STATISTICS IS THE GRAMMAR OF SCIENCE

PROBABILITY AND STATISTICS

LECTURE # 30

HYPOTHESIS TESTING

INFERENCE ABOUT DIFFERENCE OF MEANS

PREPARED BY

HAZBER SAMSON

SCIENCES & HUMANITIES DEPARTMENT

FAST ISLAMABAD CAMPUS

INFERENCE ABOUT TWO POPULATIONS

INFERENCE ABOUT DIFFERENCE OF POPULATION MEANS WHEN SIGMAS ARE KNOWN

CASE-1 $If \quad H_0 : \mu_1 - \mu_2 = \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 \neq \mu_0$

- When σ_1 *and* σ_2 are known then we use **test statistic**

$$Z = \frac{\left(\overline{x_1} - \overline{x_2}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- When σ_1 *and* σ_2 are known then we use **P-Value**

$$P = 2[1 - \Phi(|z_0|)]$$

- The $(1-\alpha)$ 100% **confidence interval** for $\mu_1 - \mu_2$ is

$$\left(\overline{x_1} - \overline{x_2}\right) - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}} \leq \mu_1 - \mu_2 \leq \left(\overline{x_1} - \overline{x_2}\right) - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}$$

CASE-2 $If \quad H_0 : \mu_1 - \mu_2 \geq \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 < \mu_0$

- When σ_1 *and* σ_2 are known then we use **test statistic**

$$Z = \frac{\left(\overline{x_1} - \overline{x_2}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- When σ_1 *and* σ_2 are known then we use **P-Value**

$$P = \Phi(z_0)$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\mu_1 - \mu_2 \leq \left(\overline{x_1} - \overline{x_2}\right) + Z_{\alpha} \cdot \sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}$$

CASE-3 $If \quad H_0 : \mu_1 - \mu_2 \leq \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 > \mu_0$

- When σ_1 *and* σ_2 are known then we use **test statistic**

$$Z = \frac{\left(\overline{x_1} - \overline{x_2}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- When σ_1 *and* σ_2 are known then we use **P-Value**

$$P = 1 - \Phi(z_0)$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\mu_1 - \mu_2 \geq \left(\overline{x_1} - \overline{x_2}\right) - Z_{\alpha} \cdot \sqrt{\frac{\sigma_1^{\,2}}{n_1} + \frac{\sigma_2^{\,2}}{n_2}}$$

EXAMPLES OF INFERENCE ABOUT DIFFERENCE OF MEANS

EXAMPLE-1 A survey found that the average hotel room rate in New Orleans is $88.42 and the average room rate in Phoenix is $80.61. Assume that the data were obtained from two samples of 50 hotels each and that the standard deviations of the populations are $5.62 and $4.83, respectively. At a 0.05, can it be concluded that there is a significant difference in the rates?

SOLUTION *Here* $n_1 = 50,\ n_2 = 50,\ \overline{x_1} = 88.42,\ \overline{x_2} = 80.61,\ \sigma_1 = 5.62, \sigma_2 = 4.83,\ \alpha = 0.05$

CRITICAL VALUE APPROACH

Step – 1: Formulation of Hypotheses

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Step – 2: Level of Significance

$$\alpha = 0.05$$

Step – 3: Test Statistic

$$Z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(88.42 - 80.61) - (0)}{\sqrt{\frac{5.62^2}{50} + \frac{4.83^2}{50}}} = 7.45$$

Step – 4: Critical Region

$$|z| \geq z_{\frac{\alpha}{2}} \quad \Rightarrow |z| \geq 1.96$$

Step – 5: Conclusion

Since calculated value of z lies in CR so reject $H_{0.}$

P-VALUE APPROACH

Step – 1: Formulation of Hypotheses

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu \neq \mu_2$$

Step – 2: Level of Significance

$$\alpha = 0.05$$

Step – 3: Test Statistic

$$Z = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(88.42 - 80.61) - (0)}{\sqrt{\frac{5.62^2}{50} + \frac{4.83^2}{50}}} = 7.45$$

