Inferences Concerning Mean

Comparisons -Two independent Large samples

Let us suppose that we are dealing with two independent random samples of size n_1 and n_2 from two normal populations having means μ_1 and μ_2 and known variances σ_1^2 and σ_2^2 respectively and that we want to test the null hypothesis $\mu_1 - \mu_2 = \delta_0$ where δ_0 is a given constant, against one of the alternatives $\mu_1 - \mu_2 \neq \delta_0$, $\mu_1 - \mu_2 > \delta_0$, $\mu_1 - \mu_2 < \delta_0$. The test procedure is as follows:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = \delta_0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 < \delta_0$

or H_1: $\mu_1 - \mu_2 > \delta_0$

or H_1: $\mu_1 - \mu_2 \neq \delta_0$

Sample sizes: n_1 and n_2 (both are greater than or equal to 30)

Sample means : $\bar{x}$ *and* $\bar{y}$

Sample variances: s_1^2 and s_2^2

Level of significance, α (1% or 5%)

Test Statistic, $Z = \dfrac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

Critical regions for testing H_0: $\mu_1 - \mu_2 = \delta_0$

Alternate Hypothesis	Reject null hypothesis if
$\mu_1 - \mu_2 < \delta_0$	$Z < -z_\alpha$
$\mu_1 - \mu_2 > \delta_0$	$Z > z_\alpha$
$\mu_1 - \mu_2 \neq \delta_0$	$Z < -z_{\alpha/2}$ or $Z > z_{\alpha/2}$

Finally we have to write the decision whether to accept or reject the null hypothesis.

Problem 18: To test the claim that the resistance of electric wire can be reduced by more than 0.05 ohm by alloying, 32 values obtained for standard wire yielded $\overline{x_1} = 0.136$ ohm and $s_1 = 0.004$ ohm, and 32 values obtained for alloyed wire yielded $\overline{x_2} = 0.083$ ohm and $s_2 = 0.005$ ohm. At the 0.05 level of significance, does this support the claim?

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0.050$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 > 0.050$

Level of significance: $\alpha = 0.05$

$\overline{x_1} = 0.136$ $\quad$ $\overline{x_2} = 0.083$ $\quad$ $n_1 = 32$

$s_1 = 0.004$ $\quad$ $s_2 = 0.005$ $\quad$ $n_2 = 32$

Test Statistic, $Z = \dfrac{(\overline{X_1} - \overline{X_2}) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

$$= \frac{(0.136 - 0.083) - 0.050}{\sqrt{\dfrac{0.004^2}{32} + \dfrac{0.005^2}{32}}}$$

$= 2.65$

$Z_{0.05} = 1.645$

Decision: Since z = 2.65 exceeds $Z_{0.05}$ = 1.645, we have to reject the null hypothesis. So, accept the alternate hypothesis.

Problem 19: A company claims that its light bulbs are superior to those of its main competitor. If a study showed that a sample of $n_1 = 40$ of its bulbs has a mean lifetime of 1647 hours of continuous use with a standard deviation of 27 hours, while a sample of $n_2 = 40$ bulbs made by its main competitor has a mean lifetime of 1638 hours of continuous use with a standard deviation of 31 hours, does this substantiate the claim at the 0.05 level of significance?

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 > 0$

Level of significance: $\alpha = 0.05$

$\overline{x}_1 = 1647$ $\quad\quad$ $\overline{x}_2 = 1638$

$s_1 = 27$ $\quad\quad$ $s_2 = 31$

$n_1 = 40$ $\quad\quad$ $n_2 = 40$

Test Statistic, $Z = \dfrac{(\overline{X}_1 - \overline{X}_2) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

$$= \frac{(1647 - 1638)}{\sqrt{\dfrac{27^2}{40} + \dfrac{31^2}{40}}}$$

$$= 1.38$$

$Z_{0.05} = 1.645$

Decision: Since z = 1.38 does not exceeds $Z_{0.05}$ = 1.645, we have to accept the null hypothesis. So, the difference between the two sample means is not significant.

Problem 20: An investigation of two kinds of photocopying equipment showed that 75 failures of the first kind of equipment took on the average 83.2 minutes to repair with a standard deviation of 19.3 minutes, while 75 failure of the second kind of equipment took on the average 90.8 minutes to repair with a standard deviation of 21.4minutes. Test the null hypothesis $\mu_1 - \mu_2 = 0$ (namely, the hypothesis that on the average it takes an equal amount of time to repair either kind of equipment) against the alternative hypothesis $\mu_1 - \mu_2 \neq 0$ at the level of significance $\alpha = 0.05$.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 \neq 0$

Level of significance: $\alpha = 0.05$

$\overline{x_1} = 83.2$ $\overline{x_2} = 90.8$

$s_1 = 19.3$ $s_2 = 21.4$

$n_1 = 75$ $n_2 = 75$

Test Statistic, $Z = \dfrac{(\overline{X_1} - \overline{X_2}) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

$$= \frac{(83.2 - 90.8)}{\sqrt{\dfrac{19.3^2}{75} + \dfrac{21.4^2}{75}}}$$

$$= -2.28$$

$Z_{0.025} = 1.96$

Decision: Since $Z = -2.28 < -Z_{0.025} = -1.96$, we have to reject the null hypothesis. So, accept the alternate hypothesis.

