MATH 503: Mathematical Statistics

Lecture 12: Final Exam Logistics & Review

Kimberly F. Sellers
Department of Mathematics and Statistics
Georgetown University

1

Final Exam Logistics

- **Where & when:**
  - Friday, December 10 (7-9pm); STM 110
  - Full list of final exam times & locations at https://registrar.georgetown.edu/scheduling/final-exams/fall-2021-final-exams/#
- **Exam format:**
  - Cumulative (over the semester)
  - 1-2 problems from recent material (red, mandatory)
  - 2-4 problems from previous material (black)

2

Final Exam Logistics (cont.)

- **What I will supply:**
  - Copy of the examination
  - Casella & Berger distribution handout
- **What you can bring:**
  - One 8.5"×11" piece of paper (front & back)
  - Calculator

3

Exam Instructions

- You have 2 hours to complete the exam.
- You must sit at least one seat-width apart from your neighbor (left, right).
- You may use only the information provided to you on your examination sheets today, one 8.5"×11" help sheet (two-sided), a calculator (not a cell phone), and the distribution handout provided to you by the instructor.
- Show your work, giving relevant reasoning and formulas.
- Keep the exam stapled together. If you need extra space to complete a problem, write on the back side of THAT problem's sheet.

4

Exam Conduct & Honor System

- "Cheating is the use or attempted use of unauthorized materials, information, study aids, or unauthorized collaboration on in-class examinations.... Cheating or assisting another student to cheat in connection with an examination is academic fraud." (modified from GU Honor Council website)
- Instructors **required** to report any suspicion of Honor Council violation; students **strongly encouraged**
- See Honor Council website for further details

5

Topics

- Point Estimation
  - Method of moments
  - Maximum likelihood estimation
  - Bayesian Point Estimation (Prior & Posterior Distributions)
- Estimator/Statistic qualities
  - Sufficiency
  - Completeness
  - Minimum variance unbiased estimators (MVUEs)
  - Efficiency (Cramér-Rao Lower Bound)
  - Ancillary statistics
- Hypothesis Testing
  - Likelihood Ratio Test
  - Most powerful & UMP tests
  - Power functions (including type I and II error)
- ANOVA (one-way, and two-way w/ and w/o interaction)
- Simple/linear regression

6

Problem 1

Find the MOM and MLE for θ in the following examples:

- $X_1, \dots, X_n$ iid $\sim \text{Unif}(0, \theta)$.
- $X_1, \dots, X_n$ iid $\sim \text{Exponential}(\theta)$.

SEE SCRAP

7

Problem 2

Let $X_1, \dots, X_n$ iid $\sim\text{Poisson}(\theta)$, $\theta > 0$. Show that the MLE is also an efficient estimator of θ.

SEE SCRAP

8

① For Unif $(0,\theta)$:

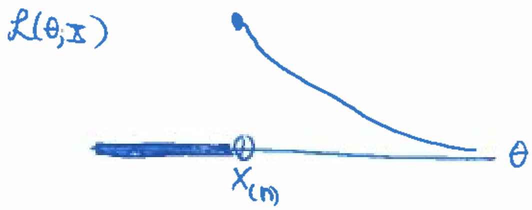
MoM for $X \sim Unif(0,\theta)$

$$E(X) = \frac{\theta}{2} \Rightarrow \bar{X} = \frac{\tilde{\theta}}{2}$$

$$\tilde{\theta} = 2\bar{X} \text{ is MoM}$$

MLE $X_1, \ldots, X_n \sim Unif(0,\theta) \Rightarrow f(x) = 1/\theta, \ 0 \le x \le \theta$

$$\mathcal{L}(\theta;\underline{x}) = \frac{1}{\theta^n} I_{(X_{(n)},\infty)}(\theta) \text{ or } \frac{1}{\theta^n}, \ X_{(n)} \le \theta$$



The likelihood has the form shown on the left. From that figure, we clearly see that $\hat{\theta} = X_{(n)}$ because the likelihood function is maximized @ $\hat{\theta} = X_{(n)}$.