Step – 4: P – value

$$p = 2[1 - \Phi(|z_0|)] = 2[1 - \Phi(|7.45|)] = 0.0002$$

Step – 5: Conclusion

Since $p - value < \alpha$ *so reject* $H_{0.}$

CONFIDENCE INTERVAL APPROACH

Step – 1: Formulation of Confidence Interval

The $(1-\alpha)100\%$ *CI for* $\mu_1 - \mu_2$ *is given by*

$$(\overline{x_1} - \overline{x_2}) - Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}} \leq \mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) + Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\sigma_1^{\ 2}}{n_1} + \frac{\sigma_2^{\ 2}}{n_2}}$$

$$(88.42 - 80.61) - (1.96)\sqrt{\frac{5.62^2}{50} + \frac{4.83^2}{50}} \leq \mu_1 - \mu_2 \leq (88.42 - 80.61) + (1.96)\sqrt{\frac{5.62^2}{50} + \frac{4.83^2}{50}}$$

$$5.76 \leq \mu \leq 9.86$$

Step – 2: Conclusion

Since $\mu_1 - \mu_2 = 0$ *does not lie in CI so reject* $H_{0.}$

EXAMPLE-2 A local college cafeteria has a self-service soft ice cream machine. The cafeteria provides bowls that can hold up to 16 ounces of ice cream. The food service manager is interested in comparing the average amount of ice cream dispensed by male students to the average amount dispensed by female students. A measurement device was placed on the ice cream machine to determine the amounts dispensed. Random samples of 85 male and 78 female students who got ice cream were selected. The sample averages were 7.23 and 6.49 ounces for the male and female students, respectively. Assume that the population standard deviations are 1.22 and 1.17 ounces, respectively. Using a 1% significance level, can you conclude that the average amount of ice cream dispensed by all male college students is larger than the average amount dispensed by all female college students?

SOLUTION *Here* $n_1 = 85,\ n_2 = 78,\ \overline{x_1} = 7.23,\ \overline{x_2} = 6.49,\ \sigma_1 = 1.22,\ \sigma_2 = 1.17,\ \alpha = 0.01$

CRITICAL VALUE APPROACH

Step − 1: Formulation of Hypotheses

$$H_0 : \mu_1 \le \mu_2$$
$$H_1 : \mu_1 > \mu_2$$

Step − 2: Level of Significance

$$\alpha = 0.01$$

Step − 3: Test Statistic

$$Z = Z = \frac{\left(\overline{x_1} - \overline{x_2}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(7.23 - 6.49) - (0)}{\sqrt{\frac{1.22^2}{85} + \frac{1.17^2}{78}}} = \frac{0.74}{0.1872} = 3.95$$

Step − 4: Critical Region

$$z > z_\alpha \quad \Rightarrow\ z > z_{0.01} \quad \Rightarrow z > 2.33$$

Step − 5: Conclusion

Since calculated value of z lies in CR so reject $H_{0.}$

P-VALUE APPROACH

Step − 1: Formulation of Hypotheses

$$H_0 : \mu_1 \le \mu_2$$
$$H_1 : \mu_1 > \mu_2$$

Step − 2: Level of Significance

$$\alpha = 0.01$$

Step − 3: Test Statistic

$$Z = Z = \frac{\left(\overline{x_1} - \overline{x_2}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(7.23 - 6.49) - (0)}{\sqrt{\frac{1.22^2}{85} + \frac{1.17^2}{78}}} = \frac{0.74}{0.1872} = 3.95$$

Step − 4: P − value

$$p = 1 - \Phi(z_0) = 1 - \Phi(3.95) = 1 - 0.9999 = 0.0001$$

Step − 5: Conclusion

Since $p - value < \alpha$ *so reject* $H_{0.}$

CONFIDENCE INTERVAL APPROACH

Step − 1: Formulation of Confidence Interval

The $(1-\alpha)100\%$ *CI for* $\mu_1 - \mu_2$ *is given by*

$$\mu_1 - \mu_2 \ge \left(\overline{x_1} - \overline{x_2}\right) - Z_\alpha . \sqrt{\frac{\sigma_1{}^2}{n_1} + \frac{\sigma_2{}^2}{n_2}}$$

