Maximum Likelihood Estimation

Statistics: about analyzing & understanding data

Common approach: use parametric model of data

$Bin(n,p)$, $Poi(\lambda)$, $Exp(\lambda)$, $N(\mu,\sigma^2)$....

How to estimate params of distn from samples?

Approach: Find MLE, most likely choice of parameters.

Ex: outcome of n coin tosses. $x_1,\dots,x_n$ each $x_i \in \{H,T\}$

assuming indep $Ber(p)$ what's p?

$$L(x_1,\dots,x_n \mid \theta) = \prod_{i=1}^{n} \underbrace{f(x_i \mid \theta)}$$

unknown parameters (in this ex, prob of H's)

$\approx$ Prob of seeing x_i if param is θ

discrete case p.m.f.

cont case p.d.f.

$x_1,\dots,x_n$ n_1 H's n_0 T's $n = n_0 + n_1$

$$L(x_1,\dots x_n \mid \theta) = \theta^{n_1}(1-\theta)^{n_0}$$

Find θ that maximizes this

$$(LL)\ \log\text{Likelihood}(x_1,\dots x_n \mid \theta) = \log L(x_1,\dots x_n \mid \theta) = n_1 \log\theta + n_0 \log(1-\theta)$$

$$\frac{d}{d\theta} LL(\quad) = \frac{n_1}{\theta} - \frac{n_0}{1-\theta}$$

at max $\hat{\theta}$

$$\frac{n_1}{\hat{\theta}} = \frac{n_0}{1-\hat{\theta}}$$

$$n_1(1-\hat{\theta}) = n_0\hat{\theta}$$

$$n_1 = (n_0+n_1)\hat{\theta}$$

$$\Rightarrow \hat{\theta} = \frac{n_1}{n_0+n_1} = \frac{n_1}{n}$$

In general

$$L(x_1,\ldots,x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

$$LL = \sum_{i=1}^{n} \log f(x_i|\theta)$$

find max $LL(x_1,\ldots x_n|\vec{\theta})$

$$\frac{d}{d\theta} LL(\vec{x}|\theta) = 0$$

solve for $\hat{\theta}$

$$\frac{\partial}{\partial\theta_1} LL(\vec{x}|\vec{\theta}) = 0$$

$$\frac{\partial}{\partial\theta_2} LL(\vec{x}|\vec{\theta}) = 0$$

$\vdots$

---

samples from a geometric distn w/ unknown param θ

$k_1, k_2, \ldots, k_n$

$$L(k_1,\ldots,k_n|\theta) = \prod_{i=1}^{n} \underbrace{(1-\theta)^{k_i-1}\theta}$$

$$LL(k_1\ldots k_n|\theta) = \sum_{i=1}^{n} \left[(k_i-1)\log(1-\theta) + \log\theta\right]$$

$$\frac{d}{d\theta} LL = \sum_{i=1}^{n}\left[-\frac{(k_i-1)}{1-\theta} + \frac{1}{\theta}\right]$$

set this $=0$ to solve $\hat{\theta}$

$$\sum_{i=1}^{n} \frac{k_i-1}{1-\theta} = \frac{n}{\theta}$$

$$\sum_{i=1}^{n} (k_i-1)\hat{\theta} = n(1-\hat{\theta})$$

$$\sum_{i=1}^{n} k_i\hat{\theta} = n$$

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n} k_i}$$

$x_1, \dots, x_n$ samples from $N(\underset{\substack{\uparrow \\ \text{mean}}}{\theta_1}, \underset{\substack{\uparrow \\ \text{variance}}}{\theta_2})$.

$$L(x_1, \dots, x_n \mid \theta_1, \theta_2) = \prod_{i=1}^{n} f(x_i \mid \theta_1, \theta_2)$$

$$= \prod_{i=1}^{n} \boxed{\frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}}}$$

density of $N(\theta_1, \theta_2)$ at x_i

$$LL(\vec{x} \mid \theta_1, \theta_2) = \sum_{i=1}^{n} \left[-\frac{1}{2}\ln(2\pi\theta_2) - \frac{(x_i-\theta_1)^2}{2\theta_2} \right]$$

$$\frac{\partial LL}{\partial \theta_1} = 0 \qquad \frac{\partial LL}{\partial \theta_2} = 0 \qquad \Rightarrow \hat{\theta}_1, \hat{\theta}_2$$

$$\frac{\partial LL}{\partial \theta_1} = \sum_{i=1}^{n} \frac{\cancel{2}(x_i-\theta_1)}{\cancel{2}\theta_2}$$

set $= 0$

$$\sum_{i=1}^{n} \frac{x_i}{\cancel{\theta_2}} = \frac{n\hat{\theta}_1}{\cancel{\theta_2}}$$

$$\hat{\theta}_1 = \frac{\sum_{i=1}^{n} x_i}{n}$$

sample mean

$$\frac{\partial LL}{\partial \theta_2} = \sum_{i=1}^{n} \left[-\frac{1}{2} \frac{\cancel{2\pi}}{\cancel{2\pi}\theta_2} + \frac{1}{2}\frac{(x_i-\theta_1)^2}{\theta_2^2} \right]$$

set $= 0$

$$\frac{n}{\cancel{2}\hat{\theta}_2} = \frac{1}{\cancel{2}} \sum_{i=1}^{n} \frac{(x_i-\hat{\theta}_1)^2}{\hat{\theta}_2^2}$$

$$\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\theta}_1)^2$$

sample variance

$$E\left(\frac{\sum X_i}{n}\right) = \text{true mean.}$$

An estimator $\hat{\theta} = g(x_1, \dots x_n)$ is called <u>unbiased</u> if

$E(\hat{\theta}) =$ true parameter

<u>Question</u>: If X_i are from dist^n w/ mean μ, var σ^2 is $E[\text{sample var}] = \sigma^2$?

$$\hat{\theta}_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\theta}_1)^2$$

sample variance

$$E(\hat{\theta}_2(X_1, \dots, X_n)) = E\Big[\underbrace{\frac{1}{n}\sum_{i=1}^{n}\Big(X_i - \frac{\sum X_i}{n}\Big)^2}_{\bar{X}}\Big]$$

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 2X_i\bar{X} + \bar{X}^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\boxed{\frac{1}{n}\sum_{i=1}^{n}X_i}\bar{X} + \bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i^2 - 2\bar{X}^2 + \bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}^2$$

$$E\Big[\frac{1}{n}\sum_{i=1}^{n}X_i^2 - \bar{X}^2\Big] = \frac{1}{n}\sum_{i=1}^{n}E(X_i^2) - \frac{1}{n^2}E\big((\sum X_i)^2\big)$$

$$\underbrace{\mathrm{Var}(X_i)}_{\sigma^2} = E(X_i^2) - \underbrace{(E(X_i))^2}_{\mu^2}$$

$$= \frac{1}{n}n(\sigma^2 + \mu^2)$$

$X_1, \dots X_n$

i.i.d. μ, σ^2

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = ?$$

random exp.

$x_1, \dots, x_n$. 1 exp

estimate variance

$$\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{\sum x_i}{n}\right)^2$$