STATISTICS 7 TUTORIAL (SOLUTIONS)
Hypothesis Testing

H2 Maths

Topic:	Hypothesis testing	***Purpose of LE:*** To enhance understanding of the concept of null and alternative hypothesis and the asymmetry of these hypotheses by relating it to a concrete analogy which is easier to appreciate.
LE:	• Relate the notion of null and alternative hypothesis to the statement "innocent until proven guilty"	
Remarks:	• Teachers can pose the question during lecture or tutorial to bring out the above idea, by asking students to compare the null and alternative hypothesis against the Singapore legal system "innocent until proven guilty". • Amount of time required: 5 – 10 minutes	

Details of Activity

Present the reading to the students.

In pairs, discuss the concept of null and alternative hypothesis that you have just learned with the idea of "innocent until proven guilty" used in the Singapore legal system.

> <u>The Singapore Legal System</u>
> In Singapore, "The presumption of innocence is widely recognised as the cornerstone of the criminal justice system. Under the presumption, the Prosecution has a "duty" to prove the accused's guilt, "subject to" the insanity defence and statutory exceptions. Hence, if there remains a "reasonable doubt … on the whole of the case", the prosecution has not established its case and the accused is acquitted."
>
> (http://www.eastasiareview.org/issues/2013/articles/2013_Low.pdf) (Available as of 11 Jul 2016)

Notes for teachers

The hypothesis test works like a criminal trial:

defendants are presumed innocent until proven guilty.

Imagine the following scenario:
A suspect was put on trial for a criminal offence. The Prosecution thought that he was guilty.
H_0: Suspect was innocent
H_1: Suspect was guilty

The Prosecution would try to show the judge, beyond reasonable doubt, that the suspect was guilty (in other words, there was enough evidence to "reject the null hypothesis"). If the Prosecution did not do his job well, then he would not have enough evidence to convict the suspect. Although the judge might still think that the suspect was guilty, but the judge could not legally prove it. In this case, with the given evidence, the jury would "fail to reject the null hypothesis".

Basic Mastery Question

1. The solution to the following problem is provided below the question. Highlight the errors in the solution and amend these errors.

   A manufacturer claims that the life span of an electrical component it produces is normally distributed with mean 22 hours and standard deviation 6 hours. A random sample of 50 such electrical components has a mean life span of 20.2 hours. Test, at the 1% significance level, whether the mean life span of the electrical components has changed.

   Solution:

   To test H_0 : $\mu_0 = 22$ against H_1 : $\mu_0 \neq 22$

   Under H_0, $Z = \dfrac{\overline{x} - 22}{6/\sqrt{50}} \sim N(0,1)$

   2-tail test at 1% level of significance.

   Using GC, $\text{p-value} = 0.0339$, we do not reject H_0.

   There is insufficient evidence that $\mu_0 \neq 22$.

Ans: Test $H_0 : \mu = 22$ against $H_1 : \mu \neq 22$

Under H_0, $Z = \dfrac{\overline{X} - 22}{6/\sqrt{50}} \sim N(0,1)$

2-tail test at 1% level of significance

From GC, p-value = $0.0339 > 0.01$, so we do not reject H_0

There is insufficient evidence, at the 1 % significance level, that the mean life span has changed.

2. A supermarket manager investigated the lengths of time that customers spent shopping in the store. The time, x minutes, spent by each of a random sample of 150 customers was measured, and it is found that $\sum x = 2871$ and $\sum x^2 = 60\,029$. Test, at the 5% level of significance, the hypothesis that the mean time spent shopping by customers is 20 minutes, against the alternative that it is less than this.