$$\mu_1 - \mu_2 \ge (7.23 - 6.49) - (2.33)\sqrt{\frac{1.22^2}{85} + \frac{1.17^2}{78}}$$

$$\mu_1 - \mu_2 \ge 0.3038$$

Step − 2: Conclusion

Since $\mu_1 - \mu_2 = 0$ *does not lie in CI so reject* $H_{0.}$

EXAMPLE-3 **REM (rapid eye movement)** sleep is sleep is during which most dreams occur. Each night a person has both REM and non-REM sleep. However it is thought that adults have less REM sleep than children. Assume that REM sleep time is normally distributed for both adults and children. A random sample of 10 adults showed that they have average REM sleep time of 2.1 hours per night with population standard deviation 0.7 hour. Another random sample of 10 children showed that they have average REM sleep time of 2.8 hours per night with population standard deviation 0.5 hour. Do the data indicate, on average, adults tend to have less REM sleep than children? Use 1% level of significance.

SOLUTION *Here* $n_1=10,\ n_2=10,\ \overline{x}_1=2.1,\ \overline{x}_2=2.8,\ \sigma_1=0.7, \sigma_2=0.5,\ \alpha=0.01$

CRITICAL VALUE APPROACH

Step – 1: Formulation of Hypotheses

$H_0:\mu_1 \ge \mu_2$

$H_1:\mu_1 < \mu_2$

Step – 2: Level of Significance

$\alpha = 0.01$

Step – 3: Test Statistic

$$Z=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(2.1-2.8)-(0)}{\sqrt{\frac{0.7^2}{10}+\frac{0.5^2}{10}}}=\frac{-0.7}{0.2720}=-2.57$$

Step – 4: Critical Region

$z<-z_\alpha \quad \Rightarrow \quad z<-2.33$

Step – 5: Conclusion

Since calculated value of z lies in CR so reject H_0.

P-VALUE APPROACH

Step – 1: Formulation of Hypotheses

$H_0:\mu_1 \ge \mu_2$

$H_1:\mu_1 < \mu_2$

Step – 2: Level of Significance

$\alpha = 0.01$

Step – 3: Test Statistic

$$Z=\frac{(\overline{x}_1-\overline{x}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}=\frac{(2.1-2.8)-(0)}{\sqrt{\frac{0.7^2}{10}+\frac{0.5^2}{10}}}=\frac{-0.7}{0.2720}=-2.57$$

Step – 4: P – value

$p=\Phi(z_0)=\Phi(-2.57)=0.0051$

Step – 5: Conclusion

Since $p-value<\alpha$ so reject H_0.

CONFIDENCE INTERVAL APPROACH

Step – 1: Formulation of Confidence Interval

The $(1-\alpha)100\%$ CI for $\mu_1-\mu_2$ is given by

$$\mu_1-\mu_2 \le (\overline{x}_1-\overline{x}_2) + Z_\alpha.\sqrt{\frac{\sigma_1^{\ 2}}{n_1}+\frac{\sigma_2^{\ 2}}{n_2}}$$

$$\mu_1-\mu_2 \le (2.1-2.8)+(2.33)\sqrt{\frac{0.7^2}{10}+\frac{0.5^2}{10}}$$

$$\mu_1-\mu_2 \le -1.33$$

Step – 2: Conclusion

Since $\mu_1-\mu_2=0$ does not lie in CI so reject H_0.

INFERENCE ABOUT DIFFERENCE OF POPULATION MEANS WHEN SIGMAS ARE UNKNOWN (UNEQUAL VARIANCES)

CASE-1 $\quad If \quad H_0 : \mu_1 - \mu_2 = \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 \neq \mu_0$

- When σ_1 *and* σ_2 are unknown then we use **test statistic**

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1) + \left(s_2^2/n_2\right)^2/(n_2-1)}$$

- When σ_1 *and* σ_2 are unknown then we use **P-Value**

$$P = 2[\Phi(|t_0|)]$$

- The $(1-\alpha)$ 100% **confidence interval** for $\mu_1 - \mu_2$ is

$$(\overline{x_1} - \overline{x_2}) - t_{\frac{\alpha}{2}(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) + t_{\frac{\alpha}{2}(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

