Module-3: Statistical Inference 1
Introduction, sampling distribution, standard error, testing of hypothesis, levels of significance, test of significances, confidence limits, simple sampling of attributes, test of significance for large samples, comparison of large samples. (12 Hours) (RBT Levels: L1, L2 and L3)

3.1 Sampling

Introduction:

- Entire group of individuals under study is called **population**. Quantity associated with population like mean(μ), SD(σ) is called **parameter**.

  Example: Set of all students in a college is a population. Mean weight of students is a parameter.

- A small part of the population is called **sample**. Quantity associated with sample like mean($\bar{x}$), SD(s) is called **statistic**.

  Example: Set of randomly selected 50 students from the college is a sample. Mean weight of these students is a statistic.

	Mean	S.D
Sample	$\bar{x}$	s
Population	μ	σ

Sampling distribution and standard error:

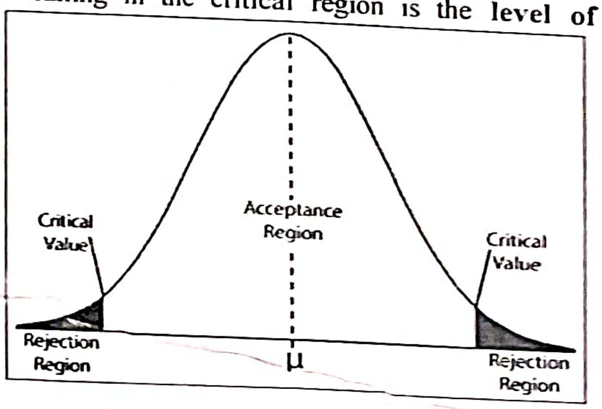
- The frequency distribution of means of different samples is called **sampling distribution** of mean. The standard deviation of the sampling distribution of mean is called **standard error** of mean. The reciprocal of the standard error is called **precision.**
- The number of units in the sample is called sample size. It is denoted by n. If $n \geq 30$, the sample is called large. Otherwise, small.

Testing of hypothesis and level of significance:

- Some assumption about the population based on sample information is called **statistical hypothesis**. This may or may not be true This is useful to reach some decision about the population.
- **Testing a hypothesis** is a process to decide whether to accept or reject the hypothesis. Initially, assume the hypothesis is correct and then find the probability of getting the observed sample. If this probability is less than some pre-assigned value, the hypothesis is rejected.
- A statistical hypothesis which we formulate to check whether it can be rejected is called **null hypothesis** (H_0). The negation of the null hypothesis is called **alternative hypothesis** (H_1).
- Rejecting H_0 when it is true is called **Type I error**. P (Type I error) is called **level of significance**. It is denoted by α. Accepting H_0 when it is false is called **Type II error**. P (Type II error) is called **power of the test**. It is denoted by β.

	True	False
Accept H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

- The region in which the calculated sample value falling is rejected is called **critical region.** The limits of the critical region are called **critical values**. Critical value splits the region in to **acceptance region** and **critical region**. These are pre-assigned values. The probability of the value of a variate falling in the critical region is the **level of significance.**



Test of significance and confidence limits:

- The procedure which enables us to decide whether to accept or reject the hypothesis is called the **test of significance.**
- An interval which is likely to contain the parameter is called **confidence interval**. The limits of the confidence interval are called **confidence limits**.
  Example: While we check BP, 80 and 120 are lower and upper limits. Interval 80-120 is called confidence interval.
- The probability that the confidence interval contains the parameter is called **confidence coefficient**. It is denoted by $1 - \alpha$.
- Confidence limits for means for large sample is $\bar{x} \pm Z_{\alpha/2} \cdot SE(\bar{x})$, where $SE(\bar{x}) = \frac{s}{\sqrt{n}}$.

Simple sampling attributes:

- An attribute means quality or characteristic such as drinking, smoking, disease, etc. An attribute may be marked by its presence (K) or absence (not K) in a member of given population. The sampling of attributes may be regarded as the selection of samples from population whose members posses the attribute K or not K.
- The presence of K is the success and its absence a failure. Suppose we draw a simple sample of size n items, it follows binomial distribution and hence the mean of this distribution is np and standard deviation of this distribution is $\sqrt{npq}$.

