Introduction to Statistics and Data Science using *eStat*

Chapter 6 Sampling Distribution and Estimation

# 6.5 Determination of Sample Size

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

# 6.5.1 Determination of Sample Size to Estimate Population Mean

- **100(1-$\alpha$)% Confidence Interval for Population Mean – Known $\sigma^2$**
  **- Assume a population is a normal distribution.**

$$[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$$

- **margin of error** $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ **= d**
  (
  ⇨ **n =** $(z_{\alpha/2}\frac{\sigma}{d})^2$

$$\hat{\sigma} = \frac{max - min}{4}$$

## 6.5.1 Determination of Sample Size to Estimate Population Mean

**[Ex 6.5.1] The standard deviation of the life of a light bulb produced at a plant is usually 100 hours. To estimate the average life of a bulb at a 95% confidence level, calculate the size of the sample to be within 20 hours of the error.**

<Answer>

- $$n = [\frac{z_{\alpha/2}\,\sigma}{d}]^2 = (\frac{1.96 \times 100}{20})^2 = 9.8^2 = 96.04$$

  Hence the sample size is 97 approximately.

## 6.5.2 Determination of Sample Size to Estimate Population Proportion

- 100(1-$\alpha$)% Confidence Interval for Population Proportion $p$ - Assume a large sample.

$$\left[\, \hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\,,\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \,\right]$$

(

- margin of error $z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ = d

⇨ n = $\hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{d}\right)^2$

## 6.5.2 Determination of Sample Size to Estimate Population Proportion

**[Ex 6.5.2] For this year's presidential election, a survey for estimating the 95% confidence interval of candidate's approval rating is conducted. Obtain the size of the sample to be within 2.5% of the error bound.**

<Answer>

- Since we do not have information on the population proportion, assume $\hat{p}$ = 0.5. Then the sample size $n$ can be calculated as follows.

$$n = \hat{p}(1-\hat{p})(\frac{z_{\alpha/2}}{d})^2$$

$$n = 0.5(1-0.5)\frac{1.96^2}{0.025^2} = 1536.6$$

- Therefore, samples must be extracted with at least 1537 persons to limit the error bound to 2.5%. Various opinion polls conducted in Korea often show that the size of the sample is around 1,500 which implies that the margin of error does not exceed 3 percent.

# Thank you