CHAPTER 8

hypothesis is a claim or statement about a property of a population.

hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population. **e.g.** Genetics – Business – Medicine - Aircraft Safety - Quality Control.

Components of a Formal Hypothesis Test:

- **Null Hypothesis: (H_0):** is a statement that the value of a population parameter (such as proportion, mean, or standard deviation) is equal to some claimed value. test the null hypothesis directly.
- **Alternative Hypothesis: (H_1):** is the statement that the parameter has a value that somehow differs from the null hypothesis. the alternative hypothesis must use one of these symbols: ≠, <, >.

Test statistic is a value used in making a decision about the null hypothesis, and is found by converting the sample statistic to a score with the assumption that the null hypothesis is true.

Formulas:

Test statistic for proportion	$z = \dfrac{\hat{p} - p}{\sqrt{\dfrac{pq}{n}}}$
Test statistic for mean	$z = \dfrac{\bar{x} - \mu}{\dfrac{\sigma}{\sqrt{n}}}$ or $t = \dfrac{\bar{x} - \mu}{\dfrac{s}{\sqrt{n}}}$
Test statistic for standard deviation	$\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

Critical region (or rejection region) is the set of all values of the test statistic that cause us to reject the null hypothesis.

Significance level (denoted by α) is the probability that the test statistic will fall in the critical region when the null hypothesis is actually true.
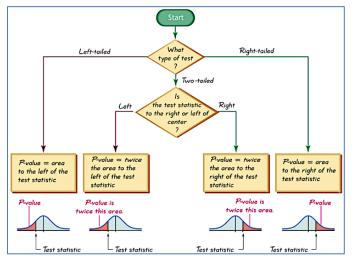
Critical value is any value that separates the critical region (where we reject the null hypothesis) from the values of the test statistic that do not lead to rejection of the null hypothesis.

The critical values depend on the:

- nature of the null hypothesis
- the sampling distribution that applies
- the significance level α.

P-value (or probability value) is the probability of getting a value of the test statistic that is at least as extreme as the one representing the sample data, assuming that the null hypothesis is true. The null hypothesis is rejected if the *P*-value is very small, such as 0.05 or less.

Critical region in the left tail:	*P*-value = area to the left of the test statistic
Critical region in the right tail:	*P*-value = area to the right of the test statistic
Critical region in two tails:	*P*-value = twice the area in the tail beyond the test statistic

P-values:

Tails in a distribution are the extreme regions bounded by critical values.

P-value method: Using the significance level α:
- If *P*-value $\leq \alpha$, reject H_0
- If *P*-value $> \alpha$, fail to reject H_0

Traditional method:
- If the test statistic falls within the critical region, reject H_0
- If the test statistic does not fall within the critical region, fail to reject H_0

Type I error is the mistake of rejecting the null hypothesis when it is actually true. (α) is used to represent the probability of a type I error.

Type II error is the mistake of failing to reject the null hypothesis when it is actually false. (β) is used to represent the probability of a type II error.

		True State of Nature	
		The null hypothesis is true	The null hypothesis is false
Decision	We decide to reject the null hypothesis	**Type I error** (rejecting a true null hypothesis) P(type I error) = α	Correct decision
	We fail to reject the null hypothesis	Correct decision	**Type II error** (failing to reject a false null hypothesis) P(type II error) = β

- For any fixed α, an increase in the sample size *n* will cause a decrease in β.

- For any fixed sample size *n*, a decrease in α will cause an increase in β. Conversely, an increase in α will cause a decrease in β.

- To decrease both α and β, increase the sample size.

Power of a hypothesis test is the probability (1 – β) of rejecting a false null hypothesis. The value of the power is computed by using a particular significance level α and a particular value of the population parameter that is an alternative to the value assumed true in the null hypothesis.

- The power of the hypothesis test is the probability of supporting an alternative hypothesis that is true.

