Midterm Exam #1

MATH 60062/70062: Mathematical Statistics II

February 24, 2022

- Please turn off your phone.
- Print your name clearly at the top of this page.
- This is a closed-book and closed-notes exam.
- This exam contains 5 questions. There are 100 points in total.
- You have 75 minutes to complete the exam.
- Please show your work and explain all of your reasoning.
- You must work by yourself. Do not communicate in any way with others.

1. (10 points) Give full definitions for the following concepts:

   a. Convergence in probability

   b. Statistic

   c. Sufficient statistic

   d. Uniformly minimum-variance unbiased estimator (UMVUE)

   e. Exponential family of distributions

   Solution:

   a. A sequence of random variables $X_1, X_2, \ldots$, converges in probability to a random variable X if, for every $\epsilon > 0$,
   $$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0.$$

   b. Suppose that $X_1, \ldots, X_n$ is an iid sample. A statistic
   $$T = T(\boldsymbol{X}) = T(X_1, \ldots, X_n),$$
   is a function of the sample $\boldsymbol{X} = (X_1, \ldots, X_n)$. The only restriction is that T cannot depend on unknown parameters.

   c. A statistic $T(\boldsymbol{X})$ is sufficient for θ if the conditional distribution of $\boldsymbol{X}$ given T does not depend on θ; i.e., the ratio
   $$f_{\boldsymbol{X}|T}(\boldsymbol{x} \mid t) = \frac{f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)}{f_T(t \mid \theta)}$$
   is free of θ, for all $\boldsymbol{x} \in \mathcal{X}$.

   d. An estimator $W^* = W^*(\boldsymbol{X})$ is a uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(\theta)$ if i) $E_\theta(W^*) = \tau(\theta)$ for all $\theta \in \Theta$, and ii) $\text{Var}_\theta(W^*) \leq \text{Var}_\theta(W)$ for all $\theta \in \Theta$, where W is any other unbiased estimator of $\tau(\theta)$.

   e. A family of PDFs or PMFs indexed by $\boldsymbol{\theta}$ is called an exponential family if it can be expressed as
   $$f_X(x \mid \boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta}) \exp\left(\sum_{i=1}^{k} w_i(\boldsymbol{\theta})t_i(x)\right), \quad \boldsymbol{\theta} \in \boldsymbol{\Theta},$$
   where $\boldsymbol{\Theta}$ is the set of all values of $\boldsymbol{\theta}$ (parameter space), $h(x) \geq 0$ and $t_1(x), \ldots, t_k(x)$ are real-valued functions of observation x (not depending on $\boldsymbol{\theta}$), and $c(\boldsymbol{\theta}) \geq 0$ and $w_1(\boldsymbol{\theta}), ..., w_k(\boldsymbol{\theta})$ are functions of the possibly vector-valued $\boldsymbol{\theta}$ (not depending on x).

2. (10 points) Suppose that $X_1, \ldots, X_n$ are iid Pois(θ),

$$f_X(x \mid \theta) = \frac{e^{-\theta}\theta^x}{x!}, \quad x = 0, 1, 2, \ldots$$

where the prior distribution on θ is Gamma(a, b),

$$\pi(\theta) = \frac{1}{\Gamma(a)b^a}\theta^{a-1}e^{-\theta/b}I(\theta > 0)$$

where the values of a and b are known. Find the posterior distribution $f(\theta \mid \boldsymbol{X} = \boldsymbol{x})$.

Solution:

The posterior distribution $f(\theta \mid \boldsymbol{x})$ is proportional to the joint distribution of θ and $\boldsymbol{X}$

$$\begin{aligned} f(\theta \mid \boldsymbol{x}) &\propto f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)\pi(\theta) \\ &= \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} \frac{1}{\Gamma(a)b^a}\theta^{a-1}e^{-\theta/b}I(\theta > 0) \\ &= \underbrace{\frac{1}{\prod_{i=1}^{n} x_i! \Gamma(a) b^a}}_{\text{free of } \theta} \underbrace{\theta^{\sum_{i=1}^{n} x_i + a - 1} e^{-\theta/\left(n + \frac{1}{b}\right)^{-1}} I(\theta > 0)}_{\text{Gamma kernel}} \end{aligned}$$

The posterior distribution is Gamma $\left(\sum_{i=1}^{n} x_i + a, \frac{1}{n + \frac{1}{b}}\right)$.