For $Exp(\theta)$:

MoM for $X \sim Exp(\theta)$

$$E(X) = \theta \Rightarrow \tilde{\theta} = \bar{X} \text{ is MoM}$$

MLE $X_1, \ldots, X_n \sim Exp(\theta) \Rightarrow f(x) = \frac{1}{\theta} e^{-x/\theta}, \ 0 \le x < \infty$

$$\mathcal{L}(\theta;\underline{x}) = \frac{1}{\theta^n} e^{-\Sigma X_i/\theta} = \theta^{-n} e^{-\Sigma X_i/\theta}$$

$$\ln \mathcal{L}(\theta;\underline{x}) = -n \ln\theta - \frac{\Sigma X_i}{\theta}$$

$$\frac{\partial \ln \mathcal{L}(\theta;\underline{x})}{\partial\theta} = \frac{-n}{\theta} + \frac{\Sigma X_i}{\theta^2} = 0$$

$$-n\theta + \Sigma X_i = 0$$

$$\Rightarrow \hat{\theta} = \frac{\Sigma X_i}{n} = \bar{X}$$

(2) $X_1, \ldots, X_n \sim \text{Poisson}(\theta) \implies P(X=x) = \dfrac{e^{-\theta}\theta^x}{x!}, \; x = 0, 1, 2, \ldots$

$$\mathcal{L}(\theta, \underline{x}) = \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{x_i}}{x_i!} = \frac{e^{-n\theta}\theta^{\sum x_i}}{\prod x_i!}$$

$$\ln \mathcal{L}(\theta, \underline{x}) = -n\theta + \left(\sum x_i\right)\ln\theta - \ln\left(\prod_{i=1}^{n} x_i!\right)$$

$$\frac{\partial \ln \mathcal{L}}{\partial \theta} = -n + \frac{\sum x_i}{\theta} = 0$$

$$\frac{\sum x_i}{\theta} = n \implies \hat{\theta} = \frac{\sum x_i}{n} = \overline{X}$$

$$E(\overline{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n}\theta = \frac{\cancel{n}\theta}{\cancel{n}} = \theta \quad \therefore \overline{X} \text{ unbiased for } \theta$$

Let $f(x) = \dfrac{e^{-\theta}\theta^x}{x!}$

$$\ln f(x) = -\theta + x\ln\theta - \ln x!$$

$$\frac{\partial \ln f(x)}{\partial \theta} = -1 + \frac{x}{\theta} = -1 + x\theta^{-1}$$

$$\frac{\partial^2 \ln f(x)}{\partial \theta^2} = -x\theta^{-2} = \frac{-x}{\theta^2}$$

$$I_f(\theta) = -E\left(\frac{\partial^2 \ln f(x)}{\partial \theta^2}\right) = -E\left(\frac{-x}{\theta^2}\right) = \frac{1}{\theta^2}E(x) = \frac{\theta}{\theta^2} = \frac{1}{\theta}$$

$\therefore \text{CRLB} = \dfrac{1}{nI_f(\theta)} = \dfrac{1}{n(1/\theta)} = \dfrac{\theta}{n}$ (because $\hat{\theta} = \overline{X}$ unbiased)

Meanwhile, ~~because $\hat{\theta} = \overline{X}$ is unbiased~~

$$\text{Var}(\hat{\theta}) = \text{Var}(\overline{X}) = \text{Var}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\text{Var}(X_i) = \frac{1}{n^2}\sum_{i=1}^{n}\theta = \frac{\cancel{n}\theta}{n^{\cancel{2}}} = \frac{\theta}{n}$$

$\text{Var}(\hat{\theta}) = \text{CRLB} \quad \therefore \hat{\theta} = \overline{X}$ is efficient estimator of θ

Problem 3

Let $X_1, \dots, X_n$ denote a random sample from a distribution that has pdf $f(x;\theta) = \theta e^{-\theta x}$, $0 < x < \infty$. Find a sufficient statistic for θ, and determine a MVUE of θ.

SEE SCRAP

9

Problem 4

Let $X_1, \dots, X_n$ be a random sample of size n from a Poisson(θ) distribution. Find a best critical region for testing $H_0\colon \theta = 0.1$ vs. $H_1\colon \theta = 0.5$. Is this region uniformly most powerful for $H_0\colon \theta = 0.1$ vs. $H_1\colon \theta > 0.1$?

SEE SCRAP

10

③ $X_1, \dots, X_n$ has pdf $f(x;\theta) = \theta e^{-\theta x}$ i.e. $X_i \sim Exp\left(\frac{1}{\theta}\right)$

By ~~Neymann Pearson~~ Neymann-Fisher Factorization Thm.,

$$\prod_{i=1}^{n} f(x_i;\theta) = \underbrace{\theta^n e^{-\theta \Sigma x_i}}_{k_1(\Sigma x_i;\theta)} \cdot \underbrace{1}_{k_2(\underline{x})} \implies Y = \sum_{i=1}^{n} X_i \text{ is sufficient statistic for } \theta$$

Because $X_i \sim Exp\left(\frac{1}{\theta}\right) = Gamma(1, 1/\theta)$, $Y = \sum_{i=1}^{n} X_i \sim Gamma(n, 1/\theta)$

