TEST OF HYPOTHESIS ABOUT A SINGLE MEAN

A hypothesis is a statement of belief, which may or may not be true about a population parameter.

The test of a hypothesis is a comparison of the statement of belief with a newly and objectively collected facts.

1) If the newly collected facts are shown to be consistent with the stated belief, the hypothesis is accepted.

2) If the newly collected facts do not support the statement of belief, the hypothesis is rejected.

H_0 : the null hypothesis, H_A : the alternative hypothesis

H_0 : the hypothesis to be tested

H_A : " " which cannot be tested.

Steps of Hypothesis Testing Procedure

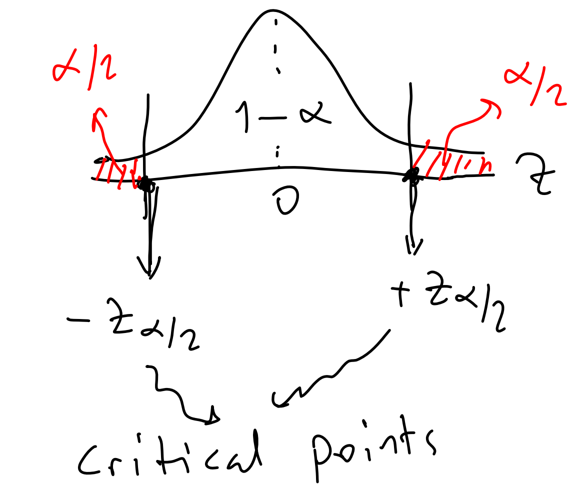
1. State the experimental goal (the problem)

2. Statement of hypothesis (write statements of H_0 and H_A)

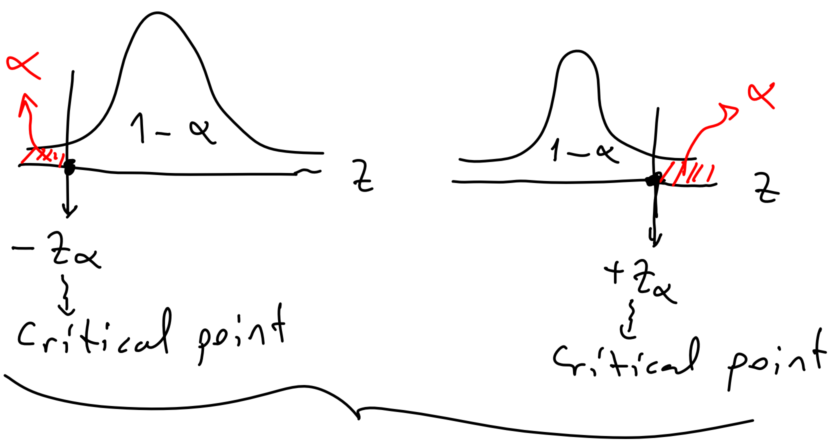
3. Decide on α (level of significance), $0 < \alpha < 1.0$

4. Use the statistic (e.g., Z, t, F, χ^2 distributions)

5. Define the critical region (use table values)



For a two-tailed test



One-tailed test

6. **Complete the computation of the test statistics (these are the calculated values: e.g., Z_{calc}, t_{calc}, F_{calc},…)**

7. **Make a statistical decision. Decision may be**

- accept H_0 (i.e., do not reject H_0 or fail to reject H_0)

OR

- reject H_0

8. **State the experimental conclusion (i.e., answer the question)**

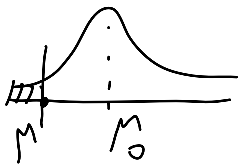
<u>Two-Tailed (Sided) Test of Hypothesis</u>

1) $H_0: \mu = \mu_0 \Rightarrow$ the true population parameter (μ) is equal to the hypothesized value (μ_0).

