EIN 5226

Hypothesis testing
- One sample means
- Sigma known

Chapter 16 Sections 16.3-4

Chapter 17 Sections 17.2-5

Need Z table

Karen E. Schmahl Ph.D., P.E.

Hypothesis Testing

Hypothesis

An assumption or theory made about a population parameter or relationship between populations

Hypothesis test

Statistical test of hypothesis based on sample data collected from the population/s of interest

Types of Hypothesis tests

- Tests for a single sample
  - One population's mean to external criterion (large sample size)
  - One population proportion to external criterion
  - One population's mean to external criterion (small sample size)
  - Populations expected distribution of outcomes to samples distribution of outcomes.
- Tests comparing two samples
  - Difference between two population means (large sample size)
  - Difference between two population proportions
  - Difference between two population means (small sample size)
  - Difference between two population variances

Statistical Inference - Using p values

Steps in Hypothesis Testing

1. Identify your objective
2. State the null hypothesis, H_0
3. State the alternative hypotheses, H_a.
4. Calculate the appropriate test statistic
5. Compute the p value of the test statistic
6. Determine the acceptable risk
7. Compare the p value of test statistic to the to the acceptable risk.

Customer complaints indicated breakage of a component during use of a product. A project to develop a new process to reduce product failure by increasing the tensile strength of a component without changing other properties was initiated.

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 86.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength? (Assume normal distribution.)

1) Identify your objective

For our problem we want to know if we could conclude that the population mean increased over the historical tensile strength of 3000. Can we conclude the project was a success?

Comparison of 1 sample mean to a defined μ

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 86.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

2) Formulate a null hypothesis.

H_0 – Null hypothesis. Always arrange the null claim such that it contains the condition of equality.

H_0: $\mu = \mu_0$ $\mu = 3000$

The null claim is what you are comparing the sample data to.

In the problem, we are comparing the sample data to 3000

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 86.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

3) Formulate an alternative hypothesis

H_a – Alternative hypothesis. The alternative will involve exactly one of three signs, <, >,or ≠. If the alternative is ≠, then the test will be two-tailed. If < or >, it will be one tailed.

H_a: $\mu > \mu_0$ $\mu > 3000$

Two tailed test	One tailed test
α split between tails,	α all in one tail,
Equal to or not equal to	Less than or greater than
Results supported in either direction	Other side ignored

7

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 86.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

H_0 : $\mu=\mu_0$ μ_0 is what is claimed, what is being tested against.

H_0: μ =3000

H_a: μ > 3000

What happens if it is equal?

Do not conclude successful

What is the alternative?

For alternative to happen, is it <, > or ≠ ?

Conclude successful

8

Note

For one-tailed tests, some textbooks will use < or > signs in the null hypothesis. This is not really wrong, but is not necessary. It makes it somewhat more difficult to establish the null and alternative hypothesis.

For purposes of this class, the null hypothesis will always be the equality.

9

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 86.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

4) Calculate a test statistic from the sample information.

$$Z_{test} = \frac{\bar{x} - u_0}{\sigma/\sqrt{n}} = \frac{3025 - 3000}{86/\sqrt{45}} = +1.95$$

Possible test statistics: Z, t, $\varkappa^2$, F

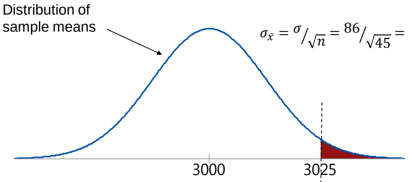
For a two tailed test, Z_{test} is ±

10

Assuming H_0 – that distribution mean is 3000

Distribution of sample means

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 86/\sqrt{45} = 12.82$$



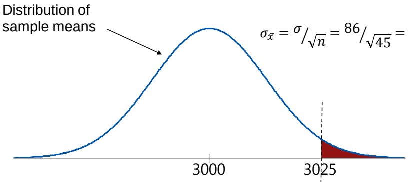
$$Z_{test} = \frac{\bar{x} - u_0}{\sigma/\sqrt{n}} = \frac{3025 - 3000}{86/\sqrt{45}} = +1.95$$

11

Assuming H_0 – that distribution mean is 3000

Distribution of sample means

$$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 86/\sqrt{45} = 12.82$$



$$Z_{test} = \frac{\bar{x} - u_0}{\sigma/\sqrt{n}} = \frac{3025 - 3000}{86/\sqrt{45}} = +1.95$$

If you have a process mean of 3000 and process standard deviation of 86, what is the probability that a sample of size 45 will have a value greater than 3025?

A. 0.0256 B. 0.0500 C. 0.0032 D. 0.0215

12

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 65.

Can we conclude with 90% certainty that the new process was effective in increasing the tensile strength?

