Introduction to Statistics and Data Science using *eStat*

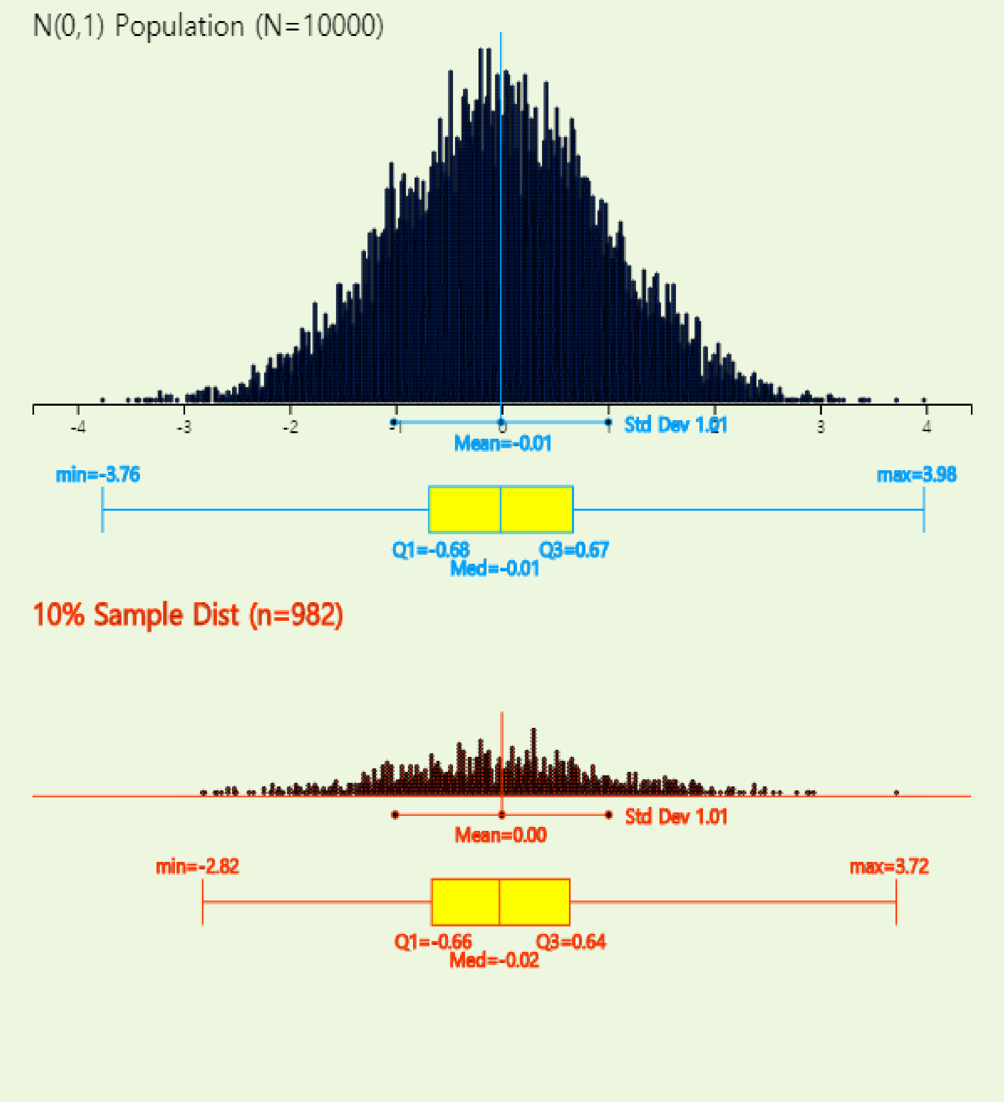
Chapter 6 Sampling Distribution and Estimation

6.2 Sampling Distribution of Sample Means and Estimation of Population Mean

Jung Jin Lee
Professor of Soongsil University, Korea
Visiting Professor of ADA University, Azerbaijan

6.2 Sampling Distribution of Sample Means and Estimation of Population Mean

- **Population parameter**
  ⇨ Characteristic values of a population,
  - population means, variance, etc
- **Sample statistic**
  ⇨ random variable of samples to estimate population parameter
  - sample mean and sample variance.
- **Sample statistic can be a good estimate of the population parameter?**



6.2.1 Sampling Distribution of Sample Means

[Ex 6.2.1] Consider a population consisting of five salespeople from a company. Consider the number of years of service at this company.

6, 2, 4, 8, 10

1) Obtain the mean and variance of this population.
2) Obtain all possible samples of size two by simple random sampling with replacement in this population and calculate each sample mean. In addition, calculate the mean and variance of all these possible sample means and compare them with the mean and variance of the population.
3) Prepare a frequency distribution of sample means and draw a bar chart. Compare this with the distribution of the population.

6.2.1 Sampling Distribution of Sample Means

<Answers of Ex6.2.1>

1) The mean and variance of the population is $\mu = 6$, $\sigma^2 = 8$

2) The number of all possible samples with replacement is $5 \times 5 = 25$.

