Hypothesis testing: A statistical hypothesis test is a method of statistical inference used to decide whether the data at hand sufficiently support a particular hypothesis. Hypothesis allows us to make probabilistic statements about population parameters.

(Ho) Null Hypothesis: It is basically choosing a statement. It is noted down in PW-skills pdf.

(H1) Alternate Hypothesis: The opposite of null hypothesis.

Question Based on Hypothesis testing and Z test:

Suppose a company is evaluating the impact of a new training program on the productivity of its employee. The company has data on the average productivity of its employees before implementing the training program. The average productivity was 50 units per day with a known population std of 5 units. After implementing the training program, the company measures the productivity of a random sample of 30 employees. The sample has an average productivity of 53 units per day. The company wants to know if the new training program has significantly increased productivity.

productivity of making average 50 units per day

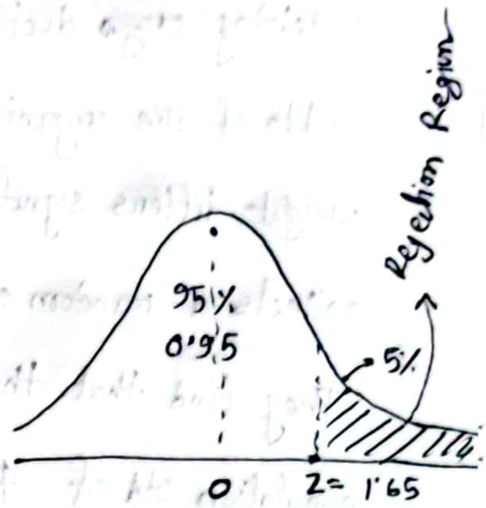
Step 1: $H_0 \rightarrow \mu = 50$. This will remain the same

$H_1 \rightarrow \mu > 50$

Step 2: α : significance level = 0.05 → 5%

Step 3: Normality valid / population std (σ) know

Step 4: As (σ) know, we will conduct Z test.



Step 5: $$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{53-50}{\frac{5}{\sqrt{30}}} = \frac{3}{5/\sqrt{30}} = 3.28$$

We will find ~~in~~ the value for 0.95 in Z table which is 1.65.

Our value is found 3.28 which > 1.65 and falls in the rejection area.

We have got strong evidence against the null hypothesis and in favor of Alternate Hypothesis. So we can reject the null hypothesis.

So it can be said that the training program significantly increases productivity

Problem 2:

Suppose a snack food company claims that their packets contain an average weight of 50 grams per packet. To verify this claim, a consumer watchdog orga decides to test a random sample of packets. The motto of the organization was to determine whether the actual average weight differs significantly from the claimed 50 grams. The organsization collects a random sample of 40 packets and measures their weights. They find that the sample has an average wait of 49 grams, with population std of 4 grams.

$\mu = 50, \quad n = 40, \quad \bar{x} = 49, \quad \sigma = 4$

Step 1: $(H_0) \rightarrow \mu = 50$ (Means Average weight is 50)

St[illegible]: $(H_1) \rightarrow \mu \,!= 50$ (Average weight is not 50)

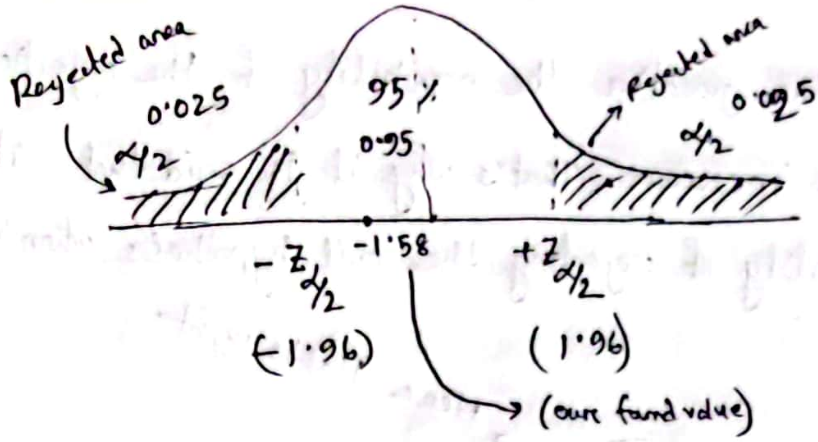
Step 2: $\alpha = 0.05$

Step 3: $n = 40 \ (>30)$. Normality is valid and σ is known.

Step 4: As population (σ) is known, we will conduct Z test.

Step 5: $z = \cancel{\mu - \bar{x}} \quad \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{49 - 50}{\frac{4}{\sqrt{40}}} = -1.58$

$\alpha = 5\%$

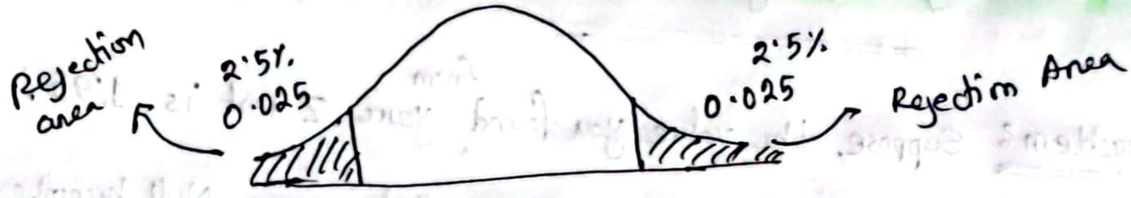


As ~~our~~ the value we got not fell in the ~~rej~~ rejected area, we cannot reject our null hypothesis.

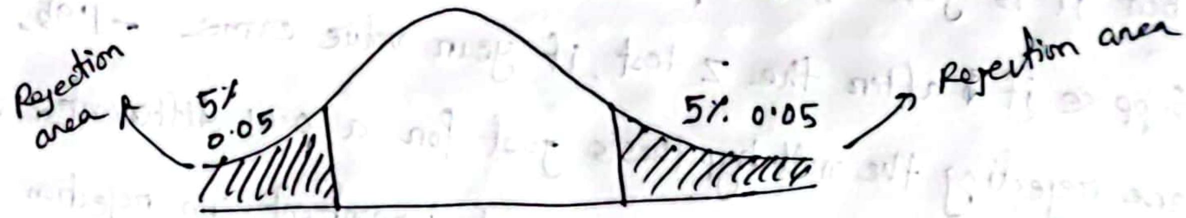
So, we couldn't reject the fact that, average chips packet weight was 50 gm.

Significance level: It represents the probability of rejecting the null hypothesis when it is actually true.

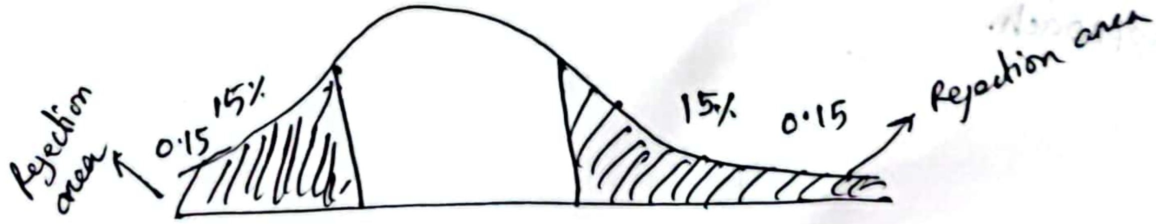
Means, Suppose, $\alpha = 5\%$, for 2 tailed test