$\mathbb{E}(Y) = \frac{n}{\theta}$, so consider

$$\mathbb{E}(1/Y) = \int_0^\infty \frac{1}{y} f(y)\,dy = \int_0^\infty \frac{1}{y} \frac{1}{\Gamma(n)(1/\theta)^n} y^{n-1} e^{-y/(1/\theta)} dy$$

$$= \frac{\theta^n}{\Gamma(n)} \int_0^\infty y^{n-2} e^{-\theta y} dy = \frac{\theta^n}{\Gamma(n)} \int_0^\infty y^{(n-1)-1} e^{-\theta y} dy$$

$$= \frac{\theta^n}{\Gamma(n)} \cdot \frac{\Gamma(n-1)}{\theta^{n-1}} \cancel{\int_0^\infty \frac{\theta^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-y/(1/\theta)}}\ 1$$

$$= \frac{\Gamma(n-1)\theta}{\Gamma(n)} = \frac{\cancel{\Gamma(n-1)}\theta}{(n-1)\cancel{\Gamma(n-1)}} = \frac{\theta}{n-1}$$

$\therefore \frac{n-1}{Y}$ is unbiased for θ

$\Rightarrow$ by the Rao-Blackwell Thm., $\frac{n-1}{Y}$ is MVUE of θ.

(4) $X_1, \dots, X_n$ iid $\sim$ Poisson (θ)

$$f(x_i) = \mathbb{P}(X_i = x_i) = \frac{e^{-\theta}\theta^{x_i}}{x_i!} \Rightarrow \mathcal{L}(\theta; \underline{x}) = \frac{e^{-n\theta}\theta^{\sum x_i}}{\prod_{i=1}^n x_i!}$$

$H_0: \theta = 0.1$ vs. $H_1: \theta = 0.5$

$$\frac{\mathcal{L}(\theta = 0.1; \underline{x})}{\mathcal{L}(\theta = 0.5; \underline{x})} = \frac{e^{-n(0.1)}(0.1)^{\sum x_i}}{\cancel{\prod x_i!}} \cdot \frac{\cancel{\prod x_i!}}{e^{-n(0.5)}(0.5)^{\sum x_i}} \leq k$$

$$= e^{n(0.5-0.1)}\left(\frac{0.1}{0.5}\right)^{\sum x_i} \leq k$$

$$(0.5-0.1)n + (\sum x_i)\underbrace{\ln\left(\frac{0.1}{0.5}\right)}_{<0} \leq k_1 = \ln k$$

$$(\sum x_i)\underbrace{\ln\left(\frac{0.1}{0.5}\right)}_{<0} \leq k_2 = k_1 - (0.5-0.1)n$$

$$\sum_1^n x_i \geq k_3 = k_2 / \ln\left(\frac{0.1}{0.5}\right)$$

$\therefore$ the best critical region for testing $H_0: \theta = 0.1$ vs. $H_1: \theta = 0.5$ is $\sum_1^n x_i \geq k_3$ where k_3 satisfies $\mathbb{P}_{H_0}(\sum x_i \geq k_3) = \alpha$ (by Neymann–Pearson Thm.)

This critical region is UMP for $H_0: \theta = 0.1$ vs. $H_1: \theta > 0.1$ because, for any $\theta_1 > 0.1$, we see that

$$\frac{\mathcal{L}(\theta = 0.1; \underline{x})}{\mathcal{L}(\theta = \theta_1; \underline{x})} \leq k \Rightarrow \sum_1^n x_i \geq k_3' \text{ for } k_3' \text{ s.t. } \mathbb{P}_{H_0}\left(\sum_1^n x_i \geq k_3'\right) = \alpha$$

Problem 5

Let $X_i \mid \theta \sim \text{Binomial}(1,\theta) = \text{Bernoulli}(\theta)$ iid, and $\Theta \sim \text{Beta}(\alpha,\beta)$ where α,β known. Find the Bayes estimator of θ using a squared-error loss function.

SEE SCRAP

11

Problem 6

The driver of a diesel-powered automobile decided to test the quality of three types of diesel fuel sold in the area based on mpg. Make the usual assumptions and take $\alpha = 0.05$.

Brand A	38.7	39.2	40.1	38.9	
Brand B	41.9	42.3	41.3		
Brand C	40.8	41.2	39.5	38.9	40.3

The ANOVA Procedure
Dependent Variable: mpg

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
fuel	_(1)_	11.78300000	___(3)___	_(4)_	0.0048
Error	9	____(2)____	0.57624074		
Corrected Total	11	16.96916667			

R-Square	Coeff Var	Root MSE	mpg Mean
0.694377	1.885585	0.759105	40.25833

1. State the appropriate hypothesis test associated with this problem.
2. Complete the ANOVA table, filling in the four spaces above.
3. Draw conclusions.

