16.3 Bayesian Credible Intervals

- Bayesian point estimates are useful, but they only describe the ______ of the posterior distribution.
- An <u>interval</u> estimate can better characterize the amount of <u>uncertainty</u> about the target parameter that is implied by the posterior.
- In classical statistics, we have

- In Bayesian statistics, our interval estimates are called ______________.

Differences in Interpretation

- In classical inference, the parameter θ is <u>fixed</u> but unknown.
- In a confidence interval <u>formula</u>

the <u>endpoints</u> of the interval are <u>random</u> (since they are functions of the data $Y_1, \ldots, Y_n$) but the parameter θ is not.

- Once we take the sample and calculate the endpoints (say, 3.7 and 5.2), it does <u>not</u> make sense to write something like

  because there is nothing random in that inequality.
- Either θ is in $(3.7, 5.2)$ or it is not.
- We interpret a 95% confidence interval thusly:

<u>Note</u>: The intervals, not θ, would be changing across samples.

- In the Bayesian context, however, θ <u>is</u> random, with posterior density

<u>Defn</u>: An interval (a, b) is a $100(1-\alpha)\%$ <u>credible interval</u> if

where $g^*(\theta)$ is the posterior density of θ.

- If we get a 95% credible interval $(3.7, 5.2)$, we <u>can</u> write

since θ is random.

- Technically, we might write

to indicate that this is a <u>posterior</u> probability.

Example 3 again: In the Army recruiting example, find a 95% credible interval for the population mean test score.

Interpretation:

Note: If we had used the noninformative prior $g(\mu) = 1$, $-\infty < \mu < \infty$, our 95% credible interval would have been which is <u>exactly</u> the formula for the classical 95% confidence interval for μ when σ is known.

- In other cases, we do not have tabulated values for the posterior distribution and the integration to find the endpoints a and b may be tedious.
- In such cases we may use the computer to help.

<u>Example 1 again</u> (Plant science example): Find a 95% credible interval for the probability that the plant type will bear fruit. Assume a beta(3,1) prior for p.

- We derived the posterior to be:

- We need a and b such that

- This integration is possible, but tedious.
- In R, the ________ function will give us these quantiles easily:

<u>Example 2 again</u>: Recall our exponential model with pdf

$$f(y) = \begin{cases} \theta e^{-\theta y} & \text{if } y > 0 \\ 0 & \text{elsewhere} \end{cases}$$

where $\theta > 0$ was the <u>reciprocal</u> of the population mean. Suppose we gather 30 waiting times $Y_1, \ldots, Y_{30}$ that follow this distribution.

- Before examining the data, the researchers believed the average waiting time was around 5 minutes. From the sample, $\sum_{i=1}^{30} y_i = 134.2$. Find a 90% credible interval for the population mean waiting time.

- Our prior for θ is

- We have derived that the posterior for θ is

- A 90% credible interval for θ can be found using R:

- Note that for any Bayesian interval, if a different prior were used (or simply different prior parameters, the resulting credible interval would be different.
- For large sample sizes, the influence of the prior is minimal.
- As with classical intervals, as the sample size increases, the credible interval tends to become __________ (__________ precise).

Restricted Parameter Spaces

- In Example 1, we found a 95% credible interval for p to be

- What is the classical z-based 95% confidence interval for p here?

- The right endpoint makes no sense since p is restricted to be

- Note that a Bayesian credible interval can never extend beyond the parameter space since the posterior only has support on the parameter space.