Statistics for Data Science - 2

Week 12 Graded Assignment Solution

1. The IQs (intelligence quotients) of 25 students from one batch of IITM students showed a mean of 110 with a standard deviation of 8, while the IQs of 25 students from another batch of IITM students showed a mean of 115 with a standard deviation of 7. Is there a significant difference between the IQs of the two groups at a 0.05 level of significance?

   a) Yes

   b) No

   Hint: Use $F_Z^{-1}(0.025) = -1.96$

   Solution:

   Let X_i and Y_i represent the IQ's of both batch of students.
   $X_1, X_2, \ldots, X_{25} \sim \text{N}(\mu_1, 8^2)$ and $Y_1, Y_2, \ldots, Y_{25} \sim \text{N}(\mu_2, 7^2)$
   $\overline{X} = 110$ and $\overline{Y} = 115$
   Consider, $H_0 : \mu_1 = \mu_2, H_A : \mu_1 \neq \mu_2$
   $T = \overline{X} - \overline{Y} \sim \text{N}(\mu_1 - \mu_2, \frac{64}{25} + \frac{49}{25})$ i.e. $\text{N}(\mu_1 - \mu_2, \frac{113}{25})$
   Test: Reject H_0 if $|T| > c$.

$$\alpha = P(|T| > c \mid H_0) = P\left(|\frac{T}{\sqrt{113/25}}| > \frac{c}{\sqrt{113/25}}\right)$$
$$= P\left(|Z| > \frac{c}{\sqrt{113/25}}\right) = 2F_Z\left(\frac{-c}{\sqrt{113/25}}\right)$$

   $\Rightarrow c = -\sqrt{\frac{113}{25}} F_Z^{-1}(\alpha/2)$
   $\Rightarrow c = -\sqrt{\frac{113}{25}} F_Z^{-1}(0.025)$
   $\Rightarrow c = -\sqrt{\frac{113}{25}} \times (-1.96) = 4.167$
   Since, $|\overline{X} - \overline{Y}| = |110 - 115| = 5 > 4.167$
   Therefore, we will reject H_0.
   This implies that there a significant difference between the IQs of the two groups at a 0.05 level of significance.

2. A sociologist focusing on popular culture and media believes that the average number of hours per week (hrs/week) spent on social media is different for men and women. The researcher knows that the standard deviations of amount of time spent on social media are 5 hrs/week and 6 hrs/week for men and women, respectively. Examining two independent random samples of 64 individuals each, if the average number of hrs/week

spent on social media for the sample of men is 1.5 hours greater than that for the sample of women, what conclusion can be made from a hypothesis test where, $H_0 : \mu_M = \mu_W$ and $H_A : \mu_M \neq \mu_W$? Take $\alpha = 0.05$.

a) Reject H_0

b) Accept H_0

Solution:
Let X_i and Y_i represent the average number of hrs/week spent on social media by men and women respectively.
$X_1, X_2, \ldots, X_{64} \sim \mathrm{N}(\mu_1, 5^2)$ and $Y_1, Y_2, \ldots, Y_{64} \sim \mathrm{N}(\mu_2, 6^2)$
$|\overline{X} - \overline{Y}| = 1.5$
Consider, $H_0 : \mu_1 = \mu_2, H_A : \mu_1 \neq \mu_2$
$T = \overline{X} - \overline{Y} \sim \mathrm{N}(\mu_1 - \mu_2, \frac{25}{64} + \frac{36}{64})$ i.e. $\mathrm{N}(\mu_1 - \mu_2, \frac{61}{64})$
Test: Reject H_0 if $|T| > c$.

$$\begin{aligned}\alpha = P(|T| > c \mid H_0) &= P\left(|\frac{T}{\sqrt{61/64}}| > \frac{c}{\sqrt{61/64}}\right)\\ &= P\left(|Z| > \frac{c}{\sqrt{61/64}}\right) = 2F_Z\left(\frac{-c}{\sqrt{61/64}}\right)\end{aligned}$$

$\Rightarrow c = -\sqrt{\frac{61}{64}}F_Z^{-1}(\alpha/2)$
$\Rightarrow c = -\sqrt{\frac{61}{64}}F_Z^{-1}(0.025)$
$\Rightarrow c = -\sqrt{\frac{61}{64}} \times (-1.96) = 1.913$
Since, $|\overline{X} - \overline{Y}| = 1.5 < 1.913$
Therefore, we will accept H_0.

