ECO220: Quantitative Methods in Economics
Lecture Notes (0201)

Tianyu Du, Instructor: Victor Yu

June 24, 2018

Contents

1 Lecture 1 May. 08 2018

1.1 Notations

Variable	Population	Sample
Size	N	n
Mean	μ	$\overline{x}$
Std	σ	s

1.2 From Sample to Population

Let p denote the percentage of qualified people in <u>population</u> and let $\hat{p}$ denote the percentage of qualified people in <u>sample</u>. Then, p has an <u>unknown</u> value and the value $\hat{p}$ can be calculated from sample data. We say $\hat{p}$ *is an **estimator** for* p, and the value of p is still unknown and can only be estimated.

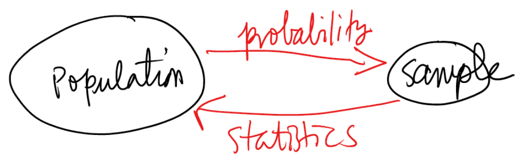
p is a **fixed value** (i.e. p is fixed once population is fixed, we can measure the exact and certain value of p if we traverse the whole population). But $\hat{p}$ will change from sample to sample. We call $\hat{p}$ an **estimator** (or **sample statistic**). The value of sample statistic will change from sample to sample. And, therefore, we call $\hat{p}$ a **random value**.

2 Lecture 2 May. 09 2018

2.1 What is Statistics?

$$\text{Statistics} \begin{cases} \text{Descriptive Statistics} \begin{cases} \text{Graphs} \\ \text{Numerical measures} \end{cases} \\ \text{Inferential Statistics} \quad \textit{Draw conclusions in a population based on sample data.} \end{cases}$$

Inferential Statistics involves uncertainties. To deal with the uncertainties, we need ***probability***



2.2 Data

$$\text{Data} \begin{cases} \text{Quantitative data} \begin{cases} \text{Discrete} \\ \text{Continuous} \end{cases} \\ \text{Qualitative data (Categorical data)} \end{cases}$$

2.3 Descriptive Statistics - Graphs

2.4 Descriptive Statistics - Numerical measures

2.4.1 Measures of centre

Mean Let $\{x_1, \ldots, x_N\}$ be measurements for the population with size N. The population mean is denoted by μ and defined as

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

Let $\{x_1, \ldots, x_n\}$ be measurements for the <u>sample</u> of size n. The sample mean is denoted by $\overline{x}$ and defined as

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Note *The mean is sensitive to extreme values.*

Median is the value in the middle when all data are put in order of magnitude. (For data with even size, median is defined as the average of the two values in the middle.)

Mode is value(s) with highest frequency.

Percentiles the k^{th} percentile is a number such that $k\%$ of data fall below this number.

3 Lecture 3 May. 15 2018

3.1 Measures of Variation(Spread)

Variance and Standard Derivation Let $\{x_1, \ldots, x_N\}$ denote the population with size N and let $\{x_1, \ldots, x_n\}$ denote the sample with size n. Then

Measures	Population	Sample
Size	N	n
Mean	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$
Std	$\sigma = +\sqrt{\sigma^2}$	$s = +\sqrt{s^2}$

Note When calculate the sample variance, use $n-1$ as denominator.

Note mathematically,

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n-1}[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2]$$

Range is defined as the difference between the largest value and the smallest value.

4 Lecture 4 May. 16 2018

4.1 Covariance and Correlation: on Populations

Consider two sets of data (population) with size N, denoted as $\{x_1, \ldots, x_N\}$ and $\{y_1, \ldots, y_N\}$, where x and y measure the age and income of observation, respectively.

Denote $\mu_x :=$ mean of x, $\mu_y :=$ mean of y
$\sigma_x :=$ std dev of x and $\sigma_y :=$ std dev of y. *When x changes, does y change?*

Covariance defined covariance between two datasets, x and y as,

$$Cov(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_x)(y_i - \mu_y)$$

Correlation coefficient the correlation coefficient ρ between datasets x and y is defined as

$$\rho = \frac{Cov(x,y)}{\sigma_x \sigma_y}$$

4.2 Covariance and Correlation: on Samples

When N is too large, we select a sample of size n.