Introduction:

- Binomial distribution tends to normal for large n. For a normal distribution, only 5% of the members lie outside $\mu \pm 1.96\sigma$ and only 1% of the members lie outside $\mu \pm 2.58\sigma$.

Working rule:

- Write the null hypothesis H_0.
- Find the calculated value using

$$|z| = \left|\frac{p-P}{S.E(p)}\right|, \text{ where } SE(P) = \sqrt{\frac{PQ}{n}}, \text{ if } P \text{ is known}$$

$$|z| = \left|\frac{\bar{x}-\mu}{S.E(\bar{x})}\right|, \text{ where } SE(\bar{x}) = \sqrt{\frac{s^2}{n}}, \text{ if } \bar{x} \text{ is known.}$$

- Find the critical value using the table.
- If calculated value < critical value, accept H_0. H_0 is the conclusion.
- If calculated value > critical value reject H_0. H_1 is the conclusion.

Note: $S.E(p) = \sqrt{\frac{pq}{n}}$, if p is known.

$$S.E(p) = \sqrt{Var(p)}$$

$$= \sqrt{Var\left(\frac{x}{n}\right)}$$

$$= \sqrt{\frac{1}{n^2}Var(x)}$$

$$= \sqrt{\frac{1}{n^2}npq} = \sqrt{\frac{pq}{n}}$$

Confidence interval:

$\bar{x} \pm 3[SE(\bar{x})]$, if $\bar{x}$ is known.

$P \pm 3[SE(P)]$, if P is known.

$p \pm 3[SE(p)]$, if P is known.

Confidence interval at α level of significance:

$\bar{x} \pm z_{\frac{\alpha}{2}}[SE(\bar{x})]$, if $\bar{x}$ is known.

$P \pm z_{\frac{\alpha}{2}}[SE(P)]$, if P is known.

$p \pm z_{\frac{\alpha}{2}}[SE(p)]$, if P is known.

VtuCode

1. **A sample of 900 members is found to have a mean of 3.4 cm. Can it be reasonably regarded as a truly random sample from a large population with mean 3.25 cm and SD 1.61 cm.**

By data, $\bar{x} = 3.4, n = 900,\ \mu = 3.25, \sigma = 1.61$

This is a large sample. Apply z test.

$$S.E(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{1.61^2}{900}} = 0.0537$$

H_0: $\mu = 3.25$, Sample is taken from the population with mean 3.25

$$|z| = \left|\frac{\bar{x} - \mu}{S.E(\bar{x})}\right| = \left|\frac{3.4 - 3.25}{0.0537}\right| = 2.8$$

Therefore, calculated value of $z = 2.8$

At $\alpha = 0.05$, critical value of $z = 1.96$

Since calculated value > critical value, Reject H_0.

Therefore, sample is not taken from the population with mean 3.25

2. **If a mean breaking strength of copper wire is 575 lbs with a standard deviation 8.3 lbs. How large a sample must be used in order that there be one chance in 100 that the mean breaking strength of the sample is less than 572 lbs.** $(Z_\alpha = 2.33)$

By data, $\bar{x} = 572,\ \mu = 575, \sigma = 8.3$.

This is a large sample. Apply z test.

$$S.E(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{8.3^2}{n}}$$

H_0: $\mu = 575$, mean breaking strength of copper wire is 575 lbs.

$$|z| = \left|\frac{\bar{x}-\mu}{S.E(\bar{x})}\right| = \left|\frac{572-575}{\sqrt{\frac{8.3^2}{n}}}\right|$$

To find: n such that $\mu < 572$

That is, find n such that H_0 is rejected

That is, find n such that Calculated value > Critical value.

$$\Rightarrow \left|\frac{572-575}{\sqrt{\frac{8.3^2}{n}}}\right| > 2.33\ ,\quad n > 41.56$$

Therefore, $n = 42$.

