# Quiz Section 8

## Review

1) **Markov's Inequality**: Let $X$ be a non-negative random variable, and $\alpha > 0$. Then, $\mathbb{P}(X \geqslant \alpha) \leqslant \frac{\mathbb{E}[X]}{\alpha}$.

2) **Chebyshev's Inequality**: Suppose $Y$ is a random variable with $\mathbb{E}[Y] = \mu$ and $\text{Var}(Y) = \sigma^2$. Then, for any $\alpha > 0$, $\mathbb{P}(|Y - \mu| \geqslant \alpha) \leqslant \frac{\sigma^2}{\alpha^2}$.

3) **Chernoff Bound (for the Binomial)**: Suppose $X \sim \text{Bin}(n, p)$ and $\mu = np$. Then, for any $0 < \delta < 1$,

   - $\mathbb{P}(|X - \mu| \geqslant \delta\mu) \leqslant e^{-\frac{\delta^2 \mu}{4}}$

4) **Maximum Likelihood Estimator (MLE)**: We denote the MLE of $\theta$ as $\hat{\theta}_{\text{MLE}}$ or simply $\hat{\theta}$, the parameter (or vector of parameters) that maximizes the likelihood function (probability of seeing the data).

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta} \mathcal{L}(x_1, \ldots, x_n \mid \theta) = \arg\max_{\theta} \ln \mathcal{L}(x_1, \ldots, x_n \mid \theta)$$

## Task 1 – Tail bounds

Suppose $X \sim \text{Binomial}(6, 0.4)$. We will bound $\mathbb{P}(X \geqslant 4)$ using the tail bounds we've learned, and compare this to the true result.

a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?

b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound.

c) Give an upper bound for this probability using the Chernoff bound.

d) Give the exact probability.

## Task 2 – Exponential Tail Bounds

Let $X \sim \text{Exp}(\lambda)$ and $k > 1/\lambda$. Recall that $\mathbb{E}[X] = \frac{1}{\lambda}$ and $\text{Var}(X) = \frac{1}{\lambda^2}$.

a) Use Markov's inequality to bound $\text{P}(X \geqslant k)$.

b) Use Chebyshev's inequality to bound $\text{P}(X \geqslant k)$.

c) What is the exact formula for $\text{P}(X \geqslant k)$?

d) For $\lambda k \geqslant 3$, how do the bounds given in parts (a), (b), and (c) compare?

## Task 3 – Mystery Dish!

A fancy new restaurant has opened up which features only 4 dishes. The unique feature of dining here is that they will serve you any of the four dishes randomly according to the following probability distribution: give dish A with probability $0.5$, dish B with probability $\theta$, dish C with probability $2\theta$, and dish D with probability $0.5 - 3\theta$. Each diner is served a dish independently. Let $x_A$ be the number of people who received dish A, $x_B$ the number of people who received dish B, etc, where $x_A + x_B + x_C + x_D = n$. Find the MLE for $\theta$, $\hat{\theta}$.

## Task 4 – A Red Poisson

Suppose that $x_1, \ldots, x_n$ are i.i.d. samples from a Poisson($\theta$) random variable, where $\theta$ is unknown. In other words, they follow the distributions $\mathbb{P}(k;\theta) = \theta^k e^{-\theta}/k!$, where $k \in \mathbb{N}$ and $\theta > 0$ is a positive real number.

Find the MLE of $\theta$.

## Task 5 – Y Me?

Let $y_1, y_2, ...y_n$ be i.i.d. samples of a random variable from the family of distributions $Y(\theta)$ with densities

$$f(y;\theta) = \frac{1}{2\theta} \exp\left(-\frac{|y|}{\theta}\right) ,$$

where $\theta > 0$. Find the MLE for $\theta$ in terms of $|y_i|$ and $n$.

## Task 6 – Pareto

The Pareto distribution was discovered by Vilfredo Pareto and is used in a wide array of fields but particularly social sciences and economics. It is a density function with a slowly decaying tail, for example it can describe the wealth distribution (a small group at the top holds most of the wealth). We consider its special form given by the family of Pareto distributions Pareto$(1, \alpha)$ with densities[1]

$$f(x;\alpha) = \frac{\alpha}{x^{\alpha+1}}$$

where $x \geqslant 1$ and the real number $\alpha \geqslant 0$ is the parameter. Moreover, $f(x;\alpha) = 0$ for $x < 1$. You are given i.i.d. samples $x_1, x_2, \ldots, x_n$ from the Pareto distribution with parameter $\alpha$. Find the MLE estimation of $\alpha$.

[1] The more general Pareto distribution depends on an additional real positive parameter $m$ and follows the density $f(x;\alpha,m) = \frac{\alpha \cdot m^\alpha}{x^{\alpha+1}}$ for $x \geqslant m$, and is 0 for $x < m$. Here, we consider the special case with $m = 1$.