Let $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_n\}$ denote the selected samples with size n, $\overline{x}, \overline{y}$ denote the sample means, and s_x, s_y denote the sample std dev.

Covariance between two sample is defined as

$$Cov(x,y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$$

Correlation coefficient the sample correlation r is defined as

$$r = \frac{Cov(x,y)}{s_x s_y}$$

4.3 Interpretations

4.3.1 Interpreting covariance

Example Consider samples x and y with

$$Cov(x,y) = -25.31$$

The <u>negative sign</u> means x and y have a <u>negative linear relationship</u>. As x increases, y tends to decrease. The <u>magnitude</u> 25.31 has **no** meaning.

4.3.2 Interpreting correlation coefficient

Example Consider samples x and y with

$$r = -0.94$$

The <u>negative sign</u> means x and y have <u>negative linear relationship</u>. As x increases, y decreases. The <u>magnitude</u> 0.94 means the <u>linear relationship is strong</u>. When r is close to 1 or -1, the string line relation is strong, when r is close to 0, the relation is weak.

Note $\rho \in [-1, 1]$ and $r \in [-1, 1]$

5 Lecture 5 May. 22 2018

5.1 Introduction to Simple Regression

Let the linear estimator to be $\hat{y} = b_0 + b_1 x$ and let y_i denote the actual value at x_i, $\hat{y}$ is the estimated y value at x_i. Then, $e_i := y_i - \hat{y_i}$ is the error of y value at x_i (a.k.a. **residual**).

Note notice that $\sum_{i=1}^{n} e_i \equiv 0$.

SSE **Sum of Squared Error**(SSE) as

$$SSE = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

OLS Minimize SSE with respect to b_0 and b_1, we have the FOC as

$$\begin{cases} \frac{\partial SSE}{\partial b_0} = 0 \\ \frac{\partial SSE}{\partial b_1} = 0 \end{cases}$$

By solving the first order conditions, we have

$$\begin{cases} b_1 = \frac{\sum (x_i - \overline{x})(y_i - \overline{y})}{\sum (x_i - \overline{x})^2} \\ b_0 = \overline{y} - b_1 \overline{x} \end{cases}$$

The above method to find b_0 and b_1 is called the <u>method of least square</u>, or <u>method of Ordinary Least Square (OLS)</u>.

5.2 Relationship between b_1 and r

$$b_1 = \frac{Cov(x, y)}{Var(x)} = \frac{Cov(x, y)}{std(x) std(y)} \frac{std(y)}{std(x)} = r \frac{s_y}{s_x}$$

5.3 Analysis of Variance (ANOVA)

Let y_i denote the actual y value at x_i and $\hat{y}_i$ denote the estimated y value at x_i.

Definition

$$SST = \sum_{i=1}^{n}(y_i - \overline{y})^2$$

$$SSR = \sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2$$

$$SSE = \sum_{i=1}^{n}(y_i - \hat{y}_1)^2$$

Notice $SST = SSR + SSE$

Anova Table

	SS	df	MS	F
Regression	SSR	1	MSR	MSR/MSE
Error(Residual)	SSE	MSE	$n-2$	
Total	SST	$n-1$		

where MS stands for **mean square** and is defined as

$$MS = \frac{SS}{df}$$

$$MSR = \frac{SSR}{1}$$

$$MSE = \frac{SSE}{n-2}$$

6 Lecture 6 May. 23 2018

6.1 OLS, continued.

R-square **coefficient of determination** is defined as

$$R^2 = \frac{SSR}{SST}$$

and notice that $R^2 \in [0, 1]$ and can be interpreted as **% of variation in** y **explained by** x **(via the linear model)**

Note in ECO220, we use R^2 or r^2 to represent the same thing.

6.2 Sample space, Event and Probability

Experiment an experiment is a process that creates <u>two or more</u> outcomes.

Random Experiment a random experiment is an experiment such that the outputs *cannot* be determined with certainty before the end of the experiment.

Sample Space a sample space is the set of all possible outcomes in a random experiment.

Event an event is a subset of a sample space.

Prob Let S be the sample space, let E be an event, then the **probability of** E, $P(E)$ is defined as

$$P(E) = \text{probability of } E = \frac{\text{Number of outcomes in } E}{\text{Number of outcomes in } S}$$

assuming that each outcome in S has equal likelihood to be chosen into E.