Table 6.2.1 All possible samples of size 2 from the population and their sample means

sample	$\bar{x}$	sample	$\bar{x}$	sample	$\bar{x}$	sample	$\bar{x}$	sample	$\bar{x}$
2,2	2	4,2	3	6,2	4	8,2	5	10,2	6
2,4	3	4,4	4	6,4	5	8,4	6	10,4	7
2,6	4	4,6	5	6,6	6	8,6	7	10,6	8
2,8	5	4,8	6	6,8	7	8,8	8	10,8	9
2,10	6	4,10	7	6,10	8	8,10	9	10,10	10

6.2.1 Sampling Distribution of Sample Means

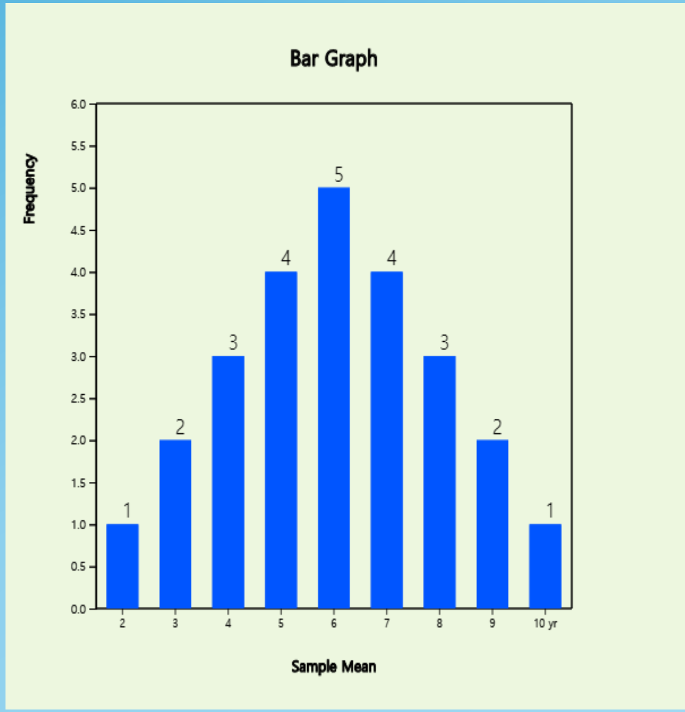
<Answers of Ex6.2.1>

- Some of these sample means are exactly the same as the population mean μ = 6, but some others such as 2 or 10 are significantly different.
- Mean of all possible 25 sample means is also 6.
  => sample mean can be a good estimate of the population mean μ
  => $\bar{x}$ is an unbiased estimator of μ
- Variance of all possible sample means is 4
  => population variance σ^2 = 8 divided by the sample size n=2,
  i.e., $\frac{\sigma^2}{n}$

6.2.1 Sampling Distribution of Sample Means

<Answers of Ex6.2.1> 3)

Sample Mean	Frequency	Relative Frequency
2	1	0.04
3	2	0.08
4	3	0.12
5	4	0.16
6	5	0.20
7	4	0.16
8	3	0.12
9	2	0.08
10	1	0.04
	25	1.00



- Mean of all 25 possible sample means is the same as population mean.
  ⇨ Sample mean is an unbiased estimator of population mean
- Variance of the sample means is the population variance divided by the sample size

6.2.1 Sampling Distribution of Sample Means

[Theorem] If a population is normally distributed with mean μ and variance σ^2,

⇨ distribution of all possible sample means is exactly a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$.

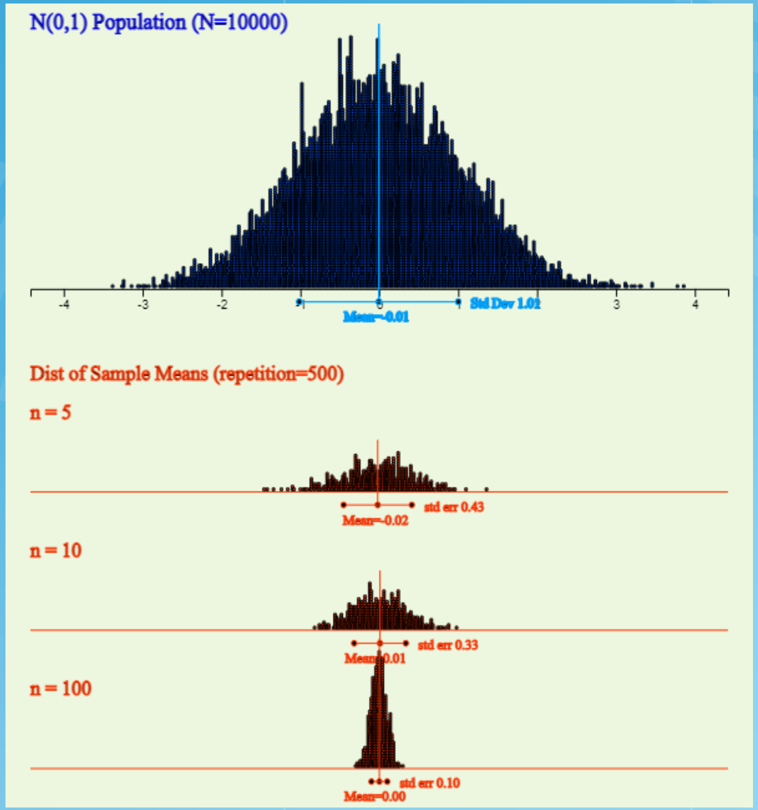
[Theorem] (Central Limit Theorem)