Problem 21: Studying the flow of traffic at two busy intersections between 4 P.M. and 6 P.M. (to determine the possible need for turn signals), it was found that on 40 week-days there were on the average 247.3 cars approaching the first intersection from the south that made left turns while on 30 weekdays there were on the average 254.1 cars approaching the second intersection from the south that made left turns. The corresponding sample standard deviations are $s_1 = 15.2$ and $s_2 = 18.7$ Test the null hypothesis $\mu_1 - \mu_2 = 0$ against the alternative hypothesis $\mu_1 - \mu_2 \neq 0$ at the level of significance $\alpha = 0.01$.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 \neq 0$

Level of significance: $\alpha = 0.01$

$\overline{x_1} = 247.3$ $\overline{x_2} = 254.1$

$s_1 = 15.2$ $s_2 = 18.7$

$n_1 = 40$ $n_2 = 30$

Test Statistic, $Z = \dfrac{(\overline{X}_1 - \overline{X}_2) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

$$= \frac{(247.3 - 254.1)}{\sqrt{\dfrac{15.2^2}{40} + \dfrac{18.7^2}{30}}}$$

$$= -1.629$$

$Z_{0.005} = 2.58$

Decision: Since Z = - 1.629 which is not less than $-Z_{0.05}$ = -2.58, we have to accept the null hypothesis.

Large sample confidence interval for the difference of means:

The (1 – α)100% large sample confidence interval for the difference means is given by

$$(\bar{x} - \bar{y}) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Problem 22: The dynamic modulus of concrete is obtained for two different concrete mixes. For the first mix, $n_1 = 33$, $\bar{x} = 115.1$ and $s_1 = 0.47$ psi. For the second mix, $n_2 = 31$, $\bar{y} = 114.6$ and $s_2 = 0.38$. Test with α = 0.05, the null hypothesis of equality of mean dynamic modulus versus the two-sided alternative. Also construct a 95% confidence interval of the difference in mean dynamic modulus.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 \neq 0$

Level of significance: α = 0.05

$\bar{x} = 115.1$ $\qquad$ $\bar{y} = 114.6$

$s_1 = 0.47$ $\qquad$ $s_2 = 0.38$

$n_1 = 33$ $\qquad$ $n_2 = 31$

Test Statistic, $Z = \dfrac{(\overline{X} - \overline{Y}) - \delta_0}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

$$= \frac{(115.1 - 114.6)}{\sqrt{\dfrac{0.47^2}{33} + \dfrac{0.38^2}{31}}}$$

$$= 4.69$$

$Z_{0.025} = 1.96$

Decision: Since Z = 4.69 > $Z_{0.025}$ = 1.96, we have to reject the null hypothesis. So, accept the alternate hypothesis

A 95% confidence interval of the difference in mean dynamic modulus is given by

$$(\overline{x} - \overline{y}) \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$\Rightarrow \quad (115.1 - 114.6) \pm z_{0.025}\sqrt{\frac{0.47^2}{33} + \frac{0.38^2}{31}}$$

$$\Rightarrow \quad (115.1 - 114.6) \pm 1.96\sqrt{\frac{0.47^2}{33} + \frac{0.38^2}{31}}$$

$$\Rightarrow \quad 0.5 \pm 1.96 * 0.10655$$

$$\Rightarrow \quad 0.5 - 1.96 * 0.10655 < \mu 1 - \mu 2 < 0.5 + 1.96 * 0.10655$$

$$\Rightarrow \quad 0.5 - 0.208838 < \mu 1 - \mu 2 < 0.5 + 0.208838$$

$$\Rightarrow \quad 0.291162 < \mu_1 - \mu_2 < 0.708838$$

Comparisons -Two independent Small samples

Let us suppose that we are dealing with two independent random samples of size n_1 and n_2 (n_1, n_2 or both less than 30) from two normal populations having means μ_1 and μ_2 and variances σ_1^2 and σ_2^2 are unknown. We want to test the null hypothesis $\mu_1 - \mu_2 = \delta_0$ where δ is a given constant, against one of the alternatives $\mu_1 - \mu_2 \neq \delta$, $\mu_1 - \mu_2 > \delta$, $\mu_1 - \mu_2 < \delta$. The test procedure is as follows:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = \delta$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 < \delta$

or H_1: $\mu_1 - \mu_2 > \delta$

or H_1: $\mu_1 - \mu_2 \neq \delta$

Sample sizes: n_1 and n_2 (n_1, n_2 or both less than 30)

Sample means : $\overline{X}$ *and* $\overline{Y}$

Sample variances: s_1^2 and s_2^2

Level of significance, α (1% or 5%)

Test Statistic, $t = \dfrac{(\overline{X} - \overline{Y}) - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ *where* $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

Has t distribution with $n_1 + n_2 - 2$ degrees of freedom.

Critical regions for testing H_0: $\mu_1 - \mu_2 = \delta$

Alternate Hypothesis	Reject null hypothesis if
$\mu_1 - \mu_2 < \delta$	$t < -t_\alpha$
$\mu_1 - \mu_2 > \delta$	$t > t_\alpha$
$\mu_1 - \mu_2 \neq \delta$	$t < -t_{\alpha/2}$ or $t > t_{\alpha/2}$

Finally we have to write the decision whether to accept or reject the null hypothesis.

