Econ 5023: Statistics for Decision Making

Univariate Statistics (IV): Continuous Variables and Mean

Le Wang

Itinerary:

1. Historical Development of Mean
2. Expected Value and Mean
3. Some properties of expectation operator
4. Sample Averages
5. Examples in Decision Anlysis
6. Further examples: Bad ones and Good ones

Our Question continued:

How should we obtain the best forecast for a continuous variabe?

We know that the complete distribution approach does not work. Conceptually, we still do not know which value is more likely.. since the probability of each possible value occuring is zero!

However, we do learn that we can use PMF and CDF to examine the probability of a range of values.

$\Longrightarrow$ **Aggregation**: If general tendencies were to be revealed, the observations must be taken as a set; they must be combined!

This thinking leads to the second approach:

This thinking leads to the second approach:

1. Instead of examining the distribution itself, we use *parameters* to represent the general tendencies.

This thinking leads to the second approach:

1. Instead of examining the distribution itself, we use *parameters* to represent the general tendencies.
2. But which parameter? Or, which part of the distribution?

Mean

Why????!!!

Use of **Mean** is a rather radical idea.

1. The earliest clearly documented use of an arithmetic mean was in 1635 (Gellibrand, 1635,on compass needle).

1. The earliest clearly documented use of an arithmetic mean was in 1635 (Gellibrand, 1635,on compass needle).
2. The mathematics of a mean was certainly known in antiquity. The Pythagoreans knew already in 280 BCE of three kinds of means: the arithmetic and the geometric means.

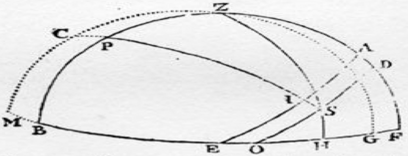
1. The earliest clearly documented use of an arithmetic mean was in 1635 (Gellibrand, 1635,on compass needle).
2. The mathematics of a mean was certainly known in antiquity. The Pythagoreans knew already in 280 BCE of three kinds of means: the arithmetic and the geometric means.
3. Why was the mean not used to combine observations in some earlier era-in astronomy, surveying, or economics?

A

DISCOVRSE MATHEMATICALL ON THE VARIATION OF THE MAGNETICALL *Needle.*

Together with Its admirable Diminution lately discovered.

By *Henry Gellibrand* Professor of Astronomie *in Gresham College.*



Veniet tempus, quo ista quæ nunc latent, in lucem dies extrahat, et longioris ævi diligentia. Sen. Nat: Quæst. lib. 7. cap. 25.

LONDON,
Printed by *William Iones*, dwelling in Red-crosse-street. 1635.

Obſervations made at Diepford An. 1634 Iunij 12 before Noone

Alt:☉ vera		*Azim: Mag*		*Azim. ☉*		*variatio*	
Gr.	*Min.*	Gr.	M.	Gr.	M,	Gr.	*M.*
44,	45.	106,	0	110	6	4.	6
46,	30,	109,	0	113	10	4,	10
48,	31,	113,	0	117	1	4.	1
50,	54,	118	0	122,	3	4.	3
54,	24,	127	0	130	55	3	55

After Noone the ſame day.

Alt. ☉ vera		Azi. *Mag*		*Azim. ☉*		*Variation*	
Gr.	*Min.*	*Gr:*	M.	G,	Mn.	*Gr,*	*Min*
44	37	114:	0	109.	53.	4:	7
40	48	108:	0	103,	50	4:	10
38	46	105.	0	100,	48	4.	12
36	43	102,	0	97.	56	4.	4
34	32	99,	0	95,	0	4:	0
32	10	96:	0	91.	55	4,	5

Theſe Concordant Obſervations can not produce a variation greater then 4 gr. 12 min. nor leſſe then 3 gr. 55 min. the Arithmeticall meane limiting it to 4 gr. and about 4 minutes.

How is it revolutionary?

More formal definition of **Population** **Mean** (or **First Moment** of the Distribution)

[Continuous Variable:] $\mathbb{E}[X] = \int xf(x)dx$
[Discrete Variable:] $\mathbb{E}[X] = \sum xf(x)$

where $f(x)$ is the **probability mass function** in the case of discrete variables, and **probability density function** in the case of continuous variables.

