Working with continuous prior distributions

Stat 340, Fall 2021

Your turn 1: Undertanding kernels

The kernel of the Beta distribution is $x^{a-1}(1-x)^{b-1}$. To help you develop your understanding of the kernel and how it relates to the PDF, we will compare plots of the beta kernel to the full PDF.

1. Pick a value for the `a` and `b` parameters.

2. Using these parameters, run the below code chunk. (This code chunk is posted on teh course webpage, so just copy and paste it to save time!)

```
a <- # insert your a value
b <- # insert your b value
x <- seq(0, 1, by = .001)
kernel <- x^{a - 1} * (1 - x)^{b - 1}
pdf <- gamma(a + b)/(gamma(a) * gamma(b)) * kernel
par(mfrow=c(1,2))
plot(x, pdf, type="l", main = "PDF")
plot(x, kernel, type = "l", main = "Kernel")
```

3. What are the similarities between the pdf and the kernel? What are the differences?

4. If we know the kernel for a distribution, why can we find the PDF?

Your turn 2: Geometric-Beta model

Let $X_1, X_2, \ldots, X_n$ be a random sample from the geometric distribution with PMF

$$f(x|p) = p(1-p)^{x-1};\ x = 1, 2, \ldots,\quad 0 < p < 1$$

1. Write down the likelihood function, $f(x_1, \ldots, x_n|\lambda)$. (Remember joint distributions!)

2. Suppose that you decide to use a Beta(a,b) prior distribution for p with PDF

$$\pi(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{a-1}(1-p)^{b-1}$$

   Find the posterior density of p.

3. Is the beta prior a conjugate family to the geometric likelihood?

Your turn 3:

Let p denote the proportion of U.S. adults that do not believe in climate change. Cards Against Humanity's "Pulse of the Nation" project (https://thepulseofthenation.com/) conducted monthly polls into people's social and political views, as well as some silly things. This data includes responses to a subset of questions included in the poll conducted in September 2017. Of 1000 survey respondents, 150 responded that it was "not real at all".

1. Using a Beta(1, 2) prior distribution, what is the posterior distribution of p?

2. Simulate 1000 draws from the posterior distribution by completing the code in the below chunk:

```
p_draws <- ___(n = ___, ___, ___)
```

3. Use your simulated draws to calculate a 93% credible interval equal-tailed for p. Interpret this interval in context.

4. Suppose you were to survey 100 more adults. Use simulated draws from part 2 to approximate the posterior predictive model of $\tilde{Y}$, the number that don't believe in climate change.

   Construct a histogram of this predictive model (distribution).

   Approximate the probability that at least 20 of the 100 people don't believe in climate change.