Chapter 9-10

Confidence Intervals and Hypothesis Testing

Type 2 Error, Power

Statistics

Mehmet Güray Güler, PhD

Last updated 22.07.2020

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?

Type 2 Error (β)

- However, what happens if you don't reject H_0
- Then you purchase the cases whose
  - μ is 66 mm
  - or whose μ is 70 mm.
- This is something you don't want!!!
- Let's illustrate this

Type 2 Error (β)

- Consider the case where you decided that $\mu = 68mm$ even if the real parameter is $\mu = 70mm$
- Hence you decided to purchase a large amount of these cases.
- Find the probability of this situation, i.e.
  - P(buy a large amount of cases when in fact $\mu = 70\ mm$)
  - =P(Do not reject H0 $|\mu = 70$)
  - =P($63 \leq \bar{X} \leq 73\ |\mu = 70\ mm$)

Type 2 Error (β)

- $= P(63 \le \bar{X} \le 73 \mid \mu = 70)$
- $= P\left(\frac{63-\mu}{0.45} \le \frac{\bar{X}-\mu}{0.45} \le \frac{73-\mu}{0.45} \middle| \mu = 70\right)$
- $= P\left(\frac{63-\mu}{0.45} \le \frac{\bar{X}-\mu}{0.45} \le \frac{73-\mu}{0.45} \middle| \mu = 70\right)$
- $= P\left(\frac{63-70}{0.45} \le \frac{\bar{X}-70}{0.45} \le \frac{73-70}{0.45} \middle| \mu = 70\right)$
- $= P(-15.55 \le Z \le 6.66) = 1$
- This probability is 1!!!

Type 2 Error (β)

- Hence, if you change your rejection region from
  - $\bar{X} > 69 \text{ or } \bar{X} < 67$
    - to
  - $\bar{X} > 73 \text{ or } \bar{X} < 63$
- Then, if the real $\mu = 70\ mm$ and if you don't reject the hypothesis, then you are almost always making an error!
- This error is indeed **type 2 error**:
  - ***Do not reject H0 when it is not true***.
- Recall that $H_0: \mu = 68$ and we calculated the following:
  - β =P(Do not reject H0 $|\mu = 70$)

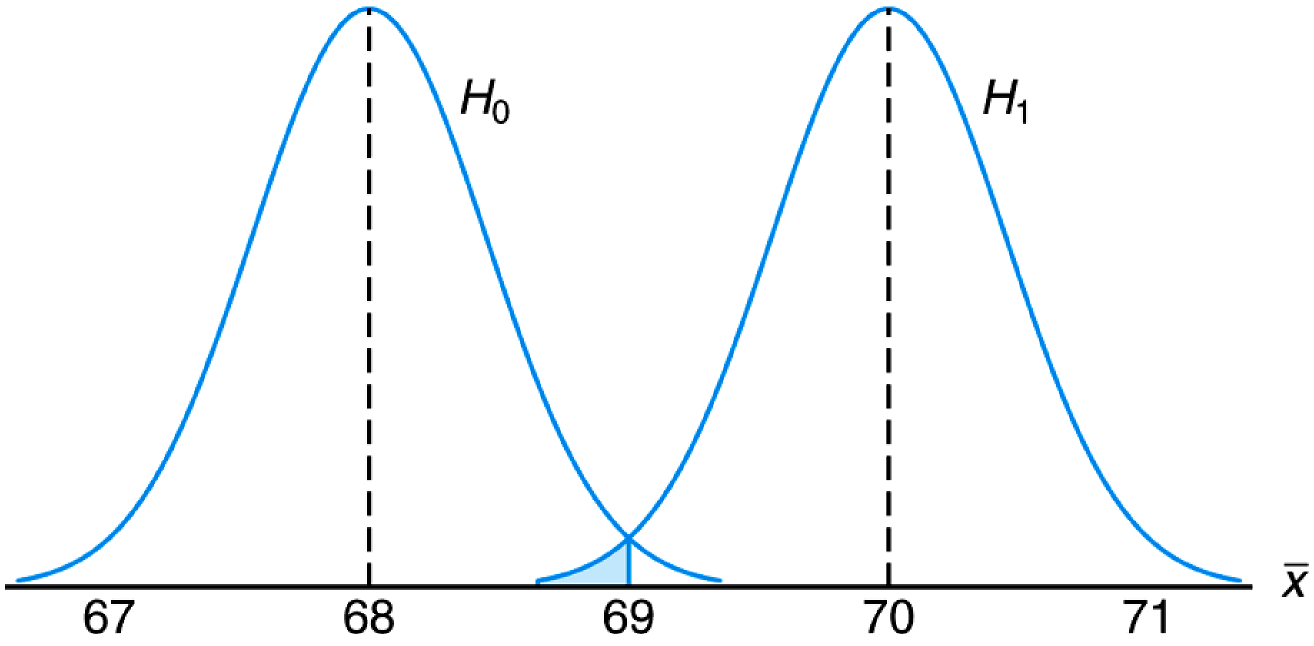
Type 2 Error (β)

