Chapter 9-10

Confidence Intervals and Hypothesis Testing

Type 1 Error

Statistics

Mehmet Güray Güler, PhD

Last updated 22.07.2020

Testing a Statistical Hypothesis: Errors

- There are four possible situations as shown in the table below:

Decision ↓	H_0 is true	H_0 is false
Do not reject H_0	Correct decision	Type II error
Reject H_0	Type I error	Correct decision

- In two of them we make a correct decision.
- But we commit an error in two other situations.
- Hence there are two types of error when we reach a decision:

Testing a Statistical Hypothesis: Errors

Decision ↓	***H_0* is true**	***H_0* is false**
Do not reject *H_0*	**Correct decision**	**Type II error**
Reject *H_0*	**Type I error**	**Correct decision**

Type I error: Reject the null hypothesis H_0 when it is true.

Type II error: Do not reject H_0 when it is false.

A good hypothesis testing procedure should attain small probabilities for these two types of error.

Testing a Statistical Hypothesis: Errors

- Let's define:
  - α = P(type I error)
  - β = P(type II error)
- Please note that Type I error is sometimes called significance level
- Now consider the following example:

Type 1 Error (α)

- **Example:**
- You have a shop in Taksim and selling mobile phone accessories, especially phone cases.
- iPhone 11 is just released and you want to purchase a large amount of iPhone 11 cases to catch the new demand.

- You know that the height of the cases are pretty standard and there are no problems associated with it.

Type 1 Error (α)

- However from your previous experiences you know that there is a serious problem with the the width of the cases:
  - they can be very wide which means it stands loose and quickly gets out of the phone
  - they can be very tight which means it is hard (or impossible) to put the phone into it.
- You decided purchase **36** many cases at first, and make a test.
  - If the width of the cases are 68 mm, then you will purchase large amounts.
  - Otherwise you won't purchase.
- Assume that $\sigma = 3.6\, mm$ is given.

Type 1 Error (α)

- Assume one of your friends say that he did such analysis before and we should not buy the cases if the average width of the cases is below or above 1 mm or more from the targated value 68mm, that is he says do not buy if :
  - $\bar{X} \leq 67$ or
  - $\bar{X} \geq 69$

Type 1 Error (α)

- Therefore, the hypothesis of our concern is:
  - $H_0: \mu = 68\, mm$
  - $H_1: \mu \neq 68\, mm$
- If we reject H_0, then we won't purchase large amounts.
- If we don't reject H_0, we will purchase large amounts.
- We will reject the hypothesis if
  - $\bar{X} \leq 67$ or
  - $\bar{X} \geq 69$

Type 1 Error (α)

- Recall, in previous HT examples we defined α and then calculated the left and right most extreme values as our extremes.
- Here we are lucky, i.e., somebody has directly given us a critical regions.
- Maybe he has an alpha and calculated the critical region from this alpha directly, who knows?

Type 1 Error (α)

- Define Type 1 error here:
  - Do not buy the cases when their true width mean is 68 mm.
- We can calculate P(type I error).
- $\alpha = P(Type\ 1\ error) = P(Reject\ H_0 | H_0\ is\ true)$
- $\alpha =$ P(do not buy the cases | The width of them is 68 mm)
- $\alpha = P(\bar{X} < 67\ or\ \bar{X} > 69 | \mu = 68)$
- $\alpha = P(\bar{X} < 67 | \mu = 68) + P(\bar{X} > 69 | \mu = 68)$ [Using CLT]
- $\alpha = P\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{67-\mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu = 68\right) + \left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{69-\mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu = 68\right)$