3. (10 points) Prove the Cramér–Rao Inequality. Suppose $\boldsymbol{X} \sim f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)$. Let $W(\boldsymbol{X})$ be any estimator satisfying the regularity condition

$$\frac{d}{d\theta} E_\theta[W(\boldsymbol{X})] = \int_{\mathcal{X}} \frac{\partial}{\partial\theta}[W(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)] d\boldsymbol{x}$$

and

$$\mathrm{Var}_\theta(W(\boldsymbol{X})) < \infty.$$

Show that

$$\mathrm{Var}_\theta(W(\boldsymbol{X})) \geq \frac{(\frac{d}{d\theta} E_\theta[W(\boldsymbol{X})])^2}{E_\theta\left[(\frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta))^2\right]},$$

where the quantity on the RHS is called the Cramér–Rao Lower Bound (CRLB) on the variance of the estimator $W(\boldsymbol{X})$.

Solution:

First, note that

$$\begin{aligned}
E_\theta\left[\frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta)\right] &= \int_{\mathcal{X}} \frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \int_{\mathcal{X}} \frac{\frac{\partial}{\partial\theta} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)}{f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \int_{\mathcal{X}} \frac{\partial}{\partial\theta} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \frac{d}{d\theta} \underbrace{\int_{\mathcal{X}} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x}}_{=1} = 0
\end{aligned}$$

The interchange of derivative and integral above is justified based on the assumptions. Next, consider

$$\begin{aligned}
\mathrm{Cov}_\theta\left(W(\boldsymbol{X}), \frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta)\right) &= E_\theta\left[W(\boldsymbol{X}) \frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta)\right] \\
&= \int_{\mathcal{X}} W(\boldsymbol{x}) \frac{\partial}{\partial\theta} \log f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \int_{\mathcal{X}} W(\boldsymbol{x}) \frac{\frac{\partial}{\partial\theta} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)}{f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta)} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \int_{\mathcal{X}} W(\boldsymbol{x}) \frac{\partial}{\partial\theta} f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \frac{d}{d\theta} \int_{\mathcal{X}} W(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x} \mid \theta) d\boldsymbol{x} \\
&= \frac{d}{d\theta} E_\theta\left[W(\boldsymbol{X})\right]
\end{aligned}$$

Applying the Cauchy-Schwarz Inequality

$$\begin{aligned}
\mathrm{Var}_\theta(W(\boldsymbol{X})) &\geq \frac{\left[\mathrm{Cov}_\theta\left(W(\boldsymbol{X}), \frac{\partial}{\partial\theta}\log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta)\right)\right]^2}{\mathrm{Var}_\theta\left(\frac{\partial}{\partial\theta}\log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta)\right)} \\
&= \frac{(\frac{d}{d\theta}E_\theta[W(\boldsymbol{X})])^2}{E_\theta\left[(\frac{\partial}{\partial\theta}\log f_{\boldsymbol{X}}(\boldsymbol{X} \mid \theta))^2\right]}.
\end{aligned}$$

4. (50 points) Suppose that $X_1, \ldots, X_n$ are iid Gamma(α_0, β),

$$f_X(x \mid \beta) = \frac{1}{\Gamma(\alpha_0)\beta^{\alpha_0}} x^{\alpha_0 - 1} e^{-x/\beta}$$

where α_0 is known and $\beta > 0$. **Useful fact:** $\Gamma(z+1) = z\Gamma(z)$. If $X \sim$ Gamma(α_1, β) and $Y \sim$ Gamma(α_2, β) are independent, then $X + Y \sim$ Gamma$(\alpha_1 + \alpha_2, \beta)$.

a. (10 points) Show that $T = T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is complete and sufficient for β.

b. (10 points) Find the CRLB on the variance of unbiased estimator of β.

c. (10 points) Find the maximum likelihood estimator (MLE) of $\tau(\beta) = 1/\beta$.

d. (20 points) Find the UMVUE for $\tau(\beta) = 1/\beta$.

