STAT 536 : Class 3

Last class : Review of Basics

→ Expected Values and Variances

→ Distributions

→ discrete

→ Continuous.

Continue review :

Law of large numbers (LLN) :

E.g : Suppose I have a coin which is biased

$$X = \begin{cases} 0 & \text{tails} \quad (1-p) \\ 1 & \text{heads} \quad p \end{cases}$$

(say, $p = 0.75$)

$E(X) = p$.

toss this coin 10 times :

$X_1 = 1,\ X_2 = 1,\ X_3 = 0 \ldots\ldots \quad X_{10} = 0$

{7 ones, 3 zeros}.

$$\hat{p} = \frac{\sum X_i}{n} = \frac{7}{10} = 0.7.$$

$n \to$ no. of tosseses

LLN: Let $X_1, X_2 \ldots X_n$ are independent and identically distributed (i.i.d) r.v.'s with $E|X_i| < \infty$, then

$$\underbrace{\bar{X} = \frac{\sum X_i}{n}}_{} \to \underbrace{E(X)}_{(\mu)} \quad \text{with probability} \to 1$$

Usefulness:

$$h(x) = (\cos 50x + \sin 50x)^2$$

$$\int_0^1 h(x)\,dx \to \text{interested in.}$$

if we define $X \sim U(0,1)$.

$$E[h(X)] = \int_0^1 h(x) f(x)\,dx$$

$=$

$$= \int_0^1 h(x)\,dx$$

if we can simulate an endless supply of $U(0,1)$ r.v.'s $[X_1, X_2 \ldots X_n]$

$$\int_0^1 h(x)\,dx = E(h(X)) \approx \frac{\sum_{i=1}^{n} h(X_i)}{n}$$

Some more motivation for this course (why do we about computing integrals and optimization?)

Maximum likelihood estimation :

E.g: In previous ex:

$X_1 = 1,\ X_2 = 1 \ldots \ldots X_{10} = 0.$

$\{\cancel{4}^{3}$ zeros, $\cancel{8}^{7}$ ones$\}$

$\hat{p} = 0.7$

we know $p = 0.2$ or 0.8

what is your guess? (0.8).

(0.8) provides a higher likelihood of observing the given data.

(This is the basic idea of an MLE)

Def: Given a random sample

$$X_1, X_2 \ldots X_n \overset{i.i.d}{\sim} f(x \mid \theta)$$

the likelihood of $X_1 \ldots X_n$

$$L(\theta \mid X) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

Def: MLE : the value of θ the maximizes the likelihood.

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \ L(\theta \mid X)$$

Note: to Simplify calculations it is convenient to instead look at

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} \ \ell(\theta \mid X)$$

$$\ell(\theta \mid X) = \log L(\theta \mid X)$$

$$=$$

Ex: Suppose we observe

$$X_1 \ldots X_n \sim Exp(\lambda)$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & , x > 0 \\ & \end{cases}$$

Compute MLE of λ :

$$L(\lambda|x) = \prod_{i=1}^{n} f(x_i|\lambda).$$

$$= \prod_{i=1}^{n} \lambda e^{-\lambda x_i}$$

$$= \lambda^n e^{-\lambda \Sigma x_i}$$

$$\log L(\lambda|x) = n \log \lambda - \lambda \Sigma x_i$$

$$\frac{\partial l}{\cancel{\partial \theta} \partial \lambda} = 0 \rightarrow \text{find solution}$$

$$\hat{\lambda} = \frac{\Sigma x_i}{n} = \bar{x}$$

$$\hat{\lambda} = \underset{\lambda}{\operatorname{argmax}}\ l(\lambda|x)$$

Ex: (when no analytical Solution exists)

(Mixture of ~~r.v.'s~~ normal r.v.'s).

Suppose $X_1, \ldots, X_n$ are generated from one of two normal distributions $N(\mu_1, \sigma_1^2)$ or $N(\mu_2, \sigma_2^2)$ with probability p or $(1-p)$, respectively

each X_i has density:

$$\rightarrow f(x) = p f_1(x) + (1-p) f_2(x).$$

$$f_1(x) \sim N(\mu_1, \sigma_1^2)$$

$$f_2(x) \sim N(\mu_2, \sigma_2^2).$$

$Z \sim N(0,1)$

$\mu + \sigma Z \sim N(\mu, \sigma^2)$.

Compute MLE of $p, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2$

$$\theta = (p, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2).$$

$$L(\theta | X) = \prod_{i=1}^{n} f(x | \theta).$$

$$\Rightarrow \left[= \prod_{i=1}^{n} \left[\frac{p}{\sigma_1} \phi\left(\frac{X_i - \mu_1}{\sigma_1} \right) + \frac{(1-p)}{\sigma_2} \phi\left(\frac{X_i - \mu_2}{\sigma_2} \right) \right] \right.$$

$\phi \rightarrow$ density of a $N(0,1)$

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left[\frac{-x^2}{2} \right].$$

$$\hat{\theta} = \underset{\theta}{\text{argmax}} \sum_{i=1}^{n} \log \left[\quad \right]$$

alternatively,

$$L(\theta | X) = \prod_{i=1}^{n} p f_1(x_i) + (1-p) f_2(x_i).$$

$$f_1(x_i) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left[- \frac{(x_i - \mu_1)^2}{2\sigma_1^2} \right].$$