STA 360/602L: Module 2.5

Frequentist vs Bayesian intervals

Dr. Olanrewaju Michael Akande

Frequentist confidence intervals

- Recall that a frequentist confidence interval $[l(y),\ u(y)]$ has 95% frequentist coverage for a population parameter θ if, before we collect the data,

$$\Pr[l(y) < \theta < u(y)|\theta] = 0.95.$$

- This means that 95% of the time, our constructed interval will cover the true parameter, and 5% of the time it won't.

- In any given sample, you don't know whether you're in the lucky 95% or the unlucky 5%.

Frequentist confidence intervals

- You just know that either the interval covers the parameter, or it doesn't (useful, but not too helpful clearly).
- There is NOT a 95% chance your interval covers the true parameter once you have collected the data.
- Asking about the definition of a confidence interval is tricky, even for those who know what they're doing.

Bayesian intervals

- An interval $[l(y),\ u(y)]$ has 95% Bayesian coverage for θ if

$$\Pr[l(y) < \theta < u(y) | Y = y] = 0.95.$$

- This describes our information about where θ lies *after* we observe the data.
- Fantastic!
- This is actually the interpretation people want to give to the frequentist confidence interval.
- Bayesian interval estimates are often generally called credible intervals.

Bayesian quantile-based interval

- The easiest way to obtain a Bayesian interval estimate is to use posterior quantiles.

- Easy since we either know the posterior densities exactly or can sample from the distributions.

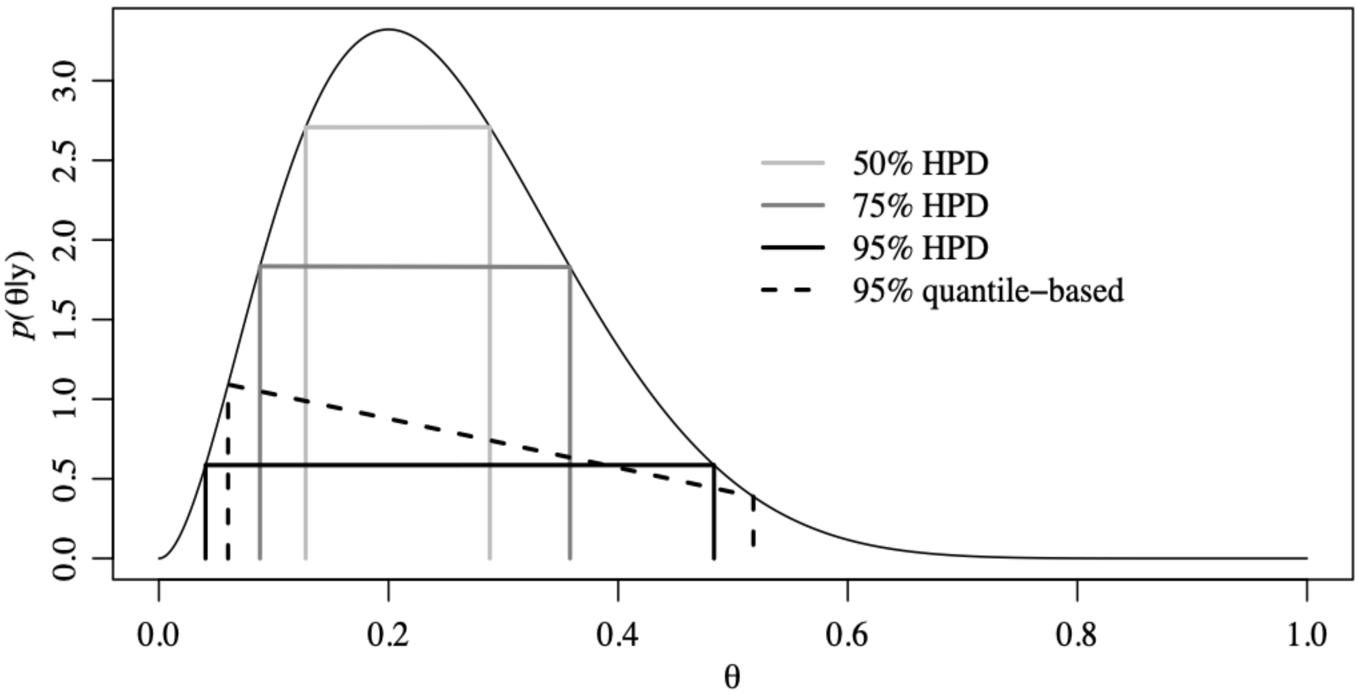
- To make a $100 \times (1-\alpha)$ quantile-based credible interval, find numbers (quantiles) $\theta_{\alpha/2} < \theta_{1-\alpha/2}$ such that

  1. $\Pr(\theta < \theta_{\alpha/2} | Y = y) = \frac{\alpha}{2}$; and

  2. $\Pr(\theta > \theta_{1-\alpha/2} | Y = y) = \frac{\alpha}{2}$.

- This is an equal-tailed interval. Often when researchers refer to a credible interval, this is what they mean.