3. **The mean of a certain normal population is equal to the standard error of the mean of the samples of 100 from that distribution. Find the probability that the mean of the sample of 25 from the distribution will be negative.**

By data, $\mu = SE\,(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{\sigma^2}{100}} = \frac{\sigma}{10}$

To find: $P(\bar{x} < 0)$ when $n = 25$.

When $n = 25$, $SE(\bar{x}) = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{\sigma^2}{25}} = \frac{\sigma}{5}$

$$\therefore P(\bar{x} < 0) = P\left(\frac{\bar{x}-\mu}{S.E(\bar{x})} < \frac{0-\mu}{S.E(\bar{x})}\right)$$

$$= P\left(z < \frac{-\mu}{S.E(\bar{x})}\right)$$

$$= P\left(z < -\frac{\frac{\sigma}{10}}{\frac{\sigma}{5}}\right)$$

$$= P\left(z < -\frac{1}{2}\right)$$

$$= 0.3085$$

4. **A coin was tossed 400 times and the head turned up 216 times. Test the hypothesis that the coin is unbiased at 5% level of significance.** $[z_{\frac{\alpha}{2}} = 1.96]$.

Since $n = 400$, apply z test.

By data, $p = \frac{216}{400}$ and $P = \frac{1}{2} = 0.5$

$$SE(P) = \sqrt{\frac{PQ}{n}} = \sqrt{\frac{(0.5)(0.5)}{400}} = 0.025$$

H_0: $P = 0.5$,The coin is unbiased.

$$|z| = \left|\frac{p-P}{SE(P)}\right| = \left|\frac{\frac{216}{400}-\frac{1}{2}}{0.025}\right| = 1.6$$

Therefore, calculated value of $z = 1.6$

At $\alpha = 0.05$, critical value of $z = 1.96$

Since calculated value < critical value, Accept H_0.

Therefore, the coin is unbiased at 5% level of significance.

5. **A die was thrown 9000 times and a throw of 5 or 6 was obtained 3240 times. On the assumption of random throwing, do the data indicate an unbiased die? ($\alpha = 0.01$)**

$[z_{\frac{\alpha}{2}} = 2.58]$

Since $n = 9000$, apply z test.

By data, $p = \frac{3240}{9000}$ and $P = \frac{2}{6} = \frac{1}{3}$

$$SE(P) = \sqrt{\frac{PQ}{n}} = \sqrt{\frac{(1/3)(2/3)}{9000}} = 0.005$$

H_0: $P = \frac{1}{3}$, The die is unbiased.

$$|z| = \left|\frac{p-P}{SE(P)}\right| = \left|\frac{\frac{3240}{9000} - \frac{1}{3}}{0.005}\right| = 5.33$$

At $\alpha = 0.01$, critical value of $z = 2.58$

Since calculated value > critical value, Reject H_0.

Therefore, the die is biased at 1% level of significance.

6. **In 324 throws of a die, an odd number turned up 181 times. Is it reasonable to think that ay 1% level of significance the die is an unbiased one?** $[z_{\frac{\alpha}{2}} = 2.58]$

Since $n = 324$, apply z test.

By data, $p = \frac{181}{324}$ and $P = \frac{3}{6} = \frac{1}{2}$

$$SE(P) = \sqrt{\frac{PQ}{n}} = \sqrt{\frac{(1/2)(1/2)}{324}} = 0.0278$$

H_0: $P = \frac{1}{3}$, The die is unbiased.

$$|z| = \left|\frac{p-P}{SE(P)}\right| = \left|\frac{\frac{181}{324} - \frac{1}{2}}{0.0278}\right| = 2.1084$$

Therefore, calculated value of $z = 2.1084$

At $\alpha = 0.01$, critical value of $z = 2.58$

Since calculated value < critical value, accept H_0.

Therefore, the die is unbiased at 1% level of significance.

Note: The die is biased at 5% level of significance. ($\because$ critical value of $z = 1.96$)

VTUCode

7. In a locality containing 18000 families, a sample of 840 families was selected at random. Of these 840 families, 206 families were found to have a monthly income of ₹ 25,000 or less. It is desired to estimate how many out of 18,000 families have a monthly income of ₹ 25,000 or less. Within what limits would you place your estimate?