CASE-2 $\quad If \quad H_0 : \mu_1 - \mu_2 \geq \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 < \mu_0$

- When σ_1 *and* σ_2 are unknown then we use **test statistic**

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1) + \left(s_2^2/n_2\right)^2/(n_2-1)}$$

- When σ_1 *and* σ_2 are unknown then we use **P-Value**

$$P = 1 - \Phi(|t_0|)$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) + t_{\alpha(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

CASE-3 $\quad If \quad H_0 : \mu_1 - \mu_2 \leq \mu_0 \quad vs \quad H_1 : \mu_1 - \mu_2 > \mu_0$

- When σ_1 *and* σ_2 are unknown then we use **test statistic**

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1) + \left(s_2^2/n_2\right)^2/(n_2-1)}$$

- When σ_1 *and* σ_2 are unknown then we use **P-Value**

$$P = \Phi(|t_0|)$$

- The $(1-\alpha)$ 100% **confidence interval** for μ is

$$\mu_1 - \mu_2 \geq (\overline{x_1} - \overline{x_2}) - t_{\alpha(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

EXAMPLE-1 The average size of a farm in Indiana County, Pennsylvania, is 191 acres. The average size of a farm in Greene County, Pennsylvania, is 199 acres. Assume the data were obtained from two samples with standard deviations of 38 and 12 acres, respectively, and sample sizes of 8 and 10, respectively. Can it be concluded at a 0.05 that the average size of the farms in the two counties is different? Assume the populations are normally distributed.

SOLUTION *Here* $n_1 = 8,\ n_2 = 10,\ \overline{x_1} = 191,\ \overline{x_2} = 199,\ s_1 = 38,\ s_2 = 12,\ \alpha = 0.05$

CRITICAL VALUE APPROACH

Step − 1: Formulation of Hypotheses

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Step − 2: Level of Significance

$\alpha = 0.05$

Step − 3: Test Statistic

$$t = \frac{\left(\overline{x_1} - \overline{x_2}\right) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{with} \quad d.f = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1 - 1) + \left(s_2^2/n_2\right)^2/(n_2 - 1)}$$

$$t = \frac{(191 - 199) - (0)}{\sqrt{\frac{38^2}{8} + \frac{12^2}{10}}} = -0.57 \quad \text{with} \quad d.f = \frac{\left(38^2/8 + 12^2/10\right)^2}{\left(38^2/8\right)^2/(8-1) + \left(12^2/10\right)^2/(10-1)} = \frac{37986}{4677.36} = 8.12 \approx 8$$

Step − 4: Critical Region

$$|t| \geq t_{\frac{\alpha}{2}(d.f)} \quad \Rightarrow |t| \geq t_{(0.025)(8)} \quad \Rightarrow |t| \geq 2.306$$

Step − 5: Conclusion

Since calculated value of t does not lie in CR so do not reject $H_{0.}$

P-VALUE APPROACH

Step − 1: Formulation of Hypotheses

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Step − 2: Level of Significance

$\alpha = 0.05$

Step − 3: Test Statistic

$$t = \frac{\left(\overline{x_1} - \overline{x_2}\right) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{with} \quad d.f = \frac{\left(s_1^2/n_1 + s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1 - 1) + \left(s_2^2/n_2\right)^2/(n_2 - 1)}$$

$$t = \frac{(191 - 199) - (0)}{\sqrt{\frac{38^2}{8} + \frac{12^2}{10}}} = -0.57 \quad \text{with} \quad d.f = \frac{\left(38^2/8 + 12^2/10\right)^2}{\left(38^2/8\right)^2/(8-1) + \left(12^2/10\right)^2/(10-1)} = \frac{37986}{4677.36} = 8.12 \approx 8$$

