Bayesian inference

Bayesian Inference

Bayesian modeling $\Rightarrow$ *posterior* distribution:

- all of the information on θ, **conditionally to both the model and the data**

Bayesian Inference

Bayesian modeling $\Rightarrow$ *posterior* distribution:

- all of the information on θ, **conditionally to both the model and the data**

Summary of this *posterior* distribution ?

- center
- spread
- ...

Decision theory

Context: estimating an unknown parameter θ

Decision: choice of an "optimal" point estimator $\widehat{\theta}$

cost function: quantify the penalty associated with the choice of a particular $\widehat{\theta}$

$\Rightarrow$ minimize the cost function to choose the optimal $\widehat{\theta}$

a large number of cost functions are available: each one yields a different point estimator based on its own minimum rule

Point estimates

- ***Posterior* mean:** $\mu_P = \mathbb{E}(\theta|\boldsymbol{y}) = \mathbb{E}_{\theta|\boldsymbol{y}}(\theta)$

  not always easy because it assumes the calculation of an integral...

  ⇒ minimize the quadratic error cost

- **Maximum *A Posteriori* (MAP):**
  easy(ier) to compute: just a simple maximization of the *posterior* $f(\boldsymbol{y}|\theta)\pi(\theta)$

- ***Posterior* median:** the median of $p(\theta|(y))$

  ⇒ minimize the absolute error cost

⚠ the Bayesian approach gives a full characterization of the *posterior* distribution that goes beyond point estimation

MAP on the historical example

Maximum *A Posteriori* on the historical example of feminine birth in Paris with a uniform prior:

$$p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\theta^S(1-\theta)^{n-S}$$

with $n = 493,472$ et $S = 241,945$

$$\widehat{\theta}_{MAP} = \frac{S}{n} = 0.4902912$$

Posterior mean on the historical example

Posterior mean on the historical example of feminine birth in Paris with a uniform prior:

$$p(\theta|\boldsymbol{y}) = \binom{n}{S}(n+1)\theta^S(1-\theta)^{n-S}$$

with $n = 493,472$ et $S = 241,945$

$$E(\theta|\boldsymbol{y}) = \int_0^1 \theta p(\theta|\boldsymbol{y})\mathrm{d}\theta$$

$$\tilde{\theta} = \binom{n}{S}(n+1)\frac{S+1}{\binom{n}{S}(n+1)(n+2)} = \frac{S+1}{n+2} = 0.4902913$$

Confidence Interval reminder

What is the interpretation of a frequentist confidence interval at a 95% level ?

...