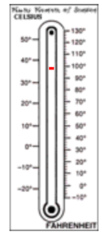
Introduction to
Hypothesis Testing

HT - 1

Is the average body temperature of healthy adults 98.6°F?



HT - 2

Scientific Method

1. State research hypotheses or questions.
2. Gather data or evidence (observational or experimental) to answer the question.
3. Summarize data and test the hypothesis.
4. Draw a conclusion.

HT - 3

Methods of Testing Hypotheses

1. *p*-value Method
2. Critical Value Method (Classical)
3. Confidence Interval Method (Approximate)

HT - 4

Steps in Hypothesis Testing

1. State statistical hypotheses: H_0 and H_a.
2. Choose a proper **test statistic**, collect data, checking the assumption and compute the value of the statistic.
3. Make decision rule based on **level of significance(α)**.
4. Draw conclusion. (Reject null hypothesis or not)

HT - 5

Statistical Hypothesis

Null hypothesis **(H_0)**:

Hypothesis of no difference or no relation, often has =, ≥, or ≤ notation when testing value of parameters.

Example: H_0: μ = 98.6°F
(average body temperature is 98.6)

HT - 6

Statistical Hypothesis

***Alternative hypothesis* (H_a)**: (or H_1)

Usually corresponds to research hypothesis and opposite to null hypothesis,

often has >, < or ≠ notation

Example: H_a: $\mu \neq 98.6$°F
(average body temperature is not 98.6°F)

HT - 7

Logic Behind Hypothesis Testing

In testing statistical hypothesis,

the null hypothesis is first assumed to be true.

We collect evidence to see if the evidence is strong enough to reject the null hypothesis and support the alternative hypothesis.

HT - 8

Evidence

***Test Statistic* (Evidence)**: A sample statistic used to decide whether to reject the null hypothesis.

HT - 9

Hypotheses Statements Example

- A researcher is interested in finding out whether average life time of male is higher than 77 years.

H_0: $\mu = 77$ (or $\mu \leq 77$)
H_a: $\mu > 77$
[Right-tailed test]

HT - 10

Hypotheses Statements Example

- A researcher is interested in finding out whether the average regular gasoline price is less than $1.45 in Mid-West region.

H_0: $\mu = 1.45$ (or $\mu \geq 1.45$)
H_a: $\mu < 1.45$
[Left-tailed test]

HT - 11

Hypotheses Statements Example

- A researcher is interested in finding out whether average hourly salary for baby sitting is different from $6.00.

H_0: $\mu = 6$
H_a: $\mu \neq 6$
[Two-tailed test]

HT - 12

One Sample Z-Test for Mean
(For large sample situations)

A Two-sided Test Examples

HT - 13

I. Hypothesis Testing

One wishes to test whether the average body temperature for healthy adults is different from 98.6°F.

H_o: μ = 98.6°F v.s. H_a: $\mu \neq$ 98.6°F

(98.6°F → μ_0)

HT - 14

Evidence

What will be the key statistic (evidence) to use for testing the hypothesis about population mean?

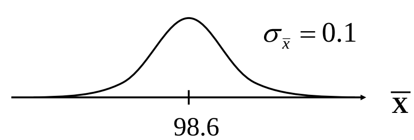
Sample mean: $\overline{X}$

A random sample of 36 subjects is chosen and the sample mean is **98.32°F** and sample standard deviation is **0.6°F**.

HT - 15

Sampling Distribution

If **H_0: μ = 98.6°F is true**, sampling distribution of mean will be approximately normally distributed with *mean* **98.6** and *standard deviation* (or standard error) $\frac{.6}{\sqrt{36}}$ = **0.1**.