Problem 23: The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines:

Mine - 1	8260	8130	8350	8070	8340	
Mine - 2	7950	7890	7900	8140	7920	7840

Use the 0.01 level of significance to test whether the difference between the means of these two samples is significant.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 \neq 0$

Level of significance: $\alpha = 0.01$

$\bar{x} = 8230$ $\qquad \bar{y} = 7940$

$s_1^2 = 15750$ $\qquad s_2^2 = 10920$

$n_1 = 5$ $\qquad n_2 = 6$

Test Statistic, $t = \dfrac{(\bar{X} - \bar{Y}) - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ where $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$= \frac{(8230 - 7940)}{114.31\sqrt{\frac{1}{5} + \frac{1}{6}}}$$

$= 4.19$

$t_{0.005}$ 9 d.f = 3.250

Decision: Since t = 4.19 > $t_{0.005}$ = 3.250, we have to reject the null hypothesis. So, accept the alternate hypothesis.

Problem 24: Measuring specimens of nylon yarn taken from two spinning machines, it was found that 8 specimens from the first machine had a mean denier of 9.67 with a standard deviation of 1.81 while 10 specimens from the second machine had a mean denier of 7.43 with a standard deviation of 1.48. Assuming that the populations sampled are normal and have the same variance, test the null hypothesis $\mu_1 - \mu_2 = 1.5$ against the alternative hypothesis $\mu_1 - \mu_2 > 1.5$ at the 0.05 level of significance.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 1.5$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 > 1.5$

Level of significance: $\alpha = 0.05$

$\bar{x} = 9.67$ $\quad\quad$ $\bar{y} = 7.43$

$s_1 = 1.81$ $\quad\quad$ $s_2 = 1.48$

$n_1 = 8$ $\quad\quad$ $n_2 = 10$

Test Statistic, $t = \dfrac{(\bar{X} - \bar{Y}) - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ *where* $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$= \frac{(9.67 - 7.43) - 1.5}{1.63260\sqrt{\frac{1}{8} + \frac{1}{10}}}$$

$= 0.9555$

$t_{0.05}$ 16 d.f = 1.746

Decision: Since t = 0.9555 is not greater than $t_{0.05}$ = 1.746, we have to accept the null hypothesis.

Problem 25: As a part of an industrial training program, some trainees are instructed by method A, which is straight computer-based instruction, and some are instructed by method B, which also involves personal attention of an instructor. If random samples of size 10 are taken from large groups of trainees instructed by each of the two methods, and the scores which they obtained in an appropriate achievements test are

Method A	71	75	65	69	73	66	68	71	74	68
Method B	72	77	84	78	69	70	77	73	65	75

Use the 0.05 level of significance to test the claim that method B is more effective. Assume that the populations sampled can be approximated closely with normal distributions having the same variance.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 < 0$

Level of significance: $\alpha = 0.05$

$\bar{x} = 70$ $\bar{y} = 74$

$s_1 = 3.3665$ $s_2 = 5.3995$

$n_1 = 10$ $n_2 = 10$

Test Statistic, $t = \dfrac{(\bar{X} - \bar{Y}) - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ *where* $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$= \frac{(70 - 74)}{4.4993\sqrt{\frac{1}{10} + \frac{1}{10}}}$$

$= -1.9879$

$t_{0.05}$ 18 d.f = 1.734

Decision: Since t = -1.9879 < - $t_{0.05}$ = 1.734, we have to reject the null hypothesis. So, accept the alternate hypothesis.

Problem 26: To compare two kinds of bumper guards, 6 of each kind, were mounted on a certain kind of compact car. Then each car was run into a concrete wall at 5 miles per hour and the following are the costs of the repairs (in rupees)

Bumper Guard 1	407	448	423	465	402	419
Bumper Guard 2	434	415	412	451	433	429

Use the 0.01 level of significance to test whether the difference between the two sample means is significant.

Solution:

Null Hypothesis, H_0 ; $\mu_1 - \mu_2 = 0$

Alternate Hypothesis, H_1: $\mu_1 - \mu_2 \neq 0$

Level of significance: $\alpha = 0.01$

$\bar{x} = 427.33$ $\quad\quad\quad$ $\bar{y} = 429$

$s_1^2 = 597.86$ $\quad\quad\quad$ $s_2^2 = 202$

$n_1 = 6$ $\quad\quad\quad$ $n_2 = 6$

Test Statistic, $t = \dfrac{(\bar{X} - \bar{Y}) - \delta}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ where $S_p^2 = \dfrac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$

$$= \frac{(427.33 - 429)}{19.998\sqrt{\dfrac{1}{6} + \dfrac{1}{6}}}$$

$= -0.1446$

$t_{0.005}$ 10 d.f = 3.169

Decision: Since t = - 0.1446 > $-t_{0.005}$ = - 3.169, we have to accept the null hypothesis.

Small sample confidence interval for the difference of means:

The $(1 - \alpha)100\%$ small sample confidence interval for the difference means is given by

$$(\bar{x} - \bar{y}) \pm t_{\alpha/2}\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Where $t_{\alpha/2}$ is based on $\upsilon = n_1 + n_2 - 2$ degrees of freedom.

Problem 27: The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines:

Mine - 1	8260	8130	8350	8070	8340	
Mine - 2	7950	7890	7900	8140	7920	7840

Construct a 99% confidence interval for the difference between means.

Solution: The $(1 - \alpha)100\%$ small sample confidence interval for the difference means is given by

$$(\bar{x} - \bar{y}) \pm t_{\alpha/2}\sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

Where $t_{\alpha/2}$ is based on $\upsilon = n_1 + n_2 - 2$ degrees of freedom.