Solution:

a. The Gamma PDF is part of the full Exponential family since

$$\begin{aligned} f_X(x \mid \beta) &= \frac{1}{\Gamma(\alpha_0)\beta^{\alpha_0}} x^{\alpha_0 - 1} e^{-x/\beta} \\ &= \frac{x^{\alpha_0 - 1} I(x > 0)}{\Gamma(\alpha_0)} \frac{1}{\beta^{\alpha_0}} \exp[(-1/\beta)x] \\ &= h(x)c(\beta)\exp[w_1(\beta)t_1(x)], \end{aligned}$$

where $d = k = 1$. Therefore,

$$T = T(\boldsymbol{X}) = \sum_{i=1}^n t_1(X_i) = \sum_{i=1}^n X_i$$

is a complete sufficient statistic for β.

b. The CRLB is

$$\frac{1}{I_n(\beta)} = \frac{1}{nI_1(\beta)},$$

where $I_1(\beta)$ is the Fisher information based on one observation

$$I_1(\beta) = -E_\beta\left[\frac{\partial^2}{\partial\beta^2} \log f_X(X \mid \beta)\right],$$

and

$$\log f_X(x \mid \beta) = -\log\Gamma(\alpha_0) - \alpha_0 \log\beta + (\alpha_0 - 1)\log x - \frac{x}{\beta}.$$

Therefore,

$$\begin{aligned} \frac{\partial}{\partial\beta} \log f_x(X \mid \beta) &= -\frac{\alpha_0}{\beta} + \frac{x}{\beta^2} \\ \frac{\partial^2}{\partial\beta^2} \log f_x(X \mid \beta) &= \frac{\alpha_0}{\beta^2} - \frac{2x}{\beta^3}. \end{aligned}$$

The Fisher information is given by

$$
\begin{aligned}
I_1(\beta) &= -E_\beta\left[\frac{\partial^2}{\partial\beta^2}\log f_X(X \mid \beta)\right] \\
&= -E_\beta\left[\frac{\alpha_0}{\beta^2} - \frac{2X}{\beta^3}\right] \\
&= \frac{\alpha_0}{\beta^2}.
\end{aligned}
$$

Therefore, the CRLB on the variance of unbiased estimator of β is

$$
\frac{1}{nI_1(\beta)} = \frac{\beta^2}{n\alpha_0}.
$$

c. The log-likelihood function is

$$
\begin{aligned}
\log L(\beta \mid \boldsymbol{x}) &= \log\prod_{i=1}^n f_X(x_i \mid \beta) \\
&= -n\log\Gamma(\alpha_0) - n\alpha_0\log\beta + \sum_{i=1}^n(\alpha_0 - 1)\log x_i - \sum_{i=1}^n \frac{x_i}{\beta}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\frac{\partial}{\partial\beta}\log L(\beta \mid \boldsymbol{x}) &= -\frac{n\alpha_0}{\beta} + \frac{\sum_{i=1}^n x_i}{\beta^2} \\
\frac{\partial^2}{\partial\beta^2}\log L(\beta \mid \boldsymbol{x}) &= \frac{n\alpha_0}{\beta^2} - \frac{2\sum_{i=1}^n x_i}{\beta^3}.
\end{aligned}
$$

The MLE of β is

$$
\hat{\beta} = \frac{\sum_{i=1}^n X_i}{n\alpha_0} = \frac{\bar{X}}{\alpha_0}.
$$

Because of the invariance property of MLE,

$$
\tau(\hat{\beta}) = \frac{1}{\hat{\beta}} = \frac{\alpha_0}{\bar{X}}
$$

is the MLE of $\tau(\beta) = 1/\beta$.

d. For a random variable $X \sim \text{Gamma}(\alpha_0, \beta)$, the expected value of $1/X$ is

$$
\begin{aligned}
E\left(\frac{1}{X}\right) &= \int_0^\infty \frac{1}{\Gamma(\alpha_0)\beta^{\alpha_0}} x^{\alpha_0 - 2} e^{-x/\beta} \\
&= \int_0^\infty \frac{1}{\alpha_*\beta}\frac{1}{\Gamma(\alpha_*)\beta^{\alpha_*}} x^{\alpha_* - 1} e^{-x/\beta},
\end{aligned}
$$

where $\alpha_* = \alpha_0 - 1$. The expected value is

$$
E\left(\frac{1}{X}\right) = \frac{1}{\alpha_*\beta} = \frac{1}{(\alpha_0 - 1)\beta}.
$$