5) Determine probability, p, that a value will be further out than the sample value.

$$P(Z > 1.95) = .0256 = 2.56\% = pvalue$$

For $H_1 = \mu \neq \mu_0$, sum the areas in the tails, cut off by Z and –Z
$H_1 = \mu > \mu_0$, p value is area to the right of Z
$H_1 = \mu < \mu_0$, p value is area to the left of Z

13

Use of P-Values in Hypothesis Testing

P-value

- *probability* associated with the test statistic, Z_{test}.
- probability of getting a value further out than your test values

14

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 65.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

6) Determine the acceptable risk.

Can we conclude, with 95% certainty, that the process change was effective in increasing the tensile strength?

$\alpha = 5\%$

- Alpha, α, is the risk of saying the process has changed when it really has not.
- Alpha is a business decision.

15

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 65.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

6) Compare your p value to the acceptable risk to make an inference about the population.

p = 2.56% $\boldsymbol{\alpha} = 5\%$

$p < \boldsymbol{\alpha}$

Therefore reject the hypothesis that $\mu = 3000$

"If p is low, null must go.

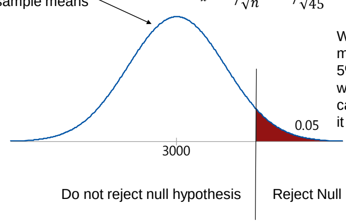
If p is high the null will fly."

16

Assuming H_0 – that distribution mean is 3000

Distribution of sample means

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 86/\sqrt{45} = 12.82$



When the population mean really is 3000, 5% of the time when we take a sample and calculate that average, it will fall in this area.

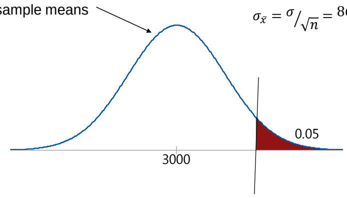
Do not reject null hypothesis   Reject Null hypothesis

17

Assuming H_0 – that distribution mean is 3000

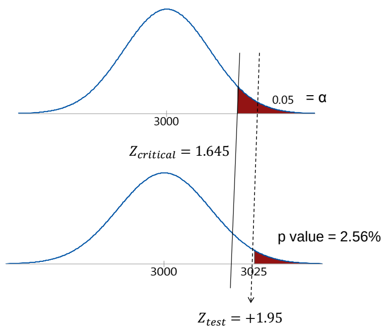
Distribution of sample means

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 86/\sqrt{45} = 12.82$



$Z_{critical} = 1.645$

18



19

Drawing Conclusions from the Results of Hypothesis Tests

- Statistically, there are only two possible conclusions
  - Reject H_0: Conclude that H_0 is false.
  - Do not reject H_0: Conclude H_0 is plausible. Our evidence is not strong enough to reject it.
- One can never conclude that H_0 is true. We can just conclude that H_0 might be true.
- The smaller the *P*-value, the stronger the evidence is against H_0.

20

Prior to process changes the component consistently averaged 3000 psi. After the process improvements, a study was conducted where 45 samples were tested and had with an average of 3025 and standard deviation of 65.

Can we conclude with 95% certainty that the new process was effective in increasing the tensile strength?

H_0: μ = 3000 H_a: μ > 3000

α = 5%

Test value: $Z_{test} = \frac{\bar{x}-u_0}{\sigma/\sqrt{n}} = \frac{3025-3000}{86/\sqrt{45}} = +1.95$

The p value associated with the Z_{test} is 0.0256.= 2.56%

Statistical conclusion: p< α, therefore reject the null hypothesis. At a 95% confidence level, the evidence is strong enough to reject the null hypothesis.

Practical conclusion: The new process was effective in increasing the tensile strength of the component.

21

More on the *P*-value

- The P-value should always be reported with the results of a hypothesis test.
- The smaller the *P*-value, the more certain we can be that H_0 is false.
- The larger the *P*-value, the more plausible H_0 becomes but we can never be certain that H_0 is true.
- If you just report the confidence level, you do not know how strong your conclusion is.
- The P-Value is considered the smallest level of α risk that would lead to rejection of the null hypothesis H_0

22

Tests on Single Population Means
Large Sample Sizes (n≥30)

	Two-tailed	One-tailed	
Null hypothesis H_0	$\mu=\mu_0$		
Alternate H_a	$\mu\neq\mu_0$	$\mu>\mu_0$	$\mu<\mu_0$
Test statistic	$Z_{test} = \frac{\bar{x} - u_o}{s/\sqrt{n}}$		
p (reject if p<α)	sum the areas in the tails, cut off by Z and –Z	Area to right of Z	Area to left of Z

23

In a power generating plant, a certain line is supposed to maintain an average of **122** psi over any 8 hour period. In one period, **36** measurements were taken with a mean of **123** and a standard deviation of **5.2**. Can we conclude, with 90% certainty, that the line is malfunctioning. (Assume normal distribution.)

