Reporting the results of a Bayesian analysis

Giorgio Corani - (IDSIA, SUPSI)

Analysis of the posterior

- Assume you want to make inference about parameter θ.
- The result of the Bayesian analysis is the whole posterior of θ, **not** a single value.
- The dispersion of the posterior distribution (posterior variance) quantifies our uncertainty.
- Uncertainty decreases when we have more data.

The impact of the prior

- If many data are available, the posterior is the same regardless the prior.
- But how many data are needed for the likelihood to *overwhelm* the prior? That is not know in advance.

The impact of the prior

- With few data, the posterior might be different using different priors; it makes sense to repeat the analysis with different priors (analysis of *sensitivity* to the prior).
- The priors impact our results. This makes sense, since the prior encodes previous knowledge and different experts have different opinions.

Discussion

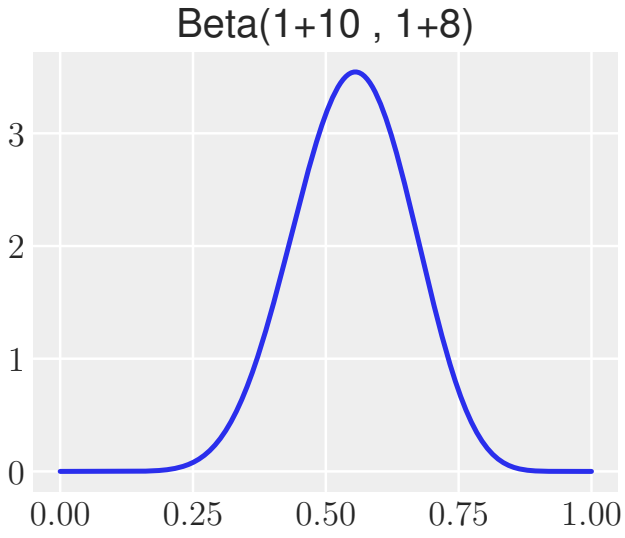
- Priors and likelihood are based on assumptions which should be justified.
- There is no single right prior, but many reasonable priors.
- *Weakly informative* priors are recommended, which for instance limit the parameter values only to the positive range, or to a certain order or magnitude, etc.
- A Beta(1,1) prior is flat but limits the values of θ between 0 and 1.
- Sensitivity analysis: evaluate if the results significantly change with the priors, or instead the data are strong enough to overwhelm the prior.

Summarizing the posterior

- Different measures can be used to summarize the posterior:
- the mean (or the mode, or the median)
- the probability of θ belonging to a certain interval
- the 95% (or 90%, 99%, etc) *credibility* interval (HDI: highest density interval)

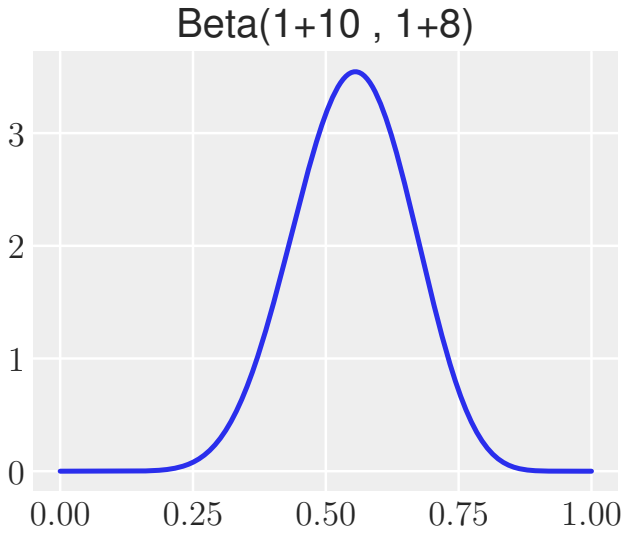
Example: inference on the bias θ of the coin

- Consider a uniform prior, Beta(1,1)
- Data: y=8 heads in n=18 tosses.