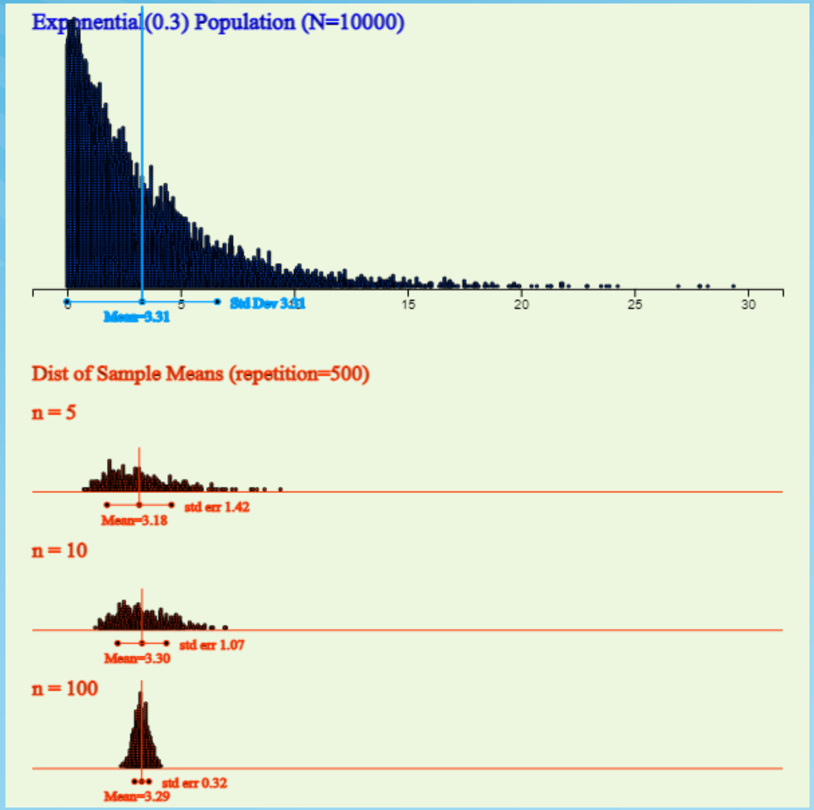
If an infinite population has mean μ and variance σ^2,

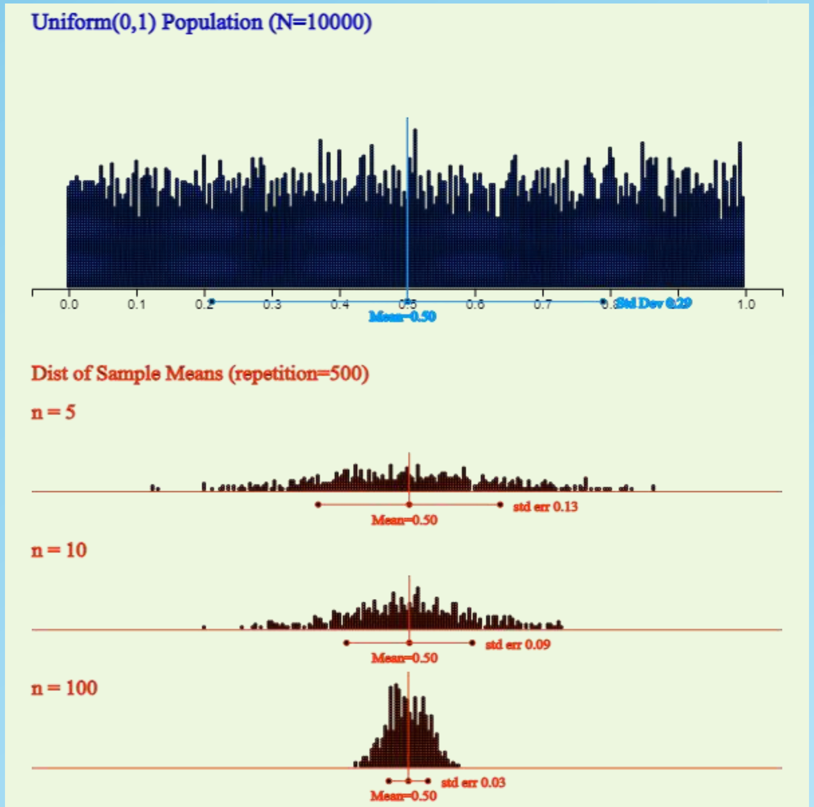
⇨ distribution of all possible sample means is approximately a normal distribution with mean μ and variance $\frac{\sigma^2}{n}$. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

6.2.1 Sampling Distribution of Sample Means

- Central Limit Theorem(CLT) - Simulation







Thank you