- In order to calculate β, we need to specify a particular value for μ
  - $''H_0$ is not true'' is not sufficient to calculate a probability.
  - In the previous example we calculate it for $\mu = 70$
- Hence we define β as $\beta(\mu)$ in order to indicate that it depends on the value of selected μ.

- Set the critical region to original one ($\bar{X} > 69 \text{ or } \bar{X} < 67$) and calculate β for the following:
  - the **true mean is $\boldsymbol{\mu = 70}$**
  - the **true mean is $\boldsymbol{\mu = 66}$**
  - the **true mean is $\boldsymbol{\mu = 68.5}$**

Probability of a Type II Error

Let's now calculate β, the probability of committing a type II error, assuming that the true value of $\mu = 70$, for the case n=64.



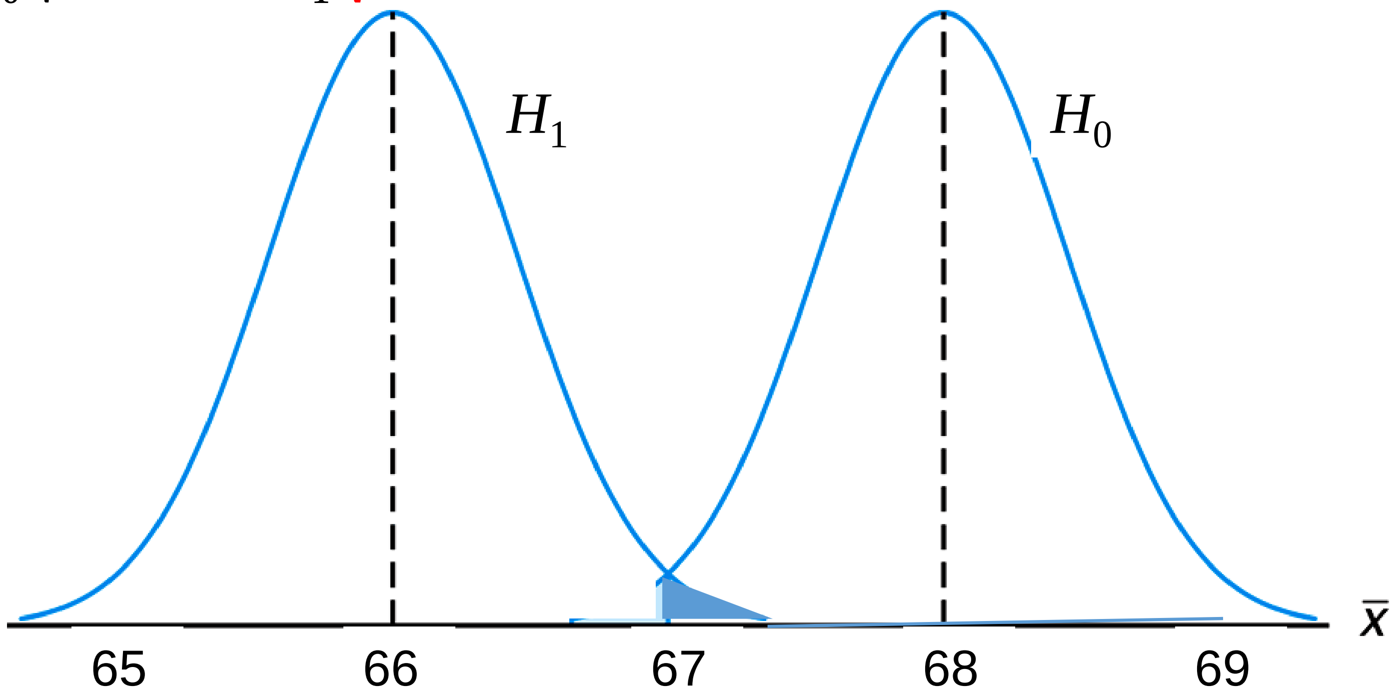
$$\beta = P(\text{type II error} | \mu = 70) = P(67 \le \bar{X} \le 69 | \mu = 70)$$