2) $H_A: \mu \neq \mu_0$

<u>One-Tailed (Sided) Test of Hypothesis</u>

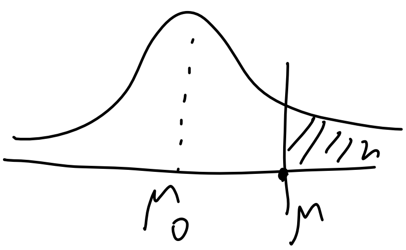
1) $H_0: \mu = \mu_0$
2) $H_A: \mu < \mu_0$



(left hand sided)

OR

1) $H_0: \mu = \mu_0$
2) $H_A: \mu > \mu_0$



(Right hand sided)

Test statistics

⊛ $Z_{calc} = \dfrac{\bar{X} - \mu \rightsquigarrow \mu_0}{\sigma/\sqrt{n}}$ ⇝ valid for the mean ($\bar{X}$) of n observations from any population when σ is known.

⊛ $Z_{calc} = \dfrac{\bar{X} - \mu \rightsquigarrow \mu_0}{S/\sqrt{n}}$ ⇝ valid for the mean ($\bar{X}$) of n observations from any population when σ is unknown and $n \geq 30$.

⊛ $Z_{calc} = \dfrac{X - \mu \rightsquigarrow \mu_0}{\sigma}$ ⇝ valid for a single value (x) from a population.

- The test statistic simply tells us in Standard deviation units (z) how far the observed mean ($\bar{x}$) is from the hypothesized mean.

- When the deviation of the observed sample mean from the hypothesized mean is small enough to be attributed to sampling chance alone (experimental variation) then the null hypothesis is accepted.

Otherwise, the null hypothesis is rejected and an alternative is accepted.

- To reject or accept a hypothesis we calculate the probability (P) of observing an outcome of the experiment.

- **If $P \leq \alpha$, then, H_0 is rejected and H_A is accepted.**
- **If $P > \alpha$, then, H_0 is accepted.**

$(1 - \alpha)$x100 % is called the confidence interval.

$(1 - \alpha)$ is the confidence coefficient or degree of confidence.

Confidence Interval (CI) For μ

$(1-\alpha)x100\ \% = CI$

A confidence interval for a parameter is an interval of numbers within which we expect the true value of the population parameter to be contained.

Confidence intervals are based on sample data and give a range of plausible values for a parameter (e.g., μ is a parameter).

Then, the confidence interval will be of the form :

$$\boxed{\text{CI : } \underbrace{\text{Point Estimate}}_{\bar{X},\ s} \pm \text{Margin of Error}}$$

e.g., $\bar{x}$ is the point estimator of μ.

Sample mean

pop. mean

	Mean	**St.dev.**
Population	μ	σ
Sample	$\bar{x}$	s

Variables about the population. We call them as parameter.

Variables about sample. We call them as statistics.

We use statistics to estimate parameters.

e.g., To estimate the population mean we use the sample mean, and so on.

$$\textbf{CI for } \mu : \quad \underbrace{\bar{x}}_{\text{point estimate}} \pm \underbrace{(z_{\alpha/2})*(\sigma/\sqrt{n})}_{\text{margin of error}}$$

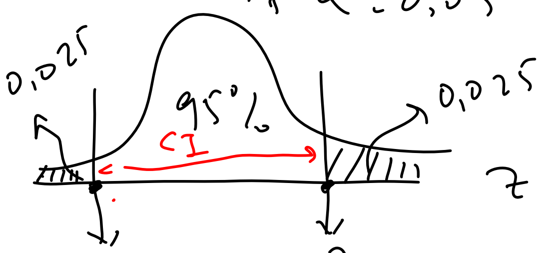
if $\alpha = 0{,}05 \Rightarrow (1-0{,}05)\times 100\% = 95\% = CI$

if $\alpha = 0{,}01 \Rightarrow (1-0{,}01)\times 100\% = 99\% = CI$

⊛ $100\% \times (1-\alpha) = CI$. If $\alpha = 0{,}05 \Rightarrow CI = 95\% \Rightarrow$

If μ is fixed, we say that μ will be in this interval with 95% confidency. We have 5% chance that our CI will not include the true mean value.

if $\alpha = 0{,}05$



lower confidence limit

upper confidence limit.