Equal-tailed quantile-based interval



- This is Figure 3.6 from the Hoff book. Focus on the quantile-based credible interval for now.
- Note that there are values of θ outside the quantile-based credible interval, with higher density than some values inside the interval.

HPD REGION

- A $100 \times (1 - \alpha)$ highest posterior density (HPD) region is a subset $s(y)$ of the parameter space Θ such that

  1. $\Pr(\theta \in s(y) | Y = y) = 1 - \alpha$; and

  2. If $\theta_a \in s(y)$ and $\theta_b \notin s(y)$, then $\Pr(\theta_a | Y = y) > \Pr(\theta_b | Y = y)$.

- $\Rightarrow$ **All** points in a HPD region have higher posterior density than points outside the region.

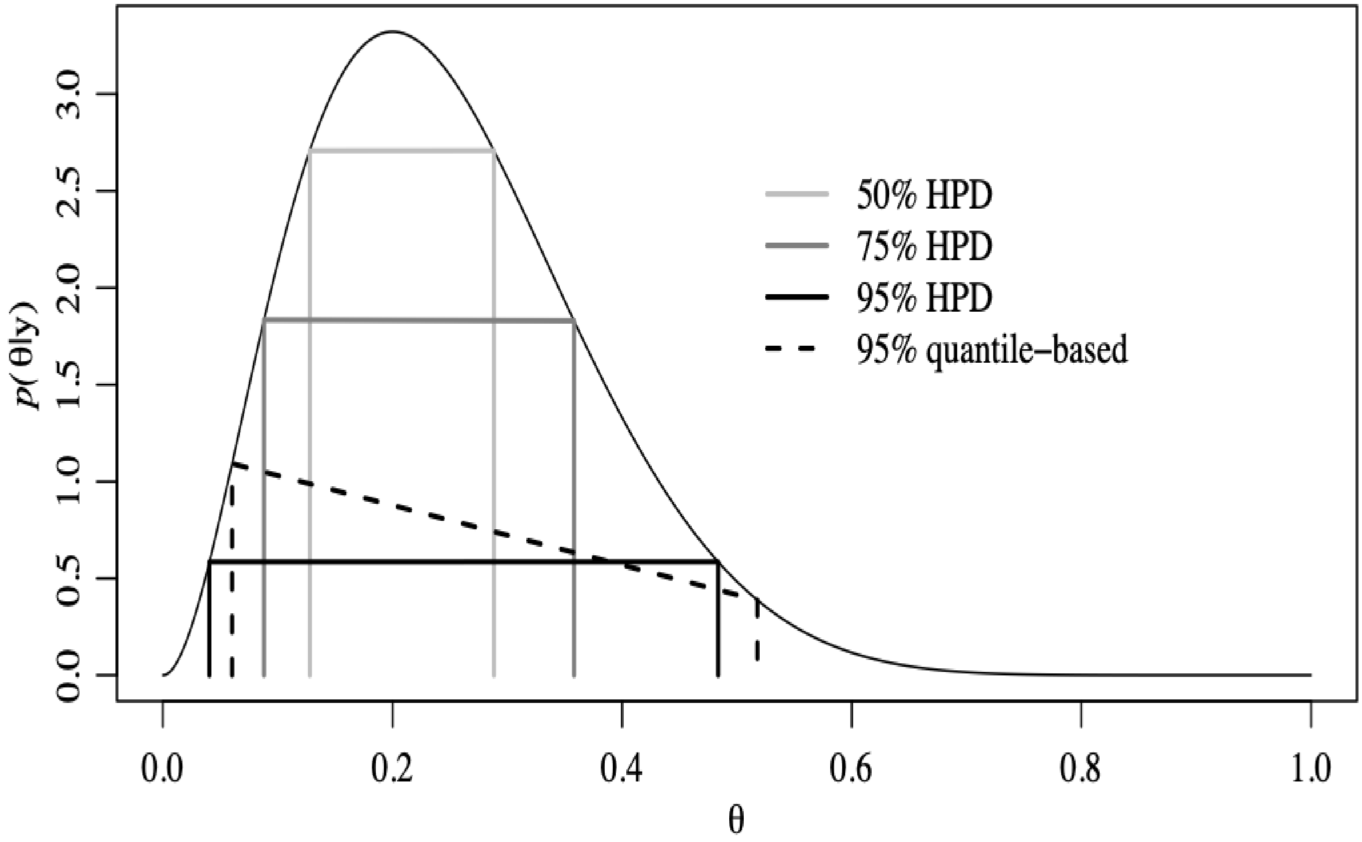
  Note this region is not necessarily a single interval (e.g., in the case of a multimodal posterior).

- The basic idea is to gradually move a horizontal line down across the density, including in the HPD region all values of θ with a density above the horizontal line.

- Stop moving the line down when the posterior probability of the values of θ in the region reaches $1 - \alpha$.

HPD REGION

Hoff Figure 3.6 shows how to construct an HPD region.



What's next?

Move on to the readings for the next module!