$$= P(\frac{67-70}{0.45} \le Z \le \frac{69-70}{0.45}) \approx P(-6.67 \le Z \le -2.22) \approx 0.0132.$$

Probability of a Type II Error

Now calculate it for $\mu = 66$, that is:

$H_0: \mu = 68 \;\; vs \;\; H_1: \boldsymbol{\mu = 66} \;\; with \; n = 64.$

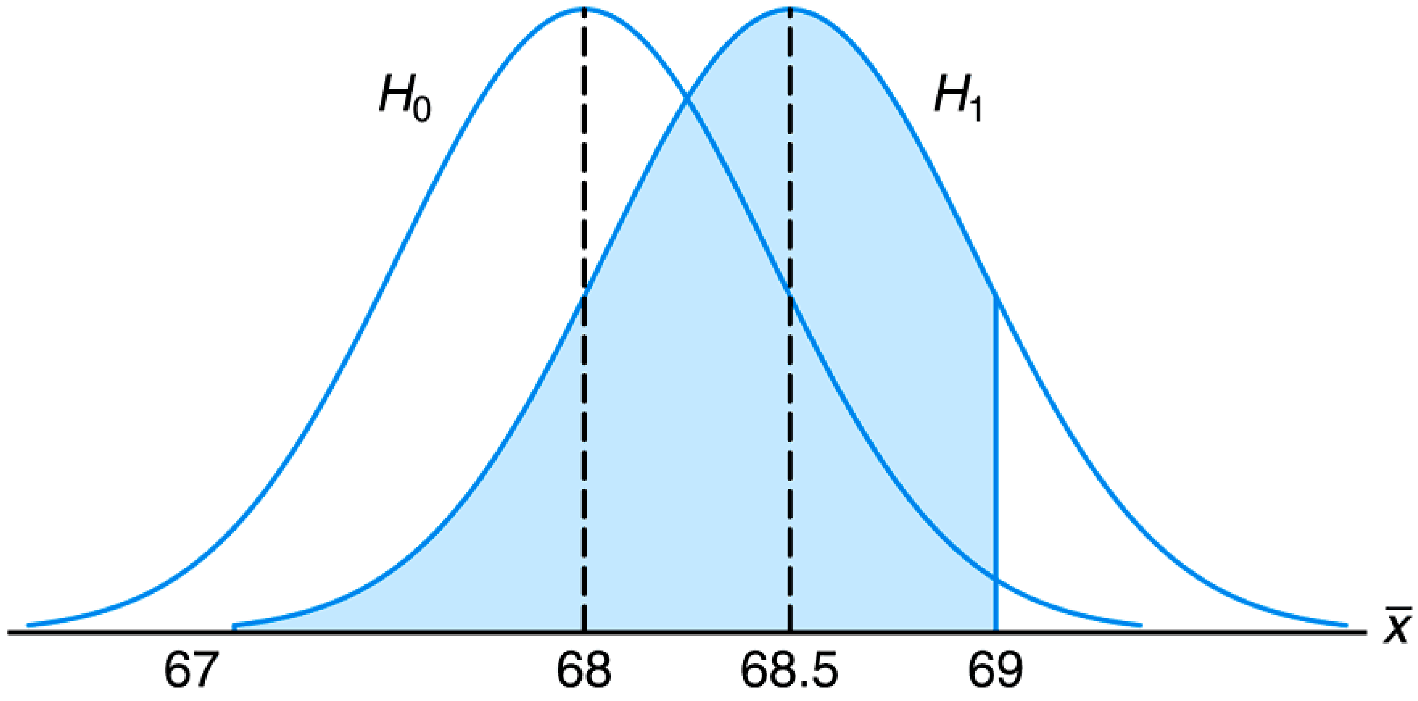


$$\beta(70) = P(\text{type II error} \mid \mu = 66) = P(67 \le \bar{X} \le 69 \mid \mu = 66)$$

$$= P\left(\frac{67-66}{0.45} \le Z \le \frac{69-66}{0.45}\right) \approx P(2.22 \le Z \le 6.67) \approx 0.0132.$$