⊛ If $\bar{X}$ is the mean of a random sample of size n from a population with known σ^2, a $(1-\alpha)\times 100\%$ confidence interval for μ is given by

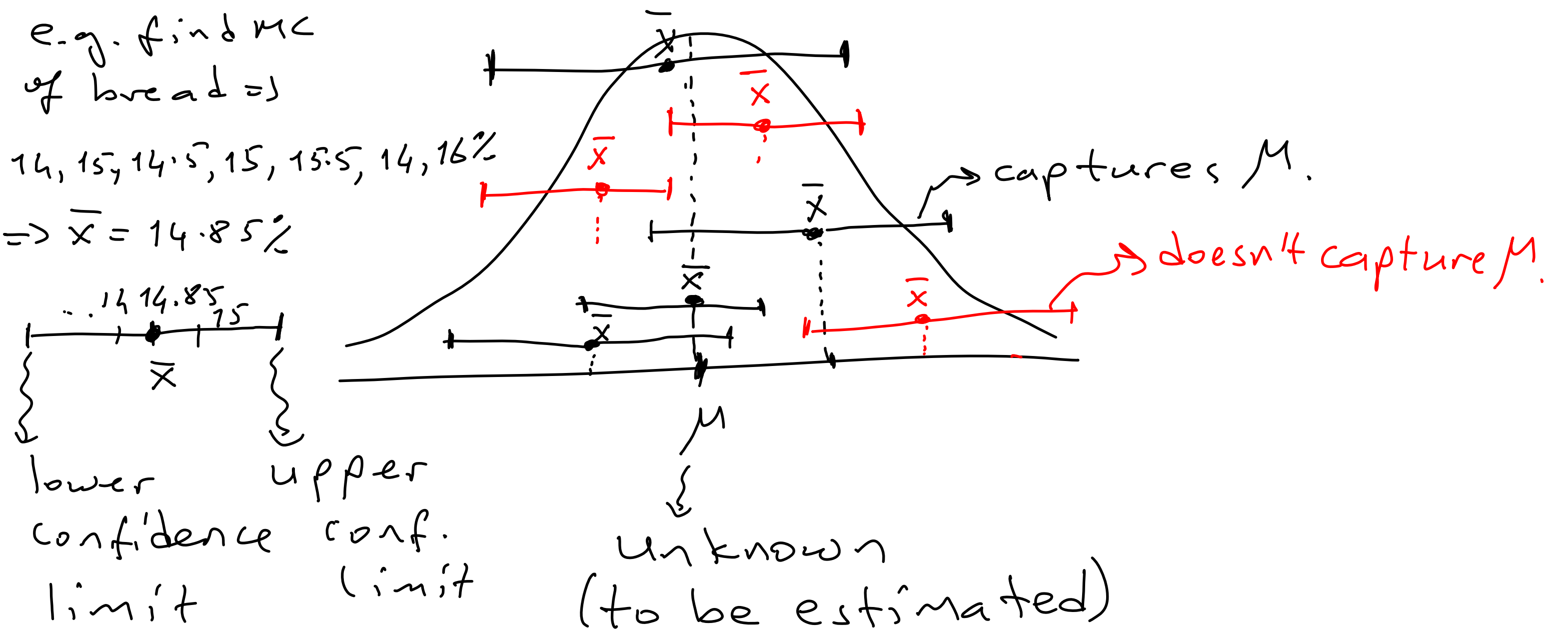
$$\boxed{\underbrace{\bar{X} - (z_{\alpha/2})\cdot \frac{\sigma}{\sqrt{n}}}_{\text{lower limit } L_1} < \mu < \underbrace{\bar{X} + (z_{\alpha/2})\cdot \frac{\sigma}{\sqrt{n}}}_{\text{upper limit } L_2}} \rightsquigarrow CI$$

here, $\mp z_{\alpha/2}$ is the z value leaving an area of $\alpha/2$ to the right and to the left.

We are trying to estimate μ with a CI.

I repeatedly draw several random samples from the population and calculate 95 % CI for μ for each sample.

