HYPOTHESIS TESTS - CONCEPT

We are interested in the population, but the sample is in our hands.

Some assumption is made on the population (e.g. the value of μ or σ), and this assumption is accepted or rejected based on the data.

May the data come from a distribution …? E.g. $\mu = \mu_0$?

$H_0 : \mu = \mu_0$ — Null hypothesis

$H_1 : \mu \neq \mu_0$ — Alternative hypothesis

Z-TEST

$\mathrm{H}_0 : \mu = \mu_0$ $\qquad$ $\mathrm{H}_1 : \mu \neq \mu_0$

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \longrightarrow z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

test statistic

If H_0 is true, $z_0 = z$

If z_0 takes its value in the usual range, H_0 is accepted.

What is the „usual range"?

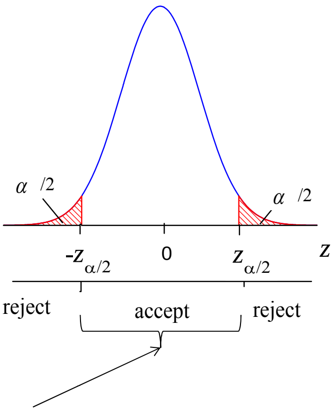
Where z can be found with $1-\alpha$ probability. (E.g. $1-\alpha=0.95$)

$$P\left(-z_{\alpha/2} < z \leq z_{\alpha/2}\right) = 1 - \alpha$$

The equality is true if H_0 is true.

$$P\left(-z_{\alpha/2} < z_0 \leq z_{\alpha/2} \;\middle|\; \mathrm{H}_0\right) = 1 - \alpha$$

REGION OF ACCEPTANCE



$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

z can be found in this is the region with 1-α probability for sure.

If z_0 is in this interval we believe that z_0 is distributed as z, thus $\mu = \mu_0$

H_0 is accepted.

CONFIDENCE INTERVAL AND HYPOTHESIS TEST

$$P\left(-z_{\alpha/2} < \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \le z_{\alpha/2} \middle| \mathrm{H}_0 \right) = 1 - \alpha$$

$$P\left(\bar{x} - z_{\alpha/2}\sigma / \sqrt{n} < \mu_0 \le \bar{x} + z_{\alpha/2}\sigma / \sqrt{n} \right) = 1 - \alpha$$

confidence interval for μ :

$$P\left(\bar{x} - z_{\alpha/2}\sigma / \sqrt{n} < \mu \le \bar{x} + z_{\alpha/2}\sigma / \sqrt{n} \right) = 1 - \alpha$$

If the confidence interval contains the hypothesised μ_0 value,
H_0 is accepted.

Example 10

The mass of an object is measured with 4 repeated measurements.
The sample mean is 5.0125 g.
From historical data the variance is known as $\sigma^2 = 10^{-4} g^2$.
May we believe (based on the data) that the expected value (the true mass of the object if the balance is unbiased) is 5.0000 g?

$$\bar{x} = 5.0125, \quad \sigma^2 = 10^{-4}, \quad n = 4, \quad \alpha = 0.05$$

1. Formulate the null and alternative hypothesis

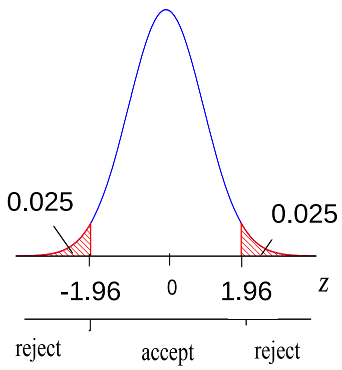
$$\mathrm{H}_0 : \mu = \mu_0 = 5.0000, \quad \mathrm{H}_1 : \mu \neq \mu_0 = 5.0000$$

2. Calculate the test statistic

$$z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{5.0125 - 5}{10^{-2} / 2} = 2.5$$

3. Find the critical values and define the reject/accept region

$$z_{a/2} = 1.96$$



4. Decision

As $2.5 = z_0 > z_{a/2} = 1.96$ we reject the null hypothesis (that the true mass of the object is 5.0000g) at α=0.05 significance level.

