14/02/21 ML_EMAP.

Q1) Let Y_i be the i.i.d random variables which follow the distribution that has density $\frac{\alpha y^{\alpha-1}}{\theta^\alpha} \exp\left(-\left(\frac{y}{\theta}\right)^\alpha\right), \alpha > 0, \theta.$

Suppose that α is known, θ is unknown, Find the maximum likelihood of θ.

Soln: Given,

Probability density function.

$$f\left(\frac{y}{\theta}\right) = \frac{(\alpha y^{\alpha-1})\left(e^{-\left(\frac{y}{\theta}\right)^\alpha}\right)}{\theta^\alpha}, \quad \theta, \alpha > 0.$$

Now, taking the log likelihood of above function w.r.t θ, we get.

$$\Rightarrow L(\theta) = \sum_{i=1}^{n} \log\left(f\left(\frac{y_i}{\theta}\right)\right)$$

$$L(\theta) = \sum_{i=1}^{n}\left(\log\left(\frac{(\alpha y_i^{\alpha-1})\left(e^{-\left(\frac{y_i}{\theta}\right)^\alpha}\right)}{\theta^\alpha}\right)\right).$$

$$L(\theta) = \sum_{i=1}^{n}\left(\log\alpha + (\alpha-1)\log(y_i) - \left(\frac{y_i}{\theta}\right)^\alpha - \alpha\log(\theta)\right)$$

Maximizing, $L(\theta)$ w.r.to θ we get,

$$\Rightarrow L'(\theta) = 0.$$

Anirudh Jakhotia Roll No: S20190010007

$$\rightarrow l'(\theta) = \sum_{i=1}^{n} \left(\theta + 0 + \alpha\left(\frac{y_i^{\alpha}}{\theta^{\alpha+1}}\right) - \left(\frac{\alpha}{\theta}\right)\right) = 0$$

$$\sum_{i=1}^{n} \left(\alpha\left(\frac{y_i^{\alpha}}{\theta^{\alpha+1}}\right) - \left(\frac{\alpha}{\theta}\right)\right) = 0$$

$$\cancel{\alpha} \sum_{i=1}^{n} \frac{y_i^{\alpha}}{\theta^{\alpha+1}} = \frac{n\cancel{\alpha}}{\theta}$$

$$\frac{\sum_{i=1}^{n} y_i^{\alpha}}{\theta^{\alpha} \cdot \cancel{\theta}} = \frac{n}{\cancel{\theta}}$$

$$\hat{\theta} = \sqrt[\alpha]{\frac{\sum_{i=1}^{n} y_i^{\alpha}}{\theta^{\alpha}}}$$

There $\hat{\theta}$, is maximum likelihood estimation of function given.

Q2) This problem is regarding the estimation of yield for a green gas process. One process of making green gasoline takes biomass in the form of sucrose and connects it into gasoline using catalytic reactions. At one step in a pilot plan process, the output includes carbon chains of length 3.

Anirudh Jakhotia Roll No:- S20190010007

Fifteen runs with same catalyst produced the yields (gal) $\{5.57, 5.76, 4.18, 4.64, 7.02, 6.62, 6.33, 7.24, 5.57, 7.89, 4.67, 7.24, 6.43, 5.59, 5.39\}$.

(a) Obtain the MLE of mean yield and the variance.

(b) obtain the MLE of the coeffecients of variation (σ/μ).

Soln The maximum likelihood estimate for the unknown population mean is just the arithmetic mean and is same for variance and coeffecient of variance.

(i) Mean $= \mu = \frac{\sum_{i=1}^{n} f(x_i)}{n}$

$$= \mu = \frac{(5.57 + 5.76 + 4.18 + 4.64 + 7.02 + 6.62 + 6.33 + 7.24 + 5.57 + 7.89 + 4.67 + 7.24 + 6.43 + 5.59 + 5.39)}{15}$$

$$\mu = \frac{90.14}{15} = 6.009.$$

Variance $= \sigma^2 = \frac{\sum_{i=1}^{n} f(x_i - \mu)^2}{n}$

$$= \frac{(16.263)}{15} = 1.0412.$$

Anirudh Jakhotia Roll No:- S20190010007

(b) Co-effecient of variance $= \frac{\sigma}{\mu} = \frac{1.0412}{6.009}$

$= 0.173.$

3Q) Suppose a unfair coin is flipped 100 times, and 52 heads are observed. what is maximum likelihood estimation (MLE)?