12

SEE SCRAP

(5) $f(x_i|\theta) = \theta^{x_i}(1-\theta)^{1-x_i} \Longrightarrow \prod_{i=1}^{n} f(x_i|\theta) = \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i} = f(\underline{x}|\theta)$

$$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

$$\pi(\theta|\underline{x}) \propto f(\underline{x}|\theta)\pi(\theta) \propto \theta^{\Sigma x_i+\alpha-1}(1-\theta)^{n+\beta-\Sigma x_i-1}$$

which is the form of a Beta pdf $\therefore$ $\theta|\underline{x} \sim \text{Beta}(\alpha+\Sigma x_i, n+\beta-\Sigma x_i)$

Under a squared-error loss function, the posterior mean is the Bayes estimator of θ

$$\therefore \mathbb{E}(\theta|\underline{x}) = \frac{\alpha+\Sigma x_i}{(\alpha+\cancel{\Sigma x_i})+(n+\beta-\cancel{\Sigma x_i})} = \frac{\alpha+\Sigma x_i}{\alpha+n+\beta}$$

(6) 1. $H_0: \mu_A = \mu_B = \mu_C$ vs. H_1: a difference exists among μ_i, $i = A, B, C$

where μ_i = average mpg associated with Brand $i = A, B, C$

2. $(1) = 11 - 9 = 2$

$(2) = 16.96916667 - 11.783 \approx 5.1862$

or $\frac{(2)}{9} = 0.57624074 \therefore (2) = 9 \times 0.57624074 \approx 5.1862$

$(3) = \frac{11.783}{2} = 5.8915$

$(4) = \frac{(3)}{.57624074} = \frac{5.8915}{0.57624074} \approx 10.224$

3. Assuming $\alpha = 0.05$, we see that the pvalue $= 0.0048 < 0.05 = \alpha$, therefore we reject H_0. Accordingly, there exists a statistically significant difference among the average mpg's for Brands A, B, C. We don't, however, know where the difference between the different brands exists.

Problem 7

For the simple regression model, $Y_i = \alpha + \beta x_i + \epsilon_i$, determine the least squares estimates for α and β.

SEE SCRAP

13

7. To find least squares estimators for α, β, we want to minimize the residual sum of squares,

$$RSS = \sum_{i=1}^{n} r_i^2 = \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{n} (Y_i - (\alpha + \beta x_i))^2 = \sum_{i=1}^{n} (Y_i - \alpha - \beta x_i)^2$$

$$\frac{\partial RSS}{\partial \alpha} = 2\sum_{i=1}^{n} (Y_i - \alpha - \beta x_i)(-1) = \frac{\cancel{-2}\sum_{i=1}^{n}(Y_i - \alpha - \beta x_i)}{-2} = \frac{0}{-2}$$

$$= \sum_{i=1}^{n} Y_i - n\alpha - \beta \sum_{i=1}^{n} X_i = 0$$

$$n\alpha = \sum Y_i - \beta \sum X_i$$

$$\therefore \boxed{\tilde{\alpha} = \frac{\sum Y_i}{n} - \beta \frac{\sum X_i}{n} = \bar{Y} - \tilde{\beta}\bar{X}}$$

$$\frac{\partial RSS}{\partial \beta} = 2\sum_{i=1}^{n} (Y_i - \alpha - \beta x_i)(-x_i) = \frac{\cancel{-2}\sum_{i=1}^{n}(x_i Y_i - \alpha x_i - \beta x_i^2)}{-2} = \frac{0}{-2}$$

$$\Rightarrow \sum_{i=1}^{n} X_i Y_i - \tilde{\alpha}\sum_{i=1}^{n} X_i - \tilde{\beta}\sum_{i=1}^{n} X_i^2 = 0 \quad \text{where } \tilde{\alpha} \text{ defined above}$$

$$= \sum X_i Y_i - \left(\frac{\sum Y_i}{n} - \tilde{\beta}\frac{\sum X_i}{n}\right)\sum X_i - \tilde{\beta}\sum X_i^2 = 0$$

$$\sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n} + \tilde{\beta}\frac{(\sum X_i)^2}{n} - \tilde{\beta}\sum X_i^2 = 0$$

$$\sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n} - \tilde{\beta}\left[\sum X_i^2 - \frac{(\sum X_i)^2}{n}\right] = 0$$

$$\therefore \boxed{\tilde{\beta} = \frac{\sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}}{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}}$$