Probability intervals
Class 16, 18.05
Jeremy Orloff and Jonathan Bloom

1 Learning Goals

1. Be able to find probability intervals given a pmf or pdf.

2. Understand how probability intervals summarize belief in Bayesian updating.

3. Be able to use subjective probability intervals to construct reasonable priors.

4. Be able to construct subjective probability intervals by systematically estimating quantiles.

2 Probability intervals

Suppose we have a pmf $p(\theta)$ or pdf $f(\theta)$ describing our belief about the value of an unknown parameter of interest θ.

Definition: A *p-probability interval* for θ is an interval $[a, b]$ with $P(a \leq \theta \leq b) = p$.

Notes.
1. In the discrete case with pmf $p(\theta)$, this means $\sum_{a \leq \theta_i \leq b} p(\theta_i) = p$.

2. In the continuous case with pdf $f(\theta)$, this means $\int_a^b f(\theta) \, d\theta = p$.

3. We may say 90%-probability interval to mean 0.9-probability interval. Probability intervals are also called credible intervals to contrast them with confidence intervals, which we will introduce in the frequentist unit.

Example 1. Between the 0.05 and 0.55 quantiles is a 0.5 probability interval. There are many 50% probability intervals, e.g. the interval from the 0.25 to the 0.75 quantiles.

In particular, notice that the *p*-probability interval for θ is not unique.

Q-notation. We can phrase probability intervals in terms of **quantiles**. Recall that the *s*-quantile for θ is the value q_s with $P(\theta \leq q_s) = s$. So for $s \leq t$, the amount of probability between the *s*-quantile and the *t*-quantile is just $t - s$. In these terms, a *p*-probability interval is any interval $[q_s, q_t]$ with $t - s = p$.

Example 2. We have 0.5 probability intervals $[q_{0.25}, q_{0.75}]$ and $[q_{0.05}, q_{0.55}]$.

Symmetric probability intervals.
The interval $[q_{0.25}, q_{0.75}]$ is symmetric because the amount of probability remaining on either side of the interval is the same, namely 0.25. If the pdf is not too skewed, the symmetric interval is usually a good default choice.

More notes.

1. Different p-probability intervals for θ may have different widths. We can make the width smaller by centering the interval under the highest part of the pdf. Such an interval is usually a good choice since it contains the most likely values. See the examples below for normal and beta distributions.

2. Since the width can vary for fixed p, a larger p does not always mean a larger width. Here's what is true: if a p_1-probability interval is fully contained in a p_2-probability interval, then p_1 is smaller than p_2.

Probability intervals for a normal distribution. The figure shows a number of probability intervals for the standard normal.

1. All of the blue bars span a 0.68-probability interval. Notice that the smallest blue bar runs between -1 and 1. This runs from the 16th percentile to the 84th percentile so it is a symmetric interval.

2. All the green bars span a 0.9-probability interval. They are longer than the blue bars because they include more probability. Note again that the shortest green bar is symmetric.

orange = 0.5, blue = 0.68, green = 0.9
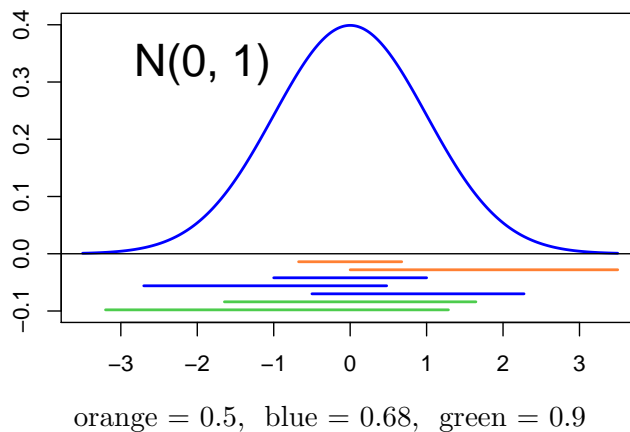
Probabilitiy intervals for a beta distribution. The following figure shows probability intervals for a beta distribution. Notice how the two blue bars have very different lengths yet cover the same probability $p = 0.68$.