Sol: The probability density fn can be written as

$$P\left(\frac{52 \text{ heads}}{\theta}\right) = {}^{100}C_{52}\, \theta^{52} (1-\theta)^{48}$$

Taking log likelihood w.r.t θ, we get.

$$L(\theta) = \log\left(P\left(\frac{52 \text{ heads}}{\theta}\right)\right)$$

$$L(\theta) = 52\log(\theta) + 48\log(1-\theta)$$

Maximizing $L(\theta)$ w.r.to θ, we get,

$\Rightarrow L'(\theta) = 0.$

$$\Rightarrow L'(\theta) = \frac{52}{\theta} - \frac{48}{1-\theta} = 0$$

$$\frac{52}{\theta} = \frac{48}{1-\theta}$$

$$\hat{\theta} = 0.52.$$

therefore, MLE $= \underline{0.52}$

Anirudh Jakhotia

Roll NO: S20190010007

Q4) For $P(\theta|D)$ with Beta and Bernoulli distributions of θ and D respectively, where $D = \{n \text{ heads and } m \text{ tails}\}$, find a maximum a posterior (MAP) of θ. i.e., $\hat{\theta}_{map}$ [Hints: Find $f_{\theta|D}(\theta = P|D)$ and then apply $\hat{\theta}_{MAP} = \arg\max_\theta f(\theta|D)$].

Sol: $X \sim$ Bernoulli (θ)

$P(X|\theta) = \theta^x (1-\theta)^{1-x} \quad (0 < \theta < 1)$

$x \sim$ Beta (θ)

$P(\theta) = \text{Beta}(\theta|a,b) = 1/\text{Beta}(a,b)$

$= \theta^{a-1}(1-\theta)^{b-1}$

$(0 < \theta < 1)$

$$P(\theta|D) = \frac{P(D|\theta) \cdot P(\theta)}{P(D)}$$

Prior: $\theta \sim \text{Beta}(a,b)$; $D = \{n \text{ heads}, m \text{ tails}\}$

$$f_{\theta|D}(\theta = P|D) = \frac{f_{(P|\theta)}\left(\left(\frac{D}{\theta}\right) = P\right) f_\theta(\theta = P)}{f_D(D)}$$

$$= \frac{{}^{n+m}C_n \, P^n (1-P)^m \, P^{a-1} (1-P)^{b-1}}{C_1 C_2}$$

where C_1, C_2 are constants.

Anirudh Jakhotia S20190010007

$C_1 = \text{Beta}(a, b)$, $C_2 = f_P(D)$

$\Rightarrow C_3 \, p^{n+a-1} (1-p)^{m+b-1}$

where C_3 is constant

C_3 is ${}^{n+m}C_n / C_1 C_2$

estimates P

$\therefore \theta_{MAP} = \arg\max_\theta f\left(\frac{\theta}{D}\right)$

$\Rightarrow \arg\max_\theta (n+a-1)\log(\theta) + (m+b-1)\log(1-\theta)$

by definition ; $f\left(\frac{\theta}{D}\right)$ is $\text{Beta}(a+n, b+m)$

5) Let X be a continuous random variable, with following PDF :

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

Also, suppose that $Y|X = x \sim \text{Geometric}(x)$.

Find MAP estimate of x given $Y = 3$.

Sol) we know that $Y|X = x \sim \text{Geometric}(x)$

so, $P_{Y|X}(y|x) = x(1-x)^{y-1}$.

Given $y = 3$

$$P_{(Y|X)}(3|x) = x(1-x)^2 \text{ at } x \in [0,1]$$

we need to find the value $x \in [0,1]$, that maximizes

Anirudh Jakhotia S20190010007.

$\Rightarrow P_{(\frac{y}{x})}(\frac{y}{x})\, f_x(x) = x(1-x)^2 \times 2x$

$\Rightarrow P(\frac{y}{x})(\frac{y}{x})\; f_x(x) = 2x^2(1-x)^2$

To maximize the above function, we have to differentiate it and equate it to zero, by doing so, we get.

$$\Rightarrow \frac{d}{dx}\left[2x^2(1-x)^2\right] = 2\left[2x(1-x)^2 - 2x^2(1-x)\right] = 0.$$

$\hat{x}_{map} = 0.5$.

$\therefore$ map estimate of X given $Y=3$ is 0.5.

Anirudh Jakhotia Roll :- S20190010007.