Course: Machine Learning - Foundations
Week 12 Questions

TUTORIALS

1. (1 point) Let X be random variable of binomial(n, p). Using Markov's inequality, find an upper bound on $P(X \geq \alpha n)$, where $p < \alpha < 1$. Evaluate the upper bound for $p = \frac{1}{3}$ and $\alpha = \frac{3}{4}$.

   A. $\frac{2}{3}$
   B. $\frac{4}{9}$ ✓ ✗
   C. $\frac{3}{5}$
   D. $\frac{4}{11}$

[**Answer:** A]

Here X is a non-negative random variable &

$E(X) = np$

Applying Markov's inequality we get,

$$P(X \geq \alpha n) \leq \frac{E(X)}{\alpha n} = \frac{np}{\alpha n}$$

$$\leq \frac{p}{\alpha}$$

$$\leq \frac{1}{3} \cdot \frac{4}{3}$$

$$\leq \frac{4}{9}$$

If Y follows $N_3(\mu, \sum)$, where Y is a vector that is, $Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$ and $\mu = \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$ and

$$\sum = \begin{pmatrix} 6 & 1 & -2 \\ 1 & 13 & 4 \\ -2 & 4 & 4 \end{pmatrix}$$

From the above information answer questions

2. (points) Suppose $Z = CY$, where $C = (2, -1, 3)$. Find the distribution of $Z = 2y_1 - y_2 + 3y_3$

   A. $N(17, 21)$
   B. $N(17, 15)$
   C. $N(15, 17)$ x
   D. $N(17, 17)$

$E(Z), V(Z)$

Sol:- $Z = CY$

$E(CY) = CE(Y)$

$= (2, -1, 3) \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}$

$= (6 - 1 + 12)$

$= 17$

Answer: A

$V(Z) = V(CY)$

$= C^2 V(Y)$

$= C V(Y) C^T$

$= \underset{1 \times 3}{(2, -1, 3)} \underset{3 \times 3}{\begin{pmatrix} 6 & 1 & -2 \\ 1 & 13 & 4 \\ -2 & 4 & 4 \end{pmatrix}} \underset{3 \times 1}{\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}}$

$= [5 \quad 1 \quad 4] \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = 21$ Ans

3. (points) In random sampling from normal distribution $N(\mu, \sigma^2)$, find the maximum likelihood estimators for σ^2 when μ is known

   A. $\sigma = \frac{\sum(x_i - \mu)^2}{n}$

   B. $\cancel{\sigma^2} = \frac{\sum(x_i - \mu)^2}{n}$

   C. $\sigma^2 = \frac{\sum(x_i - \mu)}{n}$

   D. $\sigma^2 = \frac{\sum(x_i - \mu)^2}{2n}$

Answer: B

$N(\mu, \sigma^2)$

$$L = \prod^{n}\left[\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right)\right]$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\sum_{i=1}^{n}(x_i-\mu)^2 / 2\sigma^2\right)$$

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2$$

$$\frac{\partial}{\sigma^2}\log L = 0 \Rightarrow -\frac{n}{2}\times\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum(x_i-\mu)^2 = 0$$

$$\Rightarrow n = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2 = 0$$

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

$$P(|X-2| \le 2) = P(X = 1,2,3,4) = \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 = \frac{15}{16}$$

4. For geometric distribution $p(x) = 2^{-x}; x = 1, 2, 3, \ldots\ldots.,$. Then using Chebychev's inequality calculate $P(|X-2| \le 2) \ge \frac{1}{2}$. and also calculate the actual probability and compare it's values.

$$P\{|X - E(X)| \le K\sigma\} > 1 - \frac{1}{K^2}$$

$$V(X) = E(X^2) - (E(X))^2 = 6 - 4 = 2$$

$$\Rightarrow E(X) = \sum_{1}^{\infty} x \cdot f(x) = \frac{x}{2^x} = \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3}$$

$$= \frac{1}{2}(1 + 2A + 3A^2 + \ldots)$$

$$= \frac{1}{2}(1-A)^{-2} \qquad \left(A = \frac{1}{2}\right) \quad E(X) = 2$$

$$E(X^2) = \sum_{x=1}^{\infty} \frac{x^2}{2^x} = \frac{1}{2^2} + \frac{4}{2^3} + \frac{9}{2^4} + \ldots$$

$$= \frac{1}{4}(1 + 4A + 9A^2 + \ldots)$$

$$= \frac{1}{4}(1+A)(1-A)^{-3} = 6$$

$$P(\{|X - E(X)| \le K\sigma\}) > 1 - \frac{1}{K^2}$$

$$K = \sqrt{2} \quad P((|X - E(X)| \le 2)) > 1 - \frac{1}{2} = \frac{1}{2}$$

5. (points) In random sampling from normal distribution $N(\mu, \sigma^2)$, find the maximum likelihood estimators for μ when σ^2 is known.

$$L = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} (x_i - \mu)^2 \right) \right]$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}} \right)^n \exp\left\{ -\sum_{i=1}^{n} (x_i - \mu)^2 / 2\sigma^2 \right\}$$

$$\log L = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_1^n (x_i - \mu)^2$$

$$\frac{\partial \log L}{\partial \mu} \Rightarrow -\frac{1}{2\sigma^2} \sum_1^n 2(x_i - \mu)(-1) = 0$$

$$\Rightarrow \sum_1^n (x_i - \mu) = 0$$

$$\Rightarrow \mu = \frac{1}{n} \sum x_i = \bar{x}$$

$$\boxed{\mu = \bar{x}}$$