What is the purpose of the question? Can you express the problem in mathematical terms?	Let μ minutes be the <u>mean</u> length of time that customers spent shopping in the store. Test $H_0 : \mu = 20$ against $H_1 : \mu < 20$ One-tail test at 5 % level of significance.
What are the information given? What additional information do you need?	From the sample data, $n = 150$ $\sum x = 2871$; $\sum x^2 = 60029$ $\bar{x} = \frac{\sum x}{n} = \frac{2871}{150} = 19.14$ $s^2 = \frac{1}{n-1}[\sum x^2 - \frac{(\sum x)^2}{n}] = \frac{1}{149}[60029 - \frac{(2871)^2}{150}] = 34.081$
What are the mathematical concepts used?	Test statistic: Under H_0, $Z = \frac{\bar{X} - 20}{s/\sqrt{150}} \sim N(0,1)$ approximately since n is large
What is the interpretations of the calculated test statistic and p-value?	**<u>Method 1:</u>** p – value = 0.0355 < 0.05 We reject H_0. **<u>Method 2:</u>** Critical region to reject H_0: $z < -1.645$ $z_{\text{calculated}} = \frac{19.14 - 20}{\sqrt{\frac{34.081}{150}}} = -1.804$ Since $z_{\text{calculated}}$ lies within critical region, we reject H_0. There is sufficient evidence at 5 % level of significance that the **mean** time spent shopping by customers is less than 20 minutes.
What are the assumptions? Are the assumptions valid?	There is no need to assume that X is normally distributed as $n = 150$ is large and hence $\bar{X}$ follows a normal distribution approximately.

3. The principal of a private college claims that graduates from his school have an average starting salary of \$2050. A private body checks the claim by interviewing a random sample of 60 graduates from the school.

The data obtained is summarized below, where x denotes the monthly salary per person.

$$\sum(x-2000)=2740 \text{ and } \sum(x-2000)^2=162001$$

Carry out an appropriate test at the 3% significance level to determine whether the principal is overestimating his claim. Do we need to assume that X follows a normal distribution?

Solution:

$$\bar{x}=\frac{\sum(x-2000)}{60}+2000=\frac{2740}{60}+2000=2045.66667$$

$$s^2 \quad =\frac{1}{n-1}\left[\sum(x-2000)^2-\frac{\left(\sum(x-2000)\right)^2}{60}\right]$$

$$=\frac{1}{59}\left[162001-\frac{2740^2}{60}\right]=624.98870$$

Let μ be mean starting salary of graduates from the college.

Test $\mathrm{H_o}: \mu=2050$ against $\mathrm{H_1}: \mu<2050$ at 3% significance level

Test statistic:

Under $\mathrm{H_0}$, since $n = 60$ is large, $\overline{\mathrm{X}} \sim \mathrm{N}\left(2050, \frac{62498870^2}{60}\right)$ approximately

$$\Rightarrow Z=\frac{\bar{X}-2050}{s/\sqrt{60}} \sim \mathrm{N}(0,1) \text{ approximately}$$

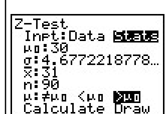
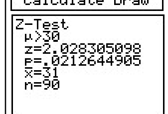
Using GC, p-value=0.089693 > 0.03, we do not reject $\mathrm{H_o}$.

There is insufficient evidence at 3% significance level that the principal is overestimating his claim.

There is no need to assume that X follows a normal distribution as $n = 60$ is large, therefore $\bar{X}$ follows a normal distribution approximately.

	Or $s^2 = \dfrac{10911}{6320}$
9(vi)	Since s^2 from Germsfree is smaller, hence the test statistic value from the test for Germsfree will be more negative. Therefore $p_1 > p_2$.

Additional Practice Questions

1(i) $\bar{x} = \dfrac{90}{90} + 30 = 31$ $s^2 = \dfrac{1}{89}\left(2037 - \dfrac{(90)^2}{90}\right) = \dfrac{1947}{89}$ or 21.9 (to 3 sig figs)	
Unbiased estimate means the mean of the estimate is expected to equal to the parameter when collected infinitely(many times). E(θ) = p where θ is estimate and p is parameter as well as having the minimum variance as compared to the other estimators.	
(ii) Test $H_0 : \mu = 30$ vs $H_1 : \mu > 30$ and perform a upper-tailed z-test at 5% significance level. Under H_0, $\overline{X} \sim (30, \dfrac{\sigma^2}{90})$. As σ^2 is unknown, estimate with $s^2 = \dfrac{1947}{89}$ The test Statistic, $Z = \dfrac{\overline{X} - \mu}{s/\sqrt{n}} \sim$ N(0,1) approx.. Using GC, p-value = 0.0212644905, Since p-value = 0.0213 < 0.05, we reject H_0 and conclude that there is sufficient evidence at 5% significance level to say that that the bus company is underestimating the mean waiting time.	 
(iii) It is not necessary for the waiting times to follow a normal distribution, as $n = 90$ is large, Central Limit Theorem states that the sampling distribution of $\overline{X}$ is approximately normal.	