By data, $p = \frac{206}{840} = 0.2452$

$$SE(p) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.2452)(0.7548)}{840}} = 0.0148$$

Required confidence limits are given by $p \pm 3[SE(p)]$

$$= 0.2452 \pm 3(0.0148)$$

$$= 0.2452 \pm 0.0444$$

$$= 0.2452 - 0.0444 \text{ and } 0.2452 + 0.0444$$

$$= 0.2008 \text{ and } 0.2896$$

8. An unbiased coin is thrown n times. It is desired that the relative frequency of the appearance of heads should lie between 0.49 and 0.51. Find the smallest value of n that will ensure this result with 90% confidence. ($Z_{\frac{\alpha}{2}}(0.1) = 1.645$)

By data, $p = 0.5,\ q = 0.5$

Standard error of proportion:

$$S.E(p) = \sqrt{\frac{pq}{n}} = \frac{1}{2\sqrt{n}}$$

90% confidence interval is given by $0.5 \pm Z_{\frac{\alpha}{2}}.SE(p)$

$$= (0.5 - Z_{\frac{\alpha}{2}}.SE(p), 0.5 + Z_{\frac{\alpha}{2}}.SE(p))$$

$$= \left(0.5 - \frac{1.645}{2\sqrt{n}}, 0.5 + \frac{1.645}{2\sqrt{n}}\right)$$

By data, $\left(0.5 - \frac{1.645}{2\sqrt{n}}, 0.5 + \frac{1.645}{2\sqrt{n}}\right) = (0.49, 0.51)$

Therefore, the smallest value of n is given by

$$0.5 + \frac{1.645}{2\sqrt{n}} = 0.51$$

$$1.645\frac{1}{2\sqrt{n}} = 0.01$$

$$n = 6765$$

9. **A random sample of 500 pineapples was taken from a large consignment and 65 were found to be bad. Show that the standard error of the proportion of bad ones in a sample of this size is 0.015 and deduce that the percentage of bad pineapples in the consignment almost certainly lies between 8.5 and 17.5.**

By data, $p = \frac{65}{500} = 0.13$

$S.E(p) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{0.13\times0.87}{500}} = 0.015$

Required confidence interval$= (p - 3[SE(p)], p + 3[SE(p)])$

$= (0.13 - 3(0.015),\ 0.13 + 3(0.015))$

$= (0.085,\ 0.175)$

$= (8.5\%, 17.5\%)$

VtuCode

3.3 Comparison of large samples

Working rule:

- Write the null hypothesis H_0.
- Find the calculated value using

$$|z| = \begin{cases} \left|\frac{\overline{x_1}-\overline{x_2}}{SE(\overline{x_1}-\overline{x_2})}\right| &, \text{ if } \overline{x_1},\overline{x_2} \text{ are known} \\ \left|\frac{p_1-p_2}{SE(p_1-p_2)}\right| &, \text{ if } \overline{x_1},\overline{x_2} \text{ are not known} \end{cases}$$

- Find the critical value using the table.
- If calculated value < critical value, accept H_0. H_0 is the conclusion.
- If calculated value > critical value reject H_0. H_1 is the conclusion.

Standard error:

$$SE(\overline{x_1}-\overline{x_2}) = \begin{cases} \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}, & If\ s_1, s_2\ are\ known \\ \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}, & If\ \sigma_1, \sigma_2\ are\ known \\ \sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}, & If\ \sigma\ is\ known \end{cases}$$

$$SE(p_1-p_2) = \begin{cases} \sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}, & If\ P_1, P_2\ are\ known \\ \sqrt{PQ\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}, & If\ p_1, p_2\ are\ known \end{cases}$$

where,

$$P = \frac{n_1p_1+n_2p_2}{n_1+n_2}$$

VtuCode

1. **The means of samples of sizes 1000 and 2000 are 67.5 and 68.0 cms respectively. Can the samples be regarded as drawn from the same population of SD 2.5 cm?**

   $[\, z_{\frac{\alpha}{2}}(0.05) = 1.96]$

Since samples sizes are $n_1 = 1000,\ n_2 = 2000$, apply z test.