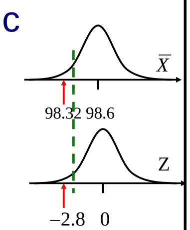


HT - 16

II. Test Statistic

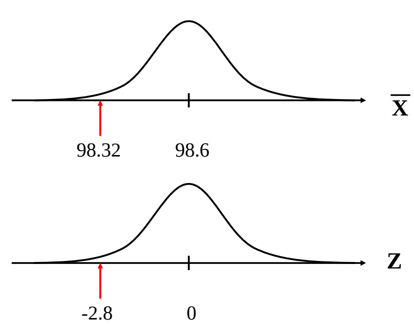
$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \quad \text{(Standardized } \bar{x} \text{ using } H_0\text{)}$$

$$= \frac{98.32 - 98.6}{\frac{0.6}{\sqrt{36}}} = \frac{-0.28}{0.1} = -2.8$$



This implies that the statistic is 2.8 standard deviations away from the mean 98.6 in H_0, and is to the left of 98.6 (or less than 98.6)

HT - 17



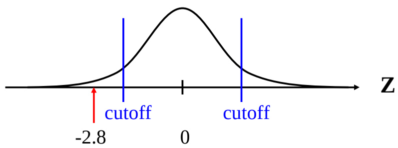
More than two standard error from the mean.

HT - 18

III. Decision Rule

Is "2.8 standard deviations away from the mean 98.6 in H_0" an extreme enough evidence to convince us that the average body temperature is different from 98.6?

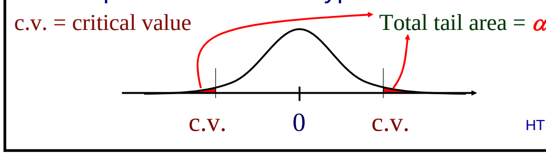
Need a cutoff for determination!



Level of Significance

***Level of significance* for the test (α)**

A probability level selected by the researcher at the beginning of the analysis that defines unlikely values of sample statistic if null hypothesis is true.



III. Decision Rule

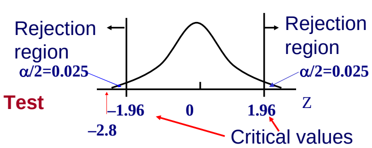
Critical value approach: Compare the test statistic with the critical values defined by significance level α, usually α = 0.05.

We reject the null hypothesis, if the test statistic

$z < -z_{\alpha/2} = -z_{0.025} = -1.96$, or $z > z_{\alpha/2} = z_{0.025} = 1.96$.

(i.e., reject H_0 if $|z| > z_{\alpha/2}$, that is if $|z| > 1.96$)

Two-sided Test



IV. Draw conclusion

Since $z = -2.8 < -z_{\alpha/2} = -1.96$ therefore we reject null hypothesis.

Therefore we conclude that **there is sufficient evidence to support the alternative hypothesis that the average body temperature is different from 98.6ºF.**

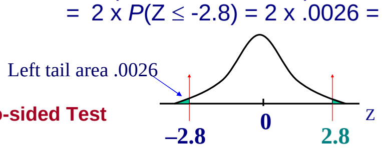
A Different Approach

III. Decision Rule

***p*-value approach:** Compare the probability of the evidence or more extreme evidence to occur when null hypothesis is true. If this probability is less than the level of significance of the test, α, then we reject the null hypothesis. (z-Table)

p-value = $P(Z \leq -2.8 \text{ or } Z \geq 2.8)$
= $2 \times P(Z \leq -2.8) = 2 \times .0026 =$ **.0052**

Two-sided Test



p-value

♥ ***p-value*** ♥ **(most popular approach)**

The probability of obtaining a test statistic that is as extreme or more extreme than actual sample statistic value observed given null hypothesis is true.

The smaller the p-value, the stronger the evidence for supporting H_a and rejecting H_0.

IV. Draw conclusion

Since p-value = .0052 < α = .05 , we reject null hypothesis.

Therefore we conclude that **there is sufficient evidence to support the alternative hypothesis that the average body temperature is different from 98.6ºF.**

HT - 25

One Sample Z-Test for Mean (For large sample situations)

An One-sided Test Examples

HT - 26

I. Hypothesis Testing

One wishes to test whether the average body temperature for healthy adults is less than 98.6°F.

H_o: μ = 98.6°F v.s. H_a: μ < 98.6°F

This is a one-sided test, **Left-sided Test**.

