Type I vs. Type II Errors

Dr. Supaporn Erjongmanee

Department of Computer Engineering
Kasetsart University
fengspe@ku.ac.th

1

Outline

- Type II Error
- Caution

2

Test Errors

- Type I Error
  - Reject null hypothesis when it is true
  - Represented by α
- Type II Error
  - Not reject null hypothesis when it is false
  - Represented by β

	H_0 is true	H_0 is false
Accept H_0	Correct	
Reject H_0		Correct

Supaporn Erjongmanee
fengspe@ku.ac.th

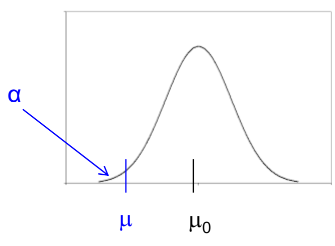
Department of Computer Engineering
Kasetsart University

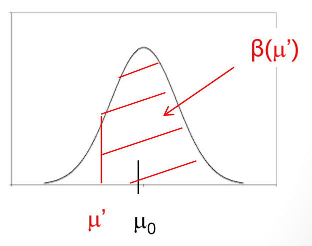
3

Significance Level

	H_0 is true	H_0 is false
Accept H_0	Correct	Type II Error
Reject H_0	Type I Error	Correct

- α is called significance level
- Typical values for α: 0.1, 0.05, 0.01
- α and β affect rejection region
  - Decrease α -> Increase β
- We cannot have both small α and β
- Generally,
  - Find largest value of α that can be tolerated
  - Then, β is a result of above





Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

4

Example

- Drying time is normally distributed
  - Average paint drying time = μ = 75 minute, σ = 9 minutes
- Company wants to decrease drying time by including new additive
  - Assume that with new additive, drying time remains normal distributed
    - H_0: μ≤ 75
    - H_a: μ < 75
- Company collects 25 drying times: $\{X_1, X_2, \ldots, X_{25}\}$ (n = 25)
    - Compute sample mean $\mu_{\bar{X}}$ and sample standard deviation $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ = 9 / $\sqrt{25}$ = 1.8

5

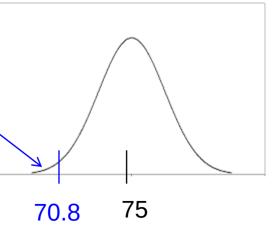
Example (cont.)

	H_0 is true	H_0 is false
Accept H_0	Correct	Type II Error
Reject H_0	Type I Error	Correct

- H_0: μ ≤ 75
- H_a: μ < 75

- If H_0 is true, $\mu_{\bar{X}}$ = 75
- If we reject H_0 , that means $\mu_{\bar{X}}$ < 75
- Let's reject H_0 when $\bar{X} \le 70.8$

α = 0.01
= Shaded area



$\alpha = P$(type I error)

= P (H_0 is rejected when it is true (μ = 75))

$= P(\bar{X} \le 70.8 \text{ when } \bar{X} \sim \mathrm{N}(\mu_{\bar{X}} = 75, \sigma_{\bar{X}} = 1.8))$

$= \Phi\left(\frac{70.8-75}{1.8}\right) = \Phi(-2.33) = 0.01$

H0 is true but rejected because of $\bar{X}$
=> Type I error

If population mean actually is $\mu = 75$,
Ho is rejected when $\bar{X} \le 70.8$ at α = 0.01

6

Example (cont.)

	H_0 is true	H_0 is false
Accept H_0		Type II Error
Reject H_0	Type I Error	

- H_0: μ ≤ 75
- H_a: μ < 75
- With $\alpha = 0.01$ reject H_0 when $\bar{X} \le 70.8$
- If actual population mean = μ = 72 (or Ha is True),
  - What is type II error?

