Course: Machine Learning - Foundations
Week 12 Questions

PRACTICE QUESTIONS

1. A biased coin, which lands heads with probability $\frac{1}{10}$ each time it is flipped, is flipped 200 times consecutively. Give an upper bound on the probability that it lands heads at least 120 times using Markov's inequality.

   A. $\frac{1}{6}$

   B. $\frac{2}{6}$

   C. $\frac{3}{6}$

   D. $\frac{4}{6}$

   Answer: A

   Solution:

   The number of heads is a binomially distributed random variable X, with parameter $p = \frac{1}{10}$ and $n = 200$.

   Thus, the expected number of heads is $E(X) = np = 200.\frac{1}{10} = 20$

   By Markov Inequality, the probability of at least 120 heads is $P(X \geq 120) \leq \frac{E(X)}{120} = \frac{20}{120} = \frac{1}{6}$

2. (1 point) Let X be random variable of binomial(n, p). Using Chebyshev'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 $\alpha = \frac{3}{4}$ and $n = 8$

   Answer: 0.16

   Solution: Solution:

   One way to obtain a bound is to write

   $P(X \geq \alpha n) = P(X - np \geq \alpha n - np)$
   $\leq (|X - np| \geq n\alpha - np)$

$\leq \dfrac{Var(X)}{(n\alpha - np)^2}$

Putting the values of α, p and n

we will get upper bound as $\dfrac{4}{n}$

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

   A. M.L.E for μ is sample mean $\bar{x}$

   B. M.L.E for μ is not sample mean $\bar{x}$

   C. Both A and B

   D. None of the above

   Answer: A

   Solution:
   $L = \prod_{i=1}^{n} \dfrac{1}{\sigma\sqrt{2\pi}} \exp -\dfrac{1}{2\sigma^2}(x_i - \mu)^2 = (\dfrac{1}{\sigma\sqrt{2\pi}})^n \exp - \sum_{i=1}^{n} -\dfrac{1}{2\sigma^2}(x_i - \mu)^2$

   Taking log on both sides,

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

   When σ^2 is known, the likelihood equation for estimating μ is

   $\dfrac{\partial logL}{\partial \mu} = 0$

   Taking partial differentiation an solving we will get, $\mu = \bar{x}_i$.

   If Y follows $\mathcal{N}(\mu, \sum)$, where Y is a vector that is, $Y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ and $\mu = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ and $\sum = \begin{pmatrix} 6 & 1 \\ 1 & 13 \end{pmatrix}$

   From the above information answer questions

4. (points) Find the distribution of y_1

   A. $\mathcal{N}(3, 6)$

   B. $\mathcal{N}(1, 13)$

C. $\mathcal{N}(3, 13)$

D. $\mathcal{N}(1, 6)$

Answer: A

Solution:
$\mathcal{N}(\mu(y_1), \sum(y_1)) = \mathcal{N}(3, 6)$

5. (points) Find the distribution of y_2

   A. $\mathcal{N}(1, 3)$

   B. $\mathcal{N}(1, 13)$

   C. $\mathcal{N}(3, 13)$

   D. $\mathcal{N}(1, 6)$

   Answer: B

   Solution:
   $\mathcal{N}(\mu(y_2), \sum(y_2)) = \mathcal{N}(1, 13)$

6. (points) Find the distribution $Z = y_1 + 3y_2$

   A. $\mathcal{N}(6, 67)$

   B. $\mathcal{N}(6, 13)$

   C. $\mathcal{N}(3, 67)$

   D. $\mathcal{N}(1, 6)$

   Answer: A

   Solution:

   $E(Z) = E(CY) = CE(Y)$
   $Var(Z) = Var(CY) = CVar(Y)\bar{C}$

7. (1 point) Find the maximum likelihood estimate of the parameter θ of a population having density function as $\frac{2}{\theta^2} \times (\theta - x)$ for $0 < x < \theta$, for a sample of unit size($n = 1$), a being the sample value.

   A. $\theta = 2x$

   B. $\theta = 4x$

   C. $\theta = 3x$

   D. $\theta = x$

Answer: A

Solution:

For a random sample of unit size, the likelihood function is:

$L(\theta) = f(x, \theta) = \frac{2}{\theta^2} \times (\theta - x)$ for $0 < x < \theta$

Likelihood equation gives: $\frac{\partial LogL}{\partial\theta} = \frac{\partial log2 - 2log\theta + log(\theta - x)}{\partial\theta} = 0$

Solving the above equation gives $\theta = 2x$

8. (points) Which of the following option is correct. (Hint: Use chebychev's inequality). Here μ and σ are the mean and standard deviation of random variable X

   A. $P(|X - \mu| \geq 2\sigma) \leq \frac{1}{4}$
   B. $P(|X - \mu| \geq 2\sigma) \geq \frac{1}{4}$
   C. $P(|X - \mu| \leq 2\sigma) \leq \frac{1}{4}$
   D. $P(|X - \mu| \leq 2\sigma) \geq \frac{1}{4}$

Answer: A

Solution:

$P(|X - \mu| \geq 2\sigma) \leq \frac{Var(X)}{4\sigma^2}$

9. (points) Find the maximum likelihood estimate for the parameter p of a Binomial(m, p) of the sample $x_1, x_2, ...x_n$.

   A. $p = \frac{\sum_{i=1}^{n} x_i}{mn}$
   B. $p = \frac{\sum_{i=1}^{n} x_i \times m}{n}$
   C. $p = \frac{\sum_{i=1}^{n} x_i \times n}{m}$
   D. M.L.E does not exist for binomial distribution

Answer: A

Solution:

$L = \prod_{i=1}^{n} mC_{xi} p^{\sum x_i} (1-p)^{m-\sum x_i}$

Taking log on both sides and then differentiating partially w.r.t p

We will get $p = \frac{\sum}{mn}$

10. Which of the following set of parameters represent the same density? (Here π_k represent what fraction of data from component k, μ_k represents the mean value of data from component k and $\sum_k$ represents the covariance matrix)

    A. $\pi_1 = 0.5, \mu_1 = 1, \sum_1 = 1$ and $\pi_1 = 0.5, \mu_1 = -1, \sum_1 = 1$.

    B. $\pi_1 = 0.5, \mu_1 = -1, \sum_1 = 1$ and $\pi_1 = 0.5, \mu_1 = 1, \sum_1 = 1$

    C. $\pi_1 = 0.8, \mu_1 = 1, \sum_1 = 1$ and $\pi_1 = 0.5, \mu_1 = 1, \sum_1 = 1$

    D. $\pi_1 = 0.5, \mu_1 = 1, \sum_1 = 1$ and $\pi_1 = 0.9, \mu_1 = 1, \sum_1 = 1$

Answer: A, B

Solution:

From the definition of gaussian model.