HT - 27

II. Test Statistic

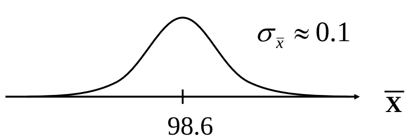
A random sample of 36 is chosen and the sample mean is **98.32°F,** with a sample **standard deviation, s, of 0.6°F**.

Assumption: Assume body temperature for healthy adults under regular environment has a **normal distribution**.

HT - 28

II. Test Statistic

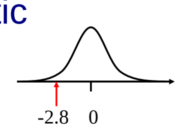
If **H_0 is true**, sampling distribution of mean will be normally distributed with *mean* **98.6** and *standard deviation* (or standard error) 0.6/6 = **0.1**. ("6" is square root of 36.)



HT - 29

II. Test Statistic

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{98.32 - 98.6}{\frac{0.6}{\sqrt{36}}} = \frac{-0.28}{0.1} = -2.8$$



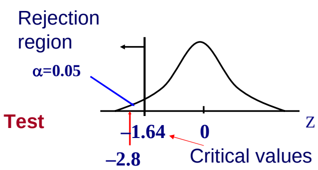
This implies that the statistic is 2.8 standard deviations away from the mean 98.6 in H_0 , and is to the left of 98.6 (or less than 98.6)

HT - 30

III. Decision Rule

Critical value approach: Compare the test statistic with the critical values defined by significance level α, usually $\alpha = 0.05$.

We reject the null hypothesis, if the test statistic
$z < -z_{\alpha} = -z_{0.05} = -1.64$.



Left-sided Test

HT - 31

IV. Draw conclusion

Since $z = -2.8 < -z_{\alpha} = -1.64$ we reject null hypothesis.

Therefore we conclude that **there is sufficient evidence to support the alternative hypothesis that the average body temperature is less than 98.6°F.**

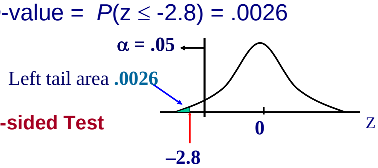
HT - 32

A Different Approach

III. Decision Rule

***p*-value approach:** Compare the probability of the evidence or more extreme evidence to occur when null hypothesis is true. If this probability is less than the level of significance of the test, α, then we reject the null hypothesis.

p-value = $P(z \leq -2.8) = .0026$



Left-sided Test

HT - 33

IV. Draw conclusion

Since *p*-value = .0026 < α = .05 , we reject null hypothesis.

Therefore we conclude that **there is sufficient evidence to support the alternative hypothesis that the average body temperature is different from 98.6ºF.**

HT - 34

Decision Rule

Critical value approach: Determine critical value(s) using α, **reject H_0** against

i) $H_a: \mu \neq \mu_0$, if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$
(or $|z| > z_{\alpha/2}$)

ii) $H_a: \mu > \mu_0$, if $z > z_{\alpha}$

iii) $H_a: \mu < \mu_0$, if $z < -z_{\alpha}$

HT - 35

Decision Rule

***p*-value approach**: Compute *p*-value,

if $H_a: \mu \neq \mu_0$, *p*-value = $2 \cdot P(Z \geq |z|)$
if $H_a: \mu > \mu_0$, *p*-value = $P(Z \geq z)$
if $H_a: \mu < \mu_0$, *p*-value = $P(Z \leq z)$

reject H_0 if *p*-value < α

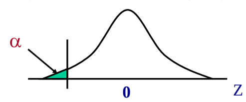
HT - 36

Errors in Hypothesis Testing

Possible statistical errors:

- ***Type I error***: The null hypothesis is true, but we reject it.
- ***Type II error***: The null hypothesis is false, but we don't reject it.

"α" is the probability of committing Type I Error.



HT - 37

Can we see data and then make hypothesis?

Data Snooping!

1. Choose a **test statistic**, co… checking the … te the …
2. St… H_0 and H_a.
3. M… decision rule based on **level of significance(α)**.
4. Draw conclusion. (Reject null hypothesis or not)

HT - 38

Is average cash carried in MATH 2625 students' pocket less than $10.00?

① Hypothesis: H_0 : ________ H_a: ________

②Test statistic: z = ?

