Statistics Comprehensive Exam II (Inference and Regression)

Examiner: Alexander Y. Shestopaloff

August 23, 2022

- Each question part is worth 5 marks for a total of 50 marks.
- This is a closed book exam.

Question 1. Suppose that $X_1, \dots, X_n$ are independent and identically distributed Exponential(λ) random variables. Let

$$g(\lambda) = P(X_i > t), \quad t > 0.$$

(a) Show that $T = X_1 + \cdots + X_n$ is independent of X_1/T.

(b) Find the minimum variance unbiased estimator of $g(\lambda)$. Hint: apply the Rao-Blackwell Theorem to the unbiased estimator $S = I(X_1 > t)$, where $I(\cdot)$ is the indicator function.

Question 2. Suppose that $X_1, \dots, X_n$ are independent random variables with density function

$$f_i(x; \beta) = \frac{1}{\beta t_i} \exp(-x/(\beta t_i)), \quad x \geq 0$$

where $t_1, \dots, t_n$ are known constants.

(a) Show that $\hat{\beta} = (1/n) \sum_{i=1}^{n} X_i/t_i$ is an unbiased estimator of β.

(b) Compute the Rao-Cramér lower bound for the variance of unbiased estimators of β.

Question 3. Suppose that $X_1, \dots, X_n$ are independent and identically distributed Bernoulli(θ) random variables.

(a) Show that $S = X_1 + \ldots + X_n$ is a complete and sufficient statistic for θ.

(b) Find the minimum variance unbiased estimator of $\theta(1 - \theta)$. Hint: start with the estimator $I(X_1 = 0, X_2 = 1)$.

Question 4. Suppose that $\boldsymbol{Y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$ where $\boldsymbol{\epsilon} \sim N_n(\mathbf{0}, \sigma^2 C)$ for a non-singular matrix C.

(a) Show that the MLE of $\boldsymbol{\beta}$ is given by $\hat{\boldsymbol{\beta}} = (X^T C^{-1} X)^{-1} X^T C^{-1} \boldsymbol{Y}$.

(b) What is the distribution of $\hat{\boldsymbol{\beta}}$?

Question 5. Consider the (univariate) normal error regression model $Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$ where β_0, β_1 are parameters, X_i are known constants and ϵ_i are independent $N(0, \sigma^2)$ with $i = 1, \ldots, n$.

(a) When testing $H_0 : \beta_1 = 5$ versus $H_a : \beta_1 \neq 5$ with a general linear test, what is the reduced model? What are the degrees of freedom of the reduced model df_R?

(b) When testing $H_0 : \beta_0 = 2, \beta_1 = 5$ versus H_a : not both $\beta_0 = 2$ and $\beta_1 = 5$ with a general linear test, what is the reduced model? What are the degrees of freedom of the reduced model df_R?