Beta Distribution - Section 4.7

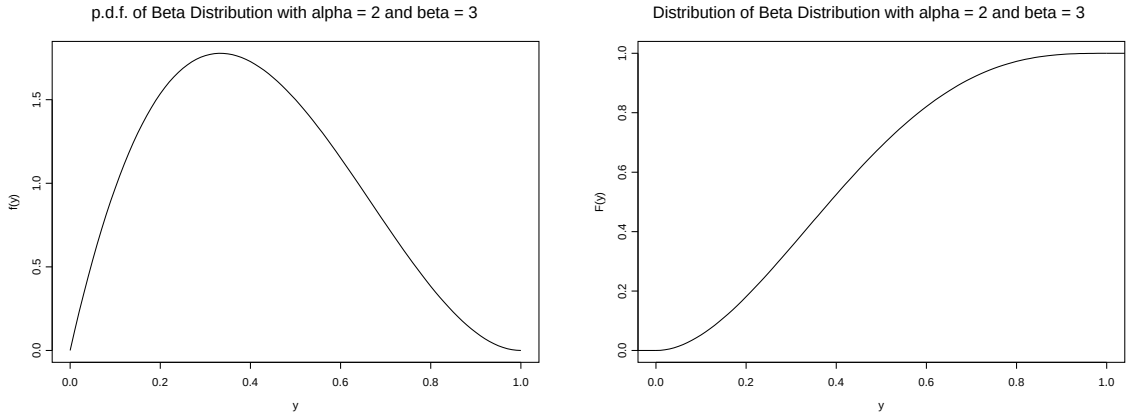
The beta distribution is used to model continuous data with values between 0 and 1. Typically, beta distributions are used to model proportions. The characteristics of the beta distribution are

- The parameters of the beta distribution are $\alpha > 0$ and $\beta > 0$. Unlike the gamma distribution, these parameters do not have special names. With different values of these parameters, the shape of the probability density curve can change significantly.
- The probability density curve for a beta distribution is

$$f(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)} \qquad 0 \le y \le 1$$

  where $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
- Under most circumstances, the distribution function $F(y)$ for the beta distribution has no closed form solution. Therefore, in order to find probabilities associated with the beta distribution, we must use tables or a statistical computer package.

  Here are graphs of the probability density function and the distribution function for a beta distribution with $\alpha = 2$ and $\beta = 3$.



- The theoretical mean of the beta distribution is

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

- The variance of the beta distribution is

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

Working with the beta distribution in R.

To find the probability $P(Y \le y)$ the command in R is

```
pbeta(y,alpha,beta)
```

To find the value of y so that $P(Y \leq y) = p$ the command in R is

```
qbeta(p,alpha,beta)
```

To generate observations from a beta distribution the command in R is

```
rbeta(numobs,alpha,beta)
```

where numobs is the number of observations you would like to generate.

Problems.

1. The relative humidity Y when measured at a location, has a beta distribution with $\alpha = 4$ and $\beta = 3$. Find the probability that the relative humidity at this location will be greater than 60%.
2. The percentage of impurities per batch in a chemical product is a random variable Y with a beta distribution with $\alpha = 3$ and $\beta = 2$. A batch with more than 40% impurities cannot be sold.
   (a) What is the probability that a randomly selected batch cannot be sold due to excess impurities?
   (b) Find the mean and variance of the percentage of impurities in a randomly selected batch of chemicals.
3. What is another name for the beta distribution when $\alpha = 1$ and $\beta = 1$?
4. The weekly repair cost Y for a machine has a beta distribution with $\alpha = 1$ and $\beta = 3$. (Measurements are in hundreds of dollars). How much money should be budgeted each week for repair costs so that the actual cost Y will exceed the budgeted amount only 10% of the time?