orange = 0.5, blue = 0.68, green = 0.9
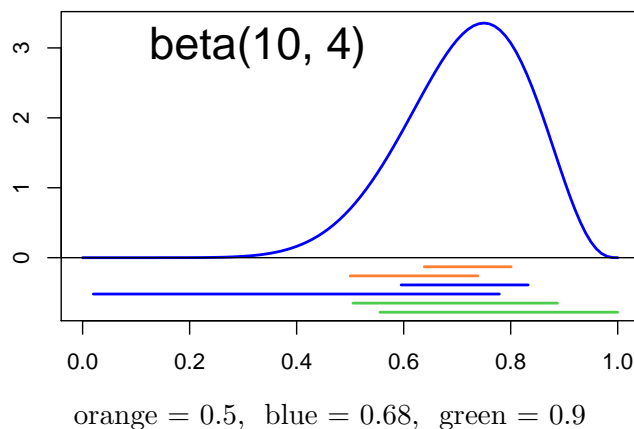
3   Uses of probability intervals

3.1   Summarizing and communicating your beliefs

Probability intervals are an intuitive and effective way to summarize and communicate your beliefs. It's hard to describe an entire function $f(\theta)$ to a friend in words. If the function isn't from a parameterized family then it's especially hard. Even with a beta distribution, it's easier to interpret "I think θ is between 0.45 and 0.65 with 50% probability" than "I think θ follows a beta(8,6) distribution". An exception to this rule of communication might be the normal distribution, but only if the recipient is also comfortable with standard deviation. Of course, what we gain in clarity we lose in precision, since the function contains more information than the probability interval.

Probability intervals also play well with Bayesian updating. If we update from the prior $f(\theta)$ to the posterior $f(\theta|x)$, then the p-probability interval for the posterior will tend to be shorter than than the p-probability interval for the prior. In this sense, the data has made us more certain. See for example the election example below.

4   Constructing a prior using subjective probability intervals

Probability intervals are also useful when we do not have a pmf or pdf at hand. In this case, subjective probability intervals give us a method for constructing a reasonable prior for θ "from scratch". The thought process is to ask yourself a series of questions, e.g., 'what is my expected value for θ?'; 'my 0.5-probability interval?'; 'my 0.9-probability interval?' Then build a prior that is consistent with these intervals.

4.1   Estimating the intervals directly

Example 3. Building priors
In 2013 there was a special election for a congressional seat in a district in South Carolina. The election pitted Republican Mark Sanford against Democrat Elizabeth Colbert Busch. Let θ be the fraction of the population who favored Sanford. Our goal in this example is to build a subjective prior for θ. We'll use the following prior evidence.

- Sanford is a former S. Carolina Congressman and Governor

- In 2009, while Governor, he had to resign after he was discovered to be having an affair in Argentina while he claimed to be hiking the Appalachian trail.

- In 2013 Sanford won the Republican primary over 15 primary opponents.

- In the district in the 2012 presidential election the Republican Romney beat the Democrat Obama 58% to 40%.

- The Colbert bump: Elizabeth Colbert Busch is the sister of well-known comedian Stephen Colbert.

Our strategy will be to use our intuition to construct some probability intervals and then find a beta distribution that approximately matches these intervals. This is subjective so someone else might give a different answer.

Step 1. Use the evidence to construct 0.5 and 0.9 probability intervals for θ.

We'll start by thinking about the 90% interval. The single strongest prior evidence is the 58% to 40% of Romney over Obama. Given the negatives for Sanford we don't expect he'll win much more than 58% of the vote. So we'll put the top of the 0.9 interval at 0.65. With all of Sanford's negatives he could lose big. So we'll put the bottom at 0.3.

$$\text{0.9 interval:} \quad [0.3, \, 0.65]$$

For the 0.5 interval we'll pull these endpoints in. It really seems unlikely Sanford will get more votes than Romney, so we can leave 0.25 probability that he'll get above 57%. The lower limit seems harder to predict. So we'll leave 0.25 probability that he'll get under 42%.