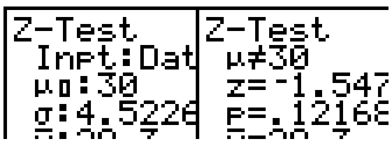
(iv) Test $H_0 : \mu = 30$ against $H_1 : \mu \neq 30$.
Perform a 2-tailed Z-test at α % significance level.

Under H_0, $\overline{X} \sim (30, \frac{\sigma^2}{100})$.

As σ^2 is unknown, estimate with $s^2 = \frac{100}{99}(4.5^2) = 20.4545454545$

The test statistic $Z = \frac{\overline{X} - \mu}{s/\sqrt{n}}$ ~N(0,1) (approximately)

$\overline{x} = 29.3$, $s^2 = 20.4545454545$
Using GC, $z = -1.548$ and p-value = 0.12168805432.



To reject H_0, p-value $< \frac{\alpha}{100}$. $\therefore$ 12.1 < α (3 sf)

2 (a)

$$\overline{x} = \frac{\sum(x-9000)}{100} + 9000$$
$$= \frac{1810}{100} + 9000$$
$$= 9018.1$$

$$s^2 = \frac{1}{99}\left(\sum(x-9000)^2 - \frac{\left(\sum(x-9000)\right)^2}{100}\right)$$
$$= \frac{1}{99}\left(737145 - \frac{1810^2}{100}\right)$$
$$= 7114.9899$$

Test $\mathrm{H_0} : \mu = 9000$ v.s $\mathrm{H_1} : \mu > 9000$
Level of significance : 1%
p-value = 0.015944 , z-value = 2.1458 , Critical region: $z > 2.3263$

Since p-value is bigger than level of significance, there is insufficient evidence at 1% level of significance to reject $\mathrm{H_0}$. We can not conclude that the mean life span of the components has increased after the change in the manufacturing procedure.

(b) $\mathrm{H_0} : \mu = 9000$ v.s $\mathrm{H_1} : \mu > 9000$
Level of significance : 5%

z-value $= \frac{\overline{x} - 9000}{\sqrt{25^2/100}}$

Critical region: $z > 1.6449$

	Since H_0 is rejected, z-value is inside the critical region, $$\frac{\bar{x}-9000}{\sqrt{25^2/100}} > 1.6449$$ $$\bar{x}-9000 > 4.11225$$ $$\bar{x} > 9004.11225$$

3 (i)

$$n=50,\ \bar{x}=\frac{\sum(x-70)}{n}+70=\frac{200.9}{50}+70=74.018=74.0$$

$$s^2=\frac{1}{n-1}\left[\sum(x-70)^2-\frac{\left(\sum(x-70)\right)^2}{n}\right]$$

$$=\frac{1}{49}\left[867.31-\frac{200.9^2}{50}\right] = 1.2264=1.23$$

(ii) $\mathrm{H}_0: \mu = 74.5$ against $\mathrm{H}_1: \mu < 74.5$ at the 4 % significance level

Test statistic, $Z=\dfrac{\bar{X}-74.5}{\sqrt{\dfrac{1.2262}{50}}}$

Using z-test, p-value = 0.00104

Since p-value = 0.00104< 0.04, we reject H_0 at 4% level of significance and conclude that there is sufficient evidence to claim that the manufacturer is overstating his claim.

Next,

$\mathrm{H}_0: \mu = \mu_o$ against $\mathrm{H}_1: \mu \neq \mu_o$ at the 4 % significance level

H_0 is rejected if

$$\mathrm{P}\left(|Z| > \frac{74.018-\mu_o}{\sqrt{\dfrac{1.2262}{50}}}\right) < 0.04 \quad \text{or} \quad \mathrm{P}\left(|\bar{X}| > 74.018\right) < 0.04$$