Basic Methods of <u>Testing Claims about a Population Proportion *p*</u>:

n = number of trials

p^ $= \frac{x}{n}$ (sample proportion)

p = population proportion (used in the null hypothesis)

q = 1 − *p*

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

- *P^* sometimes is given directly: "10% of the observed sports cars are red" is expressed as *p^* = 0.10

- *P^* sometimes must be calculated: "96 surveyed households have cable TV and 54 do not" is calculated using $\hat{p} = \frac{x}{n} = \frac{96}{(96+54)} = 0.64$

Exact Method for Testing Claims about a Proportion *p*:

- ❖ **Left-tailed test:** The *P*-value is the probability of getting *x* or <u>fewer successes</u> among *n* trials.
- ❖ **Right-tailed test:** The *P*-value is the probability of getting *x* or <u>more successes</u> among *n* trials.
- ❖ **Two-tailed test:**

$$\text{If } \hat{p} > p, \quad \text{the } P\text{-value is twice the probability of getting } x \text{ or more successes}$$

$$\text{If } \hat{p} < p, \quad \text{the } P\text{-value is twice the probability of getting } x \text{ or fewer successes}$$

Testing Claims About a Population Mean (with σ Known)

n = sample size

$\overline{X}$ = sample mean

$\mu_{\overline{x}}$ = population mean of all sample means from samples of size *n*

σ = known value of the population standard deviation

- ❖ The population is normally distributed or *n* > 30.

$$Z = \frac{\bar{x} - \mu_{\overline{x}}}{\frac{\sigma}{\sqrt{n}}}$$

Testing Claims About a Population Mean (with σ NOT Known)

n = sample size

$\overline{X}$ = sample mean

$\mu_{\overline{x}}$ = population mean of all sample means from samples of size *n*

- ❖ The population is normally distributed or *n* > 30.

$$t = \frac{\bar{x} - \mu_{\overline{x}}}{\frac{s}{\sqrt{n}}}$$

P-values and Critical Values

- Found in Table A-3

- Degrees of freedom <u>(df) = *n* − 1</u>

❖ ***np* ≥ 5 and *nq* ≥ 5**

CHAPTER 9

Two Proportions:

> **For population 1, we let:**
>
> p_1 = **population** proportion
>
> n_1 = **size of the sample**
>
> x_1 = **number of successes in the sample**
>
> $\hat{p}_1 = \dfrac{x_1}{n_1}$ (the sample proportion)
>
> $\hat{q}_1 = 1 - \hat{p}_1$
>
> The corresponding notations apply to
> $p_2, n_2, x_2, \hat{p}_2,$ and $\hat{q}_2$, which come from **population 2**.

pooled sample proportion is denoted by $\bar{p}$ and is given by:

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

The complement of $\bar{p}$ by $\bar{q}$, so $\bar{q} = 1 - \bar{p}$

Test Statistic for Two Proportions:

- ❖ For **H$_0$: p$_1$ = p$_2$**
- ❖ **H$_1$: p$_1$ ≠ p$_2$, H$_1$: p$_1$ < p$_2$, H$_1$: p$_1$ > p$_2$**

$$Z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\dfrac{\bar{p}\bar{q}}{n_1} + \dfrac{\bar{p}\bar{q}}{n_2}}}$$

Were $P_1 - P_2 = 0$

$$\hat{p}_1 = \frac{x_1}{n_1} \quad \text{and} \quad \hat{p}_2 = \frac{x_2}{n_2}$$

$$\bar{p} = \frac{x_1 + x_2}{n_1 + n_2} \quad \text{and} \quad \bar{q} = 1 - \bar{p}$$

Confidence Interval Estimate of P$_1$ – P$_2$:

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

$$\text{where} \quad E = z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

Two samples:

- ❖ **independent** if the sample values selected from one population are not related to or somehow paired or matched with the sample values from the other population.
- ❖ **dependent** if the sample values are *paired*. (That is, each pair of sample values consists of two measurements from the same subject (such as before/after data), or each pair of sample values consists of matched pairs (such as husband/wife data), where the matching is based on some inherent relationship.)