$\bar{x} = 8230$ $\bar{y} = 7940$

$s_1^2 = 15750$ $s_2^2 = 10920$

$n_1 = 5$ $n_2 = 6$

$\alpha = 0.01$ and $\alpha/2 = 0.005$

The required confidence interval is

$$(8230 - 7940) \pm t_{0.005}\sqrt{\frac{(5-1)15750+(6-1)10920}{9}}\sqrt{\frac{1}{5}+\frac{1}{6}}$$

$\Rightarrow$ $290 \pm 3.250 \times 38.103 \times 0.6055$

$\Rightarrow$ 290 ± 74.98

$\Rightarrow$ $215.02 < \mu_1 - \mu_2 < 364.98$

Matched Pairs Comparisons

When we are dealing with before and after kind of data, we will use paired sample t test. The purpose of the test is to determine whether there is statistical evidence that the mean difference between paired observations on a particular outcome is significantly different from zero. The Paired Samples *t* Test is a parametric test.

Let (X_i,Y_i); I = 1, 2,…, n. be a set of n paired observations.

Define the difference between each pair of observations as $D_i = X_i - Y_i$, i = 1, 2, …, n. Here we assume that the D_i's are normally distributed with mean μ_D. We assume $\mu_D = 0$ as the means of the two observations are the same and $\mu_D > 0$ means that the mean observation of the first is higher than that of the second.

Null Hypothesis, H_0: $\mu_D = \mu_{D,0}$

Alternate Hypothesis, H_1: $\mu_D > \mu_{D,0}$

Sample size: n

Level of significance = α

Test Statistic, $t = \frac{\overline{D} - \mu_{D,0}}{S_D / \sqrt{n}}$, where $\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$ and $S_D^2 = \frac{\sum_{i=1}^{n} (D_i - \overline{D})^2}{n-1}$

Decision: whether to accept or reject the null hypothesis.

confidence interval for μ_D:

The (1 – α)100% confidence interval for μ_D is given by

$$\overline{D} \pm t_{\alpha/2} \frac{S_D}{\sqrt{n}}$$

Problem 28: The following are the average weekly losses of worker-hours due to accidents in10 industrial plants before and after a certain safety program was put into operation:

Before	45	73	46	124	33	57	83	34	26	17
After	36	60	44	119	35	51	77	29	24	11

Use the 0.05 level of significance to test whether the safety program is effective. Also construct a 90% confidence interval for the men improvement in lost worker-hours.

Solution:

Null Hypothesis, H_0: $\mu_D = 0$

Alternate Hypothesis, H_1: $\mu_D > 0$

Sample size: 10

Level of significance = 0.05

$D_i = X_i - Y_i$, i = 1, 2, …, 10

So, D_i : 9, 13, 2, 5, -2, 6, 6, 5, 2, 6

$\overline{D}$ = 5.2, S_D = 4.08

Test Statistic, $t=\dfrac{\overline{D}-\mu_{D,0}}{S_D/\sqrt{n}}$, where $\overline{D}=\dfrac{\sum_{i=1}^{n}D_i}{n}$ *and* $S_D^2=\dfrac{\sum_{i=1}^{n}(D_i-\overline{D})^2}{n-1}$

$$\Rightarrow\ t=\frac{5.2-0}{4.08/\sqrt{10}}$$

$$\Rightarrow\ t=4.03$$

$t_{0.05}$ 9 d.f = 1.833

Decision: Since t = 4.03 exceeds $t_{0.05}$ 9 d.f = 1.833, we have to reject the null hypothesis. So, accept the alternate hypothesis and hence we conclude that the safety program is effective.

We know that The (1 – α)100% confidence interval for μ_D is given by

$$\overline{D}\pm t_{\alpha/2}\frac{S_D}{\sqrt{n}}$$

Here α = 0.1so that α/2 = 0.05

The confidence interval is $\overline{D}-t_{\alpha/2}\dfrac{S_D}{\sqrt{n}}<\mu_D<\overline{D}+t_{\alpha/2}\dfrac{S_D}{\sqrt{n}}$

$$\Rightarrow\ 5.2-t_{0.05}\frac{4.08}{\sqrt{10}}<\mu_D<5.2+t_{0.05}\frac{4.08}{\sqrt{10}}$$

$$\Rightarrow\ 5.2-1.833\frac{4.08}{\sqrt{10}}<\mu_D<5.2+1.833\frac{4.08}{\sqrt{10}}$$

$$\Rightarrow\ 5.2-2.365<\mu_D<5.2+2.365$$

$$\Rightarrow\ 2.835<\mu_D<7.565$$

Problem 29: In a study of the effectiveness of physical exercise in weight reduction, a group of 16 persons engaged in a prescribed program of physical exercise for one month showed the following results:

Weight before (pounds)	Weight after (pounds)
209	196
178	171
169	170
212	207
180	177
192	190
158	159
180	180
170	164
153	152
183	179
165	162
201	199
179	173
243	231
144	140

Use the 0.01 level of significance to test whether the prescribed program of exercise is effective.

