Introduction to Statistics and Data Science using *eStat*

Chapter 6 Sampling Distribution and Estimation

6.4 Sampling Distribution of Sample Proportions and Estimation of Population Proportion

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

6.4.1 Sampling Distribution of Sample Proportions

[Example 6.4.1] Let's call 10 employees of a company a population. When the employees' satisfaction level with the company is investigated and the satisfaction is expressed as 1, the complaint is 0 as follows.

1 0 1 1 0 1 1 0 0 1

That is, the population proportion p of the satisfaction is 0.6.
Consider all possible samples of size 5 with replacement to obtain a sampling distribution of sample proportions.

[Answer of Ex 6.4.1]

Table 6.4.1 All possible sample cases

Sample Case	Number of cases
all unsatisfactory (0,0,0,0,0)	${}_5C_0$×4×4×4×4×4 = 1024
1 satisfactory (0,0,0,0,1)	${}_5C_1$×4×4×4×4×6 = 7680
2 satisfactory (0,0,0,1,1)	${}_5C_2$×4×4×4×6×6 = 23040
3 satisfactory (0,0,1,1,1)	${}_5C_3$×4×4×6×6×6 = 34560
4 satisfactory (0,1,1,1,1)	${}_5C_4$×4×6×6×6×6 = 25920
5 satisfactory (1,1,1,1,1)	${}_5C_5$×6×6×6×6×6 = 7776
	계 100000

Table 6.4.2. Sampling Distribution of Sample Proportions

Sample case	$\hat{p}$	Frequency	Relative Frequency
all unsatisfactory	0.0	1024	0.01024
1 satisfactory	0.2	7680	0.07680
2 satisfactory	0.4	23040	0.23040
3 satisfactory	0.6	34560	0.34560
4 satisfactory	0.8	25920	0.25920
5 satisfactory	1.0	7776	0.07776
	Total	100000	1.0

6.4.1 Sampling Distribution of Sample Proportions

- **Sampling distribution of sample proportions**

Assume the population is infinite and the population proportion is p.
If $\hat{p}$ is the sample proportion and the sample size n is large, the sampling distribution of all possible sample proportions is approximately a normal distribution with the mean p and variance $p(1-p)/n$.

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$$

6.4.1 Sampling Distribution of Sample Proportions

[Ex 6.4.2] Let's say 3% of semiconductors made in a semiconductor factory are defective. When 300 samples were taken without replacement, the sample proportion for defective products was 2%. Find out where this sample rate is located among all possible sample proportions. What is the probability that the sample proportion is greater than 2%?

<Answer>

- Since the sampling distribution of the sample proportions is approximately normal distribution, $\hat{p} \sim N(0.03, \frac{0.03(1-03)}{300})$, the probability can be calculated as follows:

$$
\begin{aligned}
P(\hat{p} > 0.02) &= P(Z > (0.02-0.03)/0.00985) \\
&= P(Z > -1.02) \\
&= 1 - P(Z \leq -1.02) \\
&= 1 - 0.1539 = 0.8461
\end{aligned}
$$

Thank you