A Numerical Example

Some of you who are not familiar with the expectation operator. Note that it is just defined that way!

Consider a discrete variable that can take on only two values, either 0 or 1. And the population distribution function is characterized as follows $f(0) = \Pr[X = 0] = .25$ and

$f(1) = \Pr[X = 1] = .75$. Then, the population mean is

$$\mathbb{E}[X] = 0 \times .25 + 1 \times .75 = .75!$$

Note, however, that

1. The expected value of an unknown quantity is not necessariliy itself a possible value of the unknown quantity (especially when an outcome is discrete)
2. **Mean does not always exist, especially when you have infinite number of values!** (St. Petersburg Paradox)
3. It should be treated as a technical term, instead of common English usage of the word.
4. But it could be reasonably described as **"near the center"** of the possible values of the unknown quantity.

More on **Expectation Operator:**

1. Linear Operator (or Properties)
2. Its relations to CDF
3. Its relations to the Entire distribution (Moments)

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

$\mathbb{E}[Y] = \mu = \int_{-\infty}^{\infty} y f(y) dy$

Expectation is a **linear** operator satisfying the following properties.

1. $\mathbb{E}[c] = c$, where c is a constant.
2. $\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$, where a and b are constant.
3. $\mathbb{E}[c_1 Y_1 + c_2 Y_2 + \cdots + c_k Y_k] = \sum_1^k c_k \mathbb{E}[Y_k]$

$\mathbb{E}[c] = c$, where c is a constant.

Examples:

Suppose that you have a risk-free asset with a return of 6%. Then your expected return is?

$\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$, where a and b are constant.

$\mathbb{E}[c_1 Y_1 + c_2 Y_2 + \cdots + c_k Y_k] = \sum_1^k c_k \mathbb{E}[Y_k]$

Examples:

Suppose that the total cost of 1 million is fixed (since you've already signed the contract for renting and employees etc.), and that you know that the expected total revenue (which is determined by unknown demand) is 2 million dollars. What is your profit?

$\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$, where a and b are constant.

$\mathbb{E}[c_1 Y_1 + c_2 Y_2 + \cdots + c_k Y_k] = \sum_1^k c_k \mathbb{E}[Y_k]$

Examples:
Suppose that the total cost of 1 million is fixed (since you've already signed the contract for renting and employees etc.), and that you know that the expected total revenue (which is determined by unknown demand) is 2 million dollars. What is your profit?

We know that profit is defined as follows

$$
\begin{aligned}
\text{Profit} &= \text{Total Revenue} - \text{Total Cost} \\
\mathbb{E}[\text{Profit}] &= \mathbb{E}[\text{Total Revenue}] - \mathbb{E}[\text{Total Cost}] \\
&= 2 - 1 = 1
\end{aligned}
$$

Application in Contemporary Decision Analysis

It is generally supposed that when the relevant probabilities and utilities are known, the theory of rational decision is simple. (JOHN L. POLLOCK, 1983)

Criterion of Expected Value Maximization

> *A person should (rationally) perform an act iff the expectation value of his doing so is greater than the expectation value of his performing any alternative act.*

Theory of the Firm

Objective 1: Maximize stock prices. Maximize the present value (PV) of the expected future profits - maximize PV over time.

1. *Criterion of Expected Value Maximiation*
2. *Utility Analysis taking into account Risk Aversion*

Consider a simplistic example of investment opportunity,

1. 90% of the chance you will lose all of your money
2. 10% of the chance you will gain 11 million dollars

The fixed cost for this investment is 1 million dollars. Should you invest?

By *Criterion of Expected Value Maximiation*:

$$\mathbb{E}[\text{Total Revenue}] = .9 \times 0 + .1 \times 11 = 1.1$$

$\Longrightarrow$

$$\mathbb{E}[\text{Profit}] = 1.1 - 1 = .1$$

Your alternative: No investment. Total Profit is zero, of course.

This criterion has many good properties.

1. It takes account of all possible outcomes in a sensible way
2. It is more sensitive to outcomes that are more likely.
3. This argument is particularly compelling in games that can be repeated. In the long run (when such situation arises many times), a strategy of choosing the alternative that yields the higest expected value will almost surely maximize our long-term total payoff.

