Econ 5023: Statistics for Decision Making

Univariate Statistics (VI): Continuous Variables and Mean (Part III)

Le Wang

Itinerary:

1. **Mean is Everything! And Everything is Mean**
   1.1 Its relations to CDF
   1.2 Its relations to the Entire distribution (Moments)
2. **Why sample averages can deliver mean**?
3. **Time Series**

More on **Expectation Operator:** (Part I)

$F(y) = \mathbb{E}[I(Y_i \leq y)]$!

If you cannot see this, consider the sample analog:

More on **Expectation Operator:** (Part I)

$F(y) = \mathbb{E}[I(Y_i \leq y)]$!

If you cannot see this, consider the sample analog:

$$\widehat{F}(y) = \frac{\sum_{i=1}^{N} I(Y_i \leq y)}{N}$$

More on **Expectation Operator:** (Part I)

$F(y) = \mathbb{E}[I(Y_i \leq y)]$!

If you cannot see this, consider the sample analog:

$$\widehat{F}(y) = \frac{\sum_{i=1}^{N} I(Y_i \leq y)}{N}$$

This is nothing but an average of a random variable $I(Y_i \leq y)$

More on **Expectation Operator:** (Part I)

$F(y) = \mathbb{E}[I(Y_i \leq y)]$!

If you cannot see this, consider the sample analog:

$$\widehat{F}(y) = \frac{\sum_{i=1}^{N} I(Y_i \leq y)}{N}$$

$$\widehat{F}(y) = \frac{\sum_{i=1}^{N} \text{Something}}{N}$$

```
x<-c(1,2,3,4)
x
```

```
## [1] 1 2 3 4
```

Generate a random variable which is an indicator function of whether or the variable is smaller than a value.

How to do it?

```r
x

## [1] 1 2 3 4

indicator<- (x <= 3)
indicator

## [1]  TRUE  TRUE  TRUE FALSE

mean(indicator)

## [1] 0.75
```

R code

```r
mean(indicator)

## [1] 0.75

ecdf(x)(3)

## [1] 0.75
```

Interpretation

1. $\frac{\sum_{i=1}^{N} I(Y_i \le y)}{N}$ $\quad (\mathbb{E}[I(Y_i \le y)])$

2. $\widehat{F}(y)$

More on **Expectation Operator:** (Part II)

For a *bounded* distribution of mass or probability, the collection of all the **moments** (of all orders, from 0 to ∞) uniquely determines the distribution.

$$\text{First moment: } \mathbb{E}[X^1] \equiv \mu$$

More on **Expectation Operator:** (Part II)

n^{th} order (central) moment: $\mathbb{E}[(X-\mu)^n] \quad \text{for } n \geq 2$

Population Moments

1. $\mathbb{E}[(X - \mu)^2]$
2. $\mathbb{E}[(X - \mu)^3]$
3. $\mathbb{E}[(X - \mu)^4]$
4. $\ldots$

Sample Estimators

1. $\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$ or $\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N-1}$
2. $\frac{\sum_{i=1}^{N}(X_i - \mu)^3}{N}$
3. $\frac{\sum_{i=1}^{N}(X_i - \mu)^4}{N}$
4. $\ldots$

where N is the sample size.

Population Moments

1. $\mathbb{E}[(X-\mu)^2]$
2. $\mathbb{E}[(X-\mu)^3]$
3. $\mathbb{E}[(X-\mu)^4]$
4. $\ldots$

Sample Estimators

1. $\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}$ or $\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N-1}$
2. $\frac{\sum_{i=1}^{N}(X_i-\mu)^3}{N}$
3. $\frac{\sum_{i=1}^{N}(X_i-\mu)^4}{N}$
4. $\ldots$

where N is the sample size.

Population Moments

1. $\mathbb{E}[(X-\mu)^2]$
2. $\mathbb{E}[(X-\mu)^3]$
3. $\mathbb{E}[(X-\mu)^4]$
4. $\ldots$

Sample Estimators

1. $\frac{\sum_{i=1}^{N} \text{Something}}{N}$
2. $\frac{\sum_{i=1}^{N} \text{Something}}{N}$
3. $\frac{\sum_{i=1}^{N} \text{Something}}{N}$
4. $\ldots$

where N is the sample size.

Now, you can see if your goal is to obtain the truth about population mean of something, what you need to do is

1. College a sample of data points for this variable
2. Take a sample average of these data points

Question: Why would this work?!!

Let's consider two questions first:

1. Behavioral experiment in Economics
2. Which headphone to purchase

Question: Why would this work?!!

(Weak) Law of Large Numbers

Weak Law of Large Numbers (WLLN) Theorem: Let $\{\mathbf{w_i}\}$ be a sequence of independent, identically distributed random variables such that $\mathbb{E}|w_i| < \infty$. Then,

$$\frac{\sum_{i=1}^{N} \mathbf{w_i}}{N} \xrightarrow{p} \mathbb{E}[\mathbf{w_i}]$$

Our version

Weak Law of Large Numbers (WLLN): Let $\{\mathbf{w_i}\}$ be a sequence of independent, identically distributed random variables such that $\mathbb{E}|w_i| < \infty$. Then,

$$\frac{\sum_{i=1}^{N} \mathbf{w_i}}{N} \xrightarrow{p} \mathbb{E}[\mathbf{w_i}]$$

$\Longrightarrow$

$\frac{\sum_{i=1}^{N} I(Y_i=y)}{N} \rightarrow (\mathbb{E}[I(Y_i = y)])$
where $w_i = I(Y_i = y)$

The Book's version

Over a large number of trials, the empirical probability of an event will approach its true probability.

$\Longrightarrow$

$\frac{\sum_{i=1}^{N} I(Y_i=y)}{N} \rightarrow \Pr[Y_i = y]$

Time series

Let's think about an experiment

Our Data: (3,1,2)

Question: What is your forecast for next observation?

How about giving some order to the data

```
## [1] 1 2 3
```

Let's introduce the new *time* dimension in the study of random variables.

This is a new concept called a **stochastic process**.

There are two assumptions that will help us recover meaningful moments along the time dimension:

1. **stationarity**
2. **ergodicity**.

Stationarity

requires the underlying distribution does not change over time, which is called **strong stationarity**.

In this class we can weaken this assumption by requiring only the first two moments of the distributions (i.e., mean and variance/covariance) to remain the same over time, which is called **weak stationarity**.

Ergodicity

The second assumption (ergodicity) states that the average over time is a consistent estimator of the mean of the stochastic process.

Intuitively, we can think of ergodicity as collecting many data points over time that bring new, independent information.

Because the process is mean stationary, over time, all these data points will provide the full set of outcomes associated with the random variable. Hence pooling all data points is a good way to measure the population mean.

Note that the first assumption is imposed on the *population* mean and variance,

while the second assumption is imposed on the *relationship* between the population and sample concepts.

Mis-conception of LLN, Error and Hypothesis Testing

Tossing a fair coin: What is your sample estimate?