3. An IITM instructor conducts two live sessions for two different classes, call it A and B, in Statistics. Session A had 25 students attending while session B had 36 students. The instructor conducted a test for the two sessions. Although there was no significant difference in mean grades, session A had a standard deviation of 10 while session B had a standard deviation of 14. Can we conclude at the 0.01 level of significance that the variability in marks of class B is greater than that of A?

a) Yes

b) No

Hint: Use $F^{-1}_{F_{(35,24)}}(0.99) = 2.529$
Solution:
$H_0 : \sigma_1 = \sigma_2, H_A : \sigma_1 < \sigma_2$

Test: Reject H_0 if $\dfrac{S_B^2}{S_A^2} > 1 + c_R$

We know that, $\dfrac{S_B^2}{S_A^2} \sim F(n_2 - 1, n_1 - 1)$
$n_1 = 25, n_2 = 36$
$\Rightarrow \dfrac{S_B^2}{S_A^2} \sim F(35, 24)$
Therefore,

$$\begin{aligned} \alpha &= 1 - F_{F(35,24)}(1 + c_R) \\ \Rightarrow 1 + c_R &= F^{-1}_{F(35,24)}(1 - \alpha) = F^{-1}_{F(35,24)}(0.99) \\ \Rightarrow 1 + c_R &= 2.529 \end{aligned}$$

Since, $\dfrac{S_B^2}{S_A^2} = \frac{14^2}{10^2} = 1.96 < 2.529$
Therefore, we will accept H_0.
This implies that at the 0.01 level of significance the variability in marks of class B is not greater than that of A.

4. The manufacturer of a new car claims that a typical car gets a mileage of 40 kilometres per litre. We think that the mileage is less. To test our suspicion, we perform the hypothesis test with $H_0 : \mu = 40$ and $H_A : \mu < 40$. Suppose we take a random sample of 900 new cars and find that their average mileage is 39.8 kilometres per litre and sample standard deviation is 2, what does a t-test say about a null hypothesis with a significance level of 0.05?

   a) Reject H_0

   b) Accept H_0

Hint: Use $F^{-1}_{t_{899}}(0.05) = -1.646$
Solution:
Null hypothesis, $H_0 : \mu = 40$
Alternate hypothesis, $H_A : \mu < 40$

Test: Reject H_0 if $\overline{X} < c$

Given, $\alpha = 0.05$ and $\overline{X} = 39.8$

In this problem, we do not know the population variance, σ^2.
The sample variance $S^2 = 2^2$

$$\alpha = P(\overline{X} < c | \mu = 40)$$

$$\alpha = P\left(\frac{\overline{X} - 40}{\sqrt{S^2/n}} < \frac{c - 40}{\sqrt{S^2/n}}\right)$$

$$\alpha = P\left(\frac{\overline{X} - 40}{\sqrt{4/900}} < \frac{c - 40}{\sqrt{4/900}}\right)$$

$$\alpha = F_{t_{899}}\left(\frac{c - 40}{\sqrt{4/900}}\right)$$

$$0.05 = F_{t_{899}}\left(\frac{c - 40}{\sqrt{4/900}}\right)$$

$$c = 40 + \sqrt{\frac{4}{900}} F^{-1}_{t_{899}}(0.05)$$

$$c = 39.89$$

Since, $\overline{X} < c$, reject H_0.

5. The standard deviation of weights of 70 gram bags of white cheddar popcorn is expected to be 2.5 grams. A random sample of 20 packages showed a standard deviation of 3 grams. Is the apparent increase in variability significant at the 0.05 level.?

   a) Yes

   b) No

   Hint: Use $F^{-1}_{\chi^2_{19}}(0.95) = 30.14$
   Solution:
   As per given information, the null and alternative hypothesis are given by

$$H_0 : \sigma = 2.5, \quad H_A : \sigma > 2.5$$

   Define a test statistic T as $T = S^2$.
   We know that $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{19S^2}{2.5^2} \sim \chi^2_{19}$.

