Type I and Type II Errors in Hypothesis Testing

In §6.2 and §6.3 we learned hypothesis test for **making decisions** on **Null hypothesis** H_0 Versus **alternative hypothesis** H_1.

Decision:

When we make the decision, it is possible to have errors.

Decisions \ Facts	H_0 is True	H_1 is True
Reject H_0	**Type I Error**	Correct Decision
Fail to reject H_0	Correct Decision	**Type II Error**

Definition.

The **probability** of committing a **Type I error** is the **significance level** α:

$$\begin{aligned}\alpha &= P(H_0 \text{ is Rejected} \mid H_0 \text{ is True})\\ &= P(H_0 \text{ is Rejected} \mid \mu = \mu_0)\end{aligned}$$

We call $1-\alpha$ the **confidence level.**

Definition.

The **probability** of committing a **Type II error** is

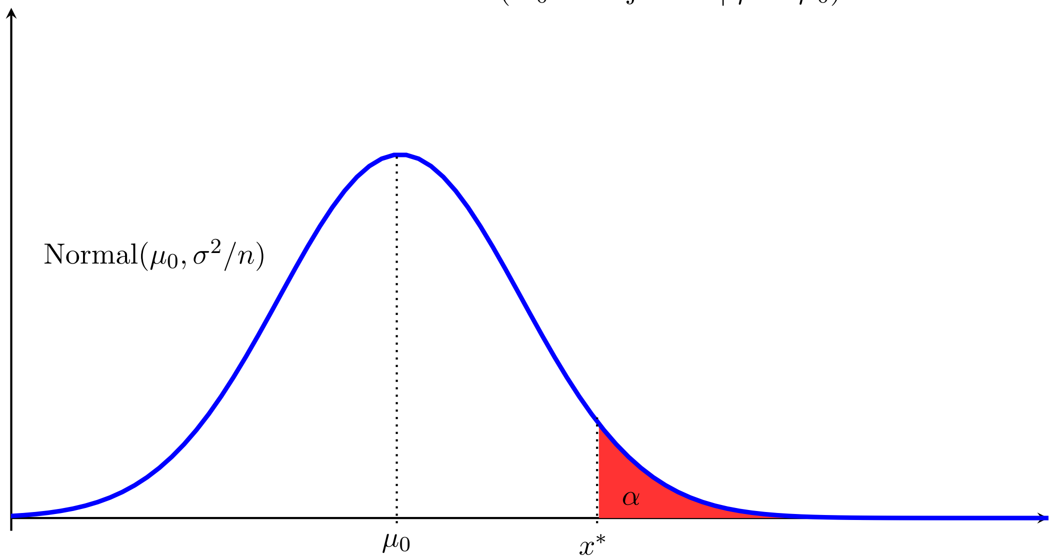
$$\begin{aligned}\beta &= P(H_0 \text{ is Accepted} \mid H_1 \text{ is True})\\ &= P(H_0 \text{ is Accepted} \mid \mu = \mu_1)\end{aligned}$$

We call $1-\beta$ the **power of the test.**

Graph for the one-sided test:

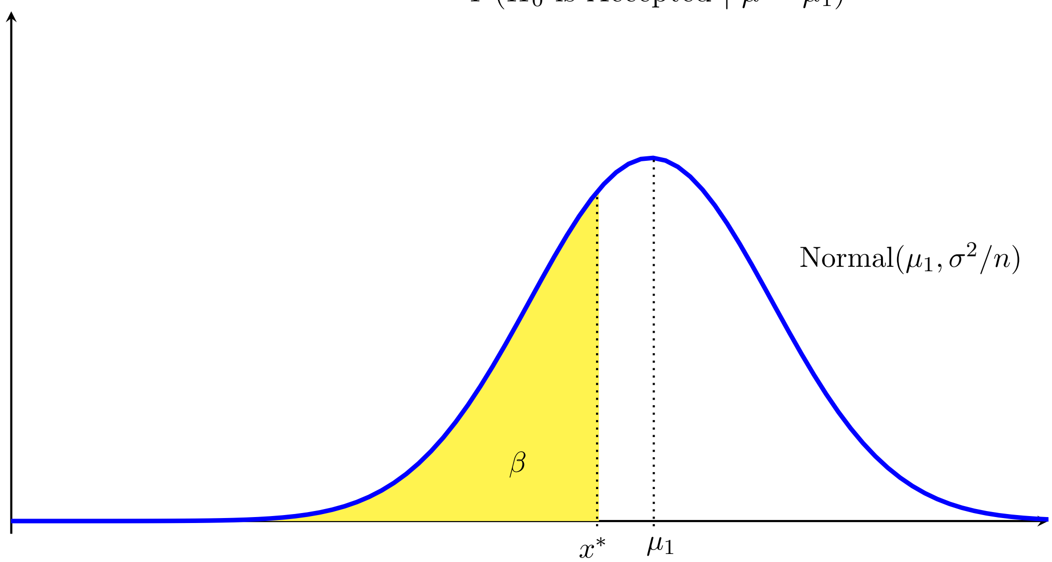
Type I error:

$$\begin{aligned}\alpha &= P(H_0 \text{ is Rejected} \mid H_0 \text{ is True})\\ &= P(H_0 \text{ is Rejected} \mid \mu = \mu_0)\end{aligned}$$



Type II error:

$$\begin{aligned}\beta &= P(H_0 \text{ is Accepted} \mid H_1 \text{ is True})\\ &= P(H_0 \text{ is Accepted} \mid \mu = \mu_1)\end{aligned}$$



Example 1. There is a sample of size 100 from a normal distribution with standard derivation $\sigma = 25$. We are testing $H_0 : \mu = \mu_0 = 134$ versus $H_1 : \mu > 134$. We will reject H_0 if the sample mean $\overline{X} \geq 140$.

