Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

Chapter 5. Estimation

Plan

Chapter 5. Estimation

Motivating example: Given an unfair coin, or p-coin, such that

$$X = \begin{cases} 1 & \text{head with probability } p, \\ 0 & \text{tail with probability } 1 - p, \end{cases}$$

how would you determine the value p?

Solutions:

1. You need to try the coin several times, say, three times. What you obtain is "HHT".
2. Draw a conclusion from the experiment you just made.

Motivating example: Given an unfair coin, or p-coin, such that

$$X = \begin{cases} 1 & \text{head with probability } p, \\ 0 & \text{tail with probability } 1 - p, \end{cases}$$

how would you determine the value p?

Solutions:

1. You need to try the coin several times, say, three times. What you obtain is "HHT".
2. Draw a conclusion from the experiment you just made.

Rationale: The choice of the parameter p should be the value that maximizes the probability of the sample.

$$\mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 0) = P(X_1 = 1)P(X_2 = 1)P(X_3 = 0)$$
$$= p^2(1 - p).$$

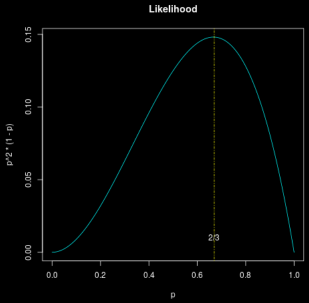
```r
# Hello, R.
p <- seq(0,1,0.01)
plot(p,p^2*(1-p),
      type="l",
      col="red")
title("Likelihood")
# add a vertical dotted (4) blue line
abline(v=0.67, col="blue", lty=4)
# add some text
text(0.67,0.01, "2/3")
```

Maximize $f(p) = p^2(1 - p)$ ....

Rationale: The choice of the parameter p should be the value that maximizes the probability of the sample.

$$\mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 0) = P(X_1 = 1)P(X_2 = 1)P(X_3 = 0) = p^2(1-p).$$

```
# Hello, R.
p <- seq(0,1,0.01)
plot(p,p^2*(1-p),
     type="l",
     col="red")
title("Likelihood")
# add a vertical dotted (4) blue line
abline(v=0.67, col="blue", lty=4)
# add some text
text(0.67,0.01, "2/3")
```



Maximize $f(p) = p^2(1-p)$ ....

A random sample of size n from the population – Bernoulli(p):

- $X_1, \cdots, X_n$ are i.i.d.[1] random variables, each following Bernoulli(p).
- Suppose the outcomes of the random sample are: $X_1 = k_1, \cdots, X_n = k_n$.
- What is your choice of p based on the above random sample?

$$p = \frac{1}{n}\sum_{i=1}^{n} k_i =: \bar{k}.$$

[1] independent and identically distributed

A random sample of size n from the population – Bernoulli(p):

- $X_1, \cdots, X_n$ are i.i.d.[1] random variables, each following Bernoulli(p).
- Suppose the outcomes of the random sample are: $X_1 = k_1, \cdots, X_n = k_n$.
- What is your choice of p based on the above random sample?

$$p = \frac{1}{n}\sum_{i=1}^{n} k_i =: \bar{k}.$$

[1] independent and identically distributed

A random sample of size n from the population – Bernoulli(p):

- $X_1, \cdots, X_n$ are i.i.d.[1] random variables, each following Bernoulli(p).
- Suppose the outcomes of the random sample are: $X_1 = k_1, \cdots, X_n = k_n$.
- What is your choice of p based on the above random sample?

$$p = \frac{1}{n} \sum_{i=1}^{n} k_i =: \bar{k}.$$

[1] independent and identically distributed

A random sample of size n from the population – Bernoulli(p):

- $X_1, \cdots, X_n$ are i.i.d.[1] random variables, each following Bernoulli(p).
- Suppose the outcomes of the random sample are: $X_1 = k_1, \cdots, X_n = k_n$.
- What is your choice of p based on the above random sample?

$$p = \frac{1}{n}\sum_{i=1}^{n} k_i =: \bar{k}.$$

[1] independent and identically distributed

A random sample of size n from the population with given pdf:

- $X_1, \cdots, X_n$ are i.i.d. random variables, each following the same given pdf.

- a **statistic** or an **estimator** is a function of the random sample.

Statistic/Estimator is a random variable!

e.g.,

$$\widehat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

- The outcome of a statistic/estimator is called an **estimate**. e.g.,

$$p_e = \frac{1}{n} \sum_{i=1}^{n} k_i.$$

A random sample of size n from the population with given pdf:

- $X_1, \cdots, X_n$ are i.i.d. random variables, each following the same given pdf.

- a **statistic** or an **estimator** is a function of the random sample.

  Statistic/Estimator is a random variable!

  e.g.,

$$\widehat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

- The outcome of a statistic/estimator is called an **estimate**. e.g.,

$$p_e = \frac{1}{n} \sum_{i=1}^{n} k_i.$$

A random sample of size n from the population with given pdf:

- $X_1, \cdots, X_n$ are i.i.d. random variables, each following the same given pdf.

- a **statistic** or an **estimator** is a function of the random sample.

  Statistic/Estimator is a random variable!

  e.g.,

$$\widehat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- The outcome of a statistic/estimator is called an **estimate**. e.g.,

$$p_e = \frac{1}{n}\sum_{i=1}^{n} k_i.$$