   Test: reject the null hypothesis if $S^2 > c^2$.
   If the significance level of the test is 0.05, then

$$
\begin{aligned}
&P(S^2 > c^2) = 0.05 \\
\Rightarrow &P\left(\frac{19S^2}{2.5^2} > \frac{19c^2}{2.5^2}\right) = 0.05 \\
\Rightarrow &P\left(\chi^2_{19} > \frac{19c^2}{2.5^2}\right) = 0.05 \\
\Rightarrow &1 - P\left(\chi^2_{19} < \frac{19c^2}{2.5^2}\right) = 0.05 \\
\Rightarrow &\frac{19c^2}{2.5^2} = 30.14 \\
\Rightarrow &c^2 = \frac{6.25 \times 30.14}{19} = 9.91
\end{aligned}
$$

Since $S^2 = 9 < 9.91$, we will not reject the null hypothesis.
Therefore, the apparent increase in variability is not significant at the 0.05 level.

6. Independent random samples of ceramic produced by two different processes were tested for hardness. The results are:

Process 1	Process 2
8.5	9.0
9.5	9.5
8.0	10.5
9.0	9.5
10.0	10.0
9.5	9.0
10.5	9.0
10.0	9.5

Table 11.1.G

Can we conclude at 5% level of significance that the variances in hardness are equal?

a) Yes

b) No

Hint: Use $F^{-1}_{F_{(7,7)}}(0.025) = 0.2$

Solution:

Let Process 1 and Process 2 values denoted by X_i and Y_i respectively.
$H_0 : \sigma_1 = \sigma_2, H_A : \sigma_1 \neq \sigma_2$

Test: Reject H_0 if $\dfrac{S_X^2}{S_Y^2} > 1 + c_R$ or $\dfrac{S_X^2}{S_Y^2} < 1 - c_L$

We know that, $\dfrac{S_X^2}{S_Y^2} \sim F(n_1-1, n_2-1)$
$n_1 = 8, n_2 = 8$
$\Rightarrow \dfrac{S_X^2}{S_Y^2} \sim F(7,7)$
Therefore,

$$\alpha/2 = F_{F(7,7)}(1-c_L)$$
$$\Rightarrow 1-c_L = F^{-1}_{F(7,7)}(\alpha/2) = F^{-1}_{F(7,7)}(0.025)$$

$\Rightarrow 1-c_L = 0.2$
Since, $\dfrac{S_X^2}{S_Y^2} = \frac{0.6964}{0.2857} = 2.437 > 0.2$
Similarly we can check for other condition.

$$\alpha/2 = 1 - F_{F(7,7)}(1+c_R)$$
$$\Rightarrow 1+c_R = F^{-1}_{F(7,7)}(1-\alpha/2) = F^{-1}_{F(7,7)}(0.975)$$

$\Rightarrow 1+c_R = 4.99$
Since, $\dfrac{S_X^2}{S_Y^2} = \frac{0.6964}{0.2857} = 2.437 < 4.99$
Therefore, we will accept H_0.

7. The standard deviation of a component in a drug is expected to be 0.00002 kg. A pharmacist suspecting the variability to be higher obtains a sample of 8 drugs and found the sample standard deviation to be 0.00005 kg.

   1. Identify the null and alternative hypothesis:
      (a) $H_0 : \sigma = 0.00002,\ H_A : \sigma > 0.00002$
      (b) $H_0 : \sigma = 0.00005,\ H_A : \sigma \neq 0.00005$
      (c) $H_0 : \sigma = 0.00002,\ H_A : \sigma < 0.00002$
      (d) $H_0 : \sigma = 0.00002,\ H_A : \sigma \neq 0.00002$
   2. What conclusion would a χ^2 test reach at a significance level of 0.01?
      (a) Accept the null hypothesis.
      (b) Accept the alternative hypothesis.

   Solution:
   As per given information, the null and alternative hypothesis are given by

   $$H_0 : \sigma = 0.00002,\ H_A : \sigma > 0.00002$$

   Define a test statistic T as $T = S^2$.

We know that $\dfrac{(n-1)S^2}{\sigma^2} = \dfrac{7 \times 0.00005^2}{0.00002^2} \sim \chi^2_7$.

Test: reject the null hypothesis if $S^2 > c^2$.