Solution:

Null Hypothesis, H_0: $\mu_D = 0$

Alternate Hypothesis, H_1: $\mu_D > 0$

Sample size: 16

Level of significance = 0.01

$D_i = X_i - Y_i$, i = 1, 2, …, 16

So, D_i : 13, 7, -1, 5, 3, 2, -1, 0, 6, 1, 4, 3, 2, 6, 12, 4

$\overline{D}$ = 4.125, S_D = 4.064

Test Statistic, $t = \dfrac{\overline{D} - \mu_{D,0}}{S_D / \sqrt{n}}$, where $\overline{D} = \dfrac{\sum_{i=1}^{n} D_i}{n}$ *and* $S_D^2 = \dfrac{\sum_{i=1}^{n} (D_i - \overline{D})^2}{n-1}$

$$\Rightarrow\ t = \frac{4.125 - 0}{4.064 / \sqrt{16}}$$

$$\Rightarrow\ t = 4.06$$

$t_{0.01}$ 15 d.f = 2.602

Decision: Since t = 4.06 exceeds $t_{0.01}$ 15 d.f = 2.602, we have to reject the null hypothesis. So, accept the alternate hypothesis and hence we conclude that the physical exercise program is effective.

Inferences Concerning Variances

The Estimation of Variances

Let $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$ be the sample variance based on a random sample from any population, discrete or continuous, having variance σ^2. The mean of the sample distribution of S^2 is given by σ^2.

So, sample variance $S^2 = \dfrac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$ is an unbiased estimator of σ^2.

For random samples from normal populations, $\dfrac{(n-1)S^2}{\sigma^2}$ is a random variable having the chi square distribution with n-1 degrees of freedom. Thus, with χ_α^2 defined with n-1 degrees of freedom, we can assert with probability 1- α that the inequality

$\chi^2_{1-\alpha/2} < \dfrac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}$ will be satisfied.

Then (1-α)100% confidence interval for population variance σ^2 is

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

Problem 1: Suppose that the refractive indices of 20 pieces of glass (randomly selected from a large shipment purchased by the optical firm) have a variance of $1.20*10^{-4}$. Construct a 95% confidence interval for σ, the standard deviation of the population sampled.

Solution: Here sample size n =20 then degrees of freedom 20-1=19

Since confidence is 95%, α=1-0.95=0.05 and hence α/2 =0.025.

From table 5, $\chi^2_{0.975} = 8.907$ and $\chi^2_{0.025} = 32.852$, then 95% confidence interval for σ^2 is

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

$$\frac{(19)(1.20\times10^{-4})}{32.852} < \sigma^2 < \frac{(19)(1.20\times10^{-4})}{8.907}$$

$$0.000069 < \sigma^2 < 0.000256$$

$$0.0083 < \sigma < 0.0160$$

Hence the 95% confidence interval from 0.0083 to 0.0160 contains σ, the true standard deviation of the refractive index.

Problem 2: A manufacturer claims that the average tar content of a certain kind of cigarette is $\mu = 14.0$. In an attempt to show that it differs from this value, five measurements are made of the tar content (mg per cigarette):

14.5, 14.2, 14.4, 14.3, 14.6

Construct a 99% confidence interval for the variance of the population sampled.

Solution:

From the sample data

Sample mean size n =5

Sample mean $\bar{x} = \frac{14.5+14.2+14.4+14.3+14.6}{5} = 14.4$

Sample variance

$$s^2 = \frac{(14.4-14.5)^2 + (14.4-14.2)^2 + (14.4-14.4)^2 + (14.4-14.3)^2 + (14.4-14.6)^2}{5-1}$$

$$= 0.025$$

s = 0.1581

Here α = 0.01, then $\chi^2_{\alpha/2} = \chi^2_{0.005} = 14.860$ and $\chi^2_{1-\alpha/2} = \chi^2_{0.995} = 0.207$ at 4 degrees of freedom.

Hence 99% confidence interval for variance is

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

$$\frac{(5-1)0.025}{14.860} < \sigma^2 < \frac{(5-1)0.025}{0.207}$$

$$0.0067 < \sigma < 0.4831$$

<u>Hypothesis concerning one variance</u>:

In hypothesis concerning one variance we shall consider the problem of testing the null hypothesis that a population variance equals a specified constant against a suitable one-sided or two-sided alternative hypothesis. The test procedure is as follows:

Null hypothesis, H_0 : $\sigma^2 = \sigma_0^2$

Alternative hypothesis, H_1 : $\sigma^2 < \sigma_0^2$

$(or)\ \sigma^2 > \sigma_0^2$

$(or)\ \sigma^2 \neq \sigma_0^2$.

Level of significance = α

Sample variance = s^2

Sample size = n

The test statistic, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$.

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 < \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha}$
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_{\alpha}$
$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$

Decision: whether to reject or accept the null hypothesis.

Table 5 Values of χ^2_α

ν	$\alpha = 0.995$	$\alpha = 0.99$	$\alpha = 0.975$	$\alpha = 0.95$	$\alpha = 0.05$	$\alpha = 0.025$	$\alpha = 0.01$	$\alpha = 0.005$	ν
1	0.0000393	0.000157	0.000982	0.00393	3.841	5.024	6.635	7.879	1
2	0.0100	0.0201	0.0506	0.103	5.991	7.378	9.210	10.597	2
3	0.0717	0.115	0.216	0.352	7.815	9.348	11.345	12.838	3
4	0.207	0.297	0.484	0.711	9.488	11.143	13.277	14.860	4
5	0.412	0.554	0.831	1.145	11.070	12.833	15.086	16.750	5
6	0.676	0.872	1.237	1.635	12.592	14.449	16.812	18.548	6
7	0.989	1.239	1.690	2.167	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	30.144	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	33.924	36.781	40.289	42.796	22
23	9.260	10.196	11.689	13.091	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	36.415	39.364	42.980	45.559	24
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928	25
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	40.113	43.195	46.963	49.645	27
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993	28
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336	29
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672	30
40	20.707	22.164	24.433	26.509	55.758	59.342	63.691	66.766	40
50	27.991	29.707	32.357	34.764	67.505	71.420	76.154	79.490	50
60	35.534	37.485	40.482	43.188	79.082	83.298	88.379	91.952	60
70	43.275	45.442	48.758	51.739	90.531	95.023	100.425	104.215	70
80	51.172	53.540	57.153	60.391	101.879	106.629	112.329	116.321	80
90	59.196	61.754	65.647	69.126	113.145	118.136	124.116	128.299	90
100	67.328	70.065	74.222	77.929	124.342	129.561	135.807	140.169	100

