

Comparación de regresiones frecuentistas y bayesianas en muestras chicas

Santiago Amena - Sergio Marchio

Maestría en Explotación de Datos y Descubrimiento de Conocimiento, Universidad de Buenos Aires

December 6, 2021

Frecuentismo vs Bayesianismo

Modelo frecuentista:

- Parámetros fijos
- Muestreo aleatorio
- $Y \sim N(\alpha + \beta X, \sigma)$
- Dado un conjunto de parámetros, ¿cuán probable es obtener los valores encontrados en los datos?

Modelo bayesiano:

- Datos fijos
- Parámetros aleatorios
- Función de verosimilitud: $p(\theta/y) = \frac{p(y/\theta)p(\theta)}{p(y)}$
- $Y \sim N(\alpha + \beta X, \sigma)$
- $\alpha \sim N(0, \tau_\alpha)$
- $\beta \sim N(0, \tau_\beta)$
- $\sigma \sim Exp(\lambda)$
- Dados los datos, ¿cuáles son los valores de los parámetros?

Modelos bayesianos en muestras pequeñas

McNeish (2016):

- Varios trabajos muestran superioridad de métodos bayesianos sobre los frecuentistas en muestras chicas
- Sin embargo, las estimaciones también pueden ser peores
- Problema frecuentista: normalidad asintótica y consistencia requieren muchos datos
- Monte Carlo Markov Chain no requiere gran cantidad de datos
- La distribución prior provee información adicional: en muestras chicas la contribución a la posterior es mayor
- Problema: prior difusas pueden llevar a sesgos de estimación

- Casella, George. Bayesian and Frequentists. Models, Assumptions and Inference. ACCP 37th Annual Meeting, Philadelphia, PA. <http://archived.stat.ufl.edu/casella/Talks/BayesRefresher.pdf>
- Farall, Andrés. "Inferencia Bayesiana". YouTube, https://www.youtube.com/watch?v=yNpPIE2rlcl&list=PLN2e9R_DoC0ST_jbiaVEiMbA5CN6PD1Da
- McNeish, Daniel. (2016). On Using Bayesian Methods to Address Small Sample Problems. Structural Equation Modeling A Multidisciplinary Journal. 23. 10.1080/10705511.2016.1186549.