Course: Machine Learning - Foundations
Week 12 Questions

GRADED QUESTIONS

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 $p = \dfrac{1}{3}$ and $\alpha = \dfrac{3}{4}$.

   A. $\dfrac{2}{3}$

   B. $\dfrac{4}{9}$

   C. $\dfrac{3}{5}$

   D. $\dfrac{4}{11}$

   Answer: B

   Solution: Note that X is a non-negative random variable and $E(X) = np$. Applying Markov's inequality, we obtain,

   $P(X \geq \alpha n) \leq \dfrac{E(X)}{\alpha n} = \dfrac{pn}{\alpha n} = \dfrac{p}{\alpha}$.

   for $p = \dfrac{1}{3}, \alpha = \dfrac{3}{4}$ we obtain,

   $P(X \geq \dfrac{3n}{4}) \leq \dfrac{4}{9}$.

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 upper the bound for $p = \dfrac{1}{3}$ and $\alpha = \dfrac{3}{4}$ and $n = 4$

   Answer: 0.32

   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 $\frac{288}{225n}$

3. (1 point) Suppose X is a non-negative random variable with expectation 60 and Standard deviation 5. What can we say about the best upper bound of $P(X \geq 70)$ (Hint: Use Chebyshev's inequality)?

Answer: 0.25

Since X is non-negative, we could just apply Chebyshev's inequality,

$P(X \geq 70) = P(X - 60 \geq 10) \leq \frac{5^2}{10^2} = \frac{1}{4} = 0.25$

If Y follows $\mathcal{N}(\mu, \sum)$, 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

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

   A. $\mathcal{N}(17, 21)$

   B. $\mathcal{N}(17, 15)$

   C. $\mathcal{N}(17, 17)$

   D. $\mathcal{N}(15, 21)$

Answer: A

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

5. (points) If $Z = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = DY$, Where $D = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$ Then Find the distribution of Z

Answer: N($\begin{pmatrix} 8 \\ 10 \end{pmatrix}$, $\begin{pmatrix} 29 & -1 \\ -1 & 9 \end{pmatrix}$)

Solution:

$E(Z) = E(DY) = DE(Y)$
$Var(Z) = Var(DY) = DVar(Y)\bar{D}$

6. (points) Find the maximum likelihood estimate for the parameter λ of a poisson distribution of sample values $x_1, x_2, ...x_n$. Here $\bar{x}$ represent the mean value of the sample values $x_1, x_2, ...x_n$?

   A. $\lambda = \bar{x}$

   B. $\lambda = 2\bar{x}$

   C. $\lambda = 3\bar{x}$

   D. $\lambda = 4\bar{x}$

Answer: A

Solution:

The probability function of the poisson distribution with parameter given by:

$P(X = x) = f(x, \lambda) = \dfrac{e^{-\lambda}\lambda^x}{x!}; x = 0, 1, 2, ...$

$L = \prod_{i=1}^{n} f(x_i, \lambda) = \dfrac{e^{-n\lambda}\lambda^{\sum_i^n x_i}}{x_1!x_2!.....x_{n!}}$

$logL = -n\lambda + n\bar{x}log\lambda - \sum_i^n log(x_i!)$

The likelihood equation for estimating λ is:

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

$-n + \dfrac{n\bar{x}}{\lambda} = 0$

$\lambda = \bar{x}$

7. (points) If X is the number scored in a throw of a fair die. Then which of the following options are correct. (Hint: Use Chebychev's inequality to solve).

   A. $P(|X - \mu| > 2.5) < 0.47$

B. $P(|X-\mu| > 2.5) > 0.47$

C. $P(|X-\mu| \leq 2.5) \leq 0.47$

D. $P(|X-\mu| \leq 2.5) < 0.47$

Answer: A

Solution:
Here X is a random variable which takes the values 1, 2, ...6 each with probability $\frac{1}{6}$.
Hence $E(X) = 3.5$

$E(X^2) = \frac{91}{6}$

$Var(X) = \frac{35}{12}$

For $k > 0$, Chebychev's inequality gives $P(|X-\mu| > k) \leq \frac{Var}{k^2}$

Choosing $k = 2.5$, we get $P(|X-\mu| > k) \leq 0.47$

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

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

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

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

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

Answer: B

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

Taking log on both sides,

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

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

$\frac{\partial logL}{\partial \sigma^2} = 0$

Taking partial differentiation an solving we will get option B as correct answer.

9. (points) The Central Limit Theorem says that the sampling distribution of the sample mean is approximately normal if

   A. all selected samples $x_1, x_2, x_3, ...$ are independent.

   B. the sample size is large.

   C. Both A and B

   D. Always

   Answer: B

   Solution:
   From definition of central limit theorem, option B is correct.