Problem 3: The lapping process which is used to grind certain silicon wafers to the proper thickness is acceptable only if σ, the population standard deviation of the thickness of dice cut from the wafers, is at most 0.50 mil. Use the 0.05 level of significance to test the null hypothesis $\sigma = 0.50$ against the alternative hypothesis $\sigma > 0.50$, if the thickness of 15 dice cut from such wafers have a standard deviation of 0.64 mil.

Solution:

Null hypothesis: $\sigma = 0.50$

Alternative hypothesis: $\sigma > 0.50$

Level of significance $\alpha = 0.05$

n = 15, $\sigma = 0.50$

s = 0.64

Test Statistic $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

$$= \frac{(15-1)(0.64)^2}{(0.50)^2}$$

$$= 22.94$$

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi_\alpha^2$

Decision: Since $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} = 22.94$ does not exceed 23.685, the value of $\chi_{0.05}^2$ for 14 degrees of freedom, the null hypothesis cannot be rejected. So, accept the null hypothesis.

Problem 4: A random sample of 6 steel beams has a mean compressive strength of 58,392 psi with a standard deviation of 648 psi. Test the null hypothesis $\sigma = 600$ psi for the compressive strength of the given kind of steel against the alternative hypothesis $\sigma > 600$ psi. Use 0.05 level of significance.

Solution:

Null hypothesis: $\sigma = 600$

Alternative hypothesis: $\sigma > 600$

Level of significance $\alpha = 0.05$

n = 6, $\sigma = 600$

s = 648

Test Statistic, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

$$= \frac{(6-1)(648)^2}{(600)^2}$$

$$= 5.832$$

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi_\alpha^2$

Decision: Since $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} = 5.832$ does not exceed 11.07, the value of $\chi_{0.05}^2$ for 5 degrees of freedom, the null hypothesis cannot be rejected. So, accept the null hypothesis.

Problem 5: If 12 determinations of the specific heat of iron have a standard deviation of 0.0086, test the null hypothesis that $\sigma = 0.010$ for such determinations. Use the alternative hypothesis $\sigma \neq 0.010$ and the level of significance $\alpha = 0.01$

Solution:

Null hypothesis: $\sigma = 0.010$

Alternative hypothesis: $\sigma \neq 0.010$

level of significance $\alpha = 0.01$

n = 12, $\sigma = 0.010$

s = 0.0086

Test statistic, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

$$= \frac{(12-1)(0.0086)^2}{(0.01)^2}$$

$$= 8.1356$$

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$

Decision: Since $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$ =8.1356 lies between 2.603 and 26.757 we can't reject the null hypothesis where $\chi^2_{0.095} = 2.603$ *and* $\chi^2_{0.005}$ =26.757 with 11 degrees of freedom.

Problem 6: Playing 10 rounds of gold on his home course, a golf professional averaged 71.3 with a standard deviation of 1.32. Test the null hypothesis that the consistency of his game on his home course is actually measured by $\sigma = 1.20$, against the alternative hypothesis that he is less consistent. Use the level of significance $\alpha = 0.05$.

Solution:

Null hypothesis: $\sigma = 1.20$

Alternative hypothesis: $\sigma > 1.20$

Level of significance $\alpha = 0.05$

n = 10, σ = 1.20

s = 1.32

Test statistic, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

$$= \frac{(10-1)(1.32)^2}{(1.20)^2}$$

$$= 10.89$$

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 > \sigma_0^2$	$\chi^2 > \chi^2_\alpha$

Decision: Since $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} = 10.89$ does not exceed 16.919, the value of $\chi^2_{0.05}$ for 9 degrees of freedom, the null hypothesis cannot be rejected. So, accept the null hypothesis.

Problem 7: Use the 0.01 level of significance to test the null hypothesis that $\sigma = 0.015$ inch for the diameters of certain bolts against the alternative hypothesis that $\sigma \neq 0.015$ inch, given that a random sample of size 15 yielded s^2=0.00011.

Solution:

Null hypothesis: $\sigma = 0.015$

Alternative hypothesis: $\sigma \neq 0.015$

Level of significance $\alpha = 0.01$

n = 15, $\sigma = 0.015$

s^2 = 0.00011

Test statistic, $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$

$$= \frac{(15-1)(0.00011)}{(0.015)^2}$$

$$= 6.8444$$

The critical region for testing $\sigma^2 = \sigma_0^2$

Alternative hypothesis	Reject the null hypothesis if
$\sigma^2 \neq \sigma_0^2$	$\chi^2 < \chi^2_{1-\alpha/2}$ or $\chi^2 > \chi^2_{\alpha/2}$

Decision: Since $\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2} = 6.8444$ lies between 4.075 and 31.319 where

$\chi^2_{0.095} = 4.075 \; and \; \chi^2_{0.005}$ =31.319 with 14 degrees of freedom, we can't reject the null hypothesis.

