STAT 514 Lecture Notes

Dr. David Hunter

PUBLISHED BY PUBLISHER

BOOK-WEBSITE.COM

First printing, March 2013

Contents

1 Point Estimation 5

1.1 Introduction 5

1.2 Mean Squared Error 6

1.3 Citation 6

1.4 Lists 6

1.4.1 Numbered List 6

1.4.2 Bullet Points 7

1.4.3 Descriptions and Definitions 7

2 In-text Elements 9

2.1 Theorems 9

2.1.1 Several equations 9

2.1.2 Single Line 9

2.2 Definitions 9

2.3 Notations 10

2.4 Remarks 10

2.5 Corollaries 10

2.6 Propositions 10

2.6.1 Several equations 10

2.6.2 Single Line 10

2.7 Examples 10

2.7.1 Equation and Text 10

2.7.2 Paragraph of Text 11

2.8 Exercises 11

2.9 Problems 11

2.10 Vocabulary 11

3 Presenting Information 13
3.1 Table 13
3.2 Figure 13

Bibliography 15
Books 15
Articles 15

Index 17

1. Point Estimation

1.1 Introduction

In the simplest case, we have n observations of data that we believe follow the same distribution.

$$X_1, \dots, X_n \overset{iid}{\sim} f_\theta(x)$$

where $f_\theta(x)$ is a density function involving a parameter θ. Our goal is to learn something about θ, which could be real or vector valued.

Definition 1.1.1 — Estimator. An *estimator* of θ is any function $W(X_1, \dots, X_n)$ of the data. That is, an estimator is a *statistic*.

Note:

1. $W(\boldsymbol{X})$ may not depend on θ.
2. $W(\boldsymbol{X})$ should resemble or "be close" to θ.
3. An estimator is *random*.
4. $W(X_1, \dots, X_n)$ is the estimator, $W(x_1, \dots, x_n)$ is the estimate.

■ **Example 1.1** Suppose we have n observations from an exponential distribution,

$$X_1, \dots, X_n \overset{iid}{\sim} f_\theta(x) = \frac{1}{\theta} \exp\left\{-\frac{x}{\theta}\right\} \mathbb{1}\{x > 0\}$$

for some $\theta > 0$. The **likelihood function** is equivalent to the joint density function, expressed as a function of θ rather than the data:

$$L(\theta) = \frac{1}{\theta^n} \exp\left\{-\frac{1}{\theta} \sum_{i=1}^{n} x_i\right\}$$

This function represents the *likelihood* of observing the data we observed assuming the parameter was a particular value of θ. If we can maximize this function, we can determine the $\hat{\theta}$ for which the likelihood of observing $\boldsymbol{X}$ was the highest. This might tell us something about the true value of θ.

To maximize $L(\theta)$, we want to take the derivative, set it equal to 0, and solve for θ. However, in many cases taking the derivative of the likelihood function will be very hard, if not impossible.

We can use the fact that taking the logarithm does not change the location of extrema. The **log-likelihood function** in this case is

$$\ell(\theta) = \log L(\theta) = -n\log\theta - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

Take the derivative with respect to the parameter and set equal to 0:

$$\begin{aligned}
\ell'(\theta) &= -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i \stackrel{\text{set}}{=} 0 \\
\hat{\theta} &= \frac{1}{n}\sum_{i=1}^{n} x_i
\end{aligned}$$

Here $\hat{\theta}$ is an estimator (the sample mean). Since it maximizes $L(\theta)$, we call it the **maximum likelihood estimator** (MLE). ■

1.2 Mean Squared Error

Definition 1.2.1 — Mean Squared Error. If $W(\boldsymbol{X})$ is an estimator of θ, then the **mean squared error** (MSE) is defined as

$$E_\theta\left[(W(\boldsymbol{X}) - \theta)^2\right].$$

Definition 1.2.2 — Unbiased estimator. If $W(\boldsymbol{X})$ is an estimator of θ, we say that $W(\boldsymbol{X})$ is **unbiased** if

$$E_\theta[W(\boldsymbol{X})] = \theta \quad \forall\theta.$$

Furthermore, the **bias** of $W(\boldsymbol{X})$ is

$$E_\theta[W(\boldsymbol{X})] - \theta.$$

Theorem 1.2.1 $\mathrm{MSE}(W) = \mathrm{bias}^2 + \mathrm{Var}(W)$

Proof:

$$\begin{aligned}
E[(W(\boldsymbol{X}) - \theta)^2] &= E[(W - E[W] + E[W] - \theta)^2] \\
&= E[(W - E[W])^2] + E[(E[W] - \theta)^2] + 2E[(W - E[W])(E[W] - \theta)] \\
&= \mathrm{Var}(W) + \mathrm{bias}^2(W) + 0
\end{aligned}$$

1.3 Citation

This statement requires citation [Smi12]; this one is more specific [Smi13, page 122].

