8.2.1   Credible Intervals

Now we will assume we are in the **Bayesian setting**, which means out unknown parameter Θ will always be some random variable, not a fixed quantity. If we give a single point estimate like we do in MAP, we will never be exactly correct. Therefore, just like we did in the Frequentist setting with confidence intervals, we might want to give an interval instead of a single number. These are called credible intervals instead, and serve the same purpose!

Actually, since Θ is a random variable, finding an interval where the probability is at least 90% for example involves just looking at the PDF/CDF! An example best illustrates this.

Example(s)

Construct a 80% credible interval for Θ (the unknown) probability of success in $\text{Ber}(\Theta)$ given iid $n = 12$ samples $\mathbf{x} = (x_1, x_2, ..., x_{12})$ where $\sum_{i=1}^{n} x_i = 11$ (observed 11 successes out of 12). Suppose our prior is $\Theta \sim \text{Beta}(\alpha = 7, \beta = 3)$ (i.e., pretend we saw 6 successes and 2 failures ahead of time).


Solution From lecture 7.5 (MAP), we showed that choosing a Beta prior for Θ leads to a Beta posterior of $\Theta \mid \mathbf{x} \sim \text{Beta}(11 + 7, 1 + 3) = \text{Beta}(18, 4)$ and our MAP was then $\frac{18-1}{(18-1)+(4-1)} = \frac{17}{20}$ (since we saw 17 total successes, and 3 total failures.)

We want an interval $[a, b]$ such that $\mathbb{P}(a \le \Theta \le b) = 0.8$
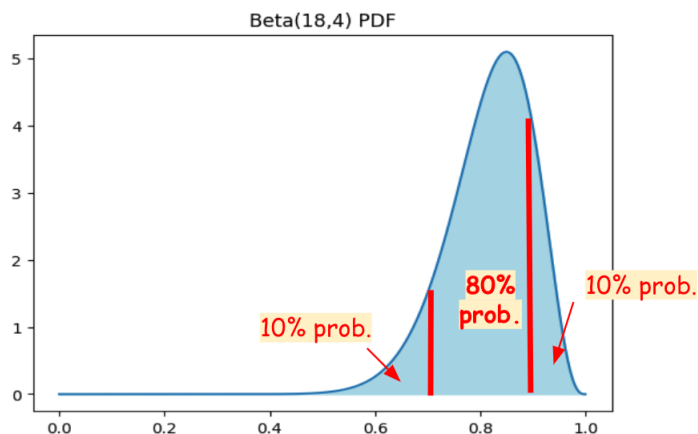
If we look at the Beta PDF, we are looking for such an interval that the probability that we fall in this area is 80%. If the area is centered, then the area to the left of that should have probability of 10%, and the area to the right of that should also have probability 10%.

This is equivalent of looking for $P(\Theta \leq a) = 0.1$ and $P(\Theta \leq b) = 0.9$. These information are given in the CDF of the Beta distribution. Note that on the x-axis we have the range of the Beta distribution $[0, 1]$, and on the y-axis, we have the cumulative probability of being to the left by integrating the PDF from above.
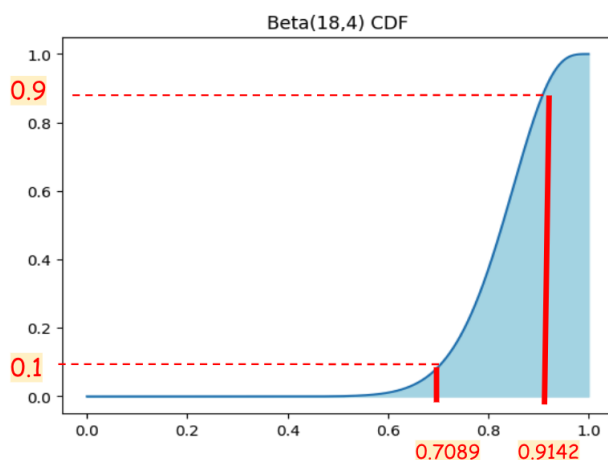
Let F_{Beta} denote the CDF of this Beta(18,4) distribution. Then, choose $a = F_{\text{Beta}}^{-1}(0.1) \approx 0.7089$ and $b = F_{\text{Beta}}^{-1}(0.9) \approx 0.9142$, so our credible interval is $[0.7089, 0.9142]$.