<u>Problem :</u> The security department of a large office building wants to test the null hypothesis that $\sigma = 2.0$ minutes for the time it takes a guard to walk his round against the alternative hypothesis that $\sigma \neq 2.0$ minutes. What can it conclude at the 0.01 level of significance if a random sample of size n = 31 yields s = 1.8 minutes?

Hypothesis concerning two variances

Consider two normal populations having the variances σ_1^2 *and* σ_2^2 respectively. To test whether these two variances are equal or not we will take two random samples of size n_1 and n_2 having the variances S_1^2 *and* S_2^2 respectively from the two normal populations. The test procedure is as follows:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 < \sigma_2^2$

or H_1: $\sigma_1^2 > \sigma_2^2$

or H_1: $\sigma_1^2 \neq \sigma_2^2$

Sample sizes: n_1 and n_2 (whatever may be the sizes)

Sample variances: S_1^2 *and* S_2^2

Level of significance = α (1% or 5%)

Test Statistic, $F = \dfrac{S_1^2}{S_2^2}$

Critical regions for testing H_0: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 < \sigma_2^2$	$F = \dfrac{S_2^2}{S_1^2}$	$F > F_{\alpha}(n_2 - 1, n_1 - 1)$
$\sigma_1^2 > \sigma_2^2$	$F = \dfrac{S_1^2}{S_2^2}$	$F > F_{\alpha}(n_1 - 1, n_2 - 1)$
$\sigma_1^2 \neq \sigma_2^2$	$F = \dfrac{S_M^2}{S_m^2}$	$F > F_{\alpha/2}(n_M - 1, n_m - 1)$

Finally we have to write the decision whether to accept or reject the null hypothesis.

Problem 8: It is desired to determine whether there is less variability in the silver plating done by Company 1 than in that done by Company 2. If independent random samples of size 12 of the two companies' work yield s_1= 0.035mil and s_2= 0.062 mil, test the null hypothesis $\sigma_1^2 = \sigma_2^2$ against the alternative hypothesis $\sigma_1^2 < \sigma_2^2$ at the 0.05 level of significance.

Solution:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 < \sigma_2^2$

Sample sizes: n_1 = 12 and n_2 = 2 (whatever may be the sizes)

Sample variances: $S_1^2 = 0.035^2$ and $S_2^2 = 0.062^2$

Level of significance, $\alpha = 0.05$

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

$= \frac{0.062^2}{0.035^2}$

= 3.14

$F_{0.05}(11, 11) = 2.82$

Critical regions for testing H_0: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 < \sigma_2^2$	$F = \frac{S_2^2}{S_1^2} = 3.14$	$F > F_{0.05}(11, 11)$

Decision:Since F = 3.14 exceeds $F_{0.05}(11, 11) = 2.82$, the null hypothesis must be rejected. That is accept the alternative hypothesis. So, The silver platting done by company 1 is less variable than that done by company 2.

Problem 9: Two different lighting techniques are compared by measuring the intensity of light at selected locations in areas lighted by two methods. If 15 measurements in the first area had a standard deviation of 2.7 foot-candles and 21 measurements in the second area had a standard deviation of 4.2 foot-candles, can it be concluded that the lighting in the second area is less uniform? Use a 0.01 level of significance.

Solution:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 < \sigma_2^2$

Sample sizes: $n_1 = 15$ and $n_2 = 21$ (whatever may be the sizes)

Sample variances: $S_1^2 = 2.7^2$ and $S_2^2 = 4.2^2$

Level of significance, $\alpha = 0.01$

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

$= \frac{4.2^2}{2.7^2}$

$= 2.41$

$F_{0.01}(20, 14) = 3.51$

Critical regions for testing H_0: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 < \sigma_2^2$	$F = \frac{S_2^2}{S_1^2} = 2.41$	$F > F_{0.01}(20, 14)$

Decision:Since F = 2.41 does not exceed $F_{0.01}(20, 14) = 3.51$, the null hypothesis can't be rejected. That is accept the Null hypothesis. So both the lighting techniques are the same

Problem 10: Studying the flow of traffic at two busy intersections between 4 P.M. and 6 P.M. (to determine the possible need for turn signals), it was found that on 40 week-days there were on the average 247.3 cars approaching the first intersection from the south that made left turns while on 30 weekdays there were on the average 254.1 cars approaching the second intersection from the south that made left turns. The corresponding sample standard deviations are s_1 =15.2 and s_2 =18.7. Use the 0.05 level of significance to test the claim that there is a greater variability in the number of cars which make left turns approaching from the south between 4 P.M and 6 P.M at the second intersection. Assume distributions are normal.

Solution:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 < \sigma_2^2$

Sample sizes: n_1 = 40 and n_2 =30 (whatever may be the sizes)

Sample variances: $S_1^2 = 2.7^2$ and S_2^2 = 4.2^2

Level of significance, α = 0.05

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

$= \frac{18.7^2}{15.2^2}$

= 1.51

$F_{0.01}(29, 39) = 1.76$

Critical regions for testing H$_0$: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 < \sigma_2^2$	$F = \frac{S_2^2}{S_1^2} = 2.41$	$F > F_{0.05}(29, 39)$

Decision: Since F = 1.51 does not exceed $F_{0.05}(29, 39) = 1.76$, the null hypothesis can't be rejected. That is accept the Null hypothesis. The variability at both the intersections are the same.

