Inferential Statistics

(N) It uses the random set of data

The Probability density of the normal distribution is,

$$\boxed{f(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}$$

Hypothesis Testing:

First step: Formulating Null Hypothesis.

A resercher thinks that if knee surgery patient go to physical therapy twice a week (instead of 3); their recovery times for knee surgery patient is 8.2 weeks.

Null hypothesis → $H_0 : \mu <= 8.2$

Alternate " → $H_1 : \mu > 8.2$

whatever we get at alternate hypothesis, every-thing else will go to Null hypothesis

Steps :

State the Null hypothesis → State the Alternate hypothesis → State the alpha level →

Find the rejection area from z-table → Find the test statistics → Support or reject Null Hypothesis

(two-tailed test)

Ex: Blood glucose levels for obese patients have a mean of 100 with a standard deviation of 15. A researcher thinks that a diet high in raw cornstarch will have a positive or negative effect on blood glucose levels. A sample of 30 patients who have tried the raw cornstarch diet have a mean glucose level of 140. Test the hypothesis that the raw cornstarch had an effect.

(if the is two condition in an hypothesis then we do two-tailed)

① $H_0 : \mu = 100$

$H_1 : \mu \neq 100$

② $\alpha = 0.05 = 5\%$

③ $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$= \dfrac{140 - 100}{15 - \sqrt{30}} = 14.6$

$\bar{x}$ = mean from the research output.

μ_0 = mean of the accepted facts

σ = standard deviation of existing facts

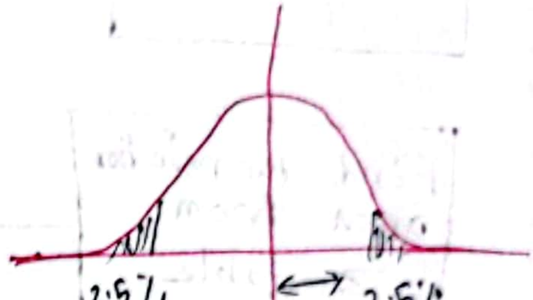
$\sqrt{n}$, n = No of sampels.

(a) If its two tailed test we devide the value of alpha by 2.

so, here,

$$\alpha \Rightarrow \frac{5\%}{2}$$

$$\Rightarrow 2.5\%$$

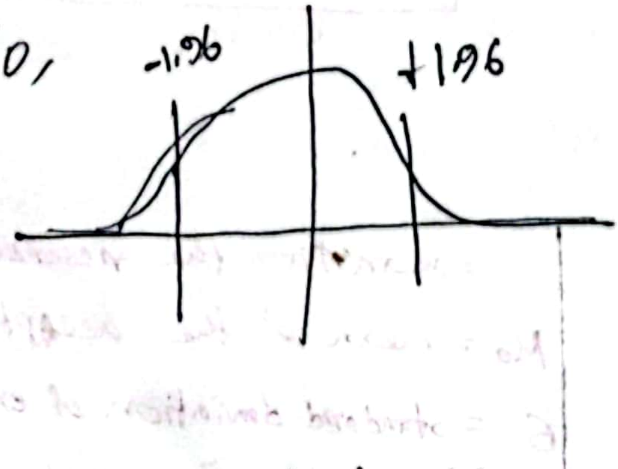


$50 - 2.5$
$= 47.5\%$
$= 0.475$

Statistical-Z value: (from the table)

for $0.475 \rightarrow 1.9 + 0.06$
$= 1.96$

so,



as, $Z = 14.6$, it will be in rejection region

so, we reject the null hypothesis.

Ex: A principal at a certain school claims that the students in his school are above average intelligence. A random sample of thirty students IQ scores have a mean score of 112. Is there sufficient evidence to support the principals claim? The mean population IQ is 100 with a standard deviation of 15.

① $H_0 = 100$

$H_1 > 100$

→ signifies percentage of error

② $\alpha = 0.05$ (5%) [user can choose if not given]

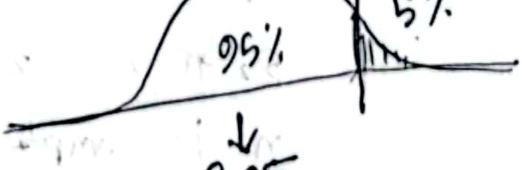
③ $z = \dfrac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$

$= \dfrac{112 - 100}{15 / \sqrt{30}}$

$= 4.38$

⑤ As $4.38 > 1.645$ it will be rejected. Null hypothesis is rejected

④



↓
0.95

[As this is one tailed test so no division]

Now find value from the table for 0.95

so, for 0.95 → $1.6 + 0.045$

$= 1.645$

T test or Z test?

if population variance (σ) is not known we have to perfome T test.

T-test:

One sample T-test

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

s = Standard daviation of sample.

all other test is same.

Two sample T-test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

x_1 = Sample mean of 1st pop

x_2 = " " " 2nd "

s_1 is sample 1 standard daviation

s_2 " " 2 " "

n is sample size

Z-test

One sample Z-test

$$Z\text{-test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Two sample Z-test

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Sample size determination:

Here we check if the sample size is enough or not.

$$n = Z^2 P(1-P)/e^2$$

Confidence interval, Chi-square statistics.

Confidence interval:

Ex: A group of 10 foot surgery patients had a mean weight of 240 pounds. The sample standard deviation was 25 pounds. Find a confidence interval for a sample for the true mean weight of all foot surgery patients. Find a 95% CI.