If the significance level of the test is 0.01, then

$$
\begin{aligned}
&P(S^2 > c^2) = 0.01 \\
\Rightarrow &P\left(\frac{7 \times 0.00005^2}{0.00002^2} > \frac{7 \times c^2}{0.00002^2}\right) = 0.01 \\
\Rightarrow &P\left(\chi^2_7 > \frac{7 \times c^2}{0.00002^2}\right) = 0.01 \\
\Rightarrow &1 - P\left(\chi^2_7 \leq \frac{7 \times c^2}{0.00002^2}\right) = 0.01 \\
\Rightarrow &F_{\chi^2_7}\left(\frac{7c^2}{0.00002^2}\right) = 0.99 \\
\Rightarrow &\frac{7c^2}{0.00002^2} = 18.475 \\
\Rightarrow &c^2 = \frac{18.475 \times 0.00002^2}{7} = 10.56 \times 10^{-8}
\end{aligned}
$$

Since $S^2 = 25 \times 10^{-8} > 10.56 \times 10^{-8}$, we will reject the null hypothesis.

Statistics for Data Science - II
Week - 12 Graded Sept 2024

Week - 12 :

1. A group of scientists is investigating the durability of two types of eco-friendly batteries, SolarCell and GreenPower, under constant usage. A random sample of 9 SolarCell batteries shows a standard deviation in battery life of 3.8 hours, while a random sample of 12 GreenPower batteries has a standard deviation in battery life of 2.9 hours. Can the scientists conclude that the variability in battery life for SolarCell is greater than that of GreenPower?
   Hint: $F^{-1}_{F(9,12)}(0.95) = 2.80, F^{-1}_{F(9,12)}(0.99) = 4.39, F^{-1}_{F(8,11)}(0.95) = 2.95, F^{-1}_{F(8,11)}(0.99) = 4.74$

   (a) There is not enough evidence at 5% level of significance to conclude that the variability in battery life is greater for SolarCell than for GreenPower batteries.

   (b) There is enough evidence at 5% level of significance to conclude that the variability in battery life is greater for SolarCell than for GreenPower batteries.

   (c) There is not enough evidence at 1% level of significance to conclude that the variability in battery life is greater for SolarCell than for GreenPower batteries.

   (d) There is enough evidence at 1% level of significance to conclude that the variability in battery life is greater for SolarCell than for GreenPower batteries.

   Answer : a, c

   Solution :
   Given :

$$\text{Number of samples of Solar Panel, } n_1 = 9$$
$$\text{Standard deviation, } S_1 = 3.8$$
$$\text{Number of samples of Solar Panel, } n_2 = 12$$
$$\text{Standard deviation, } S_2 = 2.9$$

Now,

$$\text{Null hypothesis, } H_0 : \sigma_1 = \sigma_2$$
$$\text{Alternative hypotheses, } H_A : \sigma_1 > \sigma_2$$

Test : Reject H_0 if $\dfrac{S_1^2}{S_2^2} > 1 + c_R$

We know that, $\dfrac{S_1^2}{S_2^2} \sim F(n_1 - 1, n_2 - 1)$

Therefore, $\dfrac{S_1^2}{S_2^2} \sim F(8, 11)$

$$\alpha = P(\text{Reject} H_0 | H_0)$$
$$\alpha = P(T > 1 + c_R | \sigma_1 = \sigma_2)$$
$$\alpha = 1 - P(T > 1 + c_R)$$
$$\alpha = 1 - F_{F(8,11)}(1 + c_R)$$

Now for $\alpha = 0.05$,

$$0.05 = 1 - F_{F(8,11)}(1 + c_R)$$
$$1 - 0.05 = F_{F(8,11)}(1 + c_R)$$
$$0.95 = F_{F(8,11)}(1 + c_R)$$
$$F^{-1}_{F(8,11)}(0.95) = 1 + c_R$$
$$1 + c_R = 2.95$$

Now for $\alpha = 0.01$,

$$0.01 = 1 - F_{F(8,11)}(1 + c_R)$$
$$1 - 0.01 = F_{F(8,11)}(1 + c_R)$$
$$0.99 = F_{F(8,11)}(1 + c_R)$$
$$F^{-1}_{F(8,11)}(0.99) = 1 + c_R$$
$$1 + c_R = 4.74$$

Also,

$$\frac{S_1^2}{S_2^2} = \frac{3.8}{2.9}$$
$$\frac{S_1^2}{S_2^2} = 1.31$$

Since, $\dfrac{S_1^2}{S_2^2} < 2.95$ for $\alpha = 0.05$, and $\dfrac{S_1^2}{S_2^2} < 4.74$ for $\alpha = 0.01$
Therefore, we will accept H_0.
Hence option a and c are correct options.