6.3 Some Rules of Probability

Let E be an event in sample space S, then

- $P(E) \in [0, 1]$.
- $P(S) = 1$.
- Let E^c denote the **complementary** of E, then $P(E^c) = 1 - P(E)$.
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Addition Rule)

7 Lecture 7 May. 29 2018

Mutually Exclusive Event If $A \cap B = \emptyset$, we say events A and B are **mutually exclusive/disjoint**. Then, if A, B are disjoint, we have

$$P(A \cup B) = P(A) + P(B)$$

7.1 Conditional Probability

Conditional Prob In general, if A and B are events in sample space S, the conditional probability of A given B is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Multiplication rule

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

7.2 Independent Event

Independent Event We say two events, A and B are **independent** if any of the following is true. (those definitions below are equivalent.)

- $P(A|B) = P(A)$ or
- $P(B|A) = P(B)$ or
- $P(A \cap B) = P(A)P(B)$

8 Lecture 8 May. 30 2018

8.1 Bayes Theorem

Let A and B be two events. Then,

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Proven by definition of conditional probability.

8.2 Random Variable and Prob. Distributions

Prob. distribution

Cumulative Prob. distribution

8.3 Expected Values

Expected Value Let X be a random variable with probability distribution $P(X)$. Then defined the expected value of X, $\mathbb{E}(X)$ as

$$\mu = \mathbb{E}(X) = \sum_x xP(X = x)$$

Variance of Random Variable For random variable X, we have

$$\sigma^2 = Var(X) = \sum_x (x - \mu)^2 P(X = x) = \mathbb{E}(X - \mu)^2$$

9 Lecture 9 June. 5 2018

9.1 Expected Value of a Random Variable

Mean $\mu = \mathbb{E}(X) = \sum_x xP(X = x)$.

Variance $\sigma^2 = \mathbb{E}(x - \mu)^2 = \mathbb{E}(X^2) - \mu^2$.

9.2 Laws of Expectation

In general, let X be a random variable, and let $a, c \in \mathbb{R}$, then

$$\mathbb{E}(aX + c) = a\mathbb{E}(X) + c$$

$$Var(aX + c) = Var(aX) = a^2 Var(X)$$

Let X and Y be random variables, and let $a, b, c \in \mathbb{R}$, then

$$\mathbb{E}(aX + bY + c) = a\mathbb{E}(X) + b\mathbb{E}(Y) + c$$

$$Var(aX + bY + c) = Var(aX + bY) = a^2 Var(X) + b^2 Var(Y) + 2ab\ Cov(X, Y)$$

Note if X and Y are independent, then $\rho = Cov(X, Y) = 0$ and

$$Var(aX + bY + c) = a^2 Var(X) + b^2 Var(Y)$$

9.3 Binomial Distribution

In general, let n be the number of independent trails and $p = P(\#success)$. Let X be a random variable which is the number of successes in n trails, we have

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x},\ \text{for } x = \{0, 1, 2, \ldots, n\}$$

$$\mu = \mathbb{E}(X) = np$$

$$\sigma^2 = Var(X) = npq,\ q = 1 - p$$

10 Lecture 10 June. 6 2018

10.1 Uniform Distribution

Let X be uniform from a to b. $f(x) = \frac{1}{b-a}, a \leq x \leq b$

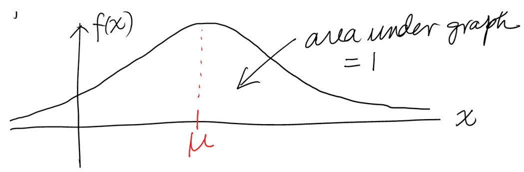
$$\mu = \mathbb{E}(X) = \int_a^b xf(x)dx = \frac{a+b}{2}$$

$$\sigma^2 = Var(X) = \mathbb{E}(X^2) - \mu^2$$

10.2 Normal Distribution

Let X be a continuous random variable, satisfying $-\infty < x < \infty$. The mean of X is μ and the variance of X is σ^2. The graph of X is The graph is symmetric at μ and the variance σ^2 determines the shape(spread) of X. We say X <u>follows a normal distribution with mean μ and variance σ^2.</u> And denote as

