**STAT 611-600: Theory of Inference** **Spring 2021**

# Prelim 1

*Professor: Tiandong Wang* *Name:* *UIN:*

**Instructions**:

- You have 90 mins to finish the exam **AND** upload your answers to eCampus.
- By **6:30 PM CST, Feb 25, 2021**, you must finish writing and uploading your answers. No late submission will be allowed.
- Please make sure your exam paper has: your name, your UIN.
- This exam is open-book and open-notes but **NO googling or other online resources**. Everything must be your own work.
- Please mark your answers **clearly**.
- The usual punishment for students caught cheating is an F* in the class. Cheating includes, but is not limited to, communicating in any form with any other student about the questions or answers on this exam before the solutions are posted.

Please affirm the Aggie Code of Honor with your signature:
"An Aggie does not lie, cheat or steal, or tolerate those who do." ______________________

## Problem 1 (45 pts) Sufficiency & Completeness

Let $X_1, \ldots, X_n$ be iid Uniform$[-\theta, \theta]$, where $\theta > 0$ is an unknown parameter. Suppose that $X_{(1)} \leq \cdots \leq X_{(n)}$ are order statistics, then

1. Show that $\max_{1 \leq i \leq n} |X_i|$ is a sufficient statistic for $\theta$. (10 pts)
2. Show that $\max_{1 \leq i \leq n} |X_i|$ is the MLE for $\theta$. Is this MLE also complete for $\theta$? (**Hint**: What is the distribution of $|X_n|$?) (15 pts)
3. Is $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$ sufficient? If so, is $T(\boldsymbol{X}) = (X_{(1)}, X_{(n)})$ minimal sufficient? (10 pts)
4. Is the statistic $X_{(n)} - X_{(1)}$ ancillary for $\theta$? (10 pts)

## Problem 2 (40 pts) Point Estimation

1. Let $X_1, \cdots, X_n$ be independent random variables, and $X_i \sim$ Poisson$(i\lambda^2)$, $\lambda > 0$. (**Notice that** $X_i$ **are not identically distributed!**).
   (a) Find a MLE for $\lambda$, i.e. $\hat{\lambda}_{MLE}$. (15 pts)
   **Hint:** Find the MLE for $\eta := \lambda^2$ first.
   (b) Show that $\mathbb{E}_\lambda(\hat{\lambda}_{MLE}) \leq \lambda$, i.e. the MLE for $\lambda$ tends to under-estimate $\lambda$. (10 pts)
   **Hint:** Use the Jensen's inequality, i.e. for a convex function $f$ (that is $f''(x) \geq 0$ for all $x$), $\mathbb{E}(f(X)) \geq f(\mathbb{E}(X))$.

2. Let $X$ and $Y$ be two independent exponential random variables with
$$f_X(x; \lambda) = \lambda e^{-\lambda x}, \qquad x \geq 0, \lambda > 0,$$
and
$$f_Y(y; \mu) = \mu e^{-\mu y}, \qquad y \geq 0, \mu > 0.$$
We observe $Z$ and $W$ with
$$Z = \min(X, Y), \qquad W = \mathbf{1}_{\{Z=X\}}.$$
Assume we have $n$ iid observations $(Z_i, W_i)$, $i = 1, \ldots, n$. Find the MLE of $\lambda$ and $\mu$. (**Hint**: Are $Z$ and $W$ independent from each other?) (15 pts)

## Problem 3 (15 pts) Bayes Estimator

The following is a generalization of the beta-binomial example discussed in class. Suppose that $\boldsymbol{\theta} := (\theta_1, \ldots, \theta_k)$ follows a *Dirichlet($\alpha_1, \ldots, \alpha_k$) distribution* with pdf
$$f_{\boldsymbol{\theta}}(\theta_1, \ldots, \theta_k; \alpha_1, \ldots, \alpha_k) = \frac{\Gamma(\sum_{i=1}^k \alpha_i)}{\prod_{i=1}^k \Gamma(\alpha_i)} \prod_{i=1}^k \theta_i^{\alpha_i - 1},$$
where $\theta_i \geq 0$, $\sum_{i=1}^k \theta_i = 1$, and $\alpha_i > 0$, $i = 1, \ldots, k$, are parameters. Conditioning on $\boldsymbol{\theta}$, let $\boldsymbol{X} := (X_1, \ldots, X_k)$ be a multinomial random variable with $\sum_{i=1}^k X_i = n$, and event probabilities $\boldsymbol{\theta}$, i.e. for $\boldsymbol{x} := (x_1, \ldots, x_k)$, the conditional pmf of $\boldsymbol{X}|\boldsymbol{\theta}$ is
$$p_{\boldsymbol{X}}(\boldsymbol{x}|\boldsymbol{\theta}) = \frac{n!}{x_1! \cdots x_k!} \prod_{i=1}^k \theta_i^{x_i}.$$

1. Find the posterior distribution of $\boldsymbol{\theta}|\boldsymbol{X}$. (6 pts)

2. In a survey 1000 English voters are asked to say for which party they would vote if there were a general election next week. The choices offered were 1: Labor, 2: Liberal, 3: Conservative, 4: Other, 5: None, 6: Undecided, i.e. $k = 6$. We assume that the population is large enough so that the responses may be considered independent given the true underlying proportions. Suppose the prior distribution is a Dirichlet$(5, 3, 5, 1, 2, 4)$ distribution, and we see 256 voters are voting for "Labor". Find the Bayes estimator for $\theta_1$. (9 pts)
   (**Hint**: The marginal (prior) distribution of $\theta_j$ is $Beta(\alpha_j, \sum_{i \neq j} \alpha_i)$.)