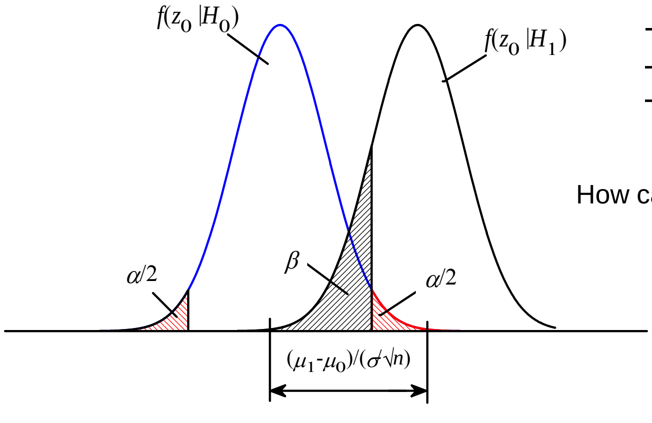
Error of first and second kind

	Decision	
	The H_0 hypothesis is	
	Accepted	Rejected
H_0 is true	Proper decision	Error of first kind (α)
H_0 is false	Error of second kind (β)	Proper decision

Power=1-β

"fail to reject"

Probability of comitting an error of second kind

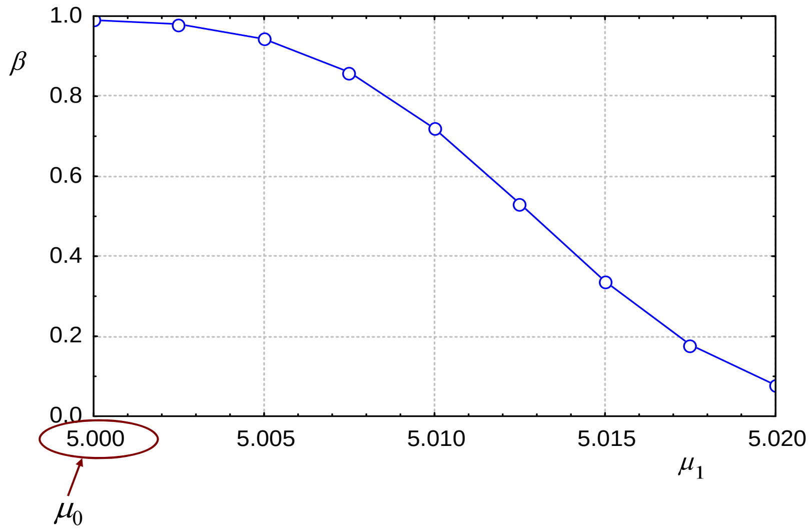


β depends on:

- α
- μ_1-μ_0
- n

How can we decrease beta?

OC (OPERATING CHARACTERISTIC) CURVE



Example 11

A raw materials' $CaSO_4$ concentration should be at least 85.0g/100g. The variance of the $CaSO_4$ content is 4.2 $(g/100g)^2$. The average $CaSO_4$ content of five 100g samples is 83.8g. We have to decide at α=0.05 significance level if the CaSO4 content reach the required 85.0g?

1. Formulate the null and alternative hypothesis

$$\mathrm{H}_0 : \mu \le \mu_0 = 85, \quad \mathrm{H}_1 : \mu > \mu_0 = 85$$

2. Calculate the test statistic

$$z_0 = \frac{\overline{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{83.8 - 85.0}{\sqrt{4.2} / \sqrt{5}} = -1.31$$

