Predictive Analytics using Statistics : (UCS 654)

Assignment → Parameter Estimation

AMAN VERMA
102103043
3-CO2

1) Given $(x_1, x_2 \ldots)$ → random sample of size n (from normal population)

- mean $= \theta_1$, variance $= \theta_2$
- To find max likelihood estimates of these two parameters.

soln) The likelihood function $L(\theta_1, \theta_2)$ is given by:

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

To maximize this function, we will take natural logarithm of $L(\theta_1, \theta_2)$

$$l(\theta_1, \theta_2) = \log_e\left(\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta_2}} e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}}\right)$$

$$\Rightarrow \sum_{i=1}^{n}\left(\ln \frac{1}{\sqrt{2\pi\theta_2}} + \ln e^{-\frac{(x_i-\theta_1)^2}{2\theta_2}}\right)$$

$$\Rightarrow \sum_{i=1}^{n} \ln\left(2\pi\theta_2\right)^{-1/2} - \frac{(x_i-\theta_1)^2}{2\theta_2}$$

$$l(\theta_1, \theta_2) \Rightarrow \sum_{i=1}^{n}\left(-\frac{1}{2}\ln(2\pi\theta_2) - \frac{(x_i-\theta_1)^2}{2\theta_2}\right)$$

- Now we need to maximise log likelihood wrt θ_1 & θ_2 separately.

Spiral

• wrt θ_1 :

• Differentiate $l(\theta_1, \theta_2)$ wrt θ_1 & equate with 0

Date...............

$$\frac{\partial l}{\partial \theta_1} = \sum_{i=1}^{n} \cancel{-}\frac{1}{\cancel{2}\theta_2}\left(\cancel{2}\times(x_i-\theta_1)\right)(\cancel{-1}) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (x_i-\theta_1) = 0$$

$$\sum_{i=1}^{n} x_i - n\theta_1 = 0$$

$$\theta_1 = \left(\sum_{i=1}^{n} x_i\right) / n$$ (max likelihood for θ_1 is sample mean.)

• wrt θ_2

• Differentiate $l(\theta_1, \theta_2)$ wrt θ_2 & equate with 0

$$\frac{\partial l}{\partial \theta_2} = \sum_{i=1}^{n}\left(\left(\frac{-1}{2} \times \frac{1}{\cancel{2\pi}\theta_2} \times \cancel{2\pi}\right) - \frac{(x_i-\theta_1)^2}{2\theta_2^2}\times(-1)\right)$$

$$\Rightarrow \sum_{i=1}^{n}\left(\frac{-1}{2\theta_2} + \frac{(x_i-\theta_1)^2}{2\theta_2^2}\right) = 0$$

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

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

$$\theta_2 = \frac{1}{n}\sum_{i=1}^{n} (x_i-\theta_1)^2$$ (sample variance)

Thus, max likelihood estimates are :

$\theta_1 = \frac{1}{n}\sum_{i=1}^{n} x_i$ (mean)	$\theta_2 = \frac{1}{n}\sum_{i=1}^{n} (x_i-\theta_1)^2$ (sample variance)

• $n \rightarrow$ unbiased estimator

Spiral

(2) Given, $X_1, X_2, \ldots X_n \to$ random sample from Binomial dist. (m, θ).

Date..................

$\to \theta \in \Theta = (0,1)$, $m \to$ +ve integer

∘ To find θ using MLE

Soln) For binomial dist. $B(m,\theta)$, likelihood func. is given by $\to$

$$L(\theta) = \prod_{i=1}^{n} \binom{m}{x_i} \theta^{x_i} (1-\theta)^{m-x_i}$$

Taking logarithm of the likelihood function $\to$

$$l(\theta) = \sum_{i=1}^{n} \left[\ln\binom{m}{x_i} + x_i \ln(\theta) + (m-x_i) \ln(1-\theta) \right]$$

To find ~~Max~~ MLE of θ, we take derivative of log likelihood w.r.t θ & equate with 0

$$\frac{dl}{d\theta} = \sum_{i=1}^{n} \left[\frac{x_i}{\theta} + \frac{(m-x_i)}{(1-\theta)} \times -1 \right]$$

$$\Rightarrow \sum_{i=1}^{n} \left[\frac{x_i}{\theta} - \frac{(m-x_i)}{1-\theta} \right] = 0$$

$$\sum_{i=1}^{n} \frac{x_i}{\theta} = \sum_{i=1}^{n} \frac{(m-x_i)}{1-\theta}$$

$$\Rightarrow \frac{\sum_{i=1}^{n} x_i}{\theta} = \frac{\sum_{i=1}^{n} m - \sum_{i=1}^{n} x_i}{1-\theta}$$

$$\frac{\sum_{i=1}^{n} x_i}{\theta} = \frac{nm - \sum_{i=1}^{n} x_i}{1-\theta} \Rightarrow (1-\theta)\sum_{i=1}^{n} x_i = \theta nm - \theta \sum_{i=1}^{n} x_i$$

Spiral

Date...................

$$\Rightarrow \sum_{i=1}^{n} x_i - \theta \cancel{\sum_{i=1}^{n} x_i} = \theta \cdot n \cdot m - \theta \cancel{\sum_{i=1}^{n} x_i}$$

$$\Rightarrow \theta = \frac{\sum x_i}{nm} \Rightarrow \frac{\bar{x}}{m}$$

$\therefore$ MLE of θ is $\bar{x}/m$, & ranges from $0, 1$

—— $*$ ——