

Study Guide for Test 1

BIOSTAT 602

Test Date

- Thursday, February 23, 1:10 PM–3:10 PM

Venue

- 3699 Med Sci II (SLH)

Syllabus

- Chapters 6 (6.1–6.2), 7 (7.1, 7.2.1, 7.2.2, 7.3.1, 7.3.2, 7.3.3) from Casella and Berger

Focus of Topics

- Sufficiency principle, Minimal sufficiency, Factorization Theorem
- Ancillary statistics, Complete Statistics, Basu's Theorem
- Method of moment estimation
- Maximum likelihood estimation
- Bias, Variance, Mean Squared Error
- Searching for best unbiased estimators
 - Cramer Rao Lower Bound
 - Rao-Blackwell Theorem
 - Lehmann-Scheffe Theorem

Some Pointers

The test will involve applications of theory of statistical inference. While you will not be asked to formally derive any theoretical result, you will be expected to identify the appropriate theorems that are applicable in working out specific examples or verify/disprove some conceptual statements. In this regard homework as well as practice problems, illustrations shown in class, the quiz questions from Friday discussions are

all good examples for studying.

Studying important shortcuts will be extremely useful. For example, establishing that a pdf/pmf belongs to exponential family significantly simplifies the problem for searching optimal estimators. Similarly, verifying that a pdf belongs to a scale or a location family facilitates the search for a ancillary statistic. As far as maximum likelihood estimation is concerned, verifying the existence of a unique local maximum will suffice. We shall skip the check for global maximum.

This is a closed book, closed note examination apart from the test aid you shall be given during the examination. You will also not be allowed to use any laptop, or other devices such as iphone, ipad etc.

Below is a selection of some problems that is designed to give you some practice. The problems are only supposed to provide you a review of the topics.

Practice Problems

- Let X_1, \dots, X_n be *i.i.d.* random variables from $\text{Uniform}(\theta, 2\theta)$ with the pdf

$$f_X(x|\theta) = \frac{1}{\theta} I(\theta \leq x \leq 2\theta), \quad \theta > 0.$$

- (a) Find a minimal sufficient statistic for θ .
- (b) Is the minimal sufficient statistic obtained in (a) also a complete statistic?
Justify your answer.
- (c) Find a maximum likelihood estimator for θ .

{ **Hint:** The pdf, mean and variance of $\text{Uniform}(a, b)$ distribution are

$$f_X(x|a, b) = \frac{1}{b-a} I(a \leq x \leq b)$$

$$\text{E}(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}$$

- Let X_1, \dots, X_n be *i.i.d.* random variables from $\text{Gamma}\left(2, \frac{1}{\theta}\right)$ with the pdf

$$f_X(x|\theta) = \theta^2 x e^{-\theta x}, \quad x \geq 0, \quad \theta > 0.$$

- (a) Show that this pdf belongs to the exponential family. Identify $h(\cdot), c(\cdot), w_j(\cdot)$, and $t_j(\cdot)$ terms in your proof.

- (b) Show that $\sum_{i=1}^n X_i$ is a complete sufficient statistic for parameter θ .
- (c) Find the maximum likelihood estimator for $\text{Var}(X)$.

{ **Hint:** The pdf, mean, and variance of $\text{Gamma}(\alpha, \beta)$ distribution are

$$f_X(x|\alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad x \geq 0, \alpha, \beta > 0$$

$$\mathbb{E}(X) = \alpha\beta, \quad \text{Var}(X) = \alpha\beta^2.$$

3. Let X_1, \dots, X_n be *i.i.d.* random variables from a $\text{Binomial}(k, p)$ distribution with pmf

$$\Pr(X = x|k, p) = \binom{k}{x} p^x (1-p)^{k-x}, \quad k \in \{1, 2, \dots\}, x \in \{0, 1, 2, \dots, k\}, 0 \leq p \leq 1$$

Assuming p is known, define

$$W(\mathbf{X}) = \frac{\sum_{i=1}^n X_i}{np}$$

as an estimator for the unknown parameter k .

- (a) Calculate the bias and MSE of $W(\mathbf{X})$.
- (b) As an estimator for k , what is the most glaring disadvantage of W ?

{ **Hint:** The mean, and variance of $\text{Binomial}(k, p)$ distribution is

$$\mathbb{E}X = kp, \quad \text{Var}(X) = kp(1-p).$$

4. Let X_1, \dots, X_n be *i.i.d.* random variables from $N(0, \theta)$ with pdf

$$f_X(x|\sigma^2) = \frac{1}{\sqrt{2\pi\theta}} \exp\left\{-\frac{x^2}{2\theta}\right\}, \quad x \in R, \theta > 0.$$

- (a) Using Corollary 7.3.15 (Attainment), find $\tau(\theta)$ for which the best unbiased estimator (UMVUE) attains its Cramer-Rao lower bound.
- (b) Find the best unbiased estimator (UMVUE) for θ^2 .

- (c) Calculate the Cramer-Rao lower bound of the variance of unbiased estimators for θ^2 , and justify whether the variance of UMVUE obtained in (b) attains the bound or not.

{ **Hint:** The pdf, mean, and variance of $N(\mu, \sigma^2)$ distribution is

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$\text{EX} = \mu, \quad \text{Var}(X) = \sigma^2$$

The pdf, mean, and variance of χ_k^2 distribution is

$$f(x|k) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{\frac{k}{2}}} x^{(k/2)-1} e^{-x/2}, \quad x \geq 0, k \in \{1, 2, \dots\}$$

$$\text{EX} = k, \quad \text{Var}(X) = 2k.$$