Reject H_0 if $|Z| \geq 2.05375$. [invNorm(0.98)= 2.053749]

$$\left|\frac{74.018-\mu_0}{\sqrt{\dfrac{1.2262}{50}}}\right| \geq 2.05375$$

$$\Rightarrow \frac{74.018-\mu_0}{\sqrt{\frac{1.2262}{50}}} \le -2.05375 \quad \text{or} \quad \frac{74.018-\mu_0}{\sqrt{\frac{1.2262}{50}}} \ge 2.05375$$

$$\Rightarrow \mu_0 \ge 74.018 + 2.05375\sqrt{\frac{1.2262}{50}} \quad \text{or} \quad \mu_0 \le 74.018 - 2.05375\sqrt{\frac{1.2262}{50}}$$

i.e. $\mu_0 \ge 74.3$ or $\mu_0 \le 73.7$

4(i)	The unbiased estimate of population mean $= \frac{\sum(x-50)}{70} + 50$ $= \frac{38990}{70} + 50$ $= 607$ The unbiased estimate of population variance $= \frac{1}{n-1}\sum(x-\overline{x})^2$ $= \frac{1}{70-1}\sum(x-\overline{x})^2$ $= \frac{1}{69}(50\ 172)$ $= 727.1304$ $= 727$ (to 3 sig fig)
(ii)	To test Ho : $\mu = 600$ Against H1 : $\mu \ne 600$ using 1 tailed test at 4% level of significance Under Ho, $\overline{X} \sim \text{N}(600, \frac{727.1304}{70})$ approximately $\mu_0 = 600$, $\sigma = \sqrt{727.1304}$, $\overline{x} = 607$, n = 70 Using GC, p – value = 0.02986 = 0.0299 (to 3 sig fig) Since p – value < 0.04, we reject Ho and there is sufficient evidence to conclude that the population mean mass of the contents differs from 600 g at 4% level of significance.
(iii)	It is not necessary to assume the mass of the contents of oatmeal cereal follows a normal distribution. Since the sample size = 70 is large, the distribution of the sample mean mass of the contents is approximately to have a normal distribution by Central Limit Theorem.
(iv)	Assume that significance level is α %

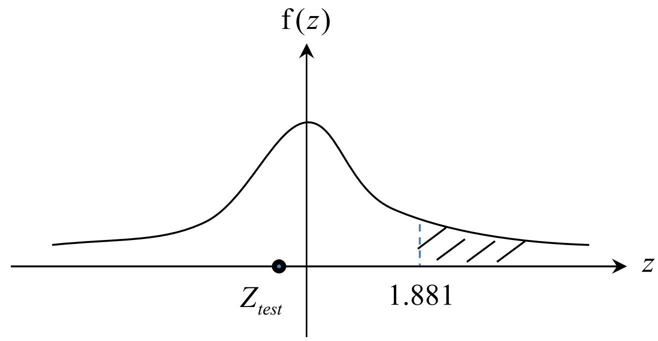
	For rejection of the null hypothesis, $p - \text{value} \le \frac{\alpha}{100}$ $0.02986 \le \frac{\alpha}{100}$ $\alpha \ge 0.02986 \times 100$ $\alpha \ge 2.986$ Least level of significance = 2.99 %

To test Ho : $\mu = \mu_0$

Against H_1 : $\mu > \mu_0$

using 1 tailed test at 3% level of significance

Under H_o, $\overline{X} \sim N(\mu_0, \dfrac{727.1304}{70})$ approximately



Since null hypothesis is not rejected, Z_{test} must not lie in the critical region

$$Z_{test} < 1.881$$

$$\frac{607-\mu_0}{\sqrt{\frac{727.1304}{70}}} < 1.881$$

$$607-\mu_0 < 1.881\sqrt{\frac{727.1304}{70}}$$

$$607-1.881\sqrt{\frac{727.1304}{70}} < \mu_0$$

$$\mu_0 > 600.94$$

$\mu_0 > 601$ (to 3 sig fig)

5.(i) Let $w = x - 25$

Then $\sum w = -27.2$, $\sum w^2 = 85.1$

$$\overline{w} = \frac{\sum w}{n} = \frac{-27.2}{80}$$

$$s_w^2 = \frac{1}{n-1}\left[\sum w^2 - \frac{\left(\sum w\right)^2}{n}\right]$$

$$= \frac{1}{79}\left[85.1 - \frac{(-27.2)^2}{80}\right] = 0.96015$$

Now $x = w + 25$

Let μ kg and σ kg be the respective mean mass and standard deviation of contents in a bag.