We always have to check it it is Sample standard deviation or Population " "

Here,

$s = 25$

$n = 10$

$\bar{x} = 240$

$CL = 95\%$
$= 0.95$

Step1: subtract 1 from your sample size to find degree of freedom
$df = 10 - 1 = 9$

Step2: Subtract the confidance level from 1; then divide by two
$\alpha = (1-0.95)/2 = .025$

Step 3: Use t-table with df and α values to find $t_\alpha = 2.262$

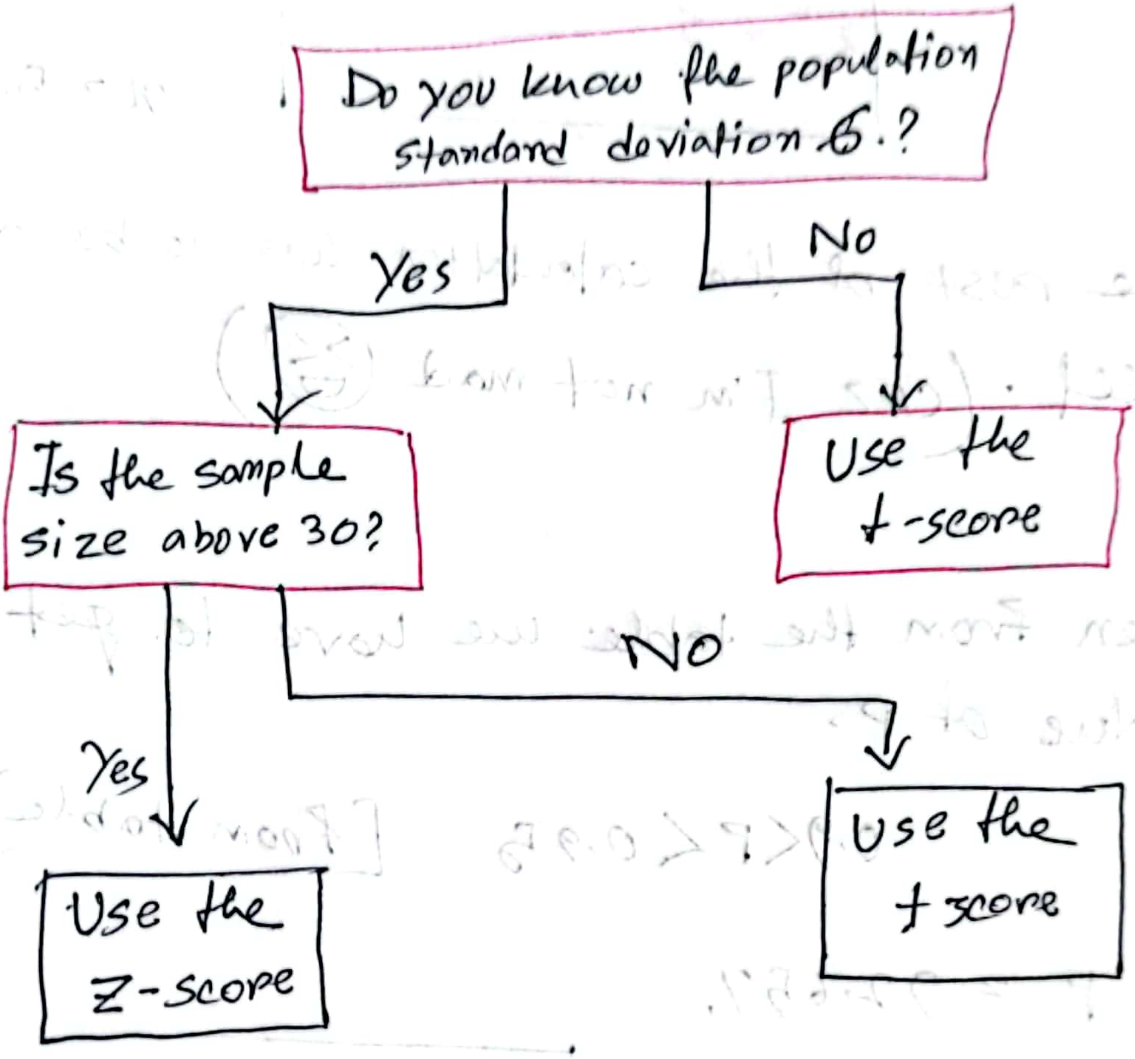
Step 4: Find out the standard error $SE = s/\sqrt{n} = \frac{25}{\sqrt{10}} = 7.90569415$

Step 5: multiply with t_α $(2.262 \times 7.90569415) = 17.8826802$

Step6: For lower end of the range $= 240 - 17.8826802$
$= 222.117$

Step7: " upper " " " " $= 240 + 17.88 = 257.883$

Things to check



with t-distribution/sample SD

$$\bar{x} \pm t \frac{s}{\sqrt{n}}$$

for mean formula/population SD

$$\bar{x} \pm z \frac{\sigma}{\sqrt{n}}$$

Chi-square test :

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$x = 5.094$

The rest of the calculation has to be done by excel. (coz I'm not mad)

Then from the table we have to get the value of P.

$0.9 < P < 0.95$ [from table]

$P = 92.65\%$

confident	$(89-0)\%$ ←	if $P > .10$ "not significant"
"	(90%)	if $P \leq .10$ "marginally " "
"	(95%)	if $P \leq .05$ " significant"
"	(99%)	if $P \leq .01$ "highly significant"

if P not significant we say null hypothesis should not be rejected.

Higher P-value null hypo not rejected
vice-verse

$\boxed{P \leq \alpha}$