- **Some of these intervals is going to capture mean μ and some of them are going to miss.**
- **Mathematically, we can say that 95 % of all these intervals will capture the mean μ and 5 % will miss.**



<u>Example</u>: Assume that we have a production factory that produces 100 mL bottles of cola. We want to estimate what is the average volume (mL) in every bottles that are produced in the factory.

Assume that the st.dev. of production (σ) is 6 mL. We take a sample of 36 bottles from the production line and notice that on the average they have 95 mL.

Can we assume that the population average is 94 mL or not ? i.e., calculate 95 % CI for μ.

<u>Solution:</u>

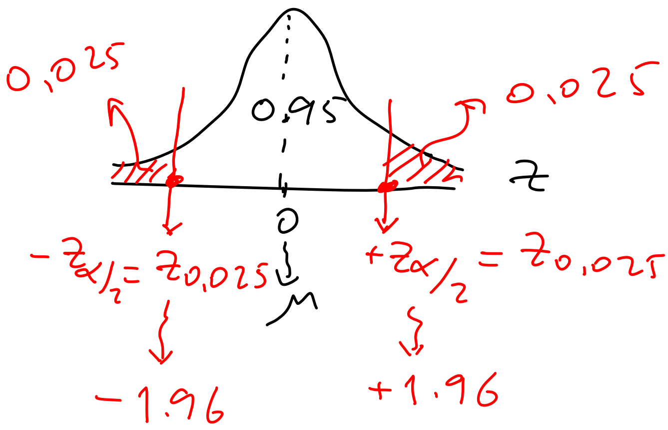
If CI is 95% $\Rightarrow \alpha = (1-0{,}95) = 0{,}05$

Find upper and lower z values $\Rightarrow$

$z_{\alpha/2} = ?$ $\quad z_{0{,}05/2} = z_{0{,}025} = ?$

it means what is the z value for $0{,}025$ area to the left and $0{,}025$ to right tails.

From z tables $\Rightarrow$ $z_{0{,}025} = \mp 1.96$



$$CI: \ \bar{X} - (z_{\alpha/2}) \times \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + (z_{\alpha/2}) \times \frac{\sigma}{\sqrt{n}} \Rightarrow$$

$\bar{X} = 95$ mL, $\sigma = 6$ mL, $n = 36$

$$95 - (1.96) \times \frac{6}{\sqrt{36}} < \mu < 95 + (1.96) \times \frac{6}{\sqrt{36}} \Rightarrow$$

95 % CI : $93.04 < \mu < 96.96$.

Yes, we can assume that population mean may be 94 mL. (i.e., we have 95 % chance that the mean of the population (cola bottles) is between 93.04 and 96.96 mL.

Example: Suppose that a new source of a cereal grain for the manufacturing of a certain product has became available. From the past records, the variance of protein content in the cereal grain is 2.25 % and it is hypothesized that protein content equals 12 %. 100 random samples are analyzed and found to have a mean protein content of 11.8 %.

Find either the protein content of new source of cereal grain is significantly different from 12 % at α = 0.05 significance level (i.e., can we use this new cereal source ?).

Solution:

Two methods

1) Compare P and α: If $P \le \alpha$, then, Reject Ho.

2) Compare z_{table} (table value) with z_{calc} (test statistic value).

Method 1:

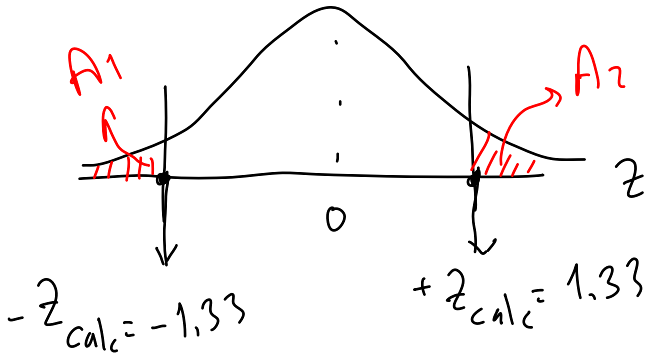
✓ $H_0: \mu = \mu_0 = 12\%$

$H_A: \mu \neq \mu_0 = 12\%$

} two-tailed

Test statistic value: z_{cal} =)

$$z_{calc} = \frac{\bar{x} - \mu_0}{(\sigma^2/n)^{1/2}} = \frac{11.8 - 12}{(2.25/100)^{1/2}} = -1.33$$



