

Lab 7 and 8 – MATH 240 – Computational Statistics

Anya Suko

27 March 2025

1 Introduction of the Beta Distribution

The Beta distribution is a continuous distribution, often used to model a random variable X that ranges from 0 to 1, making it useful for modeling proportions, probabilities, and rates. It is known for being remarkably flexible with regards to its shape- it can be symmetric or left/right skewed, depending on its parameters that define its shape: $\alpha > 0$, and $\beta > 0$.

	Statistic	Value
1	Mean	0.71
2	Variance	0.03
3	Skewness	-0.60
4	Kurtosis	-0.12

When α and β were set to .5 and .5, respectively, the population level statistics were the following:

	Statistic	Value
1	Mean	0.50
2	Variance	0.12
3	Skewness	0.00
4	Kurtosis	-1.50

2 Density Function and Parameters

The beta distribution's probability density function, which gives the likelihood that a continuous random variable takes on a specific value, is given by

$$f_x(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \mathbb{I}(x \in [0, 1])$$

where the density is zero anywhere outside of the range $[0, 1]$. Since the distribution is directly impacted by its parameters (α and β), the population level characteristics are also described by those parameters.

A series of example distributions were created with different α and β values in order to demonstrate the effect that both α and β have on the distribution, and therefore the population level characteristics.

When α and β were set to 2 and 5, respectively, the population level statistics were the following:

	Statistic	Value
1	Mean	0.29
2	Variance	0.03
3	Skewness	0.60
4	Kurtosis	-0.12

When α and β were set to 5 and 5, respectively, the population level statistics were the following:

	Statistic	Value
1	Mean	0.50
2	Variance	0.02
3	Skewness	0.00
4	Kurtosis	-0.46

When α and β were set to 5 and 2, respectively, the population level statistics were the following:

Graphical representations of these distributions with α and β values can be seen in the appendix.

Upon comparing the results of the population level characteristics for the different parameter values, it can be noted that with equal α and β values, the mean is exactly one half, and the skewness is zero. However, with increases to both α and β value, the variance is smaller, and the kurtosis is also lower. When α is greater than β , when compared to equal parameters, the mean increases, and the skewness is negative (left skewed). When α is less than β , when compared to equal parameters, the mean is smaller, and the skewness is positive (right skewed). Here it is evident that the α and β parameters significantly impact the population level characteristics.

3 Moments

The mean, variance, skewness, and kurtosis are all calculated using a moment of the distribution. In order to demonstrate the effect that using moments has on these values, the characteristics from the previous example were re-calculated:

$\alpha = 2$ and $\beta = 5$

	Statistic	Value
1	Mean	0.29
2	Variance	0.03
3	Skewness	0.03
4	Kurtosis	-0.12

$\alpha = 5$ and $\beta = 5$

	Statistic	Value
1	Mean	0.50
2	Variance	0.02
3	Skewness	-0.00
4	Kurtosis	-0.46

$\alpha = 5$ and $\beta = 2$

	Statistic	Value
1	Mean	0.71
2	Variance	0.03
3	Skewness	-0.03
4	Kurtosis	-0.12

$\alpha = .5$ and $\beta = .5$

	Statistic	Value
1	Mean	0.50
2	Variance	0.12
3	Skewness	-0.00
4	Kurtosis	-1.50

Here, it is evident that both avenues of calculating these statistics output nearly the same values. Since the formula for the moment is derived from the PDF function, it makes sense why these calculations would result in such similar values.

4 Population Distribution Approximation

The goal of summarizing data is to approximate the underlying population distribution. This section examines how graphical and numerical summaries reflect the actual population by generating random samples from known distributions and comparing sample-based estimates to theoretical values.

$\alpha = 2$ and $\beta = 5$

	Mean	Variance	Skewness	Excess_Kurtosis
1	0.29	0.03	0.03	-0.12

$\alpha = 5$ and $\beta = 5$

	Mean	Variance	Skewness	Excess_Kurtosis
1	0.72	0.02	-0.03	-0.12

$\alpha = 5$ and $\beta = 2$

	Mean	Variance	Skewness	Excess_Kurtosis
1	0.50	0.02	-0.00	-0.46

$\alpha = .5$ and $\beta = .5$

	Mean	Variance	Skewness	Excess_Kurtosis
1	0.52	0.12	-0.00	-1.50

Here, the difference in statistic calculations from the previous section examples comes from the fact that the data are samples, and therefore there is some variation. In the graphical representation of the repeated sampling, it can be seen that the outcome follows similar patterns to the actual population distribution, but that there are some places which are not exact, which is a result of the randomness of the sampling. While some variation occurs due to sampling, these summaries provide useful approximations for key distribution characteristics.

5 Sample Size and the Beta Distribution

This section demonstrates how cumulative numerical summaries—mean, variance, skewness, and excess kurtosis—change as more data points are added (sample size increases). The goal is to see how these statistics approach their true population values.

ADD GRAPHS!!!

After careful examination of all four summary statistic graphs for the $\alpha = 2$ and $\beta = 5$ beta distribution, it can be seen that as sample size increases, the summary statistics each converges towards the true value for the population for that summary statistic.

6 Modeling Variation

7 Example- Death Rates Data

???

Bibliography: Note that when you add citations to your bib.bib file *and* you cite them in your document, the bibliography section will automatically populate here.

8 Appendix

If you have anything extra, you can add it here in the appendix. This can include images or tables that don't work well in the two-page setup, code snippets you might want to share, etc.