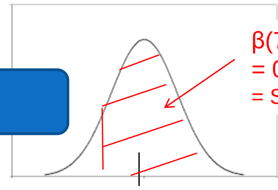
Ha is true but Ho is accepted
=> Type II error

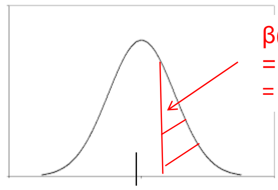
β(72) = *P*(type II error when μ = 72)

= *P* (H_0 is not rejected when it is false because μ = 72)

$= P(\bar{X} > 70.8 \text{ when } \bar{X} \sim \mathrm{N}(\mu_{\bar{X}} = 72, \sigma_{\bar{X}} = 1.8)$

$= 1 - \Phi\left(\frac{70.8-72}{1.8}\right) = 1 - \Phi(-0.67) = 1 - 0.2514 = 0.748$





β(70) = *P*(type II error when μ = 70) = $1 - \Phi\left(\frac{70.8-70}{1.8}\right) = 0.3300$

7

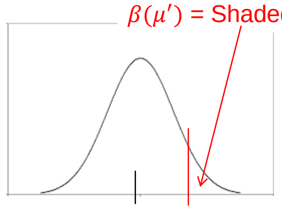
Type II Error

	H_0 is true	H_0 is false
Accept H_0	Correct	Type II Error
Reject H_0	Type I Error	Correct

H_0 : μ ≤ $μ_0$

H_a : μ > $μ_0$

$Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}$



β(μ') =*P*(type II error)

= *P* (H_0 is not rejected when μ = μ')

$= P(Z < z_\alpha$ when μ = μ')

$= P\left(\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} < z_\alpha \text{ when } \mu = \mu'\right)$

$= P(\bar{X} < \mu_0 + z_\alpha \sigma/\sqrt{n}$ when μ = μ')

$= P\left(\frac{\bar{X}-\mu'}{\sigma/\sqrt{n}} < z_\alpha + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}}\right)$

$= \Phi\left(z_\alpha + \frac{\mu_0-\mu'}{\sigma/\sqrt{n}}\right)$

With given α, find z_α.
Find $\bar{X}$ that results to reject H0 (Type I error).
With $\bar{X}$ find type II error (*β*) depending on μ'

When μ' increases,
$\mu_0 - \mu'$ is more negative
β(μ') is small
when $\mu' > \mu_0$

For lower-tailed and two-tailed can be solved in similar ways

8

Population Mean Test : Normal with Known Variance

- Null hypothesis: $\mu = \mu_0$

Alternative Hypothesis	Type II error probability $\beta(\mu')$ at α level
$H_a : \mu > \mu_0$	$\Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a : \mu < \mu_0$	$1-\Phi\left(-z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$
$H_a : \mu \neq \mu_0$	$\Phi\left(z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right) - \Phi\left(-z_{\alpha/2} + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)$

Can solve for sample size n required to have $\beta(\mu')$

One-tailed: $n = \left(\frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right)^2$

Two-tailed: $n = \left(\frac{\sigma(z_{\alpha/2} + z_\beta)}{\mu_0 - \mu'}\right)^2$

Example

- Claim: Average tire life span = μ = 30,000
- Assume tire life span is normally distributed with σ = 1,500
- Want to check tire life span
  - H_0: μ = 30,000
  - H_a: μ > 30,000
- Collect 16 sample data: $\{X_1, X_2, \ldots, X_{16}\}$ (n = 16)
- Answer the followings using significance level = 0.01 and actual population mean = μ' = 31,000:
  - Find type II error
  - If we want type II error = 0.1, how many sample size required?

Example (cont.)

- Claim: Average tire life span = μ = 30,000
- Assume tire life span is normally distributed with σ = 1,500
- Want to check tire life span
  - H_0: μ = 30,000
  - H_a: μ > 30,000
- Collect 16 sample data: $\{X_1, X_2, \ldots, X_{16}\}$ (n = 16)
- At significance level = 0.01 and actual population mean = μ' = 31,000, find type II error

$$\boxed{\beta(31{,}000) = \Phi\left(z_\alpha + \frac{\mu_0 - \mu'}{\sigma/\sqrt{n}}\right)} \Rightarrow z_\alpha = z_{0.01} = 2.33$$

$$= \Phi\left(2.33 + \frac{30{,}000 - 31{,}000}{1{,}500/\sqrt{16}}\right) = \Phi(-0.34)$$

$$= 0.3669$$

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

11

Example (cont.)