Now, consider the statistic $T = T(\boldsymbol{X}) = \sum_{i=1}^n t_1(X_i) = \sum_{i=1}^n X_i$, which is distributed as Gamma$(n\alpha_0, \beta)$. Based on the above result

$$E\left(\frac{1}{T}\right) = \frac{1}{(n\alpha_0 - 1)\beta}.$$

Therefore,

$$E\left(\frac{n\alpha_0 - 1}{T}\right) = \frac{1}{\beta}.$$

Consider

$$\phi(T) = \frac{n\alpha_0 - 1}{T},$$

a function of T, which is complete and sufficient for β, and is an unbiased estimator of $\tau(\beta) = 1/\beta$. Therefore, $\phi(T)$ must be the UMVUE for $1/\beta$.

5. (20 points) Suppose $X_1, \ldots, X_n$ are iid Pois(θ), where $\theta > 0$. Consider the function

$$\tau(\theta) = P_\theta(X = 0) = e^{-\theta}.$$

a. (5 points) Show that $W = W(\boldsymbol{X}) = I(X_1 = 0)$ is an unbiased estimator of $\tau(\theta)$.

b. (15 points) Find the UMVUE for $\tau(\theta)$. **Hint:** Rao-Blackwell Theorem. **Useful fact:** If $X \sim$ Pois(θ_1) and $Y \sim$ Pois(θ_2) are independent, then $X + Y \sim$ Pois$(\theta_1 + \theta_2)$.

Solution:

a. The expectation

$$E_\theta(W) = E_\theta[I(X_1 = 0)] = P_\theta(X_1 = 0) = e^{-\theta},$$

showing that W is an unbiased estimator.

b. The Poisson PMF is part of the full Exponential family since The Poisson PMF is a member of the Exponential family since

$$\begin{aligned} f_X(x \mid \theta) &= \frac{e^{-\theta}\theta^x}{x!} I(x = 0, 1, 2, \ldots) \\ &= \frac{I(x = 0, 1, 2, \ldots)}{x!} e^{-\theta} \exp[\log\theta \cdot x] \\ &= h(x)c(\theta)\exp[w_1(\theta)t_1(x)]. \end{aligned}$$

Therefore, the statistic

$$T = T(\boldsymbol{X}) = \sum_{i=1}^n X_i$$

is sufficient and complete. Applying the Rao-Blackwell Theorem

$$\begin{aligned} \phi(t) &= E_\theta[W \mid T = t] \\ &= E_\theta[I(X_1 = 0) \mid T = t] \\ &= P_\theta(X_1 = 0 \mid T = t) \\ &= \frac{P_\theta(X_1 = 0, T = t)}{P_\theta(T = t)} \\ &= \frac{P_\theta(X_1 = 0, \sum_{i=2}^n X_i = t)}{P_\theta(T = t)} \\ &= \frac{P_\theta(X_1 = 0)P_\theta(\sum_{i=2}^n X_i = t)}{P_\theta(T = t)} \\ &= \left(\frac{n-1}{n}\right)^t. \end{aligned}$$

where $X_1 \sim$ Pois(θ), $\sum_{i=2}^n X_i \sim$ Pois$((n-1)\theta)$, $T \sim$ Pois$(n\theta)$. Therefore,

$$\phi(t) = \frac{e^{-\theta}\frac{[(n-1)\theta]^t e^{-(n-1)\theta}}{t!}}{\frac{(n\theta)^t e^{-n\theta}}{t!}} = \left(\frac{n-1}{n}\right)^t.$$

By the Rao-Blackwell Theorem, we know $\phi(T) = [(n-1)/n]^T$ is an improved unbiased estimator of $\tau(\theta)$ over W. Furthermore, since T is sufficient and complete,

$$\phi(T) = \left(\frac{n-1}{n}\right)^T$$

is the UMVUE for $\tau(\theta) = e^{-\theta}$.