- The posterior mean is $E_{\text{post}} = \frac{1+8}{1+18} = 0.474$

Probability of the coin being almost fair



$$p_{\text{post}} = 0.49 \leq \theta \leq 0.51 = \int_{0.49}^{0.51} p_{\text{post}}(\theta) d\theta$$

```
from scipy.stats import beta
beta.cdf(0.51, 9, 19) - beta.cdf(0.49, 9, 11)
```

```
## 0.3356900915680642
```

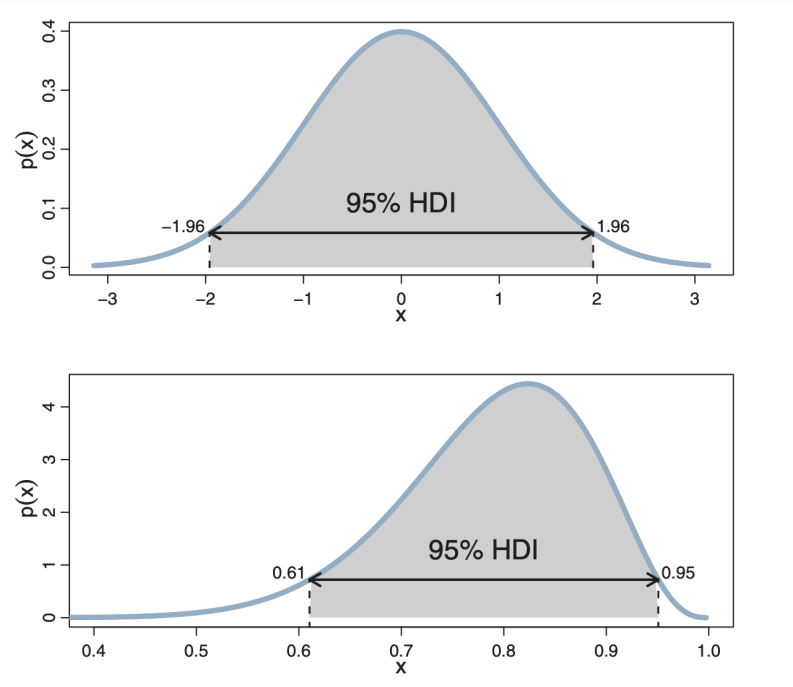
Highest Density Interval (HDI)

- The HDI is the shortest interval that contains a given portion of probability , usually 95% (or 90% or 50% are common).
- It shows the most credible points of the distribution.
- Any point within the HDI has higher density than any point outside the interval.

HDI of unimodal distributions

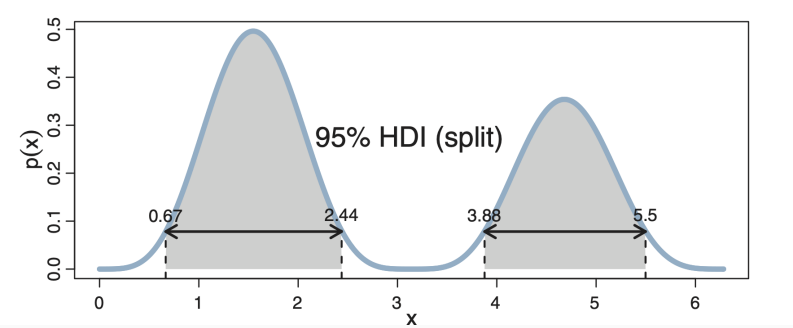
- (1-α) is the coverage of the HDI.
- If the distribution is unimodal, the HDI lies between the quantile $\alpha/2$ and $1-\alpha/2$.
- For instance, the 95% HDI lies between quantiles 0.025 and 0.975

HDI of unimodal distributions



- Any x value inside the HDI has higher density than any point outside the interval.

HDI of bimodal distribution



- The HDI can become split into two sub-intervals, one for each mode of the distribution.
- The characteristics are as before:
  - The shaded area has total area of 0.95.
  - Any x within such limits has higher probability density than any x outside the limits.
- Numerical procedures might be needed in order to compute such interval.