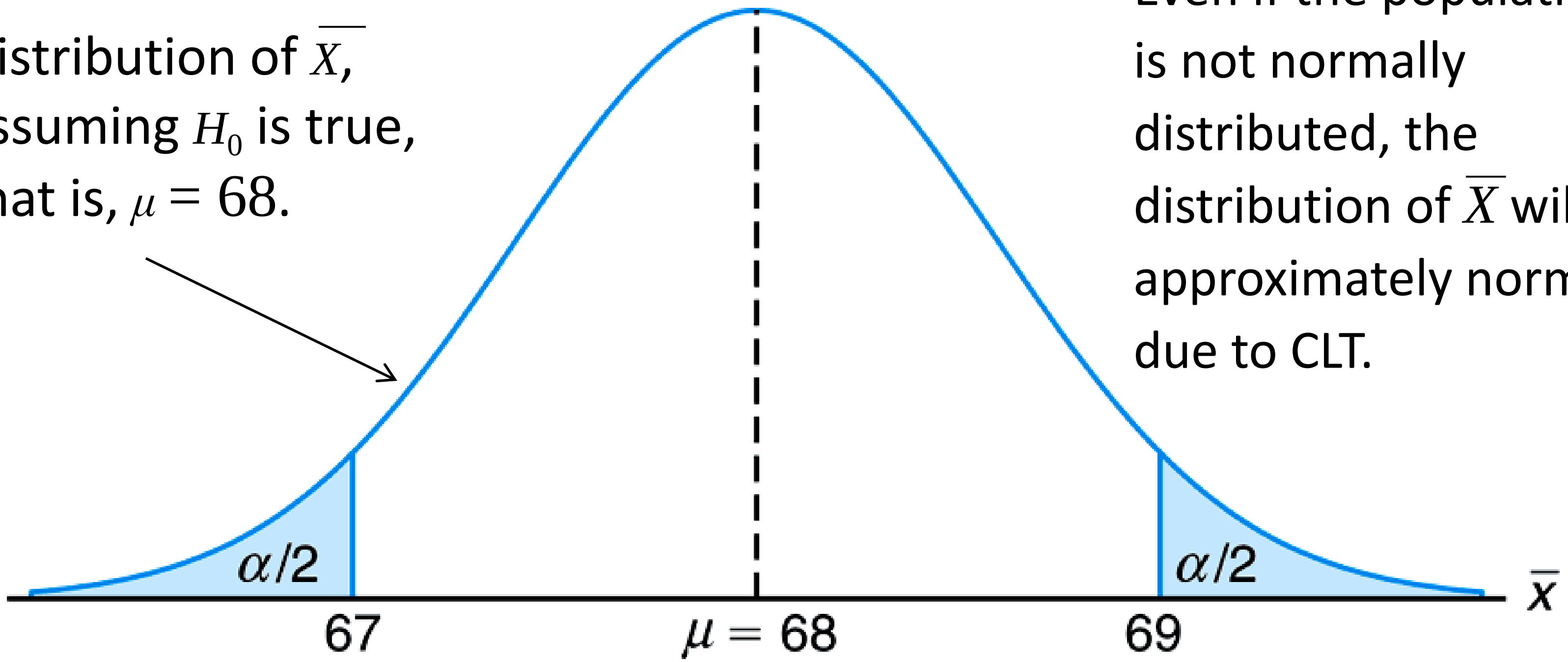
Type 1 Error (α)

- $\alpha = P\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < \frac{67-\mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu = 68\right) + P\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} > \frac{69-\mu}{\frac{\sigma}{\sqrt{n}}} \middle| \mu = 68\right)$
- $\alpha = P\left(\frac{\bar{X}-\mu}{0.6} < \frac{67-\mu}{0.6} \middle| \mu = 68\right) + P\left(\frac{\bar{X}-\mu}{0.6} > \frac{69-\mu}{0.6} \middle| \mu = 68\right)$
- $\alpha = P\left(\frac{\bar{X}-\mu}{0.6} < \frac{67-\mu}{0.6} \middle| \mu = 68\right) + P\left(\frac{\bar{X}-\mu}{0.6} > \frac{69-\mu}{0.6} \middle| \mu = 68\right)$
- $\alpha = P\left(\frac{\bar{X}-\mu}{0.6} < \frac{67-68}{0.6} \middle| \mu = 68\right) + P\left(\frac{\bar{X}-\mu}{0.6} > \frac{69-68}{0.6} \middle| \mu = 68\right)$
- α = P(Z<-1.67) + P(Z>1.67)
- $\alpha = 0.095$

Type 1 Error

The probability of committing a type I error (or the significance level of our test) is equal to the sum of the areas shaded below (for testing μ=68 against μ≠68).

Distribution of $\overline{X}$, assuming H_0 is true, that is, $\mu = 68$.



Even if the population is not normally distributed, the distribution of $\overline{X}$ will be approximately normal, due to CLT.

Type 1 Error (α)

- This means,
  - if we keep taking random samples of size n=36 from this cases, and if the width mean is indeed 68 mm, 9.5% of time we would reject the (correct) H0 and ***don't buy the cases even if their true mean is 68 mm***.
- Well, can we reduce this error?
  - Take more samples!!!
- Try the same example with n=64!

Type 1 Error (α)

- **Example (cont.d):** what is type 1 error if we make the same test with n=64 many samples?
- **Solution:** Now we have $\frac{\sigma}{\sqrt{64}} = 0.45$
- Hence

$$\alpha = P\left(\frac{\bar{X}-\mu}{0.45} < \frac{67-68}{0.45} \middle| \mu = 68\right) + \left(\frac{\bar{X}-\mu}{0.45} > \frac{69-68}{0.45} \middle| \mu = 68\right)$$
$$\alpha = 0.0264$$

- Compare with 0.095, much more smaller!!!

Type 1 Error (α)

- Any other way to decrease type 1 error? *(other than increasing n)*
- We may change the limits of the critical regions.
- For example change them to
  - $\bar{X} > 73$
  - $\bar{X} < 63$
- . Is that OK?
  - Then, you make almost no error if you reject H0
  - Calculate!
  - $\alpha = P\left(\frac{\bar{X}-\mu}{0.45} < \frac{63-68}{0.45} \middle| \mu = 68\right) + \left(\frac{\bar{X}-\mu}{0.45} > \frac{73-68}{0.45} \middle| \mu = 68\right)$ = 0 Any problem?