Hypothesis Testing

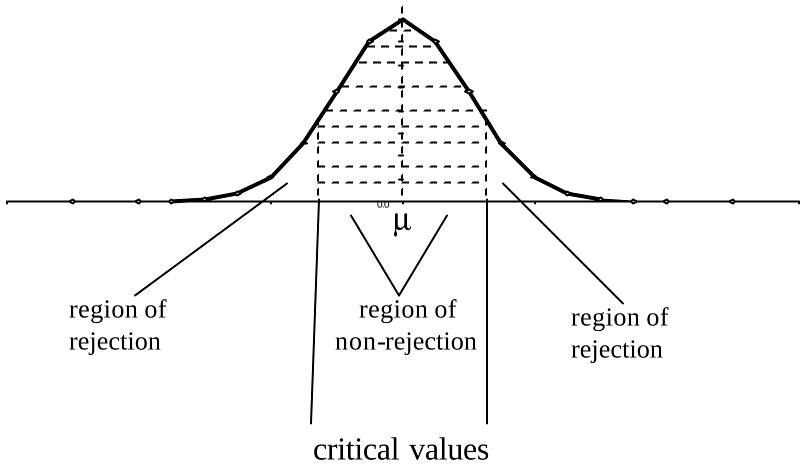
Daniel A. Menascé
Department of Computer Science
George Mason University

1

Hypothesis Testing

- Purpose: make inferences about a population parameter by analyzing differences between observed sample statistics and the results one expects to obtain if some underlying assumption is true.
- Null hypothesis: $H_0 : m = x$
- Alternative hypothesis: $H_1 : m \neq x$
- If the null hypothesis is rejected then the alternative hypothesis is accepted.

2



Test statistic: $Z = \frac{\overline{X} - m}{\frac{s}{\sqrt{n}}}$

3

Risks in Decision Making

- Type I Error occurs if H_o is rejected when it is true.
  – Pr [H_o is rejected | true] = α
- Type II Error occurs if H_o is not rejected when it is false.
  – Pr[H_o is not rejected | false] = β
- Confidence coefficient:
  – Pr [H_o not rejected | true]= 1- α
- Power of the test:
  – Pr[H_o is rejected |false]= 1-β

4

	Actual Situation	
	H_o true	H_o false
Accept H_o	Correct decision Confidence=1-α	Type II Error: Pr[Type II]=β
Reject H_o	Type I Error P[Type I]=α	Correct Decision Power=1-β

5

Example of Hypothesis Testing

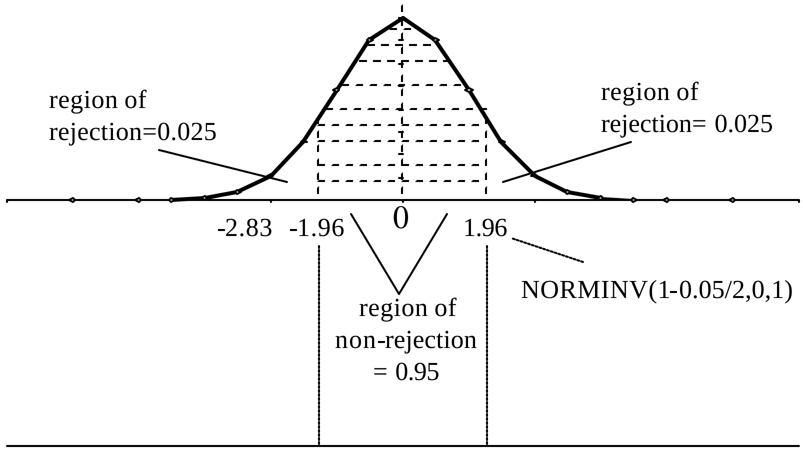
- A sample of 50 files from a file system is selected. The sample mean is 12.3Kbytes. The standard deviation is known to be 0.5 Kbytes.

  Ho: $\mu = 12.5$ Kbytes

  H_1: $\mu \neq 12.5$ Kbytes

Confidence: 0.95

6



$$Z = \frac{12.3 - 12.5}{\frac{0.5}{\sqrt{50}}} = -2.83$$

Reject H_o

Z Test of Hypothesis for the Mean	
Null Hypothesis m=	**12.5**
Level of Significance	**0.05**
Population Standard Deviation	**0.5**
Sample Size	**50**
Sample Mean	**12.3**
Standard Error of the Mean	0.070710678
Z Test Statistic	-2.828427125
Two-Tailed Test	
Lower Critical Value	**-1.959961082**
Upper Critical Value	**1.959961082**
***p*-Value**	**0.00467786**
Reject the null hypothesis	

Steps in Hypothesis Testing

1. State the null and alternative hypothesis.
2. Choose the level of significance α.
3. Choose the sample size *n*. Larger samples allow us to detect even small differences between sample statistics and true population parameters. For a given α, increasing *n* decreases β.
4. Choose the appropriate statistical technique and test statistic to use (Z or t).

