Sampling Distribution of Sample Proportion

Ansar Shahzadi

School of Electrical Engineering & Computer Science

National University of Science and Technology(NUST)

Definitions

Consider a population in which each member either has or does not have specified attributes. Then we use the following notations and terminology.

Population Proportion, p: The proportion (percentage) of the entire population that has specified attribute.

Sample Proportion, $\hat{p}$: The proportion (percentage) of a sample from the population that has specified attribute.

Examples of Proportions

Statisticians often need to determine the proportion (percentage) of a population that has a specific attribute. Some examples are:

- the percentage of Pakistan adults who have health insurance
- the percentage of cars in the Pakistan that are imports
- the percentage of students who are favor for open book paper in NUST

Population Proportion

A population proportion p may be identified with the population mean, where the means is obtained from the units whose possible values either 0's or 1's.

Mean=$\frac{Number\ of\ units\ having\ \ the\ specified\ attributes}{N}$

p=$\frac{X}{N}$, where X represents the number of units having the specified attributes.

Sample Proportion

A sample proportion, $\hat{P}$, is computed by using the formula

$$\hat{P} = \frac{x}{n},$$

where *x* denotes the number of units in the sample that have the specified attribute and, as usual, *n* denotes the sample size. For convenience, we sometimes refer to *x* as the number of successes and to *n* – *x* as the number of failures.

Sampling Distribution of Sample Proportion

The sample proportion $\hat{P}$ has different values in different samples. Obviously Sample Proportion is a random variable and has a probability distribution. This probability distribution of the proportions of all possible random samples of size n is called the sampling distribution of $\hat{P}$.

$\hat{P}$	**f($\hat{P}$)**
0	1/20
1/3	9/20
2/3	9/20
1	1/20
Sum	1

Properties of Sampling Distribution of $\hat{P}$

The sampling distribution of $\hat{P}$ has the following properties:

- The mean of the sampling distribution of proportions, denoted by $\mu_{\hat{p}}$, is equal to the population proportion p.

  that is $\mu_{\hat{p}} = \mathrm{p}$.

- The standard deviation of sampling distribution of proportion is given by as

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}},$$

  when sampling is performed with replacement or

Properties of Sampling Distribution of $\hat{P}$

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n} \cdot \frac{N-n}{N-1}}$$

When sampling is performed without replacement from a finite population of size N

- Shape of the Distribution. The sampling distribution of $\hat{P}$ is the binomial distribution. However, for sufficiently large sample sizes, the sampling distribution of $\hat{P}$ is approximately normal whenever both np and nq are equal to or greater than 5.

Example

A population consists of N=6 numbers 1, 3, 6, 8, 11, and 16. Draw all possible sample of size n=3 without replacement from this population and find the sample proportion of odd numbers in the samples. Construct the sampling distribution of sample proportion and verify

- $\mu_{\hat{p}} = p$
- $\sigma_{\hat{p}} = \sqrt{\frac{pq}{n} \cdot \frac{N-n}{N-1}}$

Solution: Sample Space

Sample No	Sample Points	Sample Proportion	Sample No	Sample Points	Sample Proportion
1	1,3,6	2/3	11	3,6,8	1/3
2	1,3,8	2/3	12	3,6,11	2/3
3	1,3, 11	3/3	13	3,6,16	1/3
4	1,3,16	2/3	14	3,8,11	2/3
5	1,6,8	1/3	15	3,8,16	1/3
6	1,6,11	2/3	16	3,11,16	2/3
7	1,6,16	1/3	17	6,8,11	1/3
8	1,8,11	2/3	18	6,8,16	0/3
9	1,8,16	1/3	19	6,11,16	1/3
10	1,11,16	2/3	20	8,11,16	1/3

Sampling Distribution of $\hat{P}$

$\hat{P}$	f	f($\hat{P}$)	$\hat{P}$f($\hat{P}$)	$\hat{P}^2$f($\hat{P}$)
0	1	1/20	0	0
1/3	9	9/20	9/60	1/20
2/3	9	9/20	18/60	4/20
1	1	1/20	1/20	1/20
sum	20	1	10/20	6/20

Now

$\mu_{\hat{p}} = \sum \hat{P}f(\hat{P})$ =10/20=0.5

$\sigma_{\hat{p}} = \sqrt{\sum \hat{P}^2 f(\hat{P}) - (\sum \hat{P} f(\hat{P}))^2}$

$\sigma_{\hat{p}} = \sqrt{6/20 - (1/2)^2}$=0.2236

Verifications

Here p=$\frac{X}{N}$= $\frac{3}{6}$=0.5 and q=0.5

Therefore

$\mu_{\hat{p}}$ = p=0.5, and

$$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n} . \frac{N-n}{N-1}}$$

$$0.2236 = \sqrt{\frac{(0.5)(0.5)}{3} . \frac{6-3}{6-1}} = \sqrt{\frac{(0.5)(0.5)}{3} . \frac{3}{5}} =$$

$$\sqrt{\frac{(0.25)}{5}} = 0.2236$$

Home Assignment

Draw all possible samples of size n=3 with replacement from the population 2, 5, and 9. Compute the sampling distribution of sample proportion of even numbers and verify that

$$\boldsymbol{E(\widehat{P}) = p \text{ And } var(\widehat{P}) = \frac{pq}{n}}$$