$$X \sim N(\mu, \sigma^2)$$



Standard Normal Distribution A standard normal distribution is a normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$. Denote the standard normal distribution as

$$Z \sim N(0, 1)$$

11 Lecture 11 June. 12 2018

11.1 Applying normal distribution

Theorem Let $X \sim N(\mu, \sigma^2)$, then

$$z = \frac{x - \mu}{\sigma} \sim N(0, 1)$$

11.2 Normal Approximation to Binomial

Consider a random variable $X \sim B(n, p)$, then we can approximate the binomial with a normal distribution $X \approx N(np, npq)$.

12 Lecture 12 June. 13 2018

12.1 Sampling Distributions

Consider population with size N has p as percentage of *success* (qualified) and sample with size n with $\hat{p} = \frac{x}{n}$ as percentage of *success*.

p is a parameter which has a fixed value. In real life, the value of p is usually unknown. $\hat{p}$ is a sample statistic, which does not have fixed value (random variable, value of $\hat{p}$ vary from sample to sample). Also, μ and σ are parameters, which are fixed but usually unknown. $\overline{x}$ is a sample statistic, and is random.

Suppose we know p for population, then we can conclude about random variables from a random sample,

1. $\mathbb{E}(\hat{p}) = p$.
2. $Var(\hat{p}) = \frac{pq}{n},\ q = 1 - p$

3. When sample size n is large, the distribution of $\hat{p}$ is approximately normal (**Central Limit Theorem in proportion**) [1]

That's

$$\hat{p} \approx\sim N(p, \frac{pq}{n}), \text{ when } n \text{ is large.}$$

Example Given $p_{success} = 0.3$ for the whole population and find the probability that at least 320 *success* found in a sample of size $n = 1000$. i.e. Let X denote the number of success in sample with $n = 1000$, find $P(X \geq 320)$.

Method 1 Use Central Limit Theorem, check $np = 300 \geq 10 \wedge nq = 700 \geq 10$, thus n is *large*. And approximate $\hat{p}$ of sample as

$$\hat{p} \sim N(p, \frac{pq}{n})$$

Soln.

$$P(X \geq 320) = P(\hat{p} \geq 0.32)$$
$$= P(\frac{\hat{p}-p}{\sqrt{\frac{pq}{n}}} \geq \frac{0.32-0.3}{\sqrt{\frac{0.3*0.7}{1000}}}) = P(z \geq \frac{0.02}{\sqrt{\frac{0.21}{1000}}})$$

Find z in z-table

■

Method 2 Use Normal Approximation to Binomial. $p = 0.3$ and $n = 1000$.

$$X \approx Y \sim N(300, 210)$$

Soln.

$$P(X \geq 320) = P(Y > 319.5)$$
$$= P(\frac{Y-\mu}{\sigma} > \frac{319.5-300}{\sqrt{210}})$$
$$= P(z > 1.35) \text{ find in z table}$$

■

Note *methods 1 and 2 do* ***not*** *give exactly same answer, but the answers should be close.*

[1] As a rule of thumb, n is considered to be large when $np \geq 10 \wedge nq \geq 10$.

12.2 Sampling distribution of $\overline{X}$, the sample mean

1. $\mathbb{E}(\overline{X}) = \mu$.
2. $Var(\overline{X}) = \frac{\sigma^2}{n}$.
3. When n is large, the distribution of $\overline{X}$ is approximately normal. (**Central Limit Theorem in Mean**).
4. When population is normal, the distribution of $\overline{X}$ is exactly normal, regardless of the sample size n.

Putting together,

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n}), \text{ when } n \text{ is large.}$$

13 Lecture 13 Jun. 19 2018

13.1 Confidence Interval

To find $100(1-\alpha)\%$ confidence interval for p estimated from $\hat{p}$ is

$$\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

13.2 Sample Size Required

When we specify the confidence level $1-\alpha$, and the margin of error, the required sample size is

$$n = \frac{z^2_{\frac{\alpha}{2}}}{(ME)^2} pq$$

If p can be estimated from previous surveys, use it to find n. Else, use $p = 0.5$ to find n.