2. Suppose $X \sim \text{Normal}(\mu_1, 3)$ and $Y \sim \text{Normal}(\mu_2, 4)$. For $n_1 = n$ i.i.d. samples of X and $n_2 = 8$ samples of Y, the observed sample means of X and Y are 10.2 and 8.2, respectively. Let the null hypothesis be $\mu_1 = \mu_2$ against an alternative hypothesis $\mu_1 \neq \mu_2$. Consider a two sample $Z-$ test that rejects H_0 if $|\overline{X} - \overline{Y}| > c$ for some constant c at 5% level of significance. If the value of c is 1.625, then find the value of n. Round off your answer to the next greatest integer.
Hint: $F_Z^{-1}(0.025) = -1.96$

Answer : 16 ; Range : (15 - 18)

Solution :
Given,

$$X \sim N(\mu_1, 3) \text{ and } Y \sim N(\mu_2, 4)$$
$$n_1 = n \text{ and } n_2 = 8$$
$$\overline{X} = 10.2 \text{ and } \overline{Y} = 8.2$$
$$\text{Significance level, } \alpha = 0.05$$
$$\text{Critical value, } c = 1.625$$

Now,

$$\text{Null hypothesis, } H_0 : \mu_1 = \mu_2$$
$$\text{Alternative hypothesis, } H_A : \mu_1 \neq \mu_2$$

Test: Reject H_0 if $|\overline{X} - \overline{Y}| > c$

We know that, $\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \dfrac{3}{n} + \dfrac{4}{8}\right)$

$$\alpha = P(\text{Reject} H_0 | H_0)$$
$$0.05 = P(|\overline{X} - \overline{Y}| > c | \mu_1 = \mu_2)$$
$$0.05 = P\left(\frac{|\overline{X} - \overline{Y}|}{\sqrt{\frac{3}{n} + \frac{4}{8}}} > \frac{1.625}{\sqrt{\frac{3}{n} + \frac{4}{8}}}\right)$$
$$0.05 = P\left(|Z| > \frac{1.625}{\sqrt{\frac{3}{n} + \frac{4}{8}}}\right)$$
$$0.05 = 2P\left(Z < \frac{-1.625}{\sqrt{\frac{3}{n} + \frac{4}{8}}}\right)$$
$$\frac{0.05}{2} = P\left(Z < \frac{-1.625}{\sqrt{\frac{3}{n} + \frac{4}{8}}}\right)$$
$$0.025 = P\left(Z < \frac{-1.625}{\sqrt{\frac{3}{n} + \frac{4}{8}}}\right)$$

$$0.025 = F_Z\left(\frac{-1.625}{\sqrt{\frac{3}{n}+\frac{4}{8}}}\right)$$

$$F_Z^{-1}(0.025) = \frac{-1.625}{\sqrt{\frac{3}{n}+\frac{4}{8}}}$$

$$-1.96 = \frac{-1.625}{\sqrt{\frac{3}{n}+\frac{4}{8}}}$$

$$\sqrt{\frac{3}{n}+\frac{4}{8}} = \frac{-1.625}{-1.96}$$

Squaring both sides,

$$\frac{3}{n}+\frac{1}{2} = 0.687$$

$$\frac{3}{n} = 0.69 - 0.5$$

$$n = \frac{3}{0.19}$$

$$n = 16$$

Statistics for Data Science - 2
Week 12 Graded Assignment Solution

Question

Use the following values if required:

$$F_{t_{24}}(1.711) = 0.95,\ F_{t_{24}}(-1.711) = 0.05,\ F_{t_{25}}(1.71) = 0.95,\ F_Z(1.645) = 0.95$$

Setup : A battery manufacturer claims that their newly designed battery lasts, on an average of 11.5 hours on a full charge. A researcher believes that the average battery life is actually higher than what the company claims. To test the claim, the researcher selects a random sample of 25 batteries and tests them under standard conditions. The mean and standard deviations of the observed sample are 12.1 hours and 2.5 hours, respectively. Assume that battery life follows a normal distribution. Consider a test that rejects the manufacturer's claim for some constant c at 5% level of significance. Find the value of c.