3. Find the critical values and define the reject/accept region

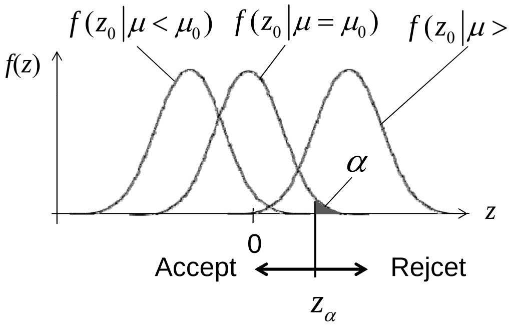
$$z_0 = \frac{\overline{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{\overline{x}-\mu}{\sigma/\sqrt{n}} + \frac{\mu-\mu_0}{\sigma/\sqrt{n}}$$

$$\mathrm{H}_0 : \mu \le \mu_0 = 85$$
$$\mathrm{H}_1 : \mu > \mu_0 = 85$$

$= z$ 0 $\mu = \mu_0$

<0 $\mu < \mu_0$

>0 $\mu > \mu_0$



$$z_\alpha = z_{0.05} = 1.65$$

4. Decision

$$z_0 = -1.31 < 1.65 = z_\alpha$$

z_0 is in the accept interval, thus we accept the null hypothesis.

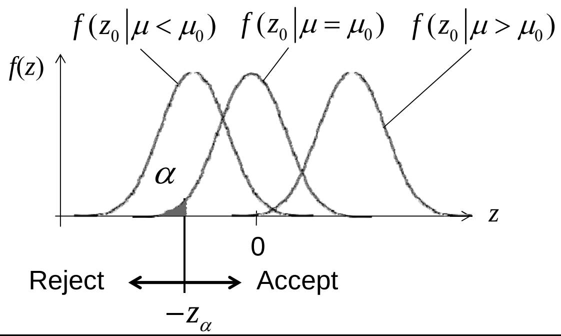
What is the decision about the raw material?

What if the null and alternative hypothesis is formulated in the opposite way?

$$\mathrm{H}_0' : \mu \ge \mu_0 = 85, \quad \mathrm{H}_1' : \mu < \mu_0 = 85$$

$$z_0 = -1.31$$

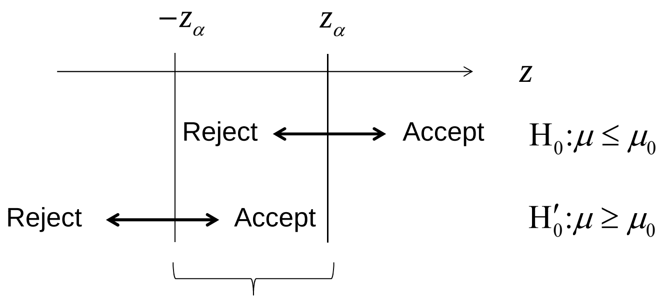
$$-z_\alpha = -z_{0.05} = -1.65$$



Decision

$$-z_\alpha = -1.65 < -1.31 = z_0$$

z_0 is in the accept interval, thus we accept the null hypothesis.
What is the decision about the raw material?



Both null hypotheses are accepted.

„FAIL TO REJECT" CONCEPT

„…it is customary to think of the decision to accept *H0 as a* **weak conclusion,** unless we know that beta is acceptably small. Therefore, rather than saying we "accept H_0," we prefer the terminology "fail to reject H_0." Failing to reject H_0 implies that we have not found sufficient evidence to reject H_0, that is, to make a strong statement. Failing to reject H_0 does not necessarily mean that there is a high probability that H_0 is true. It may simply mean that more data are required to reach a strong conclusion. This can have important implications for the formulation of hypotheses.

There is a useful analog between hypothesis testing and a jury trial. In a trial the defendant is assumed innocent (this is like assuming the null hypothesis to be true). If strong evidence is found to the contrary, the defendant is declared to be guilty (we reject the null hypothesis). If there is insufficient evidence the defendant is declared to be not guilty. This is not the same as proving the defendant innocent and so, like failing to reject the null hypothesis, it is a weak conclusion."

