Beta Distribution

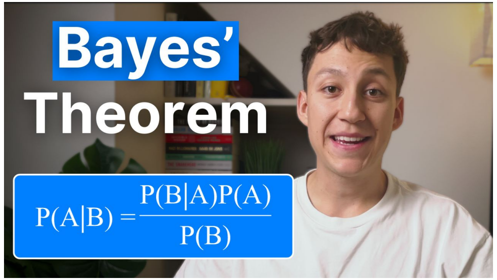
Egor Howell

Introduction

The Beta distribution is a continuous distribution that is often dubbed as the ***Probability Distribution of Probabilities***. This is because it can only take on values between 0 and 1.

It is used to infer the probability of an event when we have some information about the volumes of successes and failures.

The primary use of the Beta distribution is being the conjugate prior in Bayesian statistics to the Binomial and Bernoulli distributions.



Mathematical Detail

The probability density function looks like:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Equation generated by author in LaTeX.

Linked here is the derivation of the Beta distribution PDF.

It is parameterized by the two variables ***α*** and ***β*** and ***B(α, β)*** is the Beta function which is the normalising constant:

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

Equation generated by author in LaTeX.

Intuition

The Beta distribution numerator is very similar to the Binomial distribution PDF:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Equation generated by author in LaTeX.

$$Binomial\ PDF = \binom{n}{x} p^x (1-p)^{n-x}$$

Equation generated by author in LaTeX.

Which is the probability of achieving ***x*** successes from ***n*** events with a given probability ***p***.

For example, the probability of flipping exactly 6 heads from 10 coin flips is 0.20508.

Intuition

The Beta distribution numerator is very similar to the Binomial distribution PDF:

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Equation generated by author in LaTeX.

$$Binomial\ PDF = \binom{n}{x} p^x (1-p)^{n-x}$$

Equation generated by author in LaTeX.

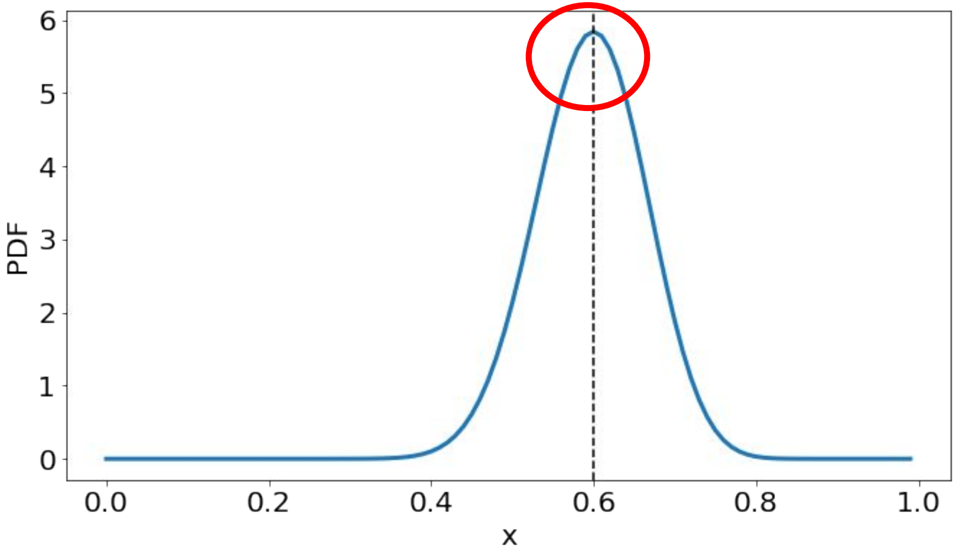
The Binomial and Beta distributions are very similar apart from one key thing:

- For the Beta distribution the probability is a random variable which we are trying to estimate.
- For the Binomial distribution the probability is a fixed parameter which we use to infer the probability of n successes.

Therefore, we can use the Beta distribution to estimate the probability of an event if we know the number of successes, α-1, and the number of failures, β -1.

Example & Plots

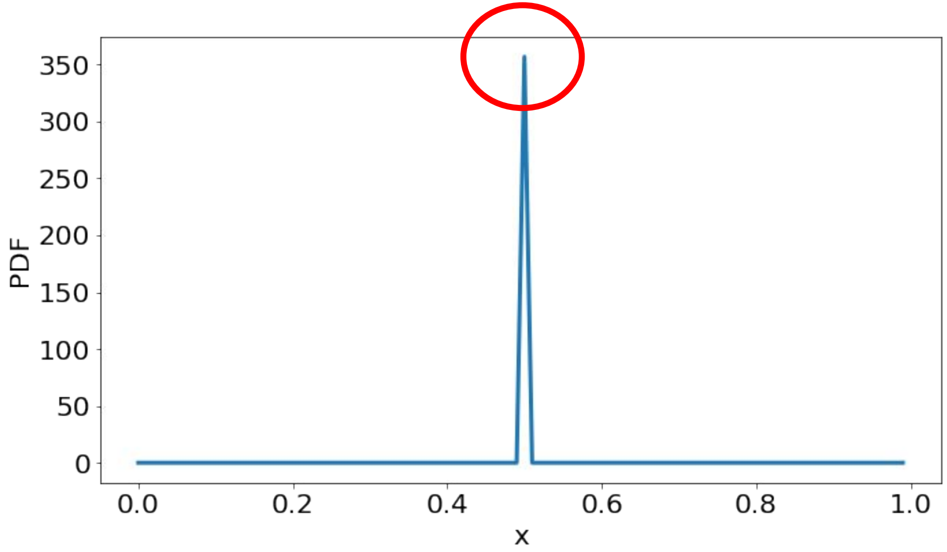
Let's assume we flip a biased coin **50** times and it lands on heads **30** (successes) times and tails **20** times (failures).




```python
# Import packages
from scipy.stats import beta as beta_dist
import matplotlib.pyplot as plt
import numpy as np

# Plot the distribution
alpha = 31
beta = 21

x = np.arange (0, 1, 0.01)
y = beta_dist.pdf(x, alpha, beta)

plt.figure(figsize=(11,6))
plt.plot(x, y, linewidth=3)
plt.xlabel('x', fontsize=20)
plt.ylabel('PDF', fontsize=20)
plt.xticks(fontsize=20)
plt.yticks(fontsize=20)
plt.axvline(0.6,  linestyle = 'dashed', color='black')
plt.show()
```

Example & Plots

Now let’s say we flip another coin and get **100,000** heads, but also **100,000** tails. The Beta distribution for this data is:




```
alpha = 100_000
beta = 100_000

x = np.arange (0, 1, 0.01)
y = beta_dist.pdf(x, alpha, beta)

plt.figure(figsize=(11,6))
plt.plot(x, y, linewidth=3)
plt.xlabel('x', fontsize=20)
plt.ylabel('PDF', fontsize=20)
plt.xticks(fontsize=20)
plt.yticks(fontsize=20)
plt.show()
```

Thanks

Beta Distribution Simply Explained

A concise and easy explanation of the Beta distribution.

Egor Howell
Published in Towards Data Science · 5 min read · Jul 22, 2022

273 1

Photo by Lucas Santos on Unsplash

@egorhowell

Newsletter

Dishing The Data