1.4 Lists

Lists are useful to present information in a concise and/or ordered way[1].

1.4.1 Numbered List

1. The first item
2. The second item
3. The third item

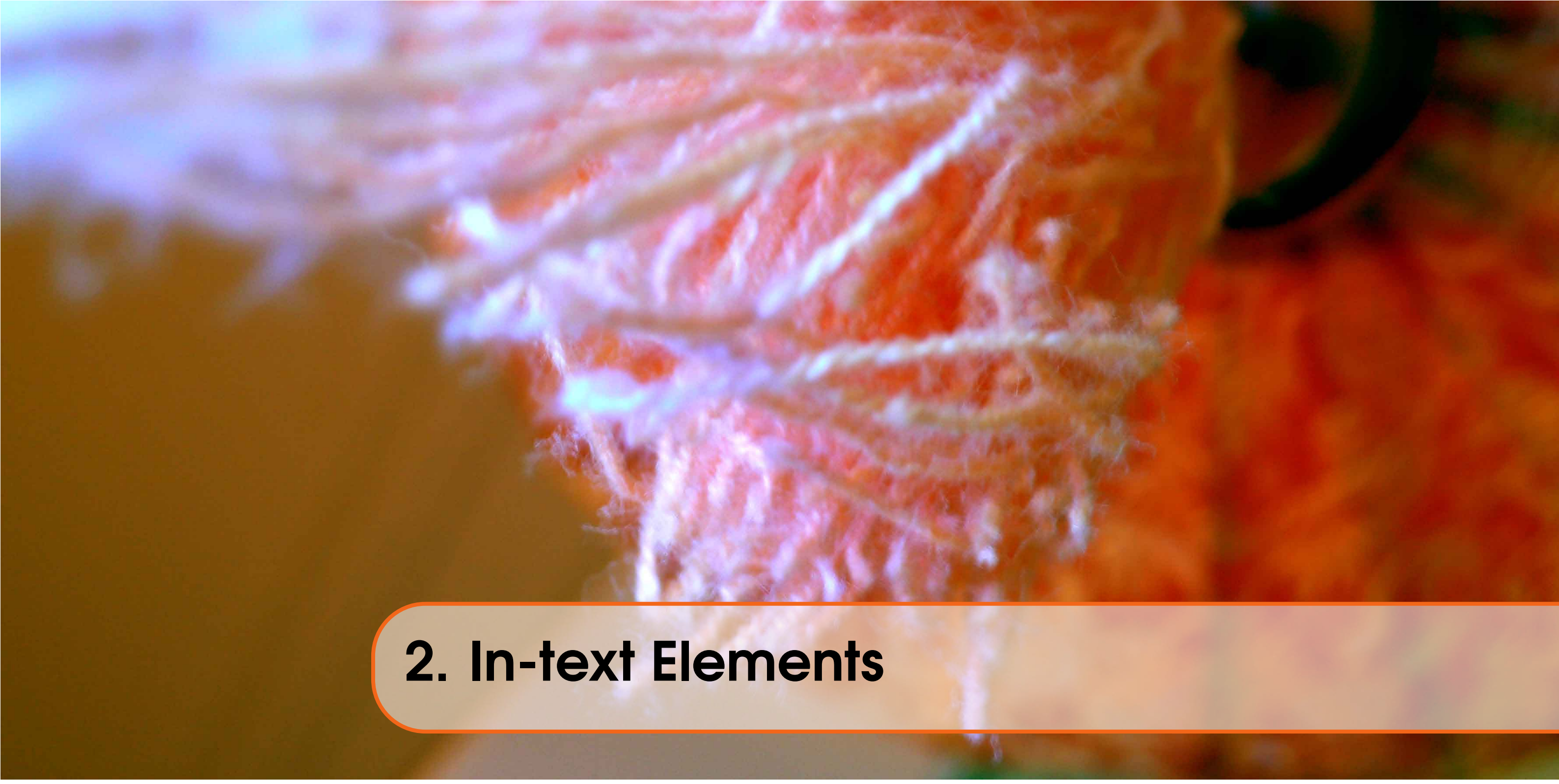
[1] Footnote example...

1.4.2 Bullet Points

- The first item
- The second item
- The third item

1.4.3 Descriptions and Definitions

Name Description
Word Definition
Comment Elaboration



2. In-text Elements

2.1 Theorems

This is an example of theorems.

2.1.1 Several equations

This is a theorem consisting of several equations.