Sample size: 36

Sample mean: $8.85

Sample standard deviation: $1.21

③ Decision rule:

④ Conclusion:

HT - 39

One Sample t-Test for Mean

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

HT - 40

One-sample Test *with* Unknown Variance σ^2

In practice, population variance is unknown most of the time. The sample standard deviation **s^2 is used instead for σ^2**. If the random sample of size *n* is from a normal distributed population and if the null hypothesis is true, the test statistic (standardized sample mean) will have a t-distribution with degrees of freedom *n*–1.

Test Statistic : $t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

HT - 41

I. State Hypothesis

One-side test example:

If one wish to test whether the body temperature is less than 98.6 or not.

H_0: μ = 98.6 v.s. H_a: μ < 98.6

(Left-sided Test)

HT - 42

II. Test Statistic

If we have a random sample of size 16 from a normal population that has a mean of 98.32°F, and a sample standard deviation 0.4. The test statistic will be a *t*-test statistic and the value will be: (standardized score of sample mean)

Test Statistic : $t = \dfrac{\bar{x} - \mu_0}{\dfrac{s}{\sqrt{n}}} = \dfrac{98.32 - 98.6}{\dfrac{0.4}{\sqrt{16}}} = \dfrac{-0.28}{0.1} = -2.8$

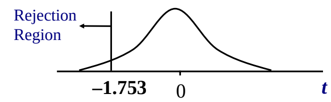
Under null hypothesis, this *t*-statistic has a *t*-distribution with degrees of freedom n – 1, that is, 15 = 16 – 1.

HT - 43

III. Decision Rule

Critical Value Approach:

To test the hypothesis at α level 0.05, df = 15, the **critical value** is $-t_{\alpha} = -t_{0.05} = -1.753$.

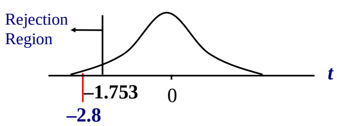


Desicion Rule: Reject null hypothesis if $t < -1.753$.

(t-Table)

HT - 44

IV. Conclusion



Decision Rule:

If $t < -1.753$, we reject the null hypothesis.

Conclusion: Since $t = -2.8 < -1.753$, we reject the null hypothesis. There is sufficient evidence to support the research hypothesis that the average body temperature is less than 98.6°F.

HT - 45

A Different Approach

III. Decision Rule

Decision Rule: Reject null hypothesis if *p*-value < α.

HT - 46

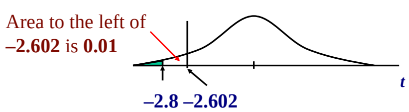
p-value Calculation

***p*-value corresponding the test statistic:**

For *t* test, unless computer program is used, *p*-value can only be approximated with a range because of the limitation of *t*-table.

***p*-value =** $P(T<-2.8)$ **<** $P(T<-2.602)$ = **0.01**

Since the area to the left of –2.602 is .01, the area to the left of –2.8 is definitely less than 0.01.



HT - 47

t-Table

Area in Upper Tail				
df	**0.10**	**0.05**	**0.025**	**0.01**
.	.	.	.	.
14	.	.	.	.
15	1.341	1.753	2.131	2.602
16	.	.	.	.
.	.	.	.	.

HT - 48

IV. Conclusion

Decision Rule:

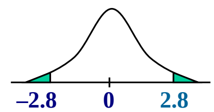
If *p*-value < 0.05, we reject the null hypothesis.

Conclusion: Since *p*-value < 0.01 < 0.05, we reject the null hypothesis. There is sufficient evidence to support the research hypothesis that the average body temperature is less than 98.6°F.

HT - 49

What if we wish to test whether the average body temperature is different from 98.6°F or not using *t*-test with the same data?

The *p*-value is equal to **twice** the *p*-value of the left-sided test which will be less than .02.



HT - 50

Statistical Significance

A statistical report shows that the average blood pressure for women in certain population is significantly *different from* a recommended level, with a *p*-value of 0.002 and the *t*-statistic of – 6.2. It generally means that the difference between the actual average and the recommended level is *statistically* significant from a two-sided test.

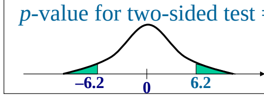
Can we use this results to check on possible one-sided test results?

