Math/Stat 310: Intro to Mathematical Statistics April 18, 2018

Lecture 18: Bayesian Estimation

Instructor: Hyunseung Kang *Scribe: Meenmo Kang*

1 Bayesian Approach to Estimation

- We want to learn about θ from n *iid* samples $X_1, ..., X_n$ from pdf_θ.
- MOM and MLE
- What if we had some prior information about θ?
- Specifically, our prior information coms in the form of a distribution function $\pi(\theta)$.
- Goal: How do we incorporate this prior information about estimator θ?

1.1 Example

Heights of UW-Madison Students n *iid* samples from $N(\mu, \sigma^2 = 3^2)$.

Our goal is to estimate μ_1 the population mean height from both n *iid* samples and prior information about μ.

Suppose we find one past distribution whose μ was 72. Then we can assume $N(72, 3^2) \Leftrightarrow$ prior distribution based on background information.

1.2 Bayes Rule

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

We can use Bayes rule to incorporate information from n *iid* samples $X_1, ..., X_n$ and prior distribution $\pi(\theta)$.

$$P(\theta|X_1, ..., X_n) = \frac{P(X_1, ..., X_n|\theta)P(\theta)}{P(X_1, ..., X_n)} \propto P(X_1, ..., X_n)P(\theta)$$

$$\text{where } P(\theta|X_1, ..., X_n) \text{ is posterior distribution or target,}$$

$$P(X_1, ..., X_n|\theta) \text{ is Likelihood function}, and$$

$$P(\theta) \text{ is prior distribution}$$

$$P(X_1, ..., X_n|\mu)P(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(3^2)}} exp\left(-\frac{(X_i-\mu)^2}{2\cdot 3^2}\right) \cdot \frac{1}{\sqrt{2\pi(3^2)}} exp\left(-\frac{(\mu-72)^2}{2\cdot 3^2}\right)$$

$$= \left(\frac{1}{\sqrt{2\pi(3^2)}}\right)^{n+1} exp\left(-\frac{1}{2\cdot 3^2}\sum_{i=1}^{n}(X_i-\mu)^2 - \frac{1}{2\cdot 3^2}(\mu-72)^2\right)$$

$$= \left(\frac{1}{\sqrt{2\pi(3^2)}}\right)^{n+1} exp\left(-\frac{1}{2\cdot 3^2}(\sum_{i=1}^{n} X_i^2 - 2\sum_{i=1}^{n}\mu X_i + n\mu^2 + (\mu^2 - 2\cdot 72\mu + 72^2)\right)$$

$$= \left(\frac{1}{\sqrt{2\pi(3^2)}}\right)^{n+1} exp\left(\frac{-\sum_{x=1}^{n} X_i^2 - 2\mu(\sum_{x=1}^{n} X_i + 72) - (n+1)\mu^2 - 72^2}{2\cdot 3^2}\right)$$

Drop all terms not involving μ

$$exp\left(\frac{-(n+1)\mu^2 + 2\mu(\sum X_i + 72)}{2\cdot 3^2}\right) = exp\left(-\frac{\mu - 2\mu(\frac{\sum X_i + 72}{n+1})}{\frac{2\cdot 3^2}{n+1}}\right)$$

$$= exp\left(-\frac{(\mu - 2\mu(\frac{\sum X_i+72}{n+1}))^2}{\frac{2\cdot 3^2}{n+1}}\right) \cdot exp\left(\frac{\sum X_i + 72}{n+1}\right)^2 / \frac{2\cdot 3^2}{n+1}$$

$$\text{Thus } P(\mu|X_1, ..., X_n) \propto exp\left(-\frac{(\mu - \frac{\sum_{i=1}^{n} X_i + 72}{n+1})^2}{\frac{2\cdot 3^2}{n+1}}\right)$$

$$\Rightarrow P(\mu|X_1, ..., X_n) \sim N\left(\frac{\sum_{i=1}^{n} X_i + 72}{n+1}, \frac{3^2}{n+1}\right)$$

Our Bayesian estimate of μ is $P(\mu|x_1, ..., X_n) \sim N(\frac{n}{n+1}\bar{X} + \frac{72}{n+1}, \frac{3^2}{n+1})$

$\Rightarrow$ Posterior parameters: $\mu = \frac{n}{n+1}\bar{X} + \frac{72}{n+1}$ $\quad\sigma^2 = \frac{3^2}{n+1}$
$\Rightarrow$ MLE: $\bar{X}$

1.3 Von-Mise Type Theorem

- As $n \to \infty$, prior information is typically irrelevant.
- As $n \to \infty$, posterior mean $\approx$ MLE
- For small n, prior information always dominates in learning about θ

1.4 A couple points about being Bayesian

- We can use distribution hyperparmeter in your prior information. (e.g. 72, 3^2)
- Bayesian estimation is computationally expensive.
- This is used a lot in consulting/marketing research.

1.5 Example

$n = 3$ *iid* samples from a distribution $P(X = 1) = \theta, P(X = 2) = 1 - \theta$
Samples are $X_1 = 1, X_2 = 2, X_3 = 2$

Suppose $\theta \sim Unif[0, 1]$ is given, but this is non-informative prior information.

What is the Bayesian estiomation of θ?

$$P(\theta|X_1, X_2, X_3) \propto P(X_1, X_2, X_3|\theta)P(\theta) = \theta(1-\theta)^2 \cdot \frac{1}{1} = \theta(1-\theta)^2$$

This is the Beta distribution

$$P(\theta|X_1, X_2, X-3,) \sim Beta(2, 3)$$

$\Rightarrow$ Posterior Mean: $\frac{\alpha}{\alpha+\beta} = \frac{2}{2+3} = \frac{2}{5}$
$\Rightarrow$ MLE & MOM : $\frac{1}{3}$