Problem 11: Measuring specimens of nylon yarn taken from two spinning machines, it was found that 8 specimens from the first machine had a mean denier of 9.67 with a standard deviation of 1.81 while 10 specimens from the second machine had a mean denier of 7.43 with a standard deviation of 1.48. Assuming that the populations sampled are normal and independent. Test the null hypothesis that the two populations have equal variances. Use the 0.02 level of significance.

Solution:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 \neq \sigma_2^2$

Sample sizes: $n_M = 8$ and $n_m = 10$ (whatever may be the sizes)

Sample variances: $S_1^2 = 1.81^2$ and $S_2^2 = 1.48^2$

Level of significance, α = 0.02

Test Statistic, $F = \frac{S_M^2}{S_m^2}$

$= \frac{1.81^2}{1.48^2}$

$= 1.495$

$F_{0.02/2}(7, 9) = F_{0.01}(7, 9) = 5.61$

Critical regions for testing H_0: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 \neq \sigma_2^2$	$F = \dfrac{S_M^2}{S_m^2} = 1.495$	$F > F_{0.01}(7, 9)$

Decision: Since F = 1.495 does not exceed $F_{0.01}(7, 9) = 5.61$, the null hypothesis can't be rejected. That is accept the Null hypothesis. So, the populations variances are equal.

Problem 12: The following random samples are measurements of the heat-producing capacity (in millions of calories per ton) of specimens of coal from two mines:

Mine - 1	8260	8130	8350	8070	8340	
Mine - 2	7950	7890	7900	8140	7920	7840

Use the 0.02 level of significance to test whether it is reasonable to assume that the variances of the two populations sampled are equal. Also construct a 98% confidence interval for $\dfrac{\sigma_2^2}{\sigma_1^2}$.

Solution:

Null Hypothesis, H_0 ; $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis, H_1: $\sigma_1^2 \neq \sigma_2^2$

Sample sizes: $n_M = 6$ and $n_m = 5$ (whatever may be the sizes)

Sample variances: $S_1^2 = 15{,}750$ and $S_2^2 = 10{,}920$

Level of significance, $\alpha = 0.02$

Test Statistic, $F = \dfrac{S_M^2}{S_m^2}$

$$= \frac{S_1^2}{S_2^2} = \frac{15750}{10920} = 1.44$$

$F_{0.02/2}(4, 5) = F_{0.01}(4, 5) = 11.392$

Critical regions for testing H_0: $\sigma_1^2 = \sigma_2^2$

Alternate Hypothesis	Test Statistic	Reject null hypothesis if
$\sigma_1^2 \neq \sigma_2^2$	$F = \frac{S_M^2}{S_m^2} = 1.495$	$F > F_{0.01}(4, 5)$

Decision: Since F = 1.44 does not exceed $F_{0.01}(4, 5) = 11.392$, the null hypothesis can't be rejected. That is accept the Null hypothesis. So, the populations variances are equal.

Confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ normal populations:

The (1 – α)100% confidence interval for $\frac{\sigma_2^2}{\sigma_1^2}$ is given by

$$F_{1-\alpha/2}(n_1 - 1, n_2 - 1)\,\frac{s_2^2}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < F_{\alpha/2}(n_1 - 1, n_2 - 1)\,\frac{s_2^2}{s_1^2}$$

Problem 13: One process of making green gasoline takes sucrose, which can be derived from biomass, and converts it into gasoline using catalytic reactions. This is not a process for making a gasoline additive but fuel itself, so research is still at the pilot plant stage. At one step in a pilot plant process, the product consists of carbon chains of length 3. Nine runs were made with each of two catalysts and the product volumes (gal) are as follows:

Catalyst 1	0.63	2.64	1.85	1.68	1.09	1.67	0.73	1.04	0.68
Catalyst 2	3.71	4.09	4.11	3.75	3.49	3.27	3.72	3.49	4.26

The sample variances are respectively, 0.4548 and 0.1089. Obtain a 98% confidence interval for σ_2^2 / σ_1^2.

Solution: Given $n_1 = n_2 = 9$, $s_1^2 = 0.4548$, $s_2^2 = 0.1089$ and $(1 - \alpha)100\% = 98\%$

We know that $F_{1-\alpha}(v_1,v_2) = 1/ F_{\alpha}(v_2,v_1)$

Now $\alpha = 0.02$ and $\alpha/2 = 0.01$

$F_{0.01}(9,9) = 6.03$ and $F_{0.99} = 1 / F_{0.01}(9,9) = 1 / 6.03 = 0.16584$

The $(1 - \alpha)100\%$ confidence interval for $\dfrac{\sigma_2^2}{\sigma_1^2}$ is given by

$$F_{1-\alpha/2}(n_1 - 1, n_2 - 1)\frac{s_2^2}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < F_{\alpha/2}(n_1 - 1, n_2 - 1)\frac{s_2^2}{s_1^2}$$

$$\Rightarrow \qquad F_{0.99}(9,9)\frac{0.1089}{0.4548} < \frac{\sigma_2^2}{\sigma_1^2} < F_{0.01}(9,9)\frac{0.1089}{0.4548}$$

$$\Rightarrow \qquad 0.16584(0.23944) < \frac{\sigma_2^2}{\sigma_1^2} < 6.03(0.23944)$$

$$\Rightarrow \qquad 0.04 < \frac{\sigma_2^2}{\sigma_1^2} < 1.44$$