9

Steps in Hypothesis Testing

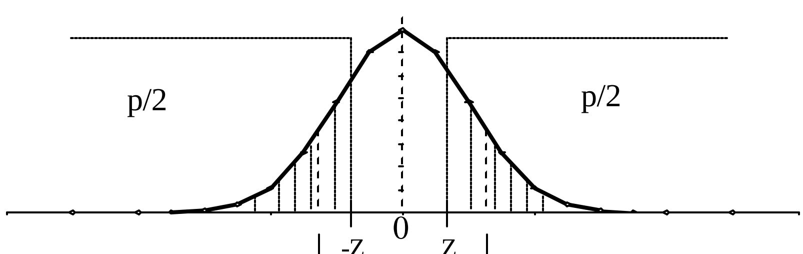
5. Determine the critical values that divide the regions of acceptance and non-acceptance.
6. Collect the data and compute the sample mean and the appropriate test statistic (e.g., Z).
7. If the test statistic falls in the non-reject region, H_o cannot be rejected. Else H_o is rejected.

10

The p-value Approach

- p-value: observed level of significance. Defined as the probability that the test statistic is equal to or more extreme than the result obtained from the sample data, given that H_o is true.

11

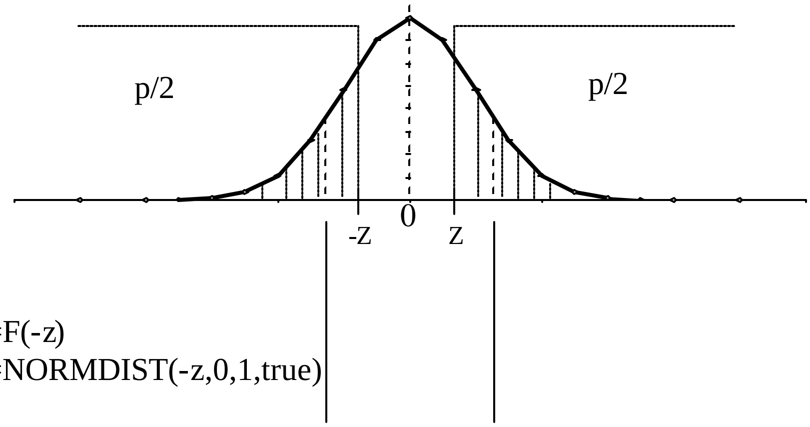


p/2=F(-z)
=NORMDIST(-z,0,1,true)

critical values

If $p \geq \alpha$ then do not reject H_o, else reject H_o.

12

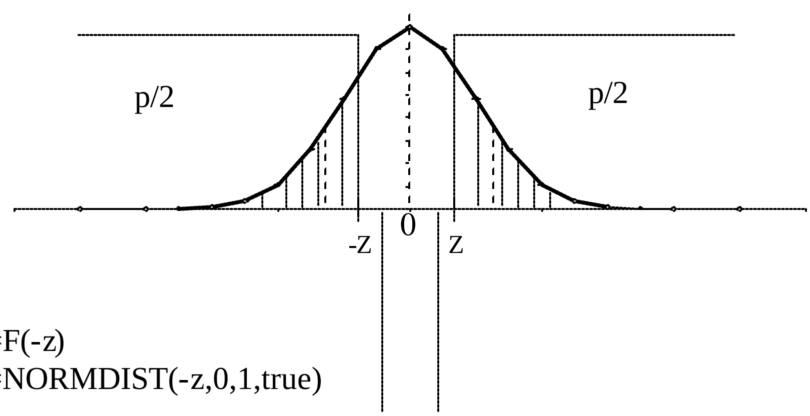


p/2=F(-z)
=NORMDIST(-z,0,1,true)

critical values

Do not reject H_o

13



p/2=F(-z)
=NORMDIST(-z,0,1,true)

critical values

Reject H_o

14

Computing p-values

Z Test of Hypothesis for the Mean	
Null Hypothesis m=	**12.5**
Level of Significance	**0.05**
Population Standard Deviation	**0.5**
Sample Size	**50**
Sample Mean	**12.3**
Standard Error of the Mean	0.070710678
Z Test Statistic	-2.828427125
Two-Tailed Test	
Lower Critical Value	**-1.959961082**
Upper Critical Value	**1.959961082**
***p*-Value**	**0.00467786**
Reject the null hypothesis	

NORMDIST(-2.828427125,0,1,TRUE)

The null hypothesis is rejected because p (0.0047) is less than the level of significance (0.05).

Steps in Determining the p-value.

1. State the null and alternative hypothesis.
2. Choose the level of significance α.
3. Choose the sample size *n*. Larger samples allow us to detect even small differences between sample statistics and true population parameters. For a given α, increasing *n* decreases β.
4. Choose the appropriate statistical technique and test statistic to use (Z or t).

Steps in Determining the p-value.

5. Collect the data and compute the sample mean and the appropriate test statistic (e.g., Z).
6. Calculate the p-value based on the test statistic
7. Compare the p-value to α.
8. If $p \geq \alpha$ then do not reject H_o, else reject H_o.