HT - 51

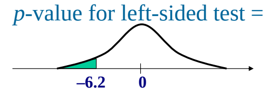
Statistical Report

$t = -6.2$

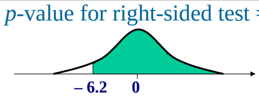
p-value for two-sided test = **.002**



p-value for left-sided test = **.001**



p-value for right-sided test = **.999**



HT - 52

Decision Rule

***p*-value approach**: Compute *p*-value,

if $H_a: \mu \neq \mu_0$, *p*-value = $2 \cdot P(T \geq |t|)$

if $H_a: \mu > \mu_0$, *p*-value = $P(T \geq t)$

if $H_a: \mu < \mu_0$, *p*-value = $P(T \leq t)$

reject H_0 if *p*-value < α

HT - 53

Decision Rule

Critical value approach: Determine critical value(s) using α, **reject H_0** against

i) $H_a: \mu \neq \mu_0$, if $t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$

(or $|t| > t_{\alpha/2}$)

ii) $H_a: \mu > \mu_0$, if $t > t_\alpha$

iii) $H_a: \mu < \mu_0$, if $t < -t_\alpha$

HT - 54

Remarks

- If the sample size is **relatively large** (>30) both z and *t* tests can be used for testing hypothesis. The number 30 is just a reference for general situations and for practicing problems. *In fact, if the sample is from a very skewed distribution, we need to increase the sample size or use nonparametric alternatives such Sign Test or Signed-Rank Test.*
- Many commercial packages only provide *t*-test since standard deviation of the population is often unknown.

HT - 55

Example

A random sample of **ten** 400-gram soil specimens were sampled in location A and analyzed for certain contaminant. The sample data are the followings:

65, 54, 66, 70, 72, 68, 64, 50, 81, 49

The contaminant levels are normally distributed. Test the hypothesis, at the level of significance **0.05**, that the true mean contaminant level in this location **exceeds 50 mg/kg**.

HT - 56

Step 1

What is the hypothesis to be tested?

Ho: $\mu = 50$

H_a: $\mu > 50$

HT - 57

Step 2

Which test can be used for testing the hypothesis above? (Check assumptions.)

One sample t-test. Why? Because the random sample was from a **normal population** and unknown variance.

Compute Test Statistic:

Test statistic: $t = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \dfrac{63.9 - 50}{\frac{10.17}{\sqrt{10}}} = 4.32$

The value of the test statistic is 4.32 with a p-value between .005 and .0005 from table.

P-value from SPSS is .00096.

HT - 58

Step 3

Decision Rule:

Specify a level of significance, α, for the test. $\alpha = .05$

Critical value approach:

Reject Ho if $t > t_{.05} = 1.833$

p-value approach:

Reject Ho if p-value < 0.05

HT - 59

Step 4

Conclusion:

Since t=4.32 > 1.833,

(or p-value = .00096 < 0.05)

we reject the null hypothesis. The data provide sufficient evidence to support the alternative hypothesis that the average contaminant level in this location exceeds 50 mg/kg.

HT - 60

<u>Question:</u> <u>Is the actual population average cholesterol level likely to be</u> **<u>higher than 211?</u>**

Evidence: A **random sample of 16** individuals from the population has a sample mean of 225 and sample standard deviation of 130.

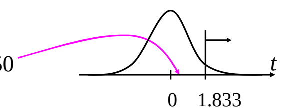
What is your conclusion?

HT - 61

Average Weight for Female Ten Years Old Children In US

Info. from a random sample: $n = 10$, $\bar{x} = 80$ lb, $s = 18.05$ lb. Is average weight greater than 78 lb at $\alpha = 0.05$ level?

Test Statistic: $t = \dfrac{80 - 78}{\dfrac{18.05}{\sqrt{10}}} = 0.350$



$t_\alpha = t_{.05}$, d.f. = 10 – 1 = 9, $t_{0.5,\,9} = 1.833$

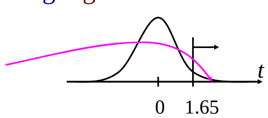
Reject H_0, if $t = 0.35 < 1.833$. **Failed to reject H_0!**

HT - 62

Average Weight for Female Ten Years Old Children In US

Info. from a random sample: $n =$ **400**, $\bar{x} = 80$ lb, $s = 18.05$ lb. Is average weight greater than 78 lb at $\alpha = 0.05$ level?

