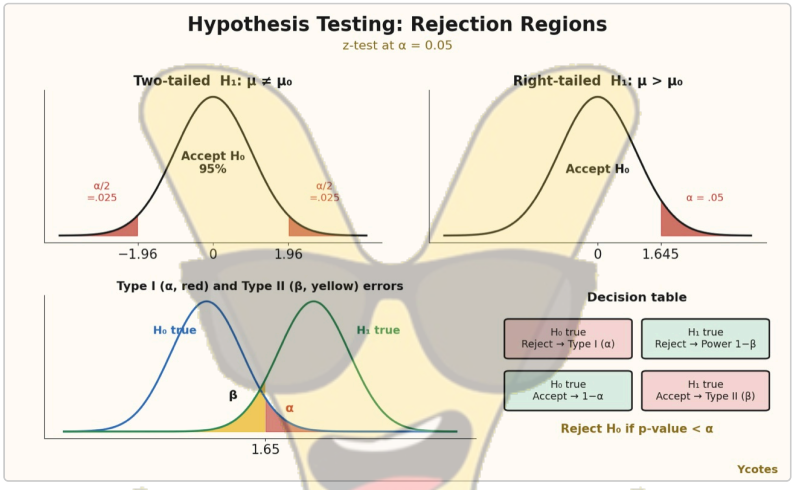
UNIT–3 : Hypothesis Testing (PYQ Solution)

PART–A Answers (≈70 Words Each)

Q1. Define Type I error, Type II error and p-value. (Expected)

- **Type I error (α): rejecting a true null hypothesis H_0. α is the level of significance (usually 0.05).**
- **Type II error (β): accepting a false H_0. Power $= 1 - \beta$.**
- **p-value: the probability, assuming H_0 is true, of getting a result at least as extreme as the one observed. Reject H_0 if p-value $< \alpha$.**



Q2. When do we use the t-test instead of the z-test? What are degrees of freedom? (Expected)

Use the t-test when the sample is small ($n < 30$), the population is normal and the population standard deviation σ is unknown (estimated by s). Use the z-test when σ is known or n is large.

Degrees of freedom is the number of independent values free to vary. For one sample $df = n - 1$ (one estimate, $\bar{x}$, is used up); for two samples with pooled variance $df = n_1 + n_2 - 2$. The t-curve approaches the normal curve as df grows.

PART–B Answers (≈250 Words Each)

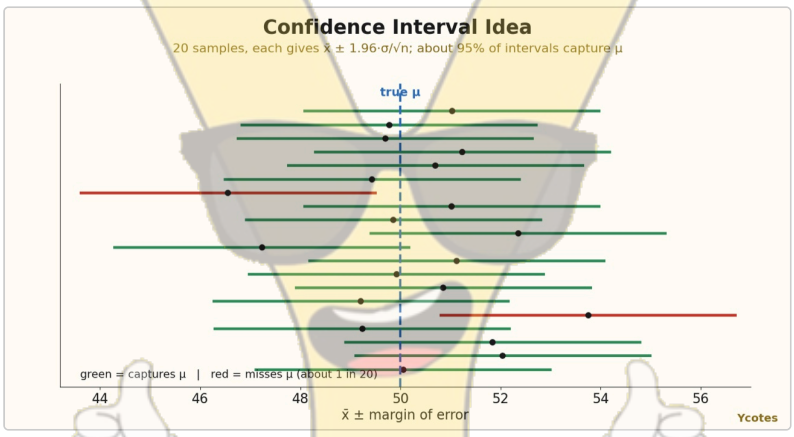
Q1. A sample of 64 items has mean 52. The population standard deviation is 8. Test H0: mu = 50 at 5% level (large-sample test) and give the 95% confidence interval. (Expected)

$H_0 : \mu = 50$, $H_1 : \mu \neq 50$ (two-tailed), $\alpha = 0.05$, critical value $z = \pm 1.96$.

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{52 - 50}{8 / \sqrt{64}} = \frac{2}{1} = 2.0$$

Since $|z| = 2.0 > 1.96$, **reject** H_0 (p-value $= 2 \times 0.0228 = 0.0455 < 0.05$). The mean differs significantly from 50.

95% CI: $\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} = 52 \pm 1.96(1) = (50.04,\ 53.96)$. The interval just excludes 50, consistent with the test.



Q2. Ten values 48, 52, 50, 47, 55, 53, 49, 51, 54, 46 are drawn from a normal population. Test H0: mu = 50 (one-sample t-test) and find the 95% confidence interval. (Expected)

$n = 10$, $\bar{x} = 505/10 = 50.5$. $\sum (x - \bar{x})^2 = 82.5$, so $s^2 = 82.5/9 = 9.167$, $s = 3.028$.

$H_0 : \mu = 50$, $H_1 : \mu \neq 50$, $\alpha = 0.05$, $df = 9$, critical $t_{0.025,9} = 2.262$.

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{50.5 - 50}{3.028 / \sqrt{10}} = \frac{0.5}{0.9574} = 0.522$$

Since $0.522 < 2.262$, **accept** H_0 (p = 0.61). There is no evidence that the mean differs from 50.

95% CI: $\bar{x} \pm t\dfrac{s}{\sqrt{n}} = 50.5 \pm 2.262(0.9574) = (48.33,\ 52.67)$. It contains 50.