By data, $\overline{x_1} = 67.5, \overline{x_2} = 68.0\ \ and\ \ \sigma = 2.5$

H₀: $\mu_1 = \mu_2$, Both the samples are drawn from the same population.

$$SE(\overline{x_1} - \overline{x_2}) = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= 2.5\sqrt{\frac{1}{1000} + \frac{1}{2000}} = 0.0968$$

$$|z| = \left|\frac{\overline{x_1} - \overline{x_2}}{SE(\overline{x_1} - \overline{x_2})}\right|$$

$$= \left|\frac{67.5 - 68.0}{0.0968}\right| = \frac{0.5}{0.0968} = 5.16$$

Therefore, calculated value of $z = 5.16$

At $\alpha = 0.05$, critical value of $z = 1.96$

Since calculated value > critical value, reject H₀.

Therefore, Both the samples are not drawn from the same population.

2. **A sample of height of 6400 soldiers has a mean of 67.85 inches and a standard deviation of 2.56 inches while a sample of height of 1600 sailors has a mean of 68.55 inches and a SD of 2.52 inches. Does the data indicate that the sailors are on an average taller than soldiers? Use 0.05 level of significance.** $[z_\alpha = 1.65]$

Since samples sizes are $n_1 = 6400, n_2 = 1600$, apply z test.

By data, $\overline{x_1} = 67.85,\ \overline{x_2} = 68.55,\ s_1 = 2.56,\ s_2 = 2.52$

H_0: $\mu_1 = \mu_2$, The sailors are not taller than soldiers.

$$SE(\overline{x_1} - \overline{x_2}) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= \sqrt{\frac{2.56^2}{6400} + \frac{2.52^2}{1600}}$$

$$= 0.0707$$

$$|z| = \left|\frac{\overline{x_1} - \overline{x_2}}{SE(\overline{x_1} - \overline{x_2})}\right|$$

$$= \left|\frac{67.85 - 68.55}{0.0707}\right| = \frac{0.7}{0.0707} = 9.9$$

Therefore, calculated value =9.9

At $\alpha = 0.05$, critical value = 1.65

Since calculated value > critical value, reject H_0.

Therefore, the sailors are taller than soldiers at 0.05 level of significance.

VTUCode

3. A sample of 100 electric bulbs produced by manufacturer A showed a mean lifetime of 1190 hours and a standard deviation of 90 hours. A sample of 75 bulbs produced by manufacturer B showed a mean lifetime of 1230 hours with a standard deviation of 120 hours. Is there a difference between the mean lifetime of two brands at significant level of 0.05? ($Z_{\alpha/2} = 1.96$)

Since samples sizes are $n_1 = 100, n_2 = 75$, apply z test.

By data, $\overline{x_1} = 1190,\ \overline{x_2} = 1230,\ s_1 = 90,\ s_2 = 120$

H₀: $\mu_1 = \mu_2$, There is no difference between the mean lifetime of two brands.

$$SE(\overline{x_1} - \overline{x_2}) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= \sqrt{\frac{90^2}{100} + \frac{120^2}{75}} = 16.5227$$

$$|z| = \left|\frac{\overline{x_1} - \overline{x_2}}{SE(\overline{x_1} - \overline{x_2})}\right|$$

$$= \left|\frac{1190-1230}{16.5227}\right| = 2.4209$$

Therefore, calculated value = 2.4209

At $\alpha = 0.05$, critical value = 1.96

Since calculated value > critical value, reject H₀.

Therefore, there is a difference between the mean lifetime of two brands at significant level of 0.05.

VTU Code

4. **One type of aircraft is found to develop engine trouble in 5 flights out of a total of 100 and another type in 7 flights out of a total of 200 flights. Is there a significant difference in the two types concerned so far as engine defects are concerned?** $(Z_{\alpha/2} = 1.96)$

Since samples sizes are $n_1 = 100$ *and* $n_2 = 200$, apply z test.

By data, $p_1 = \frac{5}{100}$, $p_2 = \frac{7}{200}$, $\alpha = 0.05$

H₀: $P_1 = P_2$. There is no significant difference in the two types concerned so far as engine defects are concerned.