D.C. Montgomery and G.C. Runger: Applied Statistics and Probability for Engineers, 5th edition, 2011, Wiley

t-TEST

$\mathrm{H}_0 : \mu = \mu_0$ $\qquad$ $\mathrm{H}_1 : \mu \neq \mu_0$

$\mathrm{H}_0' : \mu \leq \mu_0$ $\qquad$ $\mathrm{H}_1' : \mu > \mu_0$

$\mathrm{H}_0'' : \mu \geq \mu_0$ $\qquad$ $\mathrm{H}_1'' : \mu < \mu_0$

$$t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \longrightarrow t_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$ test statistic

If H_0 is true, $t_0 = t$

If t_0 takes its value in the

$(-t_{a/2}, t_{a/2})$	$\mathrm{H}_0 : \mu = \mu_0$
$(-\infty, t_\alpha)$	$\mathrm{H}_0' : \mu \leq \mu_0$
$(-t_\alpha, \infty)$	$\mathrm{H}_0'' : \mu \geq \mu_0$

range, H_0 is accepted.

CHI-SQUARE-TEST

$\mathrm{H}_0 : \sigma^2 = \sigma_0^2$ $\qquad$ $\mathrm{H}_1 : \sigma^2 \neq \sigma_0^2$

$\mathrm{H}_0' : \sigma^2 \leq \sigma_0^2$ $\qquad$ $\mathrm{H}_1' : \sigma^2 > \sigma_0^2$

$\mathrm{H}_0'' : \sigma^2 \geq \sigma_0^2$ $\qquad$ $\mathrm{H}_1'' : \sigma^2 < \sigma_0^2$

$$\chi^2 = \frac{s^2 \nu}{\sigma^2} \longrightarrow \chi_0^2 = \frac{s^2 \nu}{\sigma^2}$$ test statistic

If H_0 is true, $\chi_0^2 = \chi^2$

If t_0 takes its value in the

$(\chi^2_{1-a/2}, \chi^2_{a/2})$	$\mathrm{H}_0 : \sigma^2 = \sigma_0^2$
$(0, \chi^2_\alpha)$	$\mathrm{H}_0' : \sigma^2 \leq \sigma_0^2$
$(\chi^2_{1-\alpha}, \infty)$	$\mathrm{H}_0'' : \sigma^2 \geq \sigma_0^2$

range, H_0 is accepted.

F-TEST

$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$ $\qquad$ $\mathrm{H}_1 : \sigma_1^2 \neq \sigma_2^2$

$\mathrm{H}_0' : \sigma_1^2 \leq \sigma_2^2$ $\qquad$ $\mathrm{H}_1' : \sigma_1^2 > \sigma_2^2$

$\mathrm{H}_0'' : \sigma_1^2 \geq \sigma_2^2$ $\qquad$ $\mathrm{H}_1'' : \sigma_1^2 < \sigma_2^2$

$$F = \frac{s_1^2 / \sigma_1^2}{s_2^2 / \sigma_2^2} \longrightarrow F_0 = \frac{s_1^2}{s_2^2}$$ test statistic

If H_0 is true, $F_0 = F$

If F_0 takes its value in the $\left(F_{1-a/2}, F_{a/2}\right)$ $\quad$ $\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2$

$\left(0, F_\alpha\right)$ $\quad$ $\mathrm{H}_0' : \sigma_1^2 \leq \sigma_2^2$

$\left(F_{1-\alpha}, \infty\right)$ $\quad$ $\mathrm{H}_0'' : \sigma_1^2 \geq \sigma_2^2$

range, H_0 is accepted.