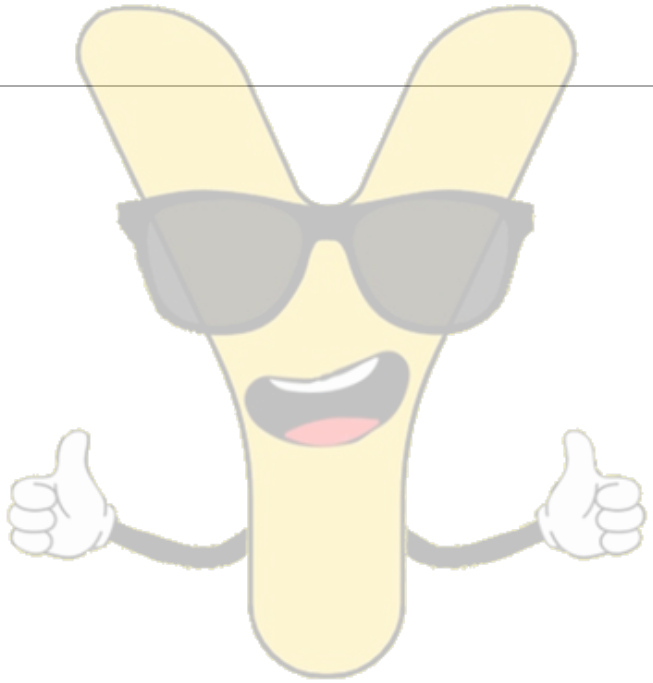
Q3. In a sample of 200 voters 60 favour a proposal. Test the claim that the true proportion is 0.25. (Expected)

$\hat{p} = 60/200 = 0.30$. $H_0 : p = 0.25$, $H_1 : p \neq 0.25$, $\alpha = 0.05$.

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}} = \frac{0.30 - 0.25}{\sqrt{0.25 \times 0.75/200}} = \frac{0.05}{0.03062} = 1.633$$

Since $1.633 < 1.96$, **accept** H_0 (p-value 0.102). The data do not contradict the claim.

95% CI for p: $0.30 \pm 1.96\sqrt{\dfrac{0.3 \times 0.7}{200}} = 0.30 \pm 0.0635 = (0.237,\ 0.364)$, which contains 0.25.



PART–C Answers (≈500 Words Each)

Q1. Two samples A: 12, 15, 14, 10, 13, 16 and B: 9, 11, 10, 12, 8, 10, 9 come from normal populations with equal variances. Test whether the means differ (two-sample pooled t-test). (Expected)

$n_1 = 6,\ \bar{x}_1 = 80/6 = 13.33,\ s_1^2 = 4.667; n_2 = 7,\ \bar{x}_2 = 69/7 = 9.857,\ s_2^2 = 1.810.$

$H_0 : \mu_1 = \mu_2, H_1 : \mu_1 \neq \mu_2, \alpha = 0.05, df = n_1 + n_2 - 2 = 11,$ critical $t = 2.201.$

Pooled variance:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{5(4.667) + 6(1.810)}{11} = \frac{34.19}{11} = 3.108, \quad s_p = 1.763$$

Test statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{3.476}{1.763\sqrt{0.1667 + 0.1429}} = \frac{3.476}{0.981} = 3.54$$

Since $3.54 > 2.201$, **reject** H_0 (p = 0.0046). The two population means differ significantly; A is higher.

95% CI for $\mu_1 - \mu_2$**:** $3.476 \pm 2.201(0.981) = (1.32,\ 5.63)$, which excludes 0.

Q2. In sample 1, 80 of 200 customers prefer a product, and in sample 2, 75 of 250. Test whether the proportions are equal and explain the large-sample test for difference of means. (Expected)

Proportions: $\hat{p}_1 = 80/200 = 0.40, \hat{p}_2 = 75/250 = 0.30.$ $H_0 : p_1 = p_2, H_1 : p_1 \neq p_2,$ $\alpha = 0.05.$

Pooled proportion: $\hat{p} = \dfrac{80 + 75}{200 + 250} = 0.3444.$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.10}{\sqrt{0.3444(0.6556)(0.009)}} = \frac{0.10}{0.04508} = 2.22$$

Since $2.22 > 1.96$, **reject** H_0 (p = 0.026): the proportions differ; sample 1 prefers the product more.

Brought to you By Ycotes : https://theycotes.com

Large-sample test for difference of means: for independent samples with $n_1, n_2 \geq 30$,

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

(replace σ^2 by s^2 if unknown). Example: $\bar{x}_1 = 82, s_1^2 = 36, n_1 = 50$; $\bar{x}_2 = 78, s_2^2 = 49, n_2 = 60$: $SE = \sqrt{0.72 + 0.8167} = 1.24$, $z = 4/1.24 = 3.23 > 1.96$, reject H_0.

Q3. Explain estimation of variance. A sample of 15 gives variance 16; find the 95% confidence interval for the population variance. (Expected)

Estimation: the sample variance s^2 is an unbiased point estimator of σ^2. For a normal population the quantity

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

follows a chi-square distribution with $n - 1$ degrees of freedom, which gives the interval

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,\,n-1}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\,n-1}}$$

Numerical: $n = 15$, $s^2 = 16$, $df = 14$, $(n-1)s^2 = 224$. $\chi^2_{0.025,14} = 26.119$, $\chi^2_{0.975,14} = 5.629$.

$$\frac{224}{26.119} = 8.58 \leq \sigma^2 \leq \frac{224}{5.629} = 39.80$$

95% CI for σ^2: (8.58, 39.80), so for σ: (2.93, 6.31). The interval is not symmetric about s^2 because the chi-square curve is skewed.

Test of a variance: $H_0 : \sigma^2 = \sigma_0^2$ uses the same statistic $\chi^2 = (n-1)s^2/\sigma_0^2$.