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$= \frac{100(0.05) + 200(0.035)}{100 + 200} = 0.04$$

$$SE(p_1 - p_2) = \sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.04 \times 0.96 \times \left(\frac{1}{100} + \frac{1}{200}\right)}$$

$$= 0.024$$

$$|z| = \left|\frac{p_1 - p_2}{SE(p_1 - p_2)}\right|$$

$$= \frac{0.015}{0.024} = 0.625$$

Therefore, calculated value = 0.625

At $\alpha = 0.05$, critical value = 1.96

Since calculated value < critical value, Accept H₀.

Therefore, there is no significant difference in the two types concerned so far as engine defects are concerned.

VtuCode

5. **A machine produces 16 imperfect articles in a sample of 500. After the machine is overhauled, it produces 3 imperfect articles in a batch of 100. Has the machine been improved? [$z_\alpha = 1.65$]**

Since samples sizes are $n_1 = 500$ *and* $n_2 = 100$, apply z test.

By data, $p_1 = \frac{16}{500} = 0.032,\ p_2 = \frac{3}{100} = 0.03$

H₀: $P_1 = P_2$, The machine has not been improved.

$$P = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{500(0.032) + 100(0.03)}{500 + 100} = 0.0317$$

$$SE(p_1 - p_2) = \sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.0317 \times 0.9683 \times \left(\frac{1}{500} + \frac{1}{100}\right)}$$

$$= 0.0192$$

$$|z| = \left|\frac{p_1 - p_2}{SE(p_1 - p_2)}\right|$$

$$= \left|\frac{0.032 - 0.03}{0.0192}\right| = \frac{0.002}{0.0192} = 0.1042$$

Therefore, calculated value = 0.1042

At $\alpha = 0.05$, critical value = 1.65

Since calculated value < critical value, Accept H₀.

Therefore, the machine has not been improved.

VtuCode

6. **In a city A 20% of a random sample of 900 schoolboys had a certain slight physical defect. In another city B 18.5% of a random sample of 1600 schoolboys had the same defect. Is the difference between the proportions significant?**

Since samples sizes are $n_1 = 900\ and\ n_2 = 1600$, apply z test.

By data, $p_1 = 20\% = 0.2,\ p_2 = 18.5\% = 0.185$

H₀: $P_1 = P_2$, the difference between the proportions in not significant.

$$P = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{900(0.2) + 1600(0.185)}{900 + 1600} = 0.1904$$

$$SE(p_1 - p_2) = \sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.19 \times 0.81 \times \left(\frac{1}{900} + \frac{1}{1600}\right)}$$

$$= 0.0163$$

$$|z| = \left|\frac{p_1 - p_2}{SE(p_1 - p_2)}\right|$$

$$= \left|\frac{0.2 - 0.185}{0.0163}\right| = \frac{0.015}{0.0163} = 0.92$$

Therefore, calculated value = 0.92

At $\alpha = 0.05$, critical value = 1.96

Since calculated value < critical value, Accept H₀.

Therefore, there is no significant difference between the proportions.

VTUCode

7. **In two large populations there are 30% and 25% respectively of fair-haired people. Is this difference likely to be hidden in samples of 1200 and 900 respectively from the two populations?**

Since samples sizes are $n_1 = 1200$ *and* $n_2 = 900$, apply z test.

By data, $P_1 = 30\% = 0.3,\ P_2 = 25\% = 0.25$

H_0: $P_1 = P_2$, the difference between the proportions in not significant.

$$SE(P_1 - P_2) = \sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}$$

$$= \sqrt{\frac{(0.3)(0.7)}{1200} + \frac{(0.25)(0.75)}{900}}$$

$$= 0.0196$$

$$|z| = \left|\frac{P_1 - P_2}{SE(P_1 - P_2)}\right|$$

$$= \left|\frac{0.3 - 0.25}{0.0196}\right| = \frac{0.05}{0.0196} = 2.5510$$

Therefore, calculated value = 2.5510

At $\alpha = 0.05$, critical value = 1.96

Since calculated value > critical value, Reject H_0.

Therefore, this difference is **unlikely** to be hidden in samples of 1200 and 900 respectively from the two populations.