- Claim: Average tire life span = μ = 30,000
- Assume tire life span is normally distributed with σ = 1,500
- Want to check tire life span
  - H_0: μ = 30,000
  - H_a: μ > 30,000
- Given significance level = 0.01 and actual population mean = μ' = 31,000
  - If we want type II error = 0.1, how many sample size required?

$$\beta(31{,}000) = 0.1 \Rightarrow z_\beta = z_{0.1} = 1.28$$

$$n = \left(\frac{\sigma(z_\alpha + z_\beta)}{\mu_0 - \mu'}\right)^2 = \left(\frac{1500(2.33+1.28)}{30{,}000 - 31{,}000}\right)^2 = (-5.42)^2 = 29.32$$

At least 30 tires are required to have type II error = 0.1

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

12

Outline

- Type II Error
- Caution

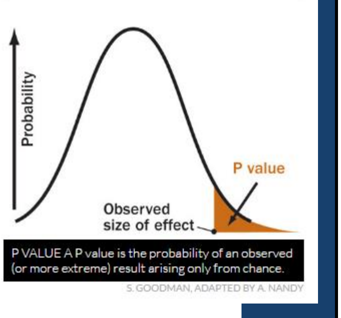
Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

Caution 1

Beware of - meaning of p-value and
- assumption in null hypothesis

- P-value often is mistaken as probability that H_0 is true.
- Actually, p-value is probability of observing results if H_0 is true.
- Example:
  - P-value = 0.05 implies that there is only 5% chance to observe extreme data if H_0 is correct.
    - Do not assume that 95% chance to observe normal data that H_0 is false.



Source: https://www.sciencenews.org/article/odds-are-its-wrong

Supaporn Erjongmanee
fengspe@ku.ac.th

Department of Computer Engineering
Kasetsart University

Caution 1 (cont.)

Beware of - meaning of p-value and
- assumption in null hypothesis

- P-value is probability of observing results if H_0 is true.
  - Do not assume that 95% chance to observe normal data that H_0 is false.

Dog is hungry.

H_0: A dog barks when it is not hungry. → H_0 assumes that dog is not hungry

H_a: A dog does not bark when it is not hungry.

Observation is done on whether the dog barks less than 5% when it is not hungry.

Q: What is p_1 = probability that H_0 is false (a dog does not bark when the dog IS hungry)?

This probability p_1 cannot be computed. Because H_0 assumes the dog is not hungry.

To compute p_1, observation must be done on the dog does not bark when it IS hungry.

Source: https://www.sciencenews.org/article/odds-are-its-wrong

Caution 2

Compare the correct data pairs

- Experiment to see impact of two anti-depression drugs: Paxil vs. Prozac
  - In experiment, give out actual drug and placebos (fake drug)
  - Found that
    - Paxil has suicidal rate in actual drug > placebos
    - Prozac has suicidal rate in actual drug < placebos

$\wedge$

Prozac is better drug.

- In fact, fewer kids on placebo in Paxil < Prozac
- It is not necessary that Prozac is better drug

Source: https://www.sciencenews.org/article/odds-are-its-wrong

Caution 3

Beware of randomization in two data sets

- Data is composed of control and experimental groups

No factor would be biased on control group over experimental group, and vice versa.

- How to make sure that both groups
  - Are randomly selected
  - Have equal variability

Source: https://www.sciencenews.org/article/odds-are-its-wrong
Image source: http://image.shutterstock.com/z/stock-vector-opinion-poll-flat-illustration-of-two-groups-of-people-and-speech-bubbles-between-them-349435901.jpg

19

References

1. J.L. Devore and K.N.Berk, Modern Mathematical Statistics with Applications, Springer, 2012.

20