Probability of a Type II Error

$$H_0: \mu = 68 \quad vs. \quad H_1: \mu = 68.5$$



$$\beta(68.5) = P(\text{type II error} \mid \mu = 68.5) = P(67 \le \bar{X} \le 69 \mid \mu = 68.5)$$

$$= P\left(\frac{67 - 68.5}{0.45} \le Z \le \frac{69 - 68.5}{0.45}\right) \approx P(-3.33 \le Z \le 1.11) \approx 0.866.$$

Type 2 Error (β)

- We can generalize this to see that β values will be the same when μ is *equal distance* from 68 ($H_0: \mu = 68$).
- We also note that β
  - becomes smaller as $|\mu - 68|$ increases, i.e., μ gets away from 68.
  - becomes larger as $|\mu - 68|$ decreases, i.e., μ gets closer to 68.
- This is understandable as it becomes more and more difficult to **identify the difference** between two means when $|\mu - 68|$ nears zero, causing the probability of a type II error to increase.

Type 1 and Type 2 Errors

- For fixed sample size, when one error type increases, the other error type decreases.
- An increase in the sample size n will reduce both α and β.

Testing a Statistical Hypothesis: Errors

- Good to control the probability of these two types of errors.
- Unfortunately, for a given sample size, we can NOT control both
- We use a hypothesis testing procedure that controls the probability of a type I error and guarantees that it cannot exceed a given small probability α.

Testing a Statistical Hypothesis: Errors

- **Example (Cont.d):** For the previous example, instead of the critical region $\bar{X} < 67$ and $\bar{X} > 69$, what is probability of type 1 error if we are given $\alpha = 0.05$.

- **Solution**
- Let the critical region be $\bar{X} < c_1$ and $\bar{X} > c_2$

- First find the critical region given $\alpha = 0.05$, i.e.,
  - Find c_1 and c_2 such that, given H0 is true, the following must hold:
  - $P(\bar{X} < c_1) + P(\bar{X} > c_2) = 0.05$

Testing a Statistical Hypothesis: Errors

- Since they are symmetric we have
  - $P(\bar{X} > c_2) = 0.025$
- Using CLT and $\sigma_{\bar{X}} = 0.45$ we have
  - $c_2 = \mu + z_{0.025} \times 0.45 = 68.74$
- Similarly we can find $c_1 = 67.26$
- Now calculate $P(type\ 1\ error)$
  - $= P(reject\ H_0 | H_0 is\ true)$
  - $= P(\bar{X} > 68.74\ or \bar{X} < 67.26 | \mu = 68)$
  - $= P(\bar{X} > 68.74 | \mu = 68) + P(\bar{X} < 67.26 | \mu = 68)$
  - =0.025 + 0.025 = 0.05

Testing a Statistical Hypothesis: Errors

- *Well, in fact α value that we used since the beginning is nothing but type 1 error.*
- Now calculate β when the real $\mu = 70$
- $\beta(70) = \; P(67.26 < \bar{X} < 68.74|\mu = 70)$
- Using CLT and $\sigma = 0.45$ we have
- $= P\left(\frac{67.26-\mu}{0.45} \leq \frac{\bar{X}-\mu}{0.45} \leq \frac{68.74-\mu}{0.45} \middle| \mu = 70\right)$
- $= P\left(\frac{67.26-70}{0.45} \leq \frac{\bar{X}-70}{0.45} \leq \frac{68.74-70}{0.45}\right) = 0.0026$
- Hence probability of type 2 error is very small for this example.

Power of a Test

- DEFINITION. $1-\beta$ = Power of the test

  $= P(\text{Reject } H_0 \text{ when } H_0 \text{ is false})$

- Power of the test = The probability of **correctly** rejecting a false H_0.
- It is the power of the test to differentiate between two hypotheses
  - If two hypotheses are close, then the power will be low.

