

Probability intervals
Class 16, 18.05, Spring 2014
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.

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 .9-probability interval. Probability intervals are also called *credible intervals* to contrast them with confidence intervals, which we'll introduce in the frequentist unit.

Example 1. Between the .05 and .55 quantiles is a .5 probability interval. Another 50% probability interval goes from the .25 to the .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 $s < t$, 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 .5 probability intervals $[q_{.25}, q_{.75}]$ and $[q_{.05}, q_{.55}]$.

Symmetric probability intervals.

The interval $[q_{.25}, q_{.75}]$ is **symmetric** because the amount of probability remaining on either side of the interval is the same, namely .25. If the pdf is not too skewed, the symmetric interval is usually a good default choice.

More notes.

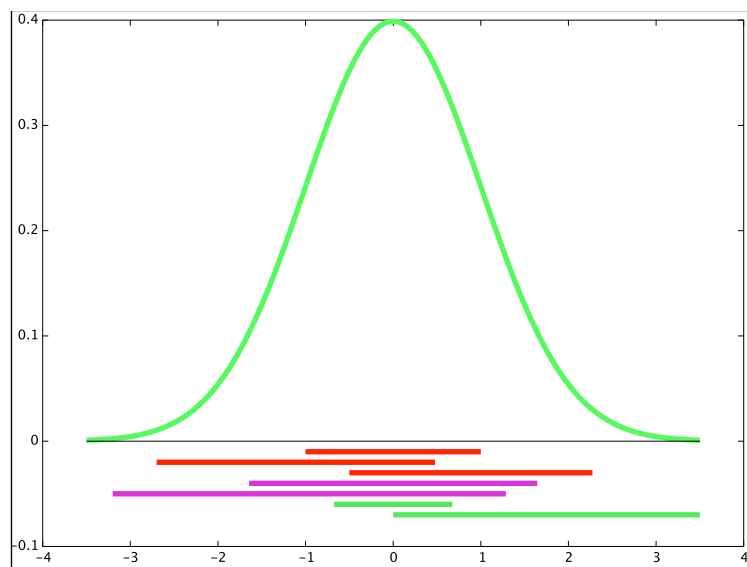
1. Different p -probability intervals for θ may have different widths. We can make the width smaller by centering the interval under the highest part of the pdf. Such an interval is usually a good choice since it contains the most likely values. See the examples below for normal and beta distributions.

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 bigger than p_2 .

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

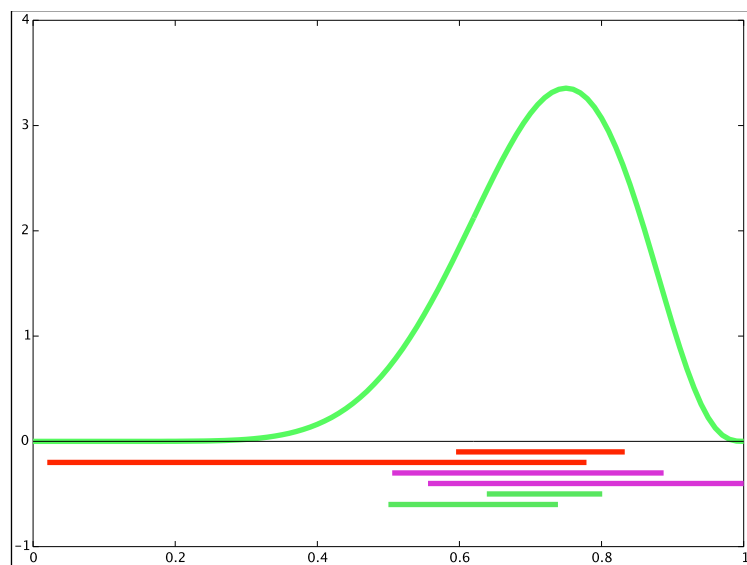
1. All of the red bars span a .68-probability interval. Notice that the smallest red bar runs between -1 and 1. This runs from the 16th percental to the 84th percentile so it is a symmetric interval.

2. All the magenta bars span a .9-probability interval. They are longer than the red bars because they include more probability. Note again that the shortest magenta bar is symmetric.



red = .68, magenta = .9, green = .5

Probability intervals for a beta distribution. The following figure shows probability intervals for a beta distribution. Notice how the two red bars have very different lengths yet cover the same probability $p = .68$.



red = .68, magenta = .9, green = .5

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 .45 and .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 the p -probability interval for the prior. In this sense, the data has made us more certain. See for example the election example below.

3.2 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". Ask yourself, what is my expected value for θ ? My .5-probability interval? My .9-probability interval? Then build a prior that is consistent with these intervals.

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 Busch. Our goal in this example is to build a subjective prior for θ . We'll use the following prior evidence.

- Sanford is former S. Carolina Congressman and Governor
- Sanford famously resigned after having an affair in Argentina while he claimed to be hiking the Appalachian trail.
- Sanford beat out 15 primary opponents.
- In the district in 2012 Romney beat Obama 58% to 40%.
- The Colbert bump: Elizabeth Colbert Busch is the sister of 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 .5 and .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 .9 interval at .65. With all of Sanford's negatives he could lose big. So we'll put the bottom at .3.

$$.9 \text{ interval: } [.3, .65]$$

For the .5 interval we'll pull these endpoints in. It really seems unlikely Sanford will get more votes than Romney, so we can leave .25 probability that he'll get above 57%. The lower limit seems harder to predict. So we'll leave .25 probability that he'll get under 42%.

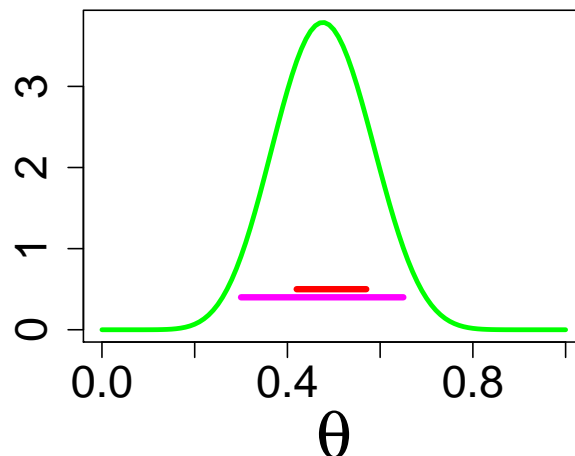
$$.5 \text{ interval: } [.42, .57]$$

Step 2. Use our .5 and 9 probability intervals to pick a beta distribution that approximates these intervals. We used the R function `pbeta` and a little trial and error to choose `beta(18,16)`. Here is our R code.

```
a = 11
b = 12
pbeta(.65,a,b) - pbeta(.3,a,b)
pbeta(.57,a,b) - pbeta(.42,a,b)
```

The computed $P([.3, .65]) = .91$ and $P([.42, .57]) = .52$. So our intervals are actually .91 and .52-probability intervals. This is pretty close to what we wanted.

At right is a graph of the density of `beta(11,12)`. The red line shows our interval `[.42, .57]` and the magenta line shows our interval `[.3, .65]`.



MIT OpenCourseWare
<http://ocw.mit.edu>

18.05 Introduction to Probability and Statistics
Spring 2014

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.