Theorem 2.1.1 — Name of the theorem. In $E = \mathbb{R}^n$ all norms are equivalent. It has the properties:

$$\big| ||\mathbf{x}|| - ||\mathbf{y}|| \big| \leq ||\mathbf{x} - \mathbf{y}|| \tag{2.1}$$

$$||\sum_{i=1}^{n} \mathbf{x}_i|| \leq \sum_{i=1}^{n} ||\mathbf{x}_i|| \quad \text{where } n \text{ is a finite integer} \tag{2.2}$$

2.1.2 Single Line

This is a theorem consisting of just one line.

Theorem 2.1.2 A set $\mathcal{D}(G)$ in dense in $L^2(G)$, $|\cdot|_0$.

2.2 Definitions

This is an example of a definition. A definition could be mathematical or it could define a concept.

Definition 2.2.1 — Definition name. Given a vector space E, a norm on E is an application, denoted $||\cdot||$, E in $\mathbb{R}^+ = [0, +\infty[$ such that:

$$||\mathbf{x}|| = 0 \;\Rightarrow\; \mathbf{x} = \mathbf{0} \tag{2.3}$$

$$||\lambda \mathbf{x}|| = |\lambda| \cdot ||\mathbf{x}|| \tag{2.4}$$

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}|| \tag{2.5}$$

2.3 Notations

Notation 2.1. *Given an open subset G of $\mathbb{R}^n$, the set of functions φ are:*

1. *Bounded support G;*
2. *Infinitely differentiable;*

a vector space is denoted by $\mathscr{D}(G)$.

2.4 Remarks

This is an example of a remark.

R The concepts presented here are now in conventional employment in mathematics. Vector spaces are taken over the field $\mathbb{K} = \mathbb{R}$, however, established properties are easily extended to $\mathbb{K} = \mathbb{C}$.

2.5 Corollaries

This is an example of a corollary.

Corollary 2.5.1 — Corollary name. The concepts presented here are now in conventional employment in mathematics. Vector spaces are taken over the field $\mathbb{K} = \mathbb{R}$, however, established properties are easily extended to $\mathbb{K} = \mathbb{C}$.

2.6 Propositions

This is an example of propositions.

2.6.1 Several equations

Proposition 2.6.1 — Proposition name. It has the properties:

$$\left| ||\mathbf{x}|| - ||\mathbf{y}|| \right| \leq ||\mathbf{x} - \mathbf{y}|| \tag{2.6}$$

$$||\sum_{i=1}^{n} \mathbf{x}_i|| \leq \sum_{i=1}^{n} ||\mathbf{x}_i|| \quad \text{where } n \text{ is a finite integer} \tag{2.7}$$

2.6.2 Single Line

Proposition 2.6.2 Let $f, g \in L^2(G)$; if $\forall \varphi \in \mathscr{D}(G)$, $(f, \varphi)_0 = (g, \varphi)_0$ then $f = g$.

2.7 Examples

This is an example of examples.

2.7.1 Equation and Text

▪ **Example 2.1** Let $G = \{x \in \mathbb{R}^2 : |x| < 3\}$ and denoted by: $x^0 = (1,1)$; consider the function:

$$f(x) = \begin{cases} e^{|x|} & \text{si } |x - x^0| \leq 1/2 \\ 0 & \text{si } |x - x^0| > 1/2 \end{cases} \tag{2.8}$$

The function f has bounded support, we can take $A = \{x \in \mathbb{R}^2 : |x - x^0| \leq 1/2 + \varepsilon\}$ for all $\varepsilon \in]0; 5/2 - \sqrt{2}[$. ▪

2.7.2 Paragraph of Text

■ **Example 2.2 — Example name.** Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris. ■

2.8 Exercises

This is an example of an exercise.

