsewhere} \end{cases}$$

Thus,

$$P(Y \geq 0.90) = \int_{0.90}^{1} 20y^3(1-y)\,dy \approx 0.081$$

Theorem:

If $Y \sim Beta(\alpha, \beta)$, then

a. $E(Y) = \frac{\alpha}{\alpha+\beta}$

b. $V(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$

PROOF:

a. Let $Y \sim Beta(\alpha, \beta)$, then

$$\begin{aligned} E(Y) &= \int_0^1 y f_Y(y)\,dy \\ &= \int_0^1 y \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1}(1-y)^{\beta-1}\,dy \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 y^{\alpha}(1-y)^{\beta-1}\,dy \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 \frac{\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}}{\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}} y^{(\alpha+1)-1}(1-y)^{\beta-1}\,dy \\ &= \frac{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}}{\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}} \int_0^1 \frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)} y^{(\alpha+1)-1}(1-y)^{\beta-1}\,dy \end{aligned}$$

$$
\begin{aligned}
E(Y) &= \frac{\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}}{\frac{\Gamma(\alpha+1+\beta)}{\Gamma(\alpha+1)\Gamma(\beta)}} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)} \\
&= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\alpha\Gamma(\alpha)\Gamma(\beta)}{(\alpha+\beta)\Gamma(\alpha+\beta)} \\
&= \frac{\alpha}{\alpha+\beta}
\end{aligned}
$$

b. The derivation of the variance is left as an exercise!

Example 2:

The percentage of LET passers is a random variable Y with density function

$$
f_Y(y) = \begin{cases} 12y^2(1-y),\ 0 < y < 1 \\ 0, \text{elsewhere} \end{cases}
$$

a) Determine the probability that less than 70% of next years batch of education graduates will pass the LET.

b) Find the mean and variance of the proportion of LET passers.

SOLUTION: Left as a classroom exercise!

Example 3:

During an eight-hour shift, the proportion of time (Y) that a sheet-metal stamping machine is down for maintenance or repairs has a beta distribution with $\alpha = 1$ and $\beta = 2$. The cost of this downtime, due to lost production and cost of maintenance and repair, is given by $C = 10 + 20Y + 4Y^2$ Find the mean and variance of C.

SOLUTION: Left as a classroom exercise!