This approach assumes *risk netural* agents. In reality, people may be averse to risk (or maybe love risk). One approach to take into account risk is *utility analysis*.

In 1947, Von Neumann and Morgensten gave an ingenious argument to show that any consistent rational decision maker should choose among risky gambles according to utility theory.

$$\mathbb{E}[U(\cdot)] = .9 \times U(-1) + .1 \times U(10)$$

Note that

$$\mathbb{E}[U(Y)] \neq U(\mathbb{E}[Y])$$

A weird way to obtain expected value

Sample Estimator: Sample Averages

Sample Estimator of Population Mean

Sample Average is an estimator of the population mean

$$\frac{\sum_{i=1}^{N} x_i}{N}$$

If you are not familiar with summation signs, you can just expand it to the form that most of you are familiar with

$$\frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 + x_3 + \ldots + x_N}{N}$$

Weighted Average:

$$\frac{\sum_{i=1}^{N} x_i}{N} = \frac{1}{N}x_1 + \frac{1}{N}x_2 + \frac{1}{N}x_3 + \ldots + \frac{1}{N}x_N$$

Version straightforward from the definition, but more difficult to see.

$$\mathbb{E}[X] = \sum x^j f(x^j)$$

$$\widehat{\mathbb{E}[X]} = \sum x^j \widehat{f(x^j)}$$

$$
\begin{aligned}
\widehat{\mathbb{E}[X]} &= \sum x^j \widehat{f(x^j)} \\
&= x^1 \cdot \frac{\sum \mathbb{I}[x_i = x^1]}{N} + \\
&\quad x^2 \cdot \frac{\sum \mathbb{I}[x_i = x^2]}{N} + \\
&\quad \cdots + \\
&\quad x^k \cdot \frac{\sum \mathbb{I}[x_i = x^k]}{N}
\end{aligned}
$$

Better to consider a numerical example:

$$(1, 2, 1, 1, 2)$$

Better to consider a numerical example:

$$(1, 2, 1, 1, 2)$$

$x^1 = 1, x^2 = 2$

$$\begin{aligned}
\widehat{\mathbb{E}[X]} &= \sum x^j \widehat{f(x^j)} \\
&= x^1 \cdot \frac{\sum \mathbb{I}[x_i = x^1]}{N} + \\
&\quad\; x^2 \cdot \frac{\sum \mathbb{I}[x_i = x^2]}{N}
\end{aligned}$$

$$
\begin{aligned}
\widehat{\mathbb{E}[X]} &= \sum x^j \widehat{f(x^j)} \\
&= x^1 \cdot \frac{\sum \mathbb{I}[x_i = x^1]}{N} + \\
&\quad\; x^2 \cdot \frac{\sum \mathbb{I}[x_i = x^2]}{N} \\
&= 1 \cdot \frac{\mathbb{I}[1=1] + \mathbb{I}[2=1] + \mathbb{I}[1=1] + \mathbb{I}[1=1] + \mathbb{I}[2=1]}{5} + \\
&\quad\; 2 \cdot \frac{\mathbb{I}[1=2] + \mathbb{I}[2=2] + \mathbb{I}[1=2] + \mathbb{I}[1=2] + \mathbb{I}[2=2]}{5}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\mathbb{E}[X]} &= \sum x^j \widehat{f(x^j)} \\
&= x^1 \cdot \frac{\sum \mathbb{I}[x_i = x^1]}{N} + \\
&\quad\; x^2 \cdot \frac{\sum \mathbb{I}[x_i = x^2]}{N} \\
&= 1 \cdot \frac{\mathbb{I}[1=1] + \mathbb{I}[2=1] + \mathbb{I}[1=1] + \mathbb{I}[1=1] + \mathbb{I}[2=1]}{5} + \\
&\quad\; 2 \cdot \frac{\mathbb{I}[1=2] + \mathbb{I}[2=2] + \mathbb{I}[1=2] + \mathbb{I}[1=2] + \mathbb{I}[2=2]}{5} \\
&= 1 \cdot \frac{1+0+1+1+0}{5} + \\
&\quad\; 2 \cdot \frac{0+1+0+0+1}{5}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\mathbb{E}[X]} &= \sum x^j \widehat{f(x^j)} \\
&= x^1 \cdot \frac{\sum \mathbb{I}[x_i = x^1]}{N} + \\
&\quad\; x^2 \cdot \frac{\sum \mathbb{I}[x_i = x^2]}{N} \\
&= 1 \cdot \frac{\mathbb{I}[1=1] + \mathbb{I}[2=1] + \mathbb{I}[1=1] + \mathbb{I}[1=1] + \mathbb{I}[2=1]}{5} + \\
&\quad\; 2 \cdot \frac{\mathbb{I}[1=2] + \mathbb{I}[2=2] + \mathbb{I}[1=2] + \mathbb{I}[1=2] + \mathbb{I}[2=2]}{5} \\
&= 1 \cdot \frac{1+0+1+1+0}{5} + \\
&\quad\; 2 \cdot \frac{0+1+0+0+1}{5} \\
&= 1 \cdot \frac{1}{5} + 1 \cdot \frac{0}{5} + 1 \cdot \frac{1}{5} + 1 \cdot \frac{0}{5} + 1 \cdot \frac{0}{5} \\
&\quad\; 2 \cdot \frac{0}{5} + 2 \cdot \frac{1}{5} + 2 \cdot \frac{0}{5} + 2 \cdot \frac{0}{5} + 2 \cdot \frac{1}{5}
\end{aligned}
$$