Solution: Let X represents the life of a newly designed battery. The null and alternative hypotheses can be defined as

$$H_0 : \mu = 11.5$$

$$H_A : \underline{\quad \mathbf{1} \quad}$$

Under H_0, the test Statistic will be

$$\frac{\overline{X} - \mu_0}{s/\sqrt{n}} \sim \underline{\quad \mathbf{2} \quad}$$

To find the value of c, we will solve:

$$\alpha = P(\text{Reject } H_0 \mid H_0)$$

$$0.05 = P(\underline{\quad \mathbf{3} \quad} \mid H_0)$$

$$0.05 = P\left(\frac{\overline{X} - \mu_0}{s/\sqrt{n}} > \underline{\quad \mathbf{4} \quad}\right)$$

$$0.05 = 1 - \underline{\quad \mathbf{5} \quad}$$

$$c = 11.5 + \underline{\quad \mathbf{6} \quad}$$

$$c = \underline{\quad \mathbf{7} \quad}$$

Consider the following options for the blanks.

(1) t_n

(2) $\mu > 11.5$

(3) $\mu < 11.5$

(4) 0.85

(5) 0.43

(6) $\overline{X} > c$

(7) $\overline{X} < c$

(8) t_{n-1}

(9) $N\left(\mu, \frac{\sigma^2}{n}\right)$

(10) $\frac{c - 11.5}{2.5/5}$

(11) $F_{t_{24}}\left(\frac{c - 11.5}{2.5/5}\right)$

(12) $\frac{c - 11.5}{(6.25/25)}$

(13) $F_{t_{25}}\left(\frac{c - 11.5}{2.5/5}\right)$

(14) $F_Z\left(\frac{c - 11.5}{2.5}\right)$

(15) $N(0, 1)$

Solution: Hypothesis Test for Battery Life

Given:

- Manufacturer claims mean battery life: $\mu_0 = 11.5$ hours
- Researcher suspects mean is higher than claimed
- Sample size: $n = 25$
- Sample mean: $\bar{X} = 12.1$ hours
- Sample standard deviation: $s = 2.5$ hours
- Battery life is normally distributed.
- Significance level: $\alpha = 0.05$ (one-sided, right-tail)

1: Defining the hypotheses

$$H_0 : \mu = 11.5 \quad Vs. \boxed{H_A : \mu > 11.5} --\text{(option number 2)}$$

2: Specifying the test statistic
Under H_0, the test statistic is

$$T = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim \boxed{t_{n-1}} --\text{(option number 8)}$$

as we are to test for mean and the population variance is unknown.

3: Stating the rejection criterion
Reject H_0 if the sample mean $\bar{X}$ exceeds a critical value c for the given alternative hypothesis is:

$$\boxed{\bar{X} > c} --\text{(option number 6)}$$

4: Expressing the rejection region in terms of probability

$$\alpha = P(\bar{X} > c \mid H_0) = P\left(\frac{\bar{X} - \mu_0}{s/\sqrt{n}} > \boxed{\frac{c - 11.5}{2.5/5}}\right) --\text{(option number 10)}$$

Since $n = 25$,

$$s/\sqrt{n} = 2.5/\sqrt{25} = 2.5/5 = 0.5$$

5: The critical value using the t-distribution
We want the upper 5% tail, so set the cumulative probability at 0.95:

$$0.05 = 1 - \boxed{F_{t_{24}}\left(\frac{c - 11.5}{2.5/5}\right)} - -(\text{option number } 11\ )$$

$$F_{t_{24}}\left(\frac{c - 11.5}{2.5/5}\right) = 0.95$$

Where $t_{24,0.95} = 1.711$ (given in question).

Solving for the critical value c

$$\frac{c - 11.5}{0.5} = 1.711$$

$$c - 11.5 = 1.711 \times 0.5$$

$$\boxed{c = 11.5 + 0.8555} - -(\text{option number } 4\ )$$

$$c = 12.3555$$

Hence
Round to two decimal places:

$$\boxed{c = 12.36}$$

The claim is rejected if the sample mean is greater than 12.36.