DATE

SML Assignment - 4

Name - Shreya Sharma   Roll no - 2015096

Ans 4. $\bar{x}$: d-dimensional binary vector
distribution of x is bernoulli -

$$P(\bar{x}/\bar{\theta}) = \prod_{i=1}^{d} \theta_i^{x_i} (1-\theta_i)^{1-x_i}$$

where $\bar{\theta} = (\theta_1, \dots, \theta_d)^T \rightarrow$ unknown parameter vector

θ_i : probability that $x_i = 1$

To prove : Maximum likelihood estimate for $\bar{\theta}$ is $\hat{\theta} = \frac{1}{n} \sum_{k=1}^{n} x_k$

Proof :
Our sample contains n different x_k's. each of which has a bernoulli distribution

$$P(x_k \mid \bar{\theta}) = \prod_{i=1}^{d} \theta_i^{x_{ki}} (1-\theta_i)^{1-x_{ki}}$$

likelihood for all the samples is given by :-

$$l(p) = \prod_{k=1}^{n} \prod_{i=1}^{d} \theta_i^{x_{ki}} (1-\theta_i)^{1-x_{ki}}$$

$$= \prod_{i=1}^{d} \theta_i^{\sum_{k=1}^{n} x_{ki}} (1-\theta_i)^{n - \sum_{k=1}^{n} x_{ki}}$$

taking log to find log-likelihood -

$$\text{Log}(l(p)) = L(p) = \sum_{k=1}^{n} \sum_{i=1}^{d} x_{ki} \ln \theta_i + \sum_{i=1}^{d} \left(n - \sum_{k=1}^{n} x_{ki}\right) \ln(1-\theta_i)$$

DATE

differentiating component by component wrt to θ_i's to get MLE :-

$$L'(\theta)_i = \frac{d}{d\theta_i}\left(\sum_{k=1}^{n} x_{ki} \ln\theta_i + \left(n - \sum_{k=1}^{n} x_{ki}\right)\ln(1-\theta_i)\right)$$

$$= \sum_{k=1}^{n} x_{ki} \frac{1}{\theta_i} + (-1)\frac{n - \sum_{k=1}^{n} x_{ki}}{(1-\theta_i)}$$

$$L'(\theta)_i = 0$$

$$\Rightarrow \frac{1}{\theta_i}\sum_{k=1}^{n} x_{ki} = \frac{n - \sum_{k=1}^{n} x_{ki}}{1-\theta_i}$$

$$\Rightarrow (1-\theta_i)\sum_{k=1}^{n} x_{ki} = \left(n - \sum_{k=1}^{n} x_{ki}\right)\theta_i$$

$$\Rightarrow \theta_i\left(n - \sum_{k=1}^{n} x_{ki} + \sum_{k=1}^{n} x_{ki}\right) = \sum_{k=1}^{n} x_{ki}$$

$$\Rightarrow n\theta_i = \sum_{k=1}^{n} x_{ki} \Rightarrow \boxed{\hat{\theta}_i = \frac{\sum_{k=1}^{n} x_{ki}}{n}}$$

$\hat{\theta}_i$: MLE of θ_i

since the result will be the same for $i=1$ to d, we can write the equation in vector from as :-

$$\boxed{\hat{\theta} = \frac{1}{n}\sum_{k=1}^{n} x_k}.$$

DATE

Ans 10.

(a) novel method to estimate mean of a same = take the 1st value

ie $M = x_1$.

Bias - difference between the estimated value of the statistic & the true value of the statistic.

~~Let E(M) estimated value, M: true value~~

let μ : true value of mean.

$E(M)$: expected value of mean.

where $M = x_1$

We assume that the random variables are a random sample from the same distribution with mean μ.

this means that expected value of each random variable is μ.

$\therefore \; E(M) = \underline{E(x_1) = \mu}$.

Since the expected value of the statistic ~~mat~~ matches the parameter that is estimated, this means it is an unbiased estimator.

DATE

(b) why is the method undesirable?

even though the method is unbiased, the variance will be large for larger values of n and that is not desirable.

as $\sigma^2 = E[(x_i-\mu)^2]$: variance.
and the root mean squared error is σ which is independent of n in this method. There is no guarantee that this error will approach 0 at all even if n is increased.

whereas, the conventional method to find mean $\Rightarrow \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ shows the desired behavior by ~~gives~~ reducing the RMS error as n approaches ∞ :-

$$E[(\bar{x}-\mu)^2] = E\left[\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \mu\right)^2\right].$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[E[(x_i-\mu)^2]\right] = \frac{\sigma^2}{n}$$

$$\text{RMS error} = \frac{\sigma}{\sqrt{n}}$$

and we can see that, as $n \to \infty$
RMS error $\longrightarrow 0$.

Scanned by CamScanner