$$
\begin{aligned}
\widehat{\mathbb{E}[X]} &= 1\cdot\frac{1}{5}+1\cdot\frac{0}{5}+1\cdot\frac{1}{5}+1\cdot\frac{0}{5}+1\cdot\frac{0}{5} \\
&\quad 2\cdot\frac{0}{5}+2\cdot\frac{1}{5}+2\cdot\frac{0}{5}+2\cdot\frac{0}{5}+2\cdot\frac{1}{5} \\
&= 1\cdot\frac{1}{5}+1\cdot\frac{1}{5}+1\cdot\frac{1}{5}+2\cdot\frac{1}{5}+2\cdot\frac{1}{5} \\
&= 1\cdot\frac{1}{5}+2\cdot\frac{1}{5}+1\cdot\frac{1}{5}+1\cdot\frac{1}{5}+2\cdot\frac{1}{5}
\end{aligned}
$$

Data:

$$(1, 2, 1, 1, 2)$$

More on **Expectation Operator:**

1. Linear Operator (or Properties) satisfied by sample averages as well.

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

$\mathbb{E}[Y] = \mu = \int_{-\infty}^{\infty} yf(y)dy$

Expectation is a **linear** operator satisfying the following properties.

1. $\mathbb{E}[c] = c$, where c is a constant.
2. $\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$, where a and b are constant.
3. $\mathbb{E}[c_1 Y_1 + c_2 Y_2 + \cdots + c_k Y_k] = \sum_1^k c_k \mathbb{E}[Y_k]$

$\mathbb{E}[c] = c$, where c is a constant.

```
x <- c(1,1,1,1)
x

## [1] 1 1 1 1

mean(x)

## [1] 1
```

$\mathbb{E}[aY + b] = a\mathbb{E}[Y] + b$, where a and b are constant.

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

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

```
mean(y)
```

```
## [1] 2.5
```

```
a <- 2
b <- 1
z <- a*y + b

mean(z)
```

```
## [1] 6
```

```
a*mean(y)+b
```

```
## [1] 6
```

$$\mathbb{E}[c_1 Y_1 + c_2 Y_2 + \cdots + c_k Y_k] = \sum_1^k c_k \mathbb{E}[Y_k]$$

```
y1 <- c(1,2,3,4)
y2 <- c(6,7,8,9)

y1

## [1] 1 2 3 4

y2

## [1] 6 7 8 9

c1 <- 2
c2 <- 1
z <- c1*y1 + c2*y2

mean(z)

## [1] 12.5

c1*mean(y1)+c2*mean(y2)

## [1] 12.5
```

Application in Social Science: Limitations of Mean

1.9 Pumpelly's composite portrait of 12 mathematicians. *(Pumpelly 1885)*