CS/CNS/EE/IDS 165 | Winter 2023
Instructor: Prof. A. Anandkumar | California Institute of Technology

Homework 1: Due Jan. 20th 2022

1 **Statistics and Sufficient Statistics** Consider experiments that produce n i.i.d. observation $Y_i \overset{i.i.d.}{\sim} p(y;\theta)$. For each of the following model, find the log-likelihood function $L(\mathbf{y};\theta)$ and a sufficient statistic of as "low dimension" as possible.

(a) Normal distribution with unknown mean $\mathcal{N}(\theta, 1)$.

(b) Exponential distribution with unknown mean $\mathcal{E}(\theta)$.

(c) Poisson distribution with unknown mean $\mathcal{P}(\theta)$.

(d) Bernoulli with unknown mean $\mathcal{B}(\theta)$.

(e) Uniform distribution $\mathcal{U}(0, \theta)$.

2 **Gaussian Mixture** Suppose that Y_i is an i.i.d. sequence drawn from $\mathcal{N}(\theta, 1)$, and $\mathbf{Y} = (Y_1, \cdots, Y_n)$. We know that $t(\mathbf{Y}) = \sum_i Y_i$ is a sufficient statistic. Consider next the model involving a Bernoulli random variable $X \sim \mathcal{B}(\frac{1}{4})$ in which

$$Y \sim \begin{cases} \mathcal{N}(\theta, 1) & X = 0 \\ \mathcal{N}(\theta, 2) & X = 1 \end{cases}$$

(a) Show that $(\sum_i Y_i, X)$ is a sufficient statistic.

(b) Is $\sum_i Y_i$ a sufficient statistic?