10-751 Second Exam

There are four problems on the exam, each worth 25 points. Write your solutions on the attached pages, clearly marking which problem your solution applies to. You have one hour and twenty minutes to work the problems. Books and notes are allowed, but no computers.

Problem 1. Consider the following simple *genetic linkage model* to describe how species are divided into groups. There are four groups, and the probability of falling into group $i \in \{1, 2, 3, 4\}$ is given by

$$\Pr(i \,|\, \theta) = \begin{cases} \frac{2+\theta}{4} & i = 1 \\ \frac{1-\theta}{4} & i = 2, 3 \\ \frac{\theta}{4} & i = 4 \end{cases}$$

where $\theta \in (0, 1)$.

1. Show that this is a valid probability model, by showing that the probability of falling into one of the four groups is one.
2. Suppose that 20 animals are classified in such a way that 14 of the animals fall into group one, 1 falls into group 3, and 5 fall into group four. Assuming that the groups of the animals are independent, what is the probability of this data under the above model?
3. Show that the maximum likelihood estimator for θ satisfies a polynomial equation. What is this polynomial for the data above, $n_1 = 14, n_2 = 0, n_3 = 1, n_4 = 5$?

Problem 2. Suppose that $X_1, X_2, \ldots, X_n$ are samples from a Uniform$(0, \theta)$ distribution. Let $\widehat{\theta}_{\min}(X_1, \ldots, X_n) = \min_{1 \leq i \leq n} X_i$.

1. What is the density function of $\widehat{\theta}_{\min}$?
2. Is $\widehat{\theta}_{\min}$ a sufficient statistic for θ? Justify your answer.
3. Give an unbiased estimator for θ (using any statistic, not necessarily $\widehat{\theta}_{\min}$).

Problem 3. A group of 100 people are given drug A, designed to cure a certain disease, and another group of 100 people are given drug B. In the first group, 10 of the people are cured, and 20 people are cured in the second group. Let p_{A} be the probability of being cured with drug A, and p_{B} the probability of being cured with drug B.

1. Give an estimate of $\theta = p_{\mathrm{A}} - p_{\mathrm{B}}$ from the available data. State any assumptions you make about the underlying models.
2. What is the estimated standard error of θ?
3. Give a 95% confidence interval for θ.

Problem 4. Give an algorithm for drawing random samples from a Poisson distribution with mean $\lambda > 0$. Analyze the computational complexity of preprocessing and running the algorithm.