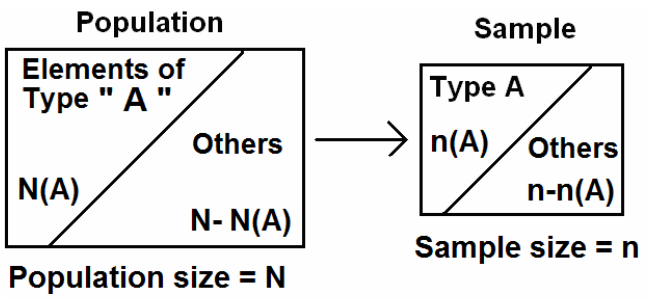
$$P(\bar{X}_1 - \bar{X}_2 > 20) = P\left(\frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} > \frac{20 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right)$$

$= P\left(Z > \frac{20-15}{4.0532}\right)$ = P(Z>1.23) =1 – P(Z<1.23)

= 1 – 0.8907

= 0.1093

5.5 Distribution of the Sample Proportion ($\hat{p}$):



- For the population:

  $N(A)$ = number of elements in the population with a specified characteristic "A"

  N = total number of elements in the population (population size)

The population proportion is

$$p = \frac{N(A)}{N} \qquad \text{(p is a parameter)}$$

- For the sample:

  $n(A)$ = number of elements in the sample with the same characteristic "A"

  n = sample size

The sample proportion is

$$\hat{p} = \frac{n(A)}{n} \qquad (\hat{p} \text{ is a statistic})$$

- The sampling distribution of $\hat{p}$ is used to make inferences

about p.

Result:

The mean of the sample proportion ($\hat{p}$) is the population proportion (p); that is:

$$\mu_{\hat{p}} = p$$

The variance of the sample proportion ($\hat{p}$) is:

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n} = \frac{pq}{n}. \qquad \text{(where q=1 −p)}$$

The standard error (standard deviation) of the sample proportion ($\hat{p}$) is:

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{pq}{n}}$$

Result:

For large sample size ($n \geq 30, np > 5, nq > 5$), the sample proportion ($\hat{p}$) has approximately a normal distribution with mean $\mu_{\hat{p}} = p$ and a variance $\sigma_{\hat{p}}^2 = pq/n$, that is:

$$\hat{p} \sim N\left(p, \frac{pq}{n}\right) \qquad \text{(approximately)}$$

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \sim N(0,1) \qquad \text{(approximately)}$$

Example:

Suppose that 45% of the patients visiting a certain clinic are females. If a sample of 35 patients was selected at random, find the probability that:

1. the proportion of females in the sample will be greater than 0.4.
2. the proportion of females in the sample will be between 0.4 and 0.5.

Solution:

- .n = 35 (large)
- p = The population proportion of females = $\frac{45}{100}$ = 0.45

- $\hat{p}$ = The sample proportion
  (proportion of females in the sample)
- The mean of the sample proportion ($\hat{p}$) is p = 0.45
- The variance of the sample proportion ($\hat{p}$) is:

$$\frac{p(1-p)}{n} = \frac{pq}{n} = \frac{0.45(1-0.45)}{35} = 0.0071.$$

- The standard error (standard deviation) of the sample proportion ($\hat{p}$) is:

$$\sqrt{\frac{p(1-p)}{n}} = \sqrt{0.0071} = 0.084$$

- $n \geq 30,\ \ np = 35 \times 0.45 = 15.75 > 5, nq = 35 \times 0.55 = 19.25 > 5$

1. The probability that the sample proportion of females ($\hat{p}$) will be greater than 0.4 is:

$$P(\hat{p} > 0.4) = 1 - P(\hat{p} < 0.4) = 1 - P\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} < \frac{0.4-p}{\sqrt{\frac{p(1-p)}{n}}}\right)$$

$$= 1 - P\left(Z < \frac{0.4-0.45}{\sqrt{\frac{0.45(1-0.45)}{35}}}\right) = 1 - P(Z < -0.59)$$

$$= 1 - 0.2776 = 0.7224$$

2. The probability that the sample proportion of females ($\hat{p}$) will be between 0.4 and 0.5 is:

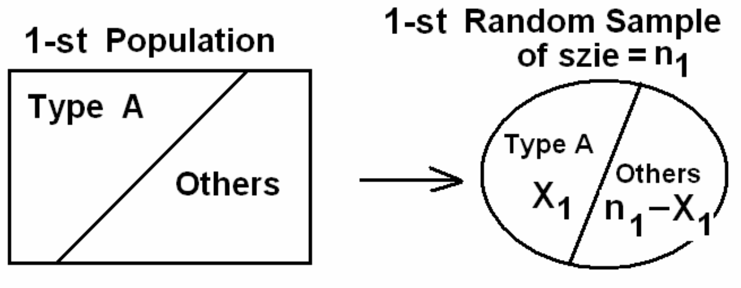
$$P(0.4 < \hat{p} < 0.5) = P(\hat{p} < 0.5) - P(\hat{p} < 0.4)$$

$$= P\left(\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} < \frac{0.5-p}{\sqrt{\frac{p(1-p)}{n}}}\right) - 0.2776$$

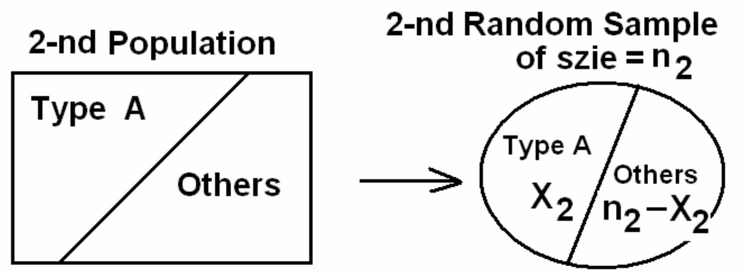
$$= P\left(Z < \frac{0.5-0.45}{\sqrt{\frac{0.45(1-0.45)}{35}}}\right) - 0.2776$$

$= P(Z < 0.59) - 0.2776$
$= 0.7224 - 0.2776$
$= 0.4448$

5.6 Distribution of the Difference Between Two Sample Proportions ($\hat{p}_1 - \hat{p}_2$):



independent



Suppose that we have two populations:

- p_1 = proportion of elements of type (A) in the 1-st population.
- p_2 = proportion of elements of type (A) in the 2-nd population.
- We are interested in comparing p_1 and p_2, or equivalently, making inferences about $p_1 - p_2$.
- We <u>independently</u> select a random sample of size n_1 from the 1-st population and another random sample of size n_2 from the 2-nd population:
- Let X_1 = no. of elements of type (A) in the 1-st sample.
- Let X_2 = no. of elements of type (A) in the 2-nd sample.
- $\hat{p}_1 = \frac{X_1}{n_1}$ = sample proportion of the 1-st sample