Note that the MAP was $\frac{17}{20} = 0.85$, which is not at the center! We could have chosen any a, b where the area between them is 80%, but we set the areas to the left and right to be equal.

In order to compute the inverse CDF, we can use the scipy.stats library as follows:

```
1 >>> from scipy.stats import beta
2 >>> beta.ppf(0.1, 18, 4) # inverse cdf of Beta(18, 4)
3 0.70898031757137514
```

☐

That's all there is to it! Just find the PDF/CDF of your posterior distribution (hopefully you chose a conjugate prior), and look up the inverse CDF at points a and b such that $b - a$ is your desired confidence level of your credible interval.

Definition 8.2.1: Credible Intervals

Suppose you have iid samples $\mathbf{x} = (x_1, ..., x_n)$ from some distribution with unknown parameter Θ. You are in the **Bayesian setting**, so you have chosen a prior distribution for the RV Θ.

A $100(1 - \alpha)\%$ **credible interval** for Θ is an interval $[a, b]$ such that the probability (over the randomness in Θ) that Θ lies in the interval is $1 - \alpha$:

$$P(\Theta \in [a, b]) = 1 - \alpha$$

If we've chosen the appropriate conjugate prior for the sampling distribution (like Beta for Bernoulli), the posterior is easy to compute. Say the CDF of the posterior is F_Y. Then, a $100(1 - \alpha)\%$ credible interval is given by

$$\left[F_Y^{-1}\left(\frac{\alpha}{2}\right), F_Y^{-1}\left(1 - \frac{\alpha}{2}\right) \right]$$

Again, this is one which has equal area to the left and right of the interval, but there are infinitely many possible credible intervals you can create.

8.2.2 Interpreting Credible Intervals

How can we interpret a 80% credible interval $[0.7089, 0.9142]$ for parameter Θ?

Correct: There is an 80% probability that Θ falls in the credible interval $[0.7089, 0.9142]$. Written out,

$$P(\Theta \in [0.7089, 0.9142]) = 0.8$$

This is correct because not Θ is a random variable, and it makes sense to say!

Contrast this with the interpretation of a confidence interval, where θ is a fixed number.

8.2.3 Exercises

1. Let $\mathbf{x} = (x_1, \ldots, x_n)$ be iid samples from $\text{Exp}(\Theta)$ where Θ is a random variable (not fixed). Recall from section 7.5 Exercise 1 that if we choose the prior distribution $\Theta \sim \text{Gamma}(r, \lambda)$, then the posterior distribution is $\Theta \mid \mathbf{x} \sim \text{Gamma}(n + r, \lambda + \sum x_i)$.

   Suppose $n = 13, \bar{x} = 0.21, r = 7, \lambda = 12$. Construct a 96% credible interval for Θ. To find the point t such that $F_T(t) = y$ for $T \sim \text{Gamma}(u, v)$, call the following function which gets the inverse CDF:

   ```
   scipy.stats.gamma.ppf(y, u, 0, 1/v)
   ```

   Then, verify that the MAP estimate is actually contained in your credible interval.

   Solution: Before we call the function, we have to identify what u and v are. Plugging in the numbers above to the general posterior we computed earlier, we find

   $$\Theta \mid \mathbf{x} \sim \text{Gamma}(13 + 7, 12 + 13 \cdot 0.21) = \text{Gamma}(20, 14.73)$$

   Since we want a 96% interval, we must look up the inverse CDF at 0.02 and 0.98 (why?).

   We write a few lines of code, calling the provided function twice:

   ```
   1 >>> from scipy.stats import gamma
   2 >>> gamma.ppf(0.02, 20, 0, 1/14.73)  # inverse cdf of Gamma(20, 14.73)
   3 0.809150510196322
   4 >>> gamma.ppf(0.98, 20, 0, 1/14.73)  # inverse cdf of Gamma(20, 14.73)
   5 2.0514641398722735
   ```

   So our 96% credible interval for Θ is approximately

   $$[0.809, 2.051]$$

   Our MAP was just the mode of the Gamma, which is

   $$\hat{\theta}_{MAP} = \frac{19}{14.73} \approx 1.28988458927$$