Unbiased estimate of the population mean, μ, is

$$\overline{x} = \overline{w} + 25 = \frac{-27.2}{80} + 25 = 24.66$$

Unbiased estimate of the population variance, σ^2, is given by

$$s^2 = s_w^2 = 0.96015 \approx 0.960$$

5 (ii)

$H_0: \ \mu = 25$

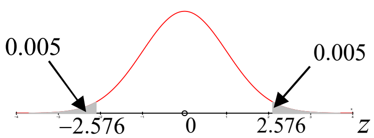
$H_1: \ \mu \neq 25$

Level of significance: 1%

Test Statistic: Since n is large, s^2 is a good estimate of σ^2 and CLT is used to approximate the distribution of $\overline{X}$ to be normal.

$$\therefore \ Z = \frac{\overline{X} - 25}{S/\sqrt{n}} \sim \mathrm{N}(0,1) \text{ when } H_0 \text{ is true}$$

Rejection Region: $z \leq -2.576$ or $z \geq 2.576$



Computation: $\bar{x} = 24.66, \ s^2 = 0.96015, \ n = 80$

$\therefore z = -3.1035$

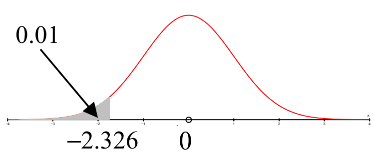
$\text{p-value} = 2\mathrm{P}(z \leq -3.1035) = 0.0019125 \approx 0.00191$

Conclusion: Since p-value = 0.00191< 0.01 (or $z = -3.1035$ <-2.576), $\therefore$ H_0 is rejected at 1% significance level. Hence there is sufficient evidence to conclude at the 1% significance level that the manufacturer's claim is incorrect.

It is not necessary to assume the distribution of the mass of the contents of a bag, is normal as $n = 80$ is large, $\overline{X}$, the sample mean mass, is approximately normal by the application of the Central Limit Theorem.

5(iii) '1% level of significance' means there is a probability of 0.01 that the test will conclude the mean mass of the contents of a bag is not 25 kg when it is actually 25 kg.

5(iv)



Rejection Region at the 1% significance level is

$z \leq -2.326$

From (ii), calculated value of $z = -3.1035 < -2.326$ $\therefore$ it is necessarily true to conclude at the 1% significance level that, on average, each bag contains less than 25 kg of salt.

Alternatively, $2\mathrm{P}(z \leq -3.1035) = 0.0019125$

$\Rightarrow \mathrm{P}(z \leq -3.1035) \approx 0.000956$

So $\text{p} = \text{value} = 0.000956 < 0.01$ $\therefore$ it is necessarily true to conclude at the 1% significance level that, on average, each bag contains less than 25 kg of salt.

Question 6	
(i)	Let T be random variable "travelling time in minutes" and μ be mean travelling time in minutes. H_0 : $\mu = 35$ H_1 : $\mu > 35$ Under H_0, $\dfrac{\overline{T}-35}{4/\sqrt{7}} \sim N(0,1)$. Using GC and the existing data set, p-value = 0.0653 > 0.05 We thus do not reject H_0 and there is insufficient evidence at the 5% significance level to conclude that his mean travelling time is more than 35 minutes.
(ii)	Level of significance is the probability of concluding that the the null hypothesis "mean travelling time is 35 minutes" is rejected when the mean time is in fact 35 minutes.
(iii)	Let X be random variable "new travelling time in minutes" and μ be new mean travelling time in minutes. H_0 : $\mu = 35$ H_1 : $\mu < 35$ Unbiased estimate for population variance = $\dfrac{60}{59} \times 3.5^2 = 12.458$ Under H_0, $\dfrac{\overline{X}-35}{12.458/\sqrt{60}} \sim N(0,1)$ approximately For H_0 to be rejected, $\dfrac{\overline{x}-35}{12.458/\sqrt{60}} < invnorm(0.05)$ $\dfrac{\overline{x}-35}{12.458/\sqrt{60}} < -1.64485$ .... $\overline{x} < 34.3$ No assumption about the population needed as n is large, $\overline{X}$ is approximately normally distributed.

