8.6 Bayesian Credible Intervals

Recall that an *interval estimator* for a parameter θ consists of a pair of statistics $A(\mathbf{X})$ and $B(\mathbf{X})$ so that $A < \theta < B$ with a particular probability γ. To find the frequentist confidence interval for θ, we had to either:

1. Find and invert the pivotal quantity
2. Use bootstrapping

We now consider interval estimates from the Bayesian perspective. As Bayesian, we treat θ as a random variable, with prior distribution $\xi(\theta)$. After observing the data $\mathbf{x}$, θ has a posterior distribution $\xi(\theta|\mathbf{x})$.

Our goal is to use a sample $\mathbf{x}$ and find $A(\mathbf{x})$ and $B(\mathbf{x})$ so that

$$P\big(A(\mathbf{x}) < \theta < B(\mathbf{x})|\mathbf{x}\big) = \gamma$$

Doing so is rather straightforward, if we have the posterior PDF! Note that the statistics A and B are just two specific quantiles of the posterior distribution $\theta|\mathbf{x}$.

There are many choices for these quantiles:

1. Equal area: A is the $\frac{1-\gamma}{2}$ quantile and B is the $\frac{1+\gamma}{2}$ quantile.
2. Mean as center point: $A, B = E[\theta|\mathbf{x}] \pm c$
3. Narrowest interval: A and B bound the highest points of the posterior density.

Of these, the first method is often most common.

The textbook calls Bayesian interval estimates **posterior intervals**, although most Bayesian literature refers to them as **credible intervals**.

Ex 1: Let's return to the mystery envelope containing red and blue tickets, where θ denotes the proportion of red tickets. We will draw 10 more tickets from the bag and use the results to make a prediction about θ. Let X denote the number of red tickets drawn and note that $X|\theta \sim \text{Bin}(10, \theta)$.

Using the Beta-Binomial conjugacy, it is reasonable to provide a Beta prior for θ. Several weeks ago, we drew 8 tickets from the bag, and obtained 6 red and 2 blue. We'll use this information to form our prior: $\theta \sim \text{Beta}(6, 2)$.

Note that before we collect any further data, we can create a prior interval for θ. Suppose we want to create 95% interval using the equal areas method. Then A and B are the 0.025 and 0.975 quantiles of $\text{Beta}(6, 2)$, which we can compute using R

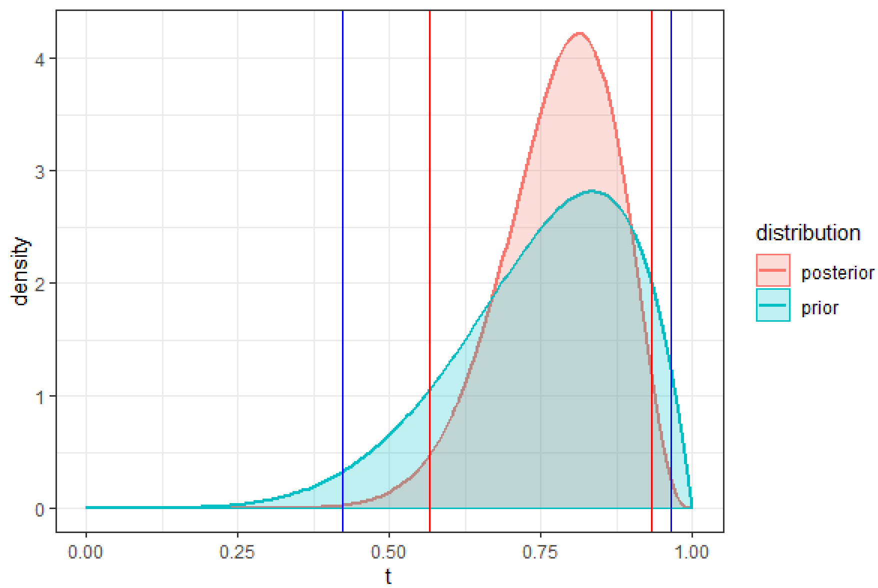
```
qbeta(c(0.025, 0.975), 6,2)
# 0.421  0.963
```

Hence, there is 95% probability that θ is between 0.421 and 0.963.

Of course, we can improve our estimate and shorter the length of this interval by collecting more data. Suppose that in 10 further draws, we get $X = 8$. Then $\theta|X = 8 \sim \text{Beta}(14, 4)$, and so the new posterior interval is $(0.566, 0.932)$, computed in R using

```
qbeta(c(0.025, 0.975), 14,4)
# 0.566 0.932
```

Visualizations of the prior and posterior distributions, along with the corresponding credible intervals, are shown below:



Samples from a Normal Distribution

Now, assume that $\mathbf{x}$ is conditionally iid Normal, given $\theta = (\mu, \sigma^2)$. Previously, we found the posterior distribution of $\mu|\mathbf{x}$ with **known** σ^2. But in a typical setting, if we don't know μ, then we also won't know σ^2. To find a joint posterior on $\mu, \sigma^2|\mathbf{x}$, we need **two** priors.