Step − 4: P − value

$$p = 2[\Phi(t_0)] = 2[0.283] = 0.566$$

Step − 5: Conclusion

Since p − value $> \alpha$ *so do not reject* $H_{0.}$

CONFIDENCE INTERVAL APPROACH

Step – 1: Formulation of Confidence Interval

The $(1-\alpha)100\%$ *CI for* $\mu_1-\mu_2$ *is given by*

$$\left(\overline{x_1}-\overline{x_2}\right)-t_{\frac{\alpha}{2}(d.f)}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}\le\mu_1-\mu_2\le\left(\overline{x_1}-\overline{x_2}\right)+t_{\frac{\alpha}{2}(d.f)}\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}$$

$$(191-199)-(2.306)\sqrt{\frac{38^2}{8}+\frac{12^2}{10}}\le\mu_1-\mu_2\le(191-199)-(2.306)\sqrt{\frac{38^2}{8}+\frac{12^2}{10}}$$

$$-40.19\le\mu\le 24.19$$

Step – 2: Conclusion

Since $\mu_1-\mu_2=0$ *lies in CI so do not reject* $H_{0.}$

EXAMPLE-2 A channel claims that the mean amount of money spent by a family at movie theaters is greater than the mean amount spent by a family at amusement parks. The results for samples of expenses for the two modes of entertainment are shown at the left. At a = 0.01, can you support the channel's claim? Assume the population variances are equal.

Sample Statistics for Amount Spent by Customers

Movies	Amusement Parks
$\bar{x}_1 = \$86$ $s_1 = \$9$ $n_1 = 26$	$\bar{x}_1 = \$82$ $s_1 = \$7$ $n_1 = 36$

SOLUTION *Here* $n_1=26,\ n_2=36,\ \overline{x_1}=86,\ \overline{x_2}=82,\ s_1=9,\ s_2=7,\ \alpha=0.01$

CRITICAL VALUE APPROACH

Step – 1: Formulation of Hypotheses

$H_0:\mu_1\le\mu_2$

$H_1:\mu_1>\mu_2$

Step – 2: Level of Significance

$\alpha=0.01$

Step – 3: Test Statistic

$$t=\frac{\left(\overline{x_1}-\overline{x_2}\right)-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}\quad\text{with}\quad d.f=\frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1)+\left(s_2^2/n_2\right)^2/(n_2-1)}$$

$$t=\frac{(86-82)-(0)}{\sqrt{\frac{9^2}{26}+\frac{7^2}{36}}}=1.89\quad\text{with}\quad d.f=\frac{\left(9^2/26+7^2/36\right)^2}{\left(9^2/26\right)^2/(26-1)+\left(7^2/36\right)^2/(36-1)}=\frac{20.04}{0.44}=45.5\approx 46$$

Step – 4: Critical Region

$$t>t_{\alpha(d.f)}\quad\Rightarrow t>t_{(0.01)(46)}\quad\Rightarrow t>2.4235$$

Step – 5: Conclusion

Since calculated value of t does not lie in CR so do not reject $H_{0.}$

P-VALUE APPROACH

$Step-1: Formulation\ of\ Hypotheses$

$H_0: \mu_1 \le \mu_2$

$H_1: \mu_1 > \mu_2$

$Step-2: Level\ of\ Significance$

$\alpha = 0.01$

$Step-3: Test\ Statistic$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)}$$

$$t = \frac{(86-82)-(0)}{\sqrt{\frac{9^2}{26} + \frac{7^2}{36}}} = 1.89 \quad with \quad d.f = \frac{(9^2/26 + 7^2/36)^2}{(9^2/26)^2/(26-1) + (7^2/36)^2/(36-1)} = \frac{20.04}{0.44} = 45.54 \approx 46$$