$A_1 = 0,0918$

$A_2 = 0,0918$

$P = 0,0918 + 0,0918$

$P = 0,1836$

Compare P and α $\Rightarrow$

$$\left.\begin{array}{l} P = 0,1836 \\ \alpha = 0,05 \end{array}\right\} \quad P > \alpha \Rightarrow \text{Accept } H_0$$

<u>Conclusion</u>: The protein content of 11.8% is not different significantly from 12% at $\alpha = 0,05$ level.

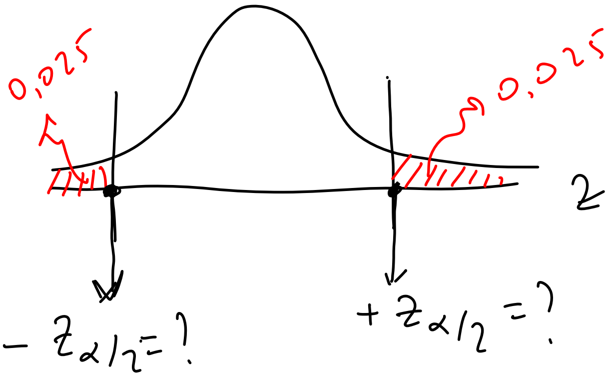
Method 2

Compare z_{calc} and z_{table} on the distribution curve.

$$\left.\begin{array}{l} H_0: \mu = \mu_0 = 12\% \\ H_A: \mu \neq \mu_0 = 12\% \end{array}\right\} \text{two-tailed.}$$

Test statistic value: z_{calc} $\Rightarrow$

$z_{calc} = -1.33$

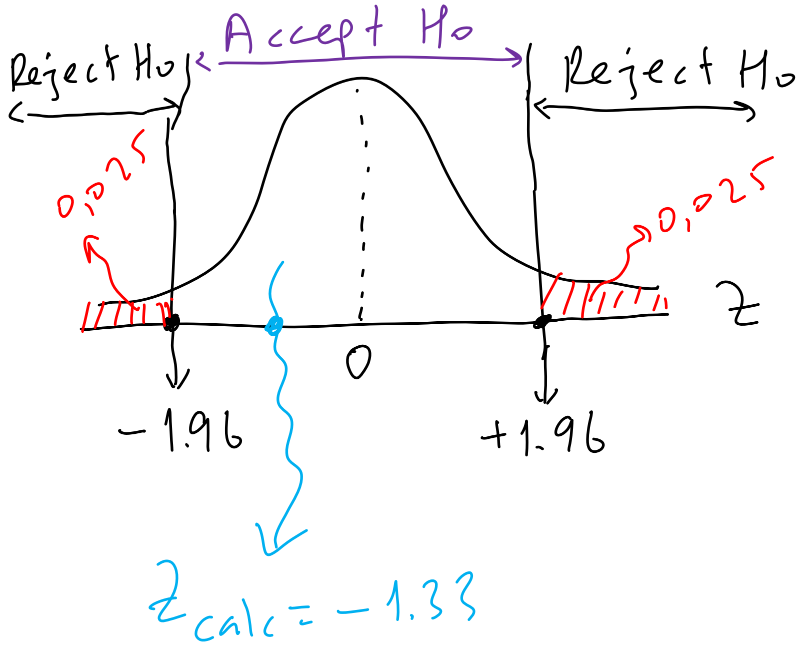


For the area 0,025 ⇒

$\mp z_{\alpha/2} = \mp z_{\frac{0,05}{2}} = \mp z_{0,025}$

$\mp z_{0,025} = \mp 1.96$ from the table.

⇒ Table value $= z_{0,025} = \mp 1.96$



Decision: Accept H_0

Conclusion: . . . . . .