17

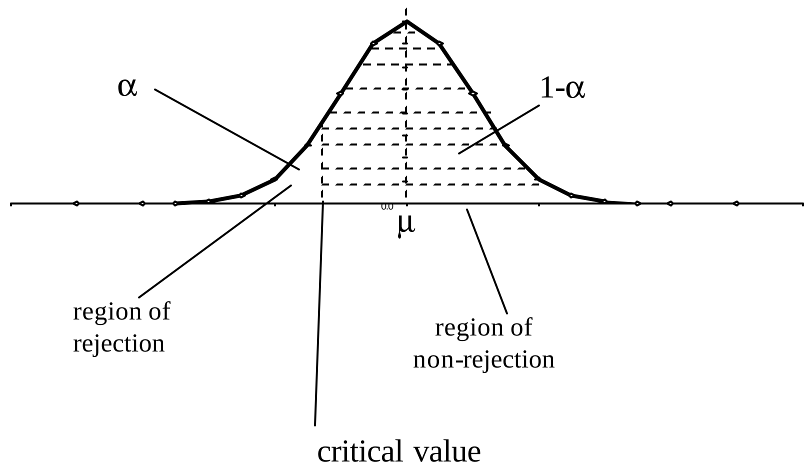
One-tailed Tests

- Null hypothesis is an inequality.

  $H_o \geq 3.5$

  $H_1 < 3.5$

18



Test statistic: $Z = \dfrac{\overline{X} - m}{\dfrac{s}{\sqrt{n}}}$

19

Example of One-Tailed Test

- A sample of 50 files from a file system is selected. The sample mean is 12.35Kbytes. The standard deviation is known to be 0.5 Kbytes.

  Ho: $\mu \geq 12.3$ Kbytes

  H_1: $\mu < 12.3$ Kbytes

Confidence: 0.95

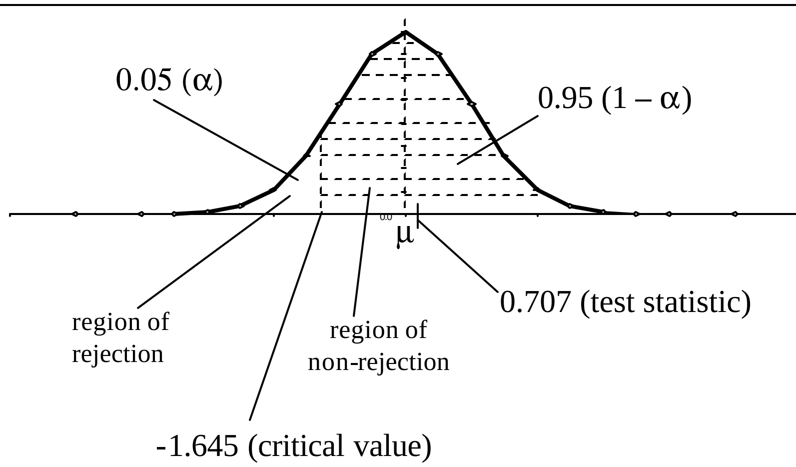
20

Example of One-Tailed Test

$$Z = \frac{\overline{X} - m}{s\ /\sqrt{n}} = \frac{12.35 - 12.3}{0.5/\sqrt{50}} = 0.707$$ (test statistic)

Critical value =NORMINV(0.05,0,1)= -1.645.
Region of non-rejection: $Z \geq -1.645$.
So, do not reject H_o. (Z exceeds critical value)

21



Test statistic: $$Z = \frac{\overline{X} - m}{\frac{s}{\sqrt{n}}}$$

22

One-tailed Test

Z Test of Hypothesis for the Mean		
Null Hypothesis	m=	**12.3**
Level of Significance		**0.05**
Population Standard Deviation		**0.5**
Sample Size		**50**
Sample Mean		**12.35**
Standard Error of the Mean		0.070710678
Z Test Statistic		0.707106781
Lower-Tail Test		
Lower Critical Value		**-1.644853**
***p*-Value**		**0.760250013**
Do not reject the null hypothesis		

Hypothesis Tests with Unknown σ

- If the population is assumed to be normally distributed the sampling distribution for the mean follows a t distribution with n-1 degrees of freedom.
- t statistic for unknown σ:

$$t = \frac{\overline{X} - m}{\frac{s}{\sqrt{n}}}$$

Use sample standard deviation

Example of Hypothesis Testing

- A sample of 50 files from a file system is selected. The sample mean is 12.3Kbytes. The sample standard deviation is 0.5 Kbytes.

  Ho: $\mu = 12.35$ Kbytes

  H_1: $\mu \neq 12.35$ Kbytes

Confidence: 0.95

25

t Test of Hypothesis for the Mean	
Null Hypothesis m=	**12.35**
Level of Significance	**0.05**
Sample Size	**50**
Sample Mean	**12.3**
Sample Standard Deviation	**0.5**
Standard Error of the Mean	0.070710678
Degrees of Freedom	49
t Test Statistic	-0.707106781
Two-Tailed Test	
Lower Critical Value	**-2.009574018**
Upper Critical Value	**2.009574018**
***p*-Value**	**0.482849571**
Do not reject the null hypothesis	

TINV(0.05,49)

The t test statistic is between the lower and critical values. So, do not reject the null hypothesis.

26