$Step-4: P-value$

$p = \Phi(|t_0|)] = \Phi[1.89] = 0.032$

$Step-5: Conclusion$

$Since\ p-value > \alpha\ so\ do\ not\ reject\ H_{0.}$

CONFIDENCE INTERVAL APPROACH

$Step-1: Formulation\ of\ Confidence\ Interval$

$The\ (1-\alpha)100\%\ CI\ for\ \mu_1 - \mu_2\ is\ given\ by$

$$\mu_1 - \mu_2 \ge (\bar{x}_1 - \bar{x}_2) - t_{\alpha(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\mu_1 - \mu_2 \ge (86-82) - (2.4235)\sqrt{\frac{9^2}{26} + \frac{7^2}{36}}$$

$$\mu_1 - \mu_2 \ge -1.12$$

$Step-2: Conclusion$

$Since\ \mu_1 - \mu_2 = 0\ lies\ in\ CI\ so\ do\ not\ reject\ H_{0.}$

EXAMPLE-3 The number of grams of carbohydrates contained in 1-ounce servings of randomly selected chocolate and non chocolate candy is listed here. Is there sufficient evidence to conclude that number of grams of carbohydrates contained in chocolate candy is less than number of grams of carbohydrates contained in non chocolate candy? $Use\ \alpha = 0.10$

Chocolate: 29 25 17 36 41 25 32 29 38 34 24 27 29

Non Chocolate: 41 41 37 29 30 38 39 10 29 55 29

SOLUTION $Here\ n_1 = 13,\ n_2 = 11,\ \bar{x}_1 = 29.69,\ \bar{x}_2 = 34.36,\ s_1 = 6.5,\ s_2 = 11.2,\ \alpha = 0.10$

$$Here\ n_1 = 13,\ n_2 = 11,\ \bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{386}{13} = 29.69,\ \bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{378}{11} = 34.36$$

$$s_1 = \frac{\sum (x - \bar{x}_1)^2}{n_1 - 1} = 6.5,\ s_2 = \frac{\sum (x - \bar{x}_2)^2}{n_2 - 1} = 11.2$$

CRITICAL VALUE APPROACH

Step – 1: Formulation of Hypotheses

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

Step – 2: Level of Significance

$$\alpha = 0.10$$

Step – 3: Test Statistic

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

$$t = \frac{(29.69 - 34.36) - (0)}{\sqrt{\frac{6.5^2}{13} + \frac{11.2^2}{11}}} = -1.22 \quad with \quad d.f = \frac{(6.5^2/13 + 11.2^2/11)^2}{(6.5^2/13)^2/(13-1) + (11.2^2/11)^2/(11-1)} = 15.46 \approx 16$$

Step – 4: Critical Region

$$t < -t_{\alpha(d.f)} \quad \Rightarrow t < -t_{(0.10)(16)} \quad \Rightarrow t < -1.337$$

Step – 5: Conclusion

Since calculated value of t does not lie in CR so do not reject $H_{0.}$

P-VALUE APPROACH

Step – 1: Formulation of Hypotheses

$$H_0 : \mu_1 \geq \mu_2$$

$$H_1 : \mu_1 < \mu_2$$

Step – 2: Level of Significance

$$\alpha = 0.10$$

Step – 3: Test Statistic

$$t = \frac{(\overline{x_1} - \overline{x_2}) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad with \quad d.f = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

$$t = \frac{(29.69 - 34.36) - (0)}{\sqrt{\frac{6.5^2}{13} + \frac{11.2^2}{11}}} = -1.22 \quad with \quad d.f = \frac{(6.5^2/13 + 11.2^2/11)^2}{(6.5^2/13)^2/(13-1) + (11.2^2/11)^2/(11-1)} = 15.46 \approx 16$$

Step – 4: P – value

$$p = 1 - \Phi(|t_0|) = 1 - \Phi(1.22) = 1 - 0.124 = 0.876$$

Step – 5: Conclusion

Since p – value > α so do not reject $H_{0.}$

CONFIDENCE INTERVAL APPROACH

Step – 1: Formulation of Confidence Interval

The $(1-\alpha)100\%$ CI for $\mu_1 - \mu_2$ is given by

$$\mu_1 - \mu_2 \leq (\overline{x_1} - \overline{x_2}) + t_{\alpha(d.f)}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\mu_1 - \mu_2 \leq (29.69 - 34.36) + (1..337)\sqrt{\frac{6.5^2}{13} + \frac{11.2^2}{11}}$$

$$\mu_1 - \mu_2 \leq 0.4480$$

Step – 2: Conclusion

Since $\mu_1 - \mu_2 = 0$ lies in CI so do not reject $H_{0.}$