$H_0: \mu=\mu_0$ μ_0 is what is claimed, what is being tested against.

H_0: μ =122

What happens if it is equal?

Do not fix machine

What is the alternative?

Fix machine

For alternative to happen, is it <, > or ≠ ?

H_a: $\mu \neq 122$

24

In a power generating plant, a certain line is supposed to maintain an average of **122** psi over any 8 hour period. In one period, **60** measurements were taken with a mean of **123** and a standard deviation of **5.2**. Can we conclude, with **90**% certainty, that the line is malfunctioning. (Assume normal distribution.)

H_0: μ = 122 H_a: μ ≠ 122

α = 10%

Test value: $Z_{test} = \frac{\bar{x} - u_0}{\sigma/\sqrt{n}} = \frac{123-122}{5.2/\sqrt{36}} = \pm 1.154$

P value = P(Z<-1.15) +P(Z>+1.15) = 0.1251 + 0.1251 = 0..2505

The p value associated with the Z_{test} is 0.2502= 25.05%

Statistical Conclusion: p>α, therefore cannot reject null hypothesis.
It is plausible that the process mean is still 122.

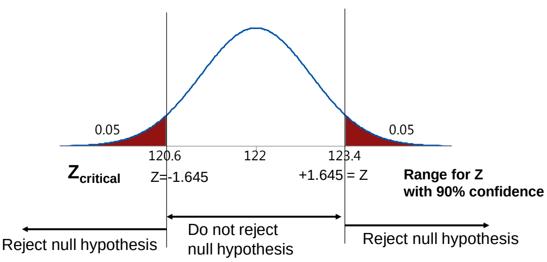
Practical Conclusion: Do not send someone to fix the machine.

25

The hypothesis test -

Distribution of Sample means where n=60 and

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 5.2/\sqrt{36} = 0.8667$

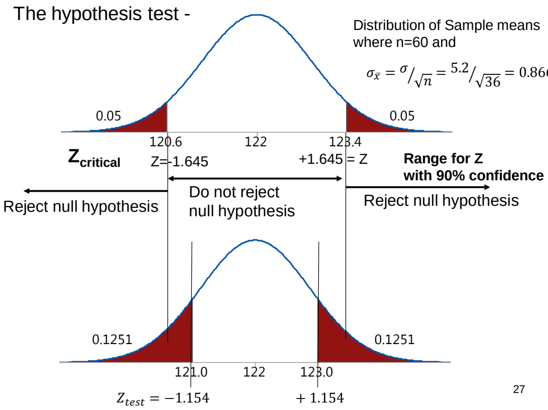


26

The hypothesis test -

Distribution of Sample means where n=60 and

$\sigma_{\bar{x}} = \sigma/\sqrt{n} = 5.2/\sqrt{36} = 0.8667$



27

Practice – Process Improvement

A company wants to reduce the assembly time for one of its products that is produced at multiple manufacturing facilities. The current assembly time averages 25.0 minutes. A new fixture was implemented at one of the facilities. 35 products were assembled with an average time of 23.5 minutes and standard deviation of 3.6 minutes. Would you recommend implementing the fixture at other facilities? (Use α=.05)

6. What are the null and alternate hypotheses for this problem?

A) H_0: $\mu = 25.0$ H_a: $\mu < 25.0$

B) H_0: $\mu = 23.5$ H_a: $\mu < 25.0$

C) H_0: $\mu = 23.5$ H_a: $\mu \neq 23.5$

D) H_0: $\mu = 25.0$ H_a: $\mu \neq 25.0$

28

Practice Process Improvement

$$Z_{test} = \frac{\bar{x} - u_o}{s / \sqrt{n}}$$

A company wants to reduce the assembly time for one of its products that is produced at multiple manufacturing facilities. The current assembly time averages 25.0 minutes. A low cost method changed was implemented at one of the facilities. 35 products were assembled with an average time of 23.5 minutes and standard deviation of 3.6 minutes. Would you recommend implementing the method at other facilities? (Use α=.05)

7. Z_{test} is

A) -2.81 B) -2.47 C) + 2.47 D) -3.26

8. The p value for this problem is

A) .0055 B) .0089 C) .0068 D) .9932

29

Practice Process Improvement

$$Z_{test} = \frac{\bar{x} - u_o}{s / \sqrt{n}}$$

A company wants to reduce the assembly time for one of its products that is produced at multiple manufacturing facilities. The current assembly time averages 25.0 minutes. A low cost method changed was implemented at one of the facilities. 35 products were assembled with an average time of 23.5 minutes and standard deviation of 3.6 minutes. Would you recommend implementing the method at other facilities? (Use α=.05)

From the hypothesis test which of the following are true statements of conclusion.