$$\text{0.5 interval:} \quad [0.42, \, 0.57]$$

Step 2. Use our 0.5 and 0.9 probability intervals to pick a beta distribution that approximats these intervals. We used the R function `pbeta` and a little trial and error to choose beta(11,12). Here is our R code.

```
a = 11
b = 12
pbeta(0.65, a, b) - pbeta(0.3, a, b)
pbeta(0.57, a, b) - pbeta(0.42, a, b)
```

This computed $P([0.3, 0.65]) = 0.91$ and $P([0.42, 0.57]) = 0.52$. So our intervals are actually 0.91 and 0.52-probability intervals. This is pretty close to what we wanted!

The plot below shows the density of beta(11,12). The horizontal orange line shows our interval [0.42, 0.57] and the blue line shows our interval [0.3, 0.65].
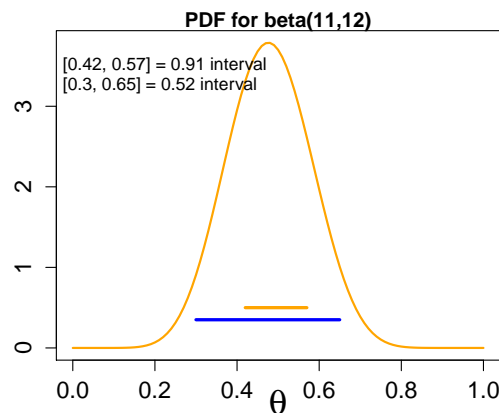
beta(11,12) fitting 0.5 and 0.9 probability intervals

4.2   Constructing a prior by estimating quantiles

The method in Example 3 gives a good feel for building priors from probability intervals. Here we illustrate a slightly different way of building a prior by estimating quantiles. The

basic strategy is to first estimate the median, then divide and conquer to estimate the first and third quantiles. Finally you choose a prior distribution that fits these estimates.

Example 4. Redo the Sanford vs. Colbert-Busch election example using quantiles.

Solution: We start by estimating the median. Just as before the single strongest evidence is the 58% to 40% victory of Romney over Obama. However, given Sanford's negatives and Busch's Colbert bump we'll estimate the median at 0.47.

In a district that went 58 to 40 for the Republican Romney it's hard to imagine Sanford's vote going a lot below 40%. So we'll estimate Sanford 25th percentile as 0.40. Likewise, given his negatives it's hard to imagine him going above 58%, so we'll estimate his 75th percentile as 0.55.

We used R to search through values of a and b for the beta distribution that matches these quartiles the best. Since the beta distribution does not require a and b to be integers we looked for the best fit to 1 decimal place. We found beta(9.9, 11.0). Above is a plot of beta(9.9,11.0) with its actual quartiles shown. These match the desired quartiles pretty well.

PDF for beta(9.9,11.0)

$q_{0.25} = 0.399$
$q_{0.5} = 0.472$
$q_{0.75} = 0.547$
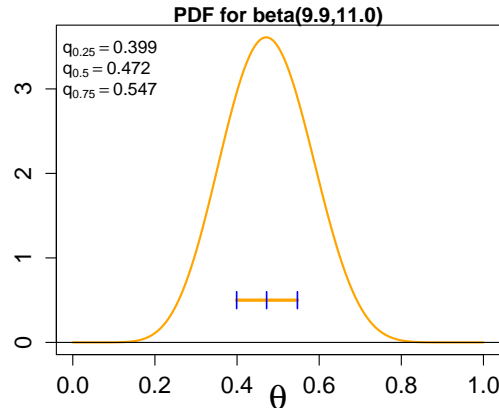
beta(9.9, 11.0) matching desired quartiles

Historic note. In the election Sanford won 54% of the vote and Busch won 45.2%. (Source: https://elections.huffingtonpost.com/2013/mark-sanford-vs-elizabeth-colbert-busch-sc1

MIT OpenCourseWare

https://ocw.mit.edu

18.05 Introduction to Probability and Statistics

Spring 2022