16.

Intro

The purpose of this exercise is to calculate the summary statistics of the Beta distribution. As we are going to see in the next chapter, the Beta distribution is very important when perfoming Bayesian analysis, so a reasonable understanding of its property is required. First of all, we should note that since the Beta distribution has two parameters a and b, the summary statistics should be functions of them. Second, it is reasonable to assume that the mean and the variance should always exist. On the other hand, it is possible that the mode doesn't exist. Let's see if this suspect holds true.

Solution

Mean

$$E[X] = \int_0^1 x \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} dx \tag{1}$$

To solve this problem, we will integrate by parts.

$$\begin{aligned} E[X] = \frac{1}{B(a,b)} \left[\frac{x^a(-1)(1-x)^b}{b} \Big|_0^1 - \int_0^1 a x^{a-1}(-1)\frac{(1-x)^b}{b} dx \right] = \\ \frac{1}{B(a,b)} \int_0^1 \frac{a}{b} x^{a-1}(1-x)^{b-1}(1-x) dx \end{aligned} \tag{2}$$

In (2), we exapanded $(1-x)^b = (1-x)^{b-1}(1-x)$. We did this because we will distribute our integrand over $(1-x)$ and separate our integral into two new ones.

$$\begin{aligned} E[X] = \frac{1}{B(a,b)} \int_0^1 \frac{a}{b} x^{a-1}(1-x)^{b-1}(1-x) dx = \\ \frac{1}{B(a,b)} \left[\int_0^1 \frac{a}{b} x^{a-1}(1-x)^{b-1} dx - \int_0^1 x \frac{a}{b} x^{a-1}(1-x)^{b-1} dx \right] = \frac{a}{b}(1 - E[X]) \end{aligned} \tag{3}$$

Therefore, the expected value of the Beta distribution is:

$$E[X] = \frac{a}{a+b} \tag{4}$$

Mode

In order to calculate the mode, we have to differentiate the Beta distribution and equal it to zero.

$$
\begin{aligned}
&mode = \frac{dBeta(a,b)}{dx} = 0 \\
&\frac{d\left(x^{a-1}(1-x)^{b-1}\right)}{dx} = (a-1)x^{a-2}(1-x)^{b-1} - (b-1)x^{a-1}(1-x)^{b-2} = 0 \\
&(a-1)(1-x) - (b-1)x = 0 \therefore x = \frac{a-1}{a+b-2}
\end{aligned} \tag{5}
$$

Variance

First of all, remember that the variance is given by:

$$var[X] = E[X^2] - E[X]^2 \tag{6}$$

We already have the mean, so we just need to determine $E[X^2]$. To compute this value, we will integrate by parts again.

$$
\begin{aligned}
&E[X^2] = \int_0^1 x^2 \frac{x^{a-1}(1-x)^{b-1}}{B(a,b)} dx = \\
&\frac{1}{bB(a,b)} \left[-x^{a+1}(1-x)^b \Big|_0^1 - \int_0^1 -(a+1)x^a(1-x)^b dx \right] = \\
&\frac{a+1}{b} \int_0^1 \frac{1}{B(a,b)} x x^{a-1}(1-x)^b
\end{aligned} \tag{7}
$$

To complete the calculation, we should note something intesting in the integrand of (7). The integral is quite close to the mean and variance of the Beta distribution. If we distribue $(1-x)$ on the rest of the integrand, we end up with the mean and the variance again!

$$
\begin{aligned}
&E[X^2] = \frac{a+1}{b} \int_0^1 \frac{1}{B(a,b)} x x^{a-1}(1-x)^b = \\
&\frac{a+1}{b} \left[\int_0^1 \frac{1}{B(a,b)} x x^{a-1}(1-x)^{b-1} - \int_0^1 \frac{1}{B(a,b)} x^2 x^{a-1}(1-x)^{b-1} \right] = \\
&\frac{a+1}{b} [E[X] - E[X^2]]
\end{aligned} \tag{8}
$$

Substituting (4) on (8) and isolating $E[X^2]$, we arrive at:

$$E[X^2] = \frac{a+1}{a+b+1} \frac{a}{a+b} \tag{9}$$

Now, substituting (9) and (4) on (6) we arrive at the final result:

$$var[X] = \frac{a+1}{a+b+1} \frac{a}{a+b} - (\frac{a}{a+b})^2 = \frac{ab}{(a+b+1)(a+b)^2} \tag{10}$$

Conclusion

After a lot of calculation, we arrived at the three results that we wanted. As we first suspected, all the results are functions of the parameteres a and b. Moreover, since in the definition of the Beta distribution we impose $a, b > 0$, the mean and the variance are always defined.

The mode, on the other hand is a bit more complicated. Note that if $a, b \geq 1$, then Beta is a continuous function (product of two polynomials) on the compact interval $[0, 1]$. From Weierstrass extreme value theorem, this means that Beta must have at least one global maximum. This theorem also states that the maximum occurs weither on the extremes of the interval (i.e. on 0 and 1) or where the derivative is null. Since the extremes values of Beta are always 0, we can conclude that the mode is achieved by our calculation of the derivative (note that our initial condition $a, b \geq 1$ implies that the mode will be finite and positive).

Now, let's assume that $a < 1$ or $b < 1$. In this situation, our function is not continious on 0 (for $a < 1$) or 1 ($b < 1$). Thus, the Weierstrass extreme value theorem is no longer valid and nothing guarantees the existence of a maximum. In fact, is not that difficult to see, that the function will go to infinity in the point of discontinuity.