Power of a Test

- $1-\beta$ = Power of the test = P(Reject H_0 when H_0 is false)
- A good testing procedure should have

1. A small level of significance (α value), and
2. A large power ($1-\beta$ value) when the deviation from the value specified by H_0 is significantly large.

- For the previous example,the power of the test is $1-\beta=0.9974$

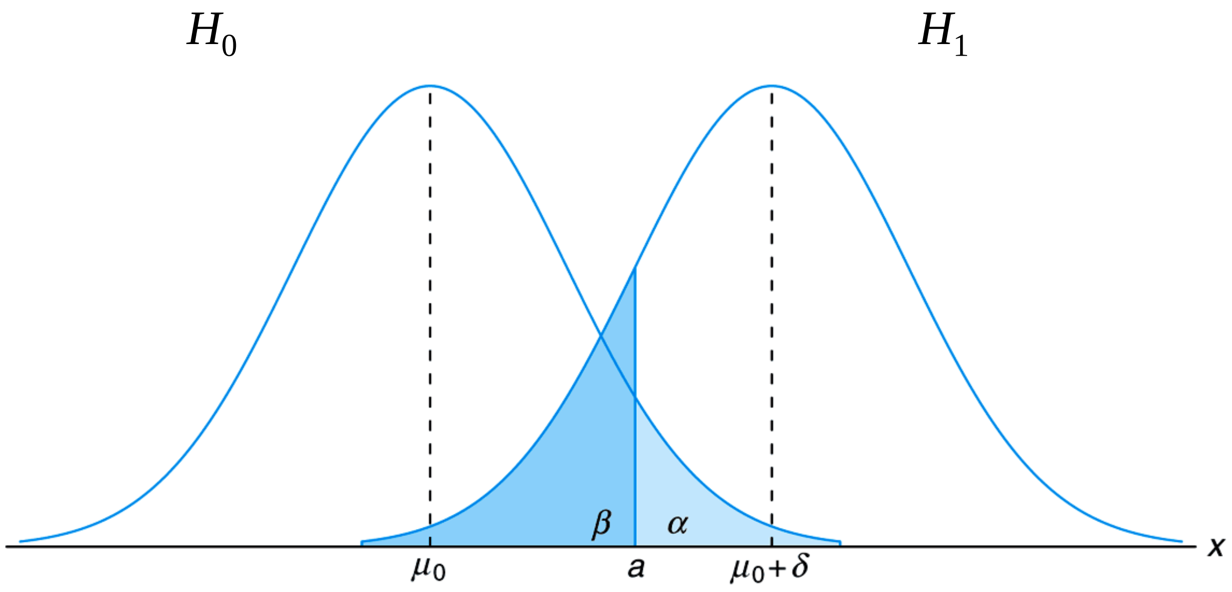
Power of a Test

- $1-\beta$ = Power of the test = P(Reject H_0 when H_0 is false)
- You should now practice the calculation of the power of this test for other values of μ under H_1.
- In particular calculate $1-\beta$(67.5) to see how it is done in EXCEL.
- EXERCISE. Prepare a simple EXCEL sheet for calculating the power of the test in Example 2 for many different values of of μ under H_1. Then graph the power function $1-\beta(\mu)$ against μ values.

Testing $\mu = \mu_0$ versus $\mu = \mu_0 + \delta$

- When testing $H_0\!: \mu = \mu_0$ against $H_1\!: \mu = \mu_0 + \delta$, for $\delta > 0$ when σ is given, we can use the Z-test.
- For given α, β, δ the needed sample size is: $$n \approx \frac{(z_\alpha + z_\beta)^2 \sigma^2}{\delta^2}$$

Sometimes we need the number of experiments to be made for a given value of type 1 and type 2 error values. For the **single** mean case, tt is given by the above formula.



Testing $\mu_1 - \mu_2 = d_0$ versus $\mu_1 - \mu_2 = d_0 + \delta$

- When testing $H_0: \mu_1 - \mu_2 = d_0$ against $H_1: \mu_1 - \mu_2 = d_0 + \delta$, when σ is given.
- we can use the Z-test, taking the sample size as: $n \text{ » } \frac{(z_a + z_b)^2 (s_1^2 + s_2^2)}{d^2}$

Sometimes we need the number of experiments to be made for a given value of type 1 and type 2 error values. For the **two population** case, tt is given by the above formula.