Test Statistic: $t = \dfrac{80 - 78}{\dfrac{18.05}{\sqrt{400}}} = 2.22$



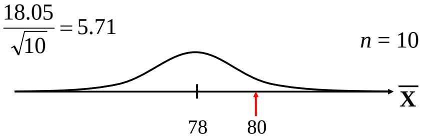
$t_\alpha = t_{.05}$, d.f. = 400 – 1 = 399, $t_{0.5,\,399} = 1.65$

Reject H_0, if $t = 2.22 > 1.65$. **Reject H_0!**

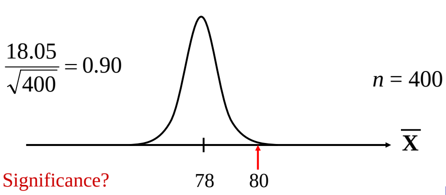
HT - 63

Sampling Distribution

S.E. $= \dfrac{18.05}{\sqrt{10}} = 5.71$



S.E. $= \dfrac{18.05}{\sqrt{400}} = 0.90$



Practical Significance?

HT - 64

Sampling Distribution of Sample Proportion

Sampling distribution of sample proportion: A random sample of size n from a large population with proportion of successes (usually represented by a value 1) p, and therefore proportion of failures (usually represented by a value 0) $1 - p$, <u>the sampling distribution of **sample proportion**</u>, $\hat{p} = x/n$, where x is the number of successes in the sample, is approximately normal with a **mean** p

and **standard deviation** $\sqrt{\dfrac{p(1-p)}{n}}$.

HT - 65

Confidence Interval

Confidence interval: The $(1-\alpha)\%$ confidence interval estimate for population proportion is

$$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

HT - 66

One-Sample *z*-test for a population proportion

z-test:

Step 1: **State Hypotheses (choose one of the three hypotheses below)**

i) $H_0: p = p_0$ v.s. $H_a: p \neq p_0$ (Two-sided test)
ii) $H_0: p = p_0$ v.s. $H_a: p > p_0$ (Right-sided test)
iii) $H_0: p = p_0$ v.s. $H_a: p < p_0$ (Left-sided test)

HT - 67

Step 2: **Compute *z* test statistic**:

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

HT - 68

Step 3. **Decision Rule**:

***p*-value approach**: Compute *p*-value,

if $H_a: p \neq p_0$, p-value $= 2 \cdot P(Z \geq |z|)$
if $H_a: p > p_0$, p-value $= P(Z \geq z)$
if $H_a: p < p_0$, p-value $= P(Z \leq z)$
reject H_0 if p-value $< \alpha$

Critical value approach: Determine critical value(s) using α, reject H_0 against

i) $H_a: p \neq p_0$, if $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$
ii) $H_a: p > p_0$, if $z > z_{\alpha}$
iii) $H_a: p < p_0$, if $z < -z_{\alpha}$

Step 4: **Draw Conclusion.**

HT - 69

Example: A researcher hypothesized that the percentage of the people living in a community who has no insurance coverage during the past 12 months is **not 10%**. In his study, 1000 individuals from the community were randomly surveyed and checked whether they were covered by any health insurance during the 12 months. Among them, 122 answered that they did not have any health insurance coverage during the last 12 months. Test the researcher's hypothesis at the level of significance of 0.05.

HT - 70

➡ **Hypothesis:** $H_0: p = .10$ v.s. $H_a: p \neq .10$ (Two-sided test)

➡ **Test Statistic:** $z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{.122 - .10}{\sqrt{\frac{.10(1-.10)}{1000}}} = 2.32$

p-value = 2 x .01 = .02

➡ **Decision Rule:** Reject null hypothesis if *p*-value < .05.

➡ **Conclusion:** There is sufficient evidence to support the alternative hypothesis that the percentage is different from 10%.

HT - 71