7(i)	Test $\quad H_0 : \mu = 14.0$ against $H_1 : \mu \neq 14.0$	
(ii)	Test statistic: Under H_0, $Z = \dfrac{\overline{X} - 14.0}{3.8/\sqrt{20}} \sim N(0,1)$ For H_0 not to be rejected at 5% level of significance, $-1.96 < \dfrac{\overline{x} - 14.0}{3.8/\sqrt{20}} < 1.96$ $12.33457 < \overline{x} < 15.66543$ $12.3 < \overline{x} < 15.7$	
(iii)	$\overline{x} = 15.8$ does not satisfy the inequality in (ii). There is sufficient evidence to indicate that the squirrels on the island do not have the same mean length as the species known to her.	

8. HCI/MYE/2018/Q11

Milk companies can have "Higher in Calcium" and/or "Lower in Fat" printed on the packaging of the milk they produce if their packets of milk meet the following criteria:

Substance	Quantity	Printed on package
Calcium	> 130 milligrams(mg) per 100 millilitres(ml) of milk	"Higher in Calcium"
Fat	< 1.5 grams(g) per 100 g of milk	"Lower in Fat"

(a) Milk Company A states on its packaging that their milk contains 1.55 g of fat in 100 ml of milk. Given that 100 ml of milk weighs 103 g, can the company print "Lower in Fat" on the packaging of their milk? Justify your answer. [2]

(b) A random sample of 40 packets of milk produced by Company A is taken, and the amount of calcium (in mg) per 100 ml of milk for each packet is summarised as follows:

Amount of calcium (in mg) per 100 ml	128	129	130	131	132	133	134	135	136	137	138
No. of packets	6	5	7	4	4	5	4	1	2	1	1

(i) Find the sample mean amount of calcium per 100 ml of milk. Suggest why it is not advisable to allow Company A to have "Higher in Calcium" printed on the packaging of their milk based solely on the value of this sample mean. [2]

(ii) Test, at 1% level of significance, to determine if Company A can have "Higher in Calcium" printed on the packaging of their milk. You should state your hypotheses and define any symbols that you use. [5]

(iii) Explain, with reason, if there is a need to know anything about the population distribution of the amount of calcium per 100 ml in the milk produced by Company A for the test to be valid. [2]

(iv) It is now given that population standard deviation of the amount of calcium per 100 ml in the milk produced is $\sigma = 3.2$ mg. Does this piece of information change the conclusion of the hypothesis test in **(iii)**? [1]

Solution:

8 (a)	**Method 1** Criterion: $\frac{1.5}{100}\times 103 = 1.545\text{g/100ml}$ Since 1.55g/100ml > 1.545g/100ml, Company A cannot print "Lower in Fat" on their packaging. **Method 2** Amount: 1.46g/103g $\frac{1.55}{103}\times 100 = 1.5049\text{g/100g}$ Since 1.5049g/100g > 1.5g/100g, Company A cannot print "Lower in Fat" on their packaging.
(b)(i)	Using GC, $\bar{x} = 131.45\,\text{mg}$ There may be outliers in the random sample that skew the mean heavily, such that a significant portion of the product actually contain less calcium than the requirement. OR It is necessary to consider the sample standard deviation (= 2.66 (3 s.f.)) of the random sample as well, as there may be a significant portion of the product with less calcium than the requirement if the variance is large.
(b)(ii)	Let X be the **mass of calcium** (in mg) in every 100ml of milk, and μ (in mg) denote the **population mean** mass of calcium per 100 ml of milk. Null Hypothesis - $H_0 : \mu = 130$ Alternative Hypothesis - $H_1 : \mu > 130$ **Unbiased estimate of population variance** $= s^2 = 7.279487179$ Under H_0 , Test statistic $Z = \frac{\bar{X} - 130}{\sqrt{\frac{7.279487179}{40}}} \sim N(0,1)$ approximately since n = 40 is large. From GC, $z = 3.39897$, p-value $= 0.000338$ (3 s.f.) Since p-value < 0.01, we **reject** H_0 at 1% **level of significance**. There is **sufficient evidence at 1% significance level** that the *population mean amount of calcium exceeds 130 mg per 100 ml of milk* and that *Company A can have "Higher in Calcium" printed on the packaging*.
(b)(iii)	Since the sample size is large, by Central Limit Theorem, the distribution of sample mean of the amount of calcium per 100 ml of milk produced by Company A follows a normal distribution approximately. Therefore there is no need to know anything about the population distribution.
(b)(iv)	No, it does not change the conclusion. The new p-value is 0.00208 (to 3 s.f.) < 0.01, so we still reject H_0.