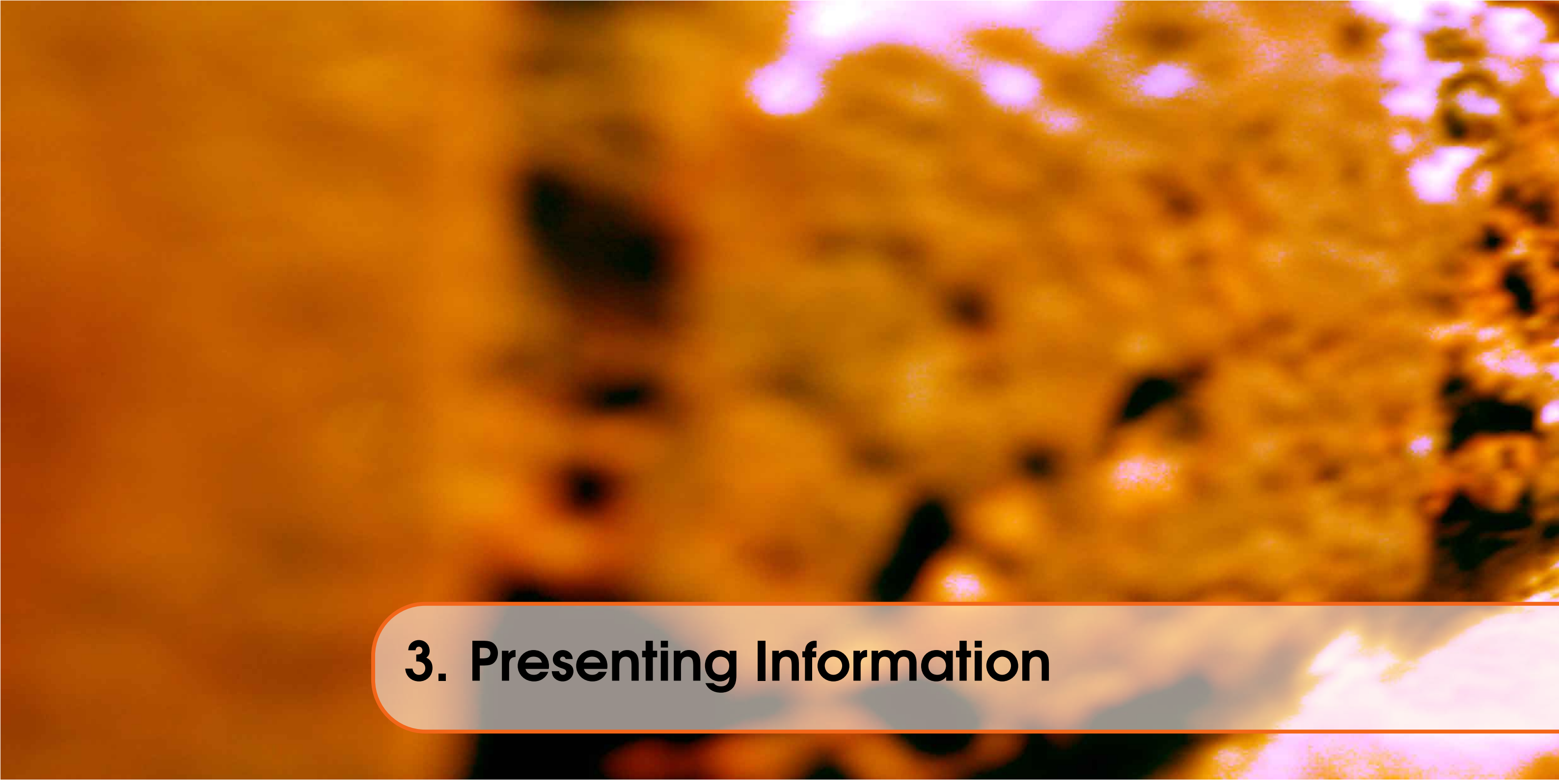
Exercise 2.1 This is a good place to ask a question to test learning progress or further cement ideas into students' minds. ■

2.9 Problems

Problem 2.1 What is the average airspeed velocity of an unladen swallow?

2.10 Vocabulary

Define a word to improve a students' vocabulary.
Vocabulary 2.1 — Word. Definition of word.



3. Presenting Information

3.1 Table

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table 3.1: Table caption

3.2 Figure

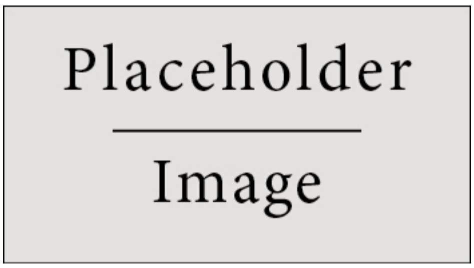


Figure 3.1: Figure caption

Bibliography

Books

[Smi12] John Smith. *Book title*. 1st edition. Volume 3. 2. City: Publisher, Jan. 2012, pages 123–200 (cited on page 6).

Articles

[Smi13] James Smith. “Article title”. In: 14.6 (Mar. 2013), pages 1–8 (cited on page 6).

Index

Citation, 6
Corollaries, 10

Definitions, 9

Examples, 10
    Equation and Text, 10
    Paragraph of Text, 11
Exercises, 11

Figure, 13

Lists, 6
    Bullet Points, 7
    Descriptions and Definitions, 7
    Numbered List, 6

Notations, 10

Paragraphs of Text, 5
Problems, 11
Propositions, 10
    Several Equations, 10
    Single Line, 10

Remarks, 10

Table, 13
Theorems, 9
    Several Equations, 9
    Single Line, 9

Vocabulary, 11