$\mu 1$ = population mean

$\sigma 1$ = population standard deviation

$n1$ = size of the first sample

$\overline{X}_1$ = sample mean

$s1$ = sample standard deviation

Corresponding notations for μ_2, σ_2, s_2, and n_2 apply to population 2.

Requirements:

- ❖ σ_1 and σ_2 are unknown and no assumption is made about the equality of σ_1 and σ_2.

- ❖ The two samples are independent.
- ❖ Both samples are simple random samples.
- ❖ Either or both of these conditions are satisfied: The two sample sizes are both large (with $n_1 > 30$ and $n_2 > 30$) or both samples come from populations having normal distributions.

Hypothesis Test for Two Means: Independent Samples:

$$t = \frac{\left(\overline{x}_1 - \overline{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

- ❖ **Degrees of freedom:** In this book we use this simple and conservative estimate:
  df = smaller of $n_1 - 1$ and $n_2 - 1$.
- ❖ **P-values:** Refer to Table A-3. Use the procedure summarized in Figure 8-5.
- ❖ **Critical values:** Refer to Table A-3.

Confidence Interval Estimate of $\mu_1 - \mu_2$: Independent Samples

$$(\overline{x}_1 - \overline{x}_2) - E \; < \; (\mu_1 - \mu_2) \; < \; (\overline{x}_1 - \overline{x}_2) + E$$

$$\text{where} \quad E = t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Where df = smaller of $n_1 - 1$ and $n_2 - 1$.

Hypothesis Test for Two Means: Independent Samples with σ_1 and σ_2 both known:

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

P-values and critical values: Refer to Table A-2.

Confidence Interval: Independent Samples with σ_1 and σ_2 both known:

$$(\bar{x}_1 - \bar{x}_2) - E \; < \; (\mu_1 - \mu_2) \; < \; (\bar{x}_1 - \bar{x}_2) + E$$

$$\text{where} \quad E = z_{\alpha/2} \quad \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Hypothesis Test Statistic for Two Means: Independent Samples and $\sigma_1 = \sigma_2$:

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

Where

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

and the number of degrees of freedom is df = $n_1 + n_2$ - 2

Confidence Interval Estimate of $\mu_1 - \mu_2$: Independent Samples with $\sigma_1 = \sigma_2$

$$(\bar{x}_1 - \bar{x}_2) - E \; < \; (\mu_1 - \mu_2) \; < \; (\bar{x}_1 - \bar{x}_2) + E$$

$$\text{where} \quad E = t_{\alpha/2} \quad \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

and number of degrees of freedom **is df = $n_1 + n_2 - 2$**

Notation for Dependent Samples:

d = individual difference between the two values of a single matched pair.

μ_d = mean value of the differences d for the population of paired data.

$\bar{d}$ = mean value of the differences d for the paired sample data (equal to the mean of the $x - y$ values).

S_d = standard deviation of the differences d for the paired sample data.

n = number of pairs of data.

Hypothesis Test Statistic for Matched Pairs:

$$t = \frac{\bar{d} - \mu_d}{\dfrac{s_d}{\sqrt{n}}}$$

where degrees of freedom = $n - 1$

Confidence Intervals for Matched Pairs:

$$\bar{d} - E < \mu_d < \bar{d} + E$$

$$\text{where } E = t_{\alpha/2} \; \frac{s_d}{\sqrt{n}}$$

Critical values of $t_{a/2}$: Use Table A-3 with n – 1 degrees of freedom.

F test for Comparing Variances:

Notation for Hypothesis Tests with Two Variances or Standard Deviations:

S^2_1 = *larger* of two sample variances

n_1 = size of the sample with the *larger* variance

σ^2_1 = variance of the population from which the sample with the *larger* variance is drawn

S^2_2, n_2, and σ^2_2 are used for the other sample and population

Requirements:
- ❖ The two populations are independent.
- ❖ The two samples are simple random samples.
- ❖ The two populations are each normally distributed.