TWO SAMPLE *t*-TEST

Two independent samples $\qquad n_1, n_2; \quad s_1^2, s_2^2; \quad \bar{x}_1, \bar{x}_2$

$$\mathrm{H}_0 : E\left(\bar{x}_1 - \bar{x}_2\right) = \mu_1 - \mu_2 = 0$$

Assuming the equality of variance for the two populations (to be checked through F-test):

$$\sigma_1^2 = \sigma_2^2$$

TWO SAMPLE *t*-TEST

$$\mathrm{H}_0 : \; E(\bar{x}_1 - \bar{x}_2) \; = \mu_1 - \mu_2 = 0 \qquad\qquad \mathrm{H}_1 : \mu_1 - \mu_2 \neq 0$$

$$t = \frac{\bar{x}_1 - \bar{x}_2 - E(\bar{x}_1 - \bar{x}_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad\qquad \nu = n_1 + n_2 - 2$$

$$s^2 = \frac{1}{n_1 + n_2 - 2}\left[s_1{}^2(n_1 - 1) + s_2{}^2(n_2 - 1)\right]$$

TWO SAMPLE *t*-TEST

$$\mathrm{H}_0 : \; E(\bar{x}_1 - \bar{x}_2) \; = \mu_1 - \mu_2 = 0 \qquad\qquad \mathrm{H}_1 : \mu_1 - \mu_2 \neq 0$$

The test statistic:

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \;, \qquad\qquad \nu = (n_1 - 1) + (n_2 - 1)$$

The assumption $\sigma_1^2 = \sigma_2^2$ is checked through *F* test

Example 12

(Box-Hunter-Hunter: Statistics for Experimenters, J. Wiley, 1978, p.97)

The wear of two kinds of raw material (A and B) is compared as shoe soles on the foot of 10-10 boys.

Is the difference of means significant at α=0.05 level?

	n	mean	sample variance
A	10	10.61	6.063
B	10	11.04	6.343

PAIRED *t*-TEST

$$\mathrm{H}_0 : E(x_i) = E(y_i) \qquad \mathrm{H}_1 : E(x_i) \neq E(y_i)$$

$$\mathrm{H}_0 :\ E(d_i) = E(x_i) - E(y_i) = 0$$

$$d_i = x_i - y_i$$

one-sample *t*-test for the differences

$$\bar{d} = \frac{\sum_i d_i}{n} \qquad s_d^2 = \frac{\sum_i \left(d_i - \bar{d}\right)^2}{n-1}$$

$$t_0 = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad \nu = n-1$$

Example 12

(Box-Hunter-Hunter: Statistics for Experimenters, J. Wiley, 1978, p.97)

The wear of two kinds of raw material (A and B) is compared as shoe soles on the foot of 10-10 boys.

Is the difference of means significant at α=0.05 level?

TABLE 4.3. Data on the amount of wear measured with two different materials A and B, boy's shoes example*

boy	material A	material B	B – A difference *d*
1	13.2(L)	14.0(R)	0.8
2	8.2(L)	8.8(R)	0.6
3	10.9(R)	11.2(L)	0.3
4	14.3(L)	14.2(R)	-0.1
5	10.7(R)	11.8(L)	1.1
6	6.6(L)	6.4(R)	-0.2
7	9.5(L)	9.8(R)	0.3
8	10.8(L)	11.3(R)	0.5
9	8.8(R)	9.3(L)	0.5
10	13.3(L)	13.6(R)	0.3
		average difference	0.41

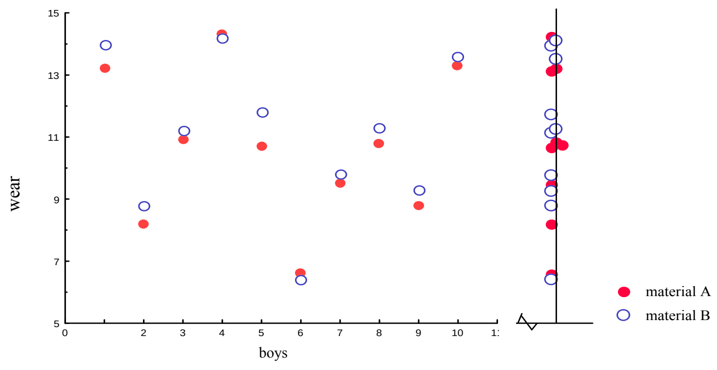


Figure 4.2. Data on two different materials A and B, used for making soles of boy’s shoes.