Noninformative Prior Distributions

Donald E. Brown

School of Data Science
University of Virginia
Charlottesville, VA 22904

Agenda

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$
- Uniform prior $[0, 1]$ will do this.

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$
- Uniform prior $[0, 1]$ will do this.
- Laplace's Principle of Insufficient Reason - without further knowledge, assume a uniform distribution

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$
- Uniform prior $[0, 1]$ will do this.
- Laplace's Principle of Insufficient Reason - without further knowledge, assume a uniform distribution
  - Improper

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$
- Uniform prior $[0, 1]$ will do this.
- Laplace's Principle of Insufficient Reason - without further knowledge, assume a uniform distribution
  - Improper
  - Not coherent

Priors with No Information

- Bayesian approach is still appropriate with no information to inform the prior
- Suppose we are uncertain when an event will occur: prior should give $p = 0.5$
- Uniform prior $[0, 1]$ will do this.
- Laplace's Principle of Insufficient Reason - without further knowledge, assume a uniform distribution
  - Improper
  - Not coherent
  - Not invariant under transformation

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?
- Uniform? But $f(\theta) = c$ implies $\int_{-\infty}^{\infty} f(\theta) d\theta = \infty$, i.e., improper

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?
- Uniform? But $f(\theta) = c$ implies $\int_{-\infty}^{\infty} f(\theta) d\theta = \infty$, i.e., improper
- An improper prior does not matter:

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?
- Uniform? But $f(\theta) = c$ implies $\int_{-\infty}^{\infty} f(\theta) d\theta = \infty$, i.e., improper
- An improper prior does not matter:
  - Suppose $X_1, \ldots, X_N$ sampled from $N(\mu, \sigma^2)$ with known σ^2

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?
- Uniform? But $f(\theta) = c$ implies $\int_{-\infty}^{\infty} f(\theta) d\theta = \infty$, i.e., improper
- An improper prior does not matter:
  - Suppose $X_1, \ldots, X_N$ sampled from $N(\mu, \sigma^2)$ with known σ^2
  - If $f(\mu) = 1, -\infty < \mu < \infty$, then from Bayes rule:

$$f(\mu | x_1, \ldots, x_N) \propto exp\left[\frac{-N}{2\sigma^2}(\theta - \bar{x})^2\right]$$

The M.L. solution

Improper Priors

- Suppose we have a parameter, $\theta \in \Re$. What noninformative prior will work?
- Uniform? But $f(\theta) = c$ implies $\int_{-\infty}^{\infty} f(\theta) d\theta = \infty$, i.e., improper
- An improper prior does not matter:
  - Suppose $X_1, \ldots, X_N$ sampled from $N(\mu, \sigma^2)$ with known σ^2
  - If $f(\mu) = 1, -\infty < \mu < \infty$, then from Bayes rule:

$$f(\mu | x_1, \ldots, x_N) \propto exp\left[\frac{-N}{2\sigma^2}(\theta - \bar{x})^2\right]$$

    The M.L. solution
  - Can also use a proper prior and let $\sigma^2 \to \infty$

Incoherent

- Solutions with uniform or noninformative prior do not make sense

Incoherent

- Solutions with uniform or noninformative prior do not make sense
- Example: Does life in the universe exist outside of the earth?

Incoherent

- Solutions with uniform or noninformative prior do not make sense
- Example: Does life in the universe exist outside of the earth?
  - Uniform prior would suggest $p = 0.5$.

Incoherent

- Solutions with uniform or noninformative prior do not make sense
- Example: Does life in the universe exist outside of the earth?
  - Uniform prior would suggest $p = 0.5$.
  - Need to correctly define the sample space from two events to many possible planetary events

Incoherent

- Solutions with uniform or noninformative prior do not make sense
- Example: Does life in the universe exist outside of the earth?
  - Uniform prior would suggest $p = 0.5$.
  - Need to correctly define the sample space from two events to many possible planetary events
- When well defined, noninformative priors are coherent

Not Invariant under Transformation

- Technical issue, of only modest practical importance

Not Invariant under Transformation

- Technical issue, of only modest practical importance
- Solutions

Not Invariant under Transformation

- Technical issue, of only modest practical importance
- Solutions
  - Jeffrey's prior

Not Invariant under Transformation

- Technical issue, of only modest practical importance
- Solutions
  - Jeffrey's prior
  - Reference prior

Agenda

Summary

- Prior information is not the same as a prior distribution

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions
  - Base rate - always use, but may modify

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions
  - Base rate - always use, but may modify
  - Subjective - use only when you have very good experts

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions
  - Base rate - always use, but may modify
  - Subjective - use only when you have very good experts
  - Maximum entropy - use when you need to justify your work

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions
  - Base rate - always use, but may modify
  - Subjective - use only when you have very good experts
  - Maximum entropy - use when you need to justify your work
  - Conjugate - general purpose approach that can deliver fast results

Summary

- Prior information is not the same as prior distribution
- Approaches to prior distributions
  - Base rate - always use, but may modify
  - Subjective - use only when you have very good experts
  - Maximum entropy - use when you need to justify your work
  - Conjugate - general purpose approach that can deliver fast results
  - Noninformative - good starting point