Test Statistic for Hypothesis Tests with Two Variances:

$$F = \frac{S_1^{\,2}}{S_2^{\,2}}$$

Where S^2_1 is the larger of the two sample variances

Critical Values: Using Table A-5, we obtain critical *F* values that are determined by the following three values:
- ❖ The significance level α
- ❖ Numerator degrees of freedom = $n_1 - 1$
- ❖ Denominator degrees of freedom = $n_2 - 1$

Properties of the *F* Distribution:
- ❖ If the two populations do have equal variances, then F = $\frac{S^2_1}{S^2_2}$ will be close to 1 because S^2_1 and S^2_2 are close in value.
- ❖ If the two populations have radically different variances, then *F* will be a large number.

Remember, the larger sample variance will be S^2_1.

Conclusions from the *F* Distribution:
a value of *F* near 1 will be evidence in favor of the conclusion that $\sigma^2_1 = \sigma^2_2$
But a large value of *F* will be evidence against the conclusion of equality of the population variances.

CHAPTER 10

Linear correlation coefficient *r* is a numerical measure of the strength of the relationship between two variables representing quantitative data.

❖ measures the strength of the linear relationship between the paired quantitative *x*- and *y*-values in a sample.

Correlation exists between two variables when the values of one are somehow associated with the values of the other in some way.

Requirements:
❖ The sample of paired (*x, y*) data is a simple random sample of quantitative data.
❖ Visual examination of the scatterplot must confirm that the points approximate a straight-line pattern.
❖ The outliers must be removed if they are known to be errors. The effects of any other outliers should be considered by calculating *r* with and without the outliers included.

Notation for the Linear Correlation Coefficient:

n = number of pairs of sample data

$\sum$ = denotes the addition of the items indicated.

$\sum x$ = denotes the sum of all *x*-values.

$\sum x^2$ = indicates that each *x*-value should be squared and then those squares added.

$(\sum x)^2$ = indicates that the *x*-values should be added and then the total squared.

$\sum xy$ = indicates that each *x*-value should be first multiplied by its corresponding *y*-value. After obtaining all such products, find their sum.

r = linear correlation coefficient for sample data.

ρ = linear correlation coefficient for population data.

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

Properties of the Linear Correlation Coefficient *r*:
❖ $-1 \leq r \leq 1$
❖ if all values of either variable are converted to a different scale, the value of *r* does not change.
❖ The value of *r* is not affected by the choice of *x* and *y*. Interchange all *x*- and *y*-values and the value of *r* will not change.
❖ *r* measures strength of a linear relationship.
❖ *r* is very sensitive to outliers, they can dramatically affect its value.

Interpreting *r*: Explained Variation

The value of r^2 is the proportion of the variation in *y* that is explained by the linear relationship between *x* and *y*.

Common Errors Involving Correlation:

❖ **Causation:** It is wrong to conclude that correlation implies causality.
❖ **Averages:** Averages suppress individual variation and may inflate the correlation coefficient.
❖ **Linearity:** There may be some relationship between x and y even when there is no linear correlation.

Hypothesis Test for Correlation *P*-Value from a *t* Test:

❖ **H$_0$: p = 0**   **(There is no linear correlation.)**

❖ **H$_1$: p $\neq$ 0**   **(There is a linear correlation.)**

Test Statistic: *t*

$$t = \dfrac{r}{\sqrt{\dfrac{1 - r^2}{n - 2}}}$$

The regression equation expresses a relationship between

❖ x (called the explanatory variable, predictor variable or independent variable)
❖ y (called the response variable or dependent variable).

Regression Equation Given a collection of paired data.

$$\hat{y} = b_0 + b_1 x$$

algebraically describes the relationship between the two variables.

Regression Line The graph of the regression equation is called the regression line (or line of best fit, or least squares line).

marginal change in a variable is the amount that it changes when the other variable changes by exactly one unit.

residual is the difference between the *observed* sample value of y and the y-value that is *predicted* by using the regression equation.