T / F 9. The null hypothesis, u = 25.0, is rejected.

T / F 10.The time for assembly is now less than 25 minutes.

T / F 11.The time for assembly is now 23.5 minutes

T / F 12.Assuming the cost to implement is sufficiently low, we should implement at the other facilities.

30

Significance

- Even though a result may be statistically significant, common sense needs to be applied in taking action based on the result.

.

- Sometimes statistically significant results do not have any scientific or practical importance.

31

Process Control

Control charts established for a process indicated the process was stable with a mean of 12.5 and standard deviation of 1.2.

Samples of size 9 are taken from the process. Sample #8 had an average of 11.7. Assuming the standard deviation was still 1.2, Can you conclude with 99.74 confidence that the process mean has changed? (2 sided test)

What are the null and alternate hypotheses for this problem?

A) H_0: $\mu = 12.5$ H_a: $\mu < 12.5$

B) H_0: $\mu = 11.7$ H_a: $\mu > 11.7$

C) H_0: $\mu = 11.7$ H_a: $\mu \neq 11.7$

D) H_0: $\mu = 12.5$ H_a: $\mu \neq 12.5$

32

Process Control

$$Z_{test} = \frac{\bar{x} - u_o}{s/\sqrt{n}}$$

Control charts established for a process indicated the process was stable with a mean of 12.5 and standard deviation of 1.2.

Samples of size 9 are taken from the process. Sample #8 had an average of 11.7. Assuming the standard deviation was still 1.2, Can you conclude with 99.74 confidence that the process mean has changed? (2 sided test)

Z_{test} is

A) -1.25 B) -1.45 C) -0.68 D) -3.75

33

Process Control

$$Z_{test} = \frac{\bar{x} - u_o}{s/\sqrt{n}}$$

Control charts established for a process indicated the process was stable with a mean of 12.5 and standard deviation of 1.2.

Samples of size 10 are taken from the process. Sample #9 had an average of 11.7. Assuming the standard deviation was still 1.2, Can you conclude with 99.74 confidence that the process mean has changed? (2 sided test)

The p value for this problem is

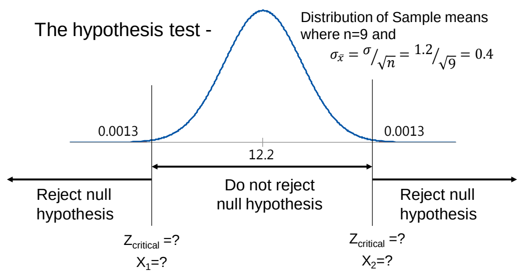
A) 0.1648 B) 0.2112 C) 0.1056 D) 0.0968

T / F p < α, therefore reject the null hypothesis.

T / F At 99.7 % confidence, it is plausible that the process mean is still 12.5.

34

The hypothesis test -



If I just wanted to compare my Z_{test} value to a $Z_{critical}$ value, I would compare it to

A) ± 3.00 B) ± .026 C) ± 2.80

What are the values of X_1 and X_2 that correspond to the Z values?

A) 11.34, 13.06 B) 12.00, 12.40 C) 11.00, 13.40

35

Process Control

$$UCL_{\bar{x}} = u_{\bar{x}} + 3\sigma_{\bar{x}} = \bar{\bar{x}} + 3\sigma_{\bar{x}} = \bar{\bar{x}} + A_3\bar{S}$$

$$CL_{\bar{x}} = u_{\bar{x}} = \bar{\bar{x}}$$

$$LCL_{\bar{x}} = u_{\bar{x}} - 3\sigma_{\bar{x}} = \bar{\bar{x}} - 3\sigma_{\bar{x}} = \bar{\bar{x}} - A_3\bar{S}$$

Control charts established for a process indicated the process was stable with a mean of 12.5 and standard deviation of 1.2.

Samples of size 9 are taken from the process.

What are the control limits for the $\bar{x}$ chart for this process?

A) 11.34, 13.06 B) 12.00, 12.40 C) 11.00, 13.40

A sample is taken and has an average of 11.7. What action should be taken relative to the $\bar{x}$ chart ?

A) Since it is outside the control limits, look for a special cause.
B) Since it is inside the control limits, continue running.
C) Since it is inside the control limits, look to see if there is any other signs of out-of-control process behavior.

36

Related Assignments

Please see Blackboard for related assignments

37

FLORIDA INTERNATIONAL UNIVERSITY COLLEGE OF ENGINEERING