(1.) Find the probability of committing a Type I error.

Solution: If H_0 is true, by CLT, $\overline{X} \sim \text{Normal}(134, \frac{25^2}{100})$. Hence,

$$\begin{aligned} \alpha &= P(H_0 \text{ is Rejected} \mid H_0 \text{ is True}) = P(\overline{X} \geq 140 \mid \mu = \mu_0) \\ &= \textbf{normalcdf}(140, \infty, 134, 25/\sqrt{100}) \approx 0.0082 \end{aligned}$$

(2.) Suppose the true $\mu = \mu_1 = 139$. Find the probability of committing a Type II error and the power of the test.

If H_1 is true, by CLT, $\overline{X} \sim \text{Normal}(139, \frac{25^2}{100})$. Hence,

$$\begin{aligned} \beta &= P(H_0 \text{ is Accepted} \mid H_1 \text{ is True}) = P(\overline{X} \leq 140 \mid \mu = \mu_1) \\ &= \textbf{normalcdf}(-\infty, 140, 139, 25/\sqrt{100}) \approx 0.6554 \end{aligned}$$

The power of the test is $1 - \beta = 0.3446$.

Example 2. The average score of a standardized math test for 5th graders in Massachusetts is 134. The school is experimenting with a new method of teaching. A committee has been set up to decide if the average score has remained the same (null hypothesis H_0 or increased (alternative hypothesis H_1). The committee members decided that they would reject the null hypothesis if a sample of 100 students gives a mean score of more than 140. They assume that the score distribution is normal with standard derivation 25. (1.) Find the probability that the committee making a wrong decision. (2.) Suppose the true mean after the change is 139. Find the probability of committing a Type II error and the power of the test.

Remark: Solution is the same as Example 1. we find that the decision rule gives rise to a tiny Type I error α, but a huge Type II error β and a small power. The committee should consider to lower the critical value, and increase the sample size. (e.g., $x^* = 138$ (example 3), or $n = 500$)

Example 3. There is a sample of size 100 from a normal distribution with standard derivation $\sigma = 25$. We are testing $H_0 : \mu = \mu_0 = 134$ versus $H_1 : \mu > 134$. We will reject H_0 if the sample mean $\overline{X} \geq 138$.

(1.) Find the probability of committing a Type I error.

Solution: If H_0 is true, by CLT, $\overline{X} \sim \text{Normal}(134, \frac{25^2}{100})$. Hence,

$$\alpha = P(H_0 \text{ is Rejected} \mid H_0 \text{ is True}) = P(\overline{X} \geq 138 \mid \mu = \mu_0)$$
$$= \textbf{normalcdf}(138, \infty, 134, 25/\sqrt{100}) \approx 0.0548$$

(2.) Suppose the true $\mu = \mu_1 = 139$. Find the probability of committing a Type II error and the power of the test.

If H_1 is true, by CLT, $\overline{X} \sim \text{Normal}(139, \frac{25^2}{100})$. Hence,

$$\beta = P(H_0 \text{ is Accepted} \mid H_1 \text{ is True}) = P(\overline{X} \leq 138 \mid \mu = \mu_1)$$
$$= \textbf{normalcdf}(-\infty, 138, 139, 25/\sqrt{100}) \approx 0.3446$$

The power of the test is $1 - \beta = 0.6554$.

Example 4. People in ages 20–50 have blood pressure with mean 120 mm Hg with standard deviation of 12 mm Hg. A team in a medicine company claims that they have discovered a new medicine that can maintain patients' blood pressure. The company test 100 patients for the new medicine. The company decide to reject the Null hypothesis, if the average blood pressure is higher than 126 or lower than 116.

(Hint: Two sided Test $H_0 : \mu = 120$ versus $H_1 : \mu \neq 120$.)

(1.) Find the probability of committing a Type I error.

Solution: If H_0 is true, by CLT, $\overline{X} \sim \text{Normal}(120, \frac{12^2}{100})$. Hence,

$$\begin{aligned}\alpha &= P(H_0 \text{ is Rejected} \mid H_0 \text{ is True}) = P(\overline{X} \leq 116 \text{ or } \overline{X} \geq 126 \mid \mu = \mu_0) \\ &= P(\overline{X} \leq 116 \mid \mu = \mu_0) + P(\overline{X} \geq 126 \mid \mu = \mu_0) \\ &= \textbf{normalcdf}(\infty, 116, 120, 12/\sqrt{100}) + \textbf{normalcdf}(126, \infty, 120, 12/\sqrt{100}) \\ &\approx 4.29 \times 10^{-4} + 2.87 \times 10^{-7}\end{aligned}$$

(2.) Suppose the true $\mu = \mu_1 = 124$. Find the probability of committing a Type II error and the power of the test.

If H_1 is true, by CLT, $\overline{X} \sim \text{Normal}(124, \frac{12^2}{100})$. Hence,

$$\begin{aligned}\beta &= P(H_0 \text{ is Accepted} \mid H_1 \text{ is True}) = P(116 \leq \overline{X} \leq 126 \mid \mu = \mu_1) \\ &= \textbf{normalcdf}(116, 126, 124, 12/\sqrt{100}) \approx 0.9522\end{aligned}$$

The power of the test is $1 - \beta = 0.0478$.

(3.) Suppose the true $\mu = \mu_1 = 127$. Find the probability of committing a Type II error.

$$\beta = \textbf{normalcdf}(116, 126, 127, 12/\sqrt{100}) \approx 0.2$$