<u>Def:</u> The **precision** of a Normal distribution is $\tau = \frac{1}{\sigma^2}$.

Then

$$f(x|\mu,\tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp\left\{-\frac{\tau}{2}(x-\mu)^2\right\} \qquad f(\mathbf{x}|\mu,\tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp\left\{-\frac{\tau}{2}\sum(x_i-\mu)^2\right\}$$

In order to implement Bayesian inference procedures, we need to find a family of conjugate joint priors for these parameters.

<u>Thm:</u> Let $\mathbf{X} \sim N(\mu, 1/\tau)$, with priors $\mu|\tau$ and τ

$$\begin{aligned} \mu|\tau \sim & N(\mu_0, 1/(\lambda_0\tau)) \\ \tau \sim & \text{Gamma}(\alpha_0, \beta_0) \end{aligned}$$

Then the posteriors on $\mu|\tau$ and τ are

$$\begin{aligned} \mu|\tau \sim & N(\mu_1, 1/(\lambda_1\tau)) \\ \tau \sim & \text{Gamma}(\alpha_1, \beta_1) \end{aligned}$$

where

$$\mu_1 = \frac{\lambda_0\mu_0 + n\bar{x}}{\lambda_0 + n} \quad \lambda_1 = \lambda_0 + n \quad \alpha_1 = \lambda_0 + \frac{n}{2} \quad \beta_1 = \beta_0 + \frac{1}{2}\sum(x_i - \bar{x})^2 + \frac{n\lambda_0(\bar{x}-\mu_0)^2}{2(\lambda_0 + n}$$

Proof. Use Bayes Theorem and complete the square in the exponent. Proceed analogous to the case when μ is unknown and σ is known. □

<u>Def:</u> We say that a vector (X, T) has Normal-gamma distribution with parameters $(\mu, \lambda, \alpha, \beta)$ if $X|Y$ is conditional $N(\mu, 1/(\lambda Y))$ and Y is marginally Gamma(α, β). The PDF for (X, Y) is therefore

$$f_{X,T}(x,\tau) = f_{X|T}(x|t)f_T(\tau) = \frac{\beta^\alpha\sqrt{\lambda}}{\Gamma(\alpha)\sqrt{2\pi}}\tau^{\alpha-1/2}e^{-\beta\tau}\exp\left\{-\frac{\lambda\tau(x-\mu)^2}{2}\right\}$$

The preceding theorem says that the Normal-Gamma distribution is a conjugate prior for samples from Normal distribution with unknown mean and precision.

<u>Thm:</u> Suppose $\mathbf{X}$ are conditionally iid $N(\mu, 1/\tau)$ with priors from a Normal-Gamma distribution as in the preceding theorem. Then the marginal distribution of μ can be expressed as

$$\sqrt{\frac{\lambda_0\alpha_0}{\beta_0}}(\mu-\mu_0) \sim t(2\alpha_0)$$

Proof. Let $Z = \sqrt{\lambda_0\tau}(\mu-\mu_0)$ and observe that $Z|\tau \sim N(0,1)$. We can express the joint density of Z and τ in terms of the conditional density $f_{Z|\tau}$ of $Z|\tau$ and the marginal density f_τ of τ:

$$f(z,\tau) = f_{Z|\tau}(z|\tau)f_\tau(\tau) = \varphi(z)f_\tau(\tau)$$

which shows that Z and τ are independent. Let $Y = 2\beta_1\tau$ and note that $\sim \chi^2(2\alpha_0)$. Define a variable U as

$$U = \sqrt{\frac{\lambda_0\alpha_0}{\beta_0}}(\mu-\mu_0) = \frac{\sqrt{\lambda_0\tau}(\mu-\mu_0)}{\sqrt{\frac{2\beta_0\tau}{2\alpha_0}}} = \frac{Z}{\sqrt{\frac{Y}{2\alpha_0}}}$$

which is t distributed with $2\alpha_0$ degrees of freedom. □

Note that the marginal posterior distribution of μ can be obtained in a similar fashion by using the parameters $\lambda_1, \alpha_1, \beta_1, \mu_1$.

Now, to obtain a γ probability credible interval, let $U = \sqrt{\frac{\lambda_1\alpha_1}{\beta_1}}(\mu-\mu_1)$ and let c be such that $P(-c < U < c|\mathbf{x}) = \gamma$. Then

$$\begin{aligned}\gamma =& P(-c < U < c)\\ =& P\left(-c < \sqrt{\frac{\lambda_1\alpha_1}{\beta_1}}(\mu-\mu_1) < c\right)\\ =& P\left(\mu_1 - c\sqrt{\frac{\lambda_1\alpha_1}{\beta_1}} < \mu < \mu_1 + c\sqrt{\frac{\lambda_1\alpha_1}{\beta_1}}\right)\end{aligned}$$

which gives the γ probability Credible Interval for μ.