❖ residual = observed y – predicted y = $y - y$

Coefficient of determination is the amount of the variation in y that is explained by the regression line.

Notation for Regression Equation (Sample Statistic): b$_0$ and b$_1$

$$b_1 = r \dfrac{s_y}{s_x} \qquad \text{(slope)}$$

$$b_0 = \bar{y} - b_1 \bar{x} \qquad (y\text{-intercept})$$

CHAPTER 11

Goodness-of-fit test is used to test the hypothesis that an observed frequency distribution fits or conforms to some claimed distribution.

❖ Goodness-of-fit hypothesis tests are always *right-tailed.*

Notation:
- ❖ **O**   represents the observed frequency of an outcome.
- ❖ **E**   represents the expected frequency of an outcome.
- ❖ **k**   represents the number of different categories or outcomes.
- ❖ **n**   represents the total number of trials.

Goodness-of-fit (Test Statistic):

$$\chi^2 = \sum \frac{(O-E)^2}{E} \qquad E = \frac{(\text{row total})\,(\text{column total})}{(\text{grand total})}$$

Expected Frequencies:
- ❖ If all expected frequencies are equal:

$$E = \frac{n}{k}$$

the sum of all observed frequencies divided by the number of categories

- ❖ If expected frequencies are not all equal:

$E = np$

Each expected frequency is found by multiplying the sum of all observed frequencies by the probability for the category.

Close agreement between observed and expected values will lead to a small value of X^2 and a large *P*-value.

Large disagreement between observed and expected values will lead to a large value of X^2 and a small *P*-value.

Significantly large value of X^2 will cause a rejection of the null hypothesis of no difference between the observed and the expected.

contingency table (or two-way frequency table) is a table in which frequencies correspond to two variables. (One variable is used to categorize rows, and a second variable is used to categorize columns.)

- ❖ Contingency tables have at least two rows and at least two columns.

Test of independence tests the null hypothesis that in a contingency table, the row and column variables are independent.

Null and Alternative Hypotheses
- ❖ **H_0:** The row and column variables are *independent*.

- ❖ **H_1:** The row and column variables are dependent.

Critical Values:

❖ Found in Table A-4 using **(degrees of freedom = (r – 1) (c – 1))**

r is the number of rows and c is the number of columns

❖ Tests of Independence are always *right-tailed*.

Dependence means only there is a relationship between the two variables.

Test of homogeneity we test the claim that *different populations* have the same proportions of some characteristics.

CHAPTER 12

Analysis of variance (ANOVA) is a method for testing the hypothesis that three or more population means are equal.

❖ For example: **H₀**: $\mu_1 = \mu_2 = \mu_3 = ... \mu_k$
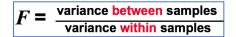❖ **H**: At least one mean is different

ANOVA Methods Require the *F*-Distribution:
1. The *F*- distribution is not symmetric; it is skewed to the right.
2. The values of *F* can be 0 or positive; they cannot be negative.
3. There is a different *F*-distribution for each pair of degrees of freedom for the numerator and denominator.

One-way analysis of variance (ANOVA) is a method of testing the equality of three or more population means by analyzing sample variances.
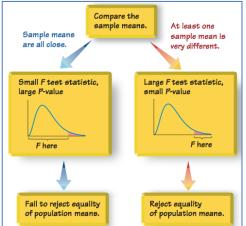
❖ One-way analysis of variance is used with data categorized with *one* treatment (or factor), which is a characteristic that allows us to distinguish the different populations from one another.
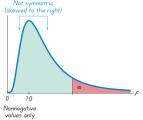
Test Statistic for One-Way ANOVA:

$$F = \frac{\text{variance between samples}}{\text{variance within samples}}$$

An excessively large *F* test statistic is evidence against equal population means.

Relationship Between *F* Test Statistic / *P*-Value:

Estimate the common value of σ^2:

1. The variance <u>between</u> samples (also called variation due to treatment) is an estimate of the common population variance σ^2 that is based on the variability among the sample <u>means</u>.

2. The variance <u>within</u> samples (also called variation due to error) is an estimate of the common population variance σ^2 based on the sample <u>variances</u>.

Calculations with <u>Equal</u> Sample Sizes:

Variance <u>between</u> samples = n $s_{\bar{x}}^2$

❖ Where $s_{\bar{x}}^2$ = variance of sample means

Variance <u>within</u> samples = s_p^2

❖ Where s_p^2 = pooled variance (or the mean of the sample variances)

Critical Value of F:

<u>Right-tailed test</u>

Degree of freedom with *k* samples of the same size *n*

❖ numerator *df* = *k* − 1
❖ denominator *df* = *k*(*n* − 1)

Calculations with <u>Unequal</u> Sample Sizes:

$$F = \frac{\text{variance within samples}}{\text{variance between samples}} = \frac{\left[\dfrac{\sum n_i(\bar{x}_i - \bar{\bar{x}})^2}{k-1}\right]}{\left[\dfrac{\sum(n_i - 1)s_i^2}{\sum(n_i - 1)}\right]}$$

$\bar{\bar{x}}$ = mean of all sample scores combined

k = number of population means being compared

n_i = number of values in the ith sample

$\bar{x}_1$ = mean of values in the ith sample

s_i^2 = variance of values in the ith sample

SS(total) or total sum of squares, is a measure of the total variation (around $\bar{\bar{x}}$) in all the sample data combined.

$$SS(\text{total}) = \sum \left(x - \bar{\bar{x}}\right)^2$$

❖ **SS(total) = SS(treatment) + SS(error)**

SS(treatment) also referred to as SS(factor) or SS(between groups) or SS(between samples), is a measure of the variation between the sample means.

$$SS(\text{treatment})$$
$$= n_1\left(\bar{x}_1 - \bar{\bar{x}}\right)^2 + n_2\left(\bar{x}_2 - \bar{\bar{x}}\right)^2 + \boxed{?} + n_k\left(\bar{x}_k - \bar{\bar{x}}\right)^2$$
$$= \sum n_i\left(x_i - \bar{\bar{x}}\right)^2$$

علامة الإستفهام بالقانونين عباره عن "نقاط" (.....)

SS(error) also referred to as SS(within groups) or SS(within samples), is a sum of squares
  representing the variability that is assumed to be common to all the populations being
  considered.

$$SS(\text{error})$$
$$= (n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \boxed{?} + (n_k - 1)s_k^2$$
$$= \sum (n_i - 1)s_i^2$$

Mean Squares (MS):

MS(treatment) is a mean square for treatment.
obtained as follows:

Formula 12-5
$$MS(\text{treatment}) = \frac{SS\ (\text{treatment})}{k - 1}$$

N = total number of values in all samples combined

MS(error) is a mean square for error.
obtained as follows:

Formula 12-6
$$MS(\text{error}) = \frac{SS\ (\text{error})}{N - k}$$

MS(total) is a mean square for the total variation.
obtained as follows:

Formula 12-7
$$MS(\text{total}) = \frac{SS(\text{total})}{N - 1}$$

Test Statistic for ANOVA with Unequal Sample Sizes:

Formula 12-8
$$F = \frac{MS\ (\text{treatment})}{MS\ (\text{error})}$$

❖ Numerator $df = k - 1$

❖ Denominator $df = N - k$

An Approach to Understanding ANOVA:

Understand that

❖ **a small *P*-value** (such as 0.05 or less) leads to rejection of the null hypothesis of equal
  means.
❖ **a large *P*-value** (such as greater than 0.05), fail to reject the null hypothesis of equal means.

Testing H$_0$: $\mu_1 = \mu_2 = \mu_3 = $ …….:

❖ If the ***P*-value ≤ α**, reject the null hypothesis of equal means and conclude that at least one
  of the population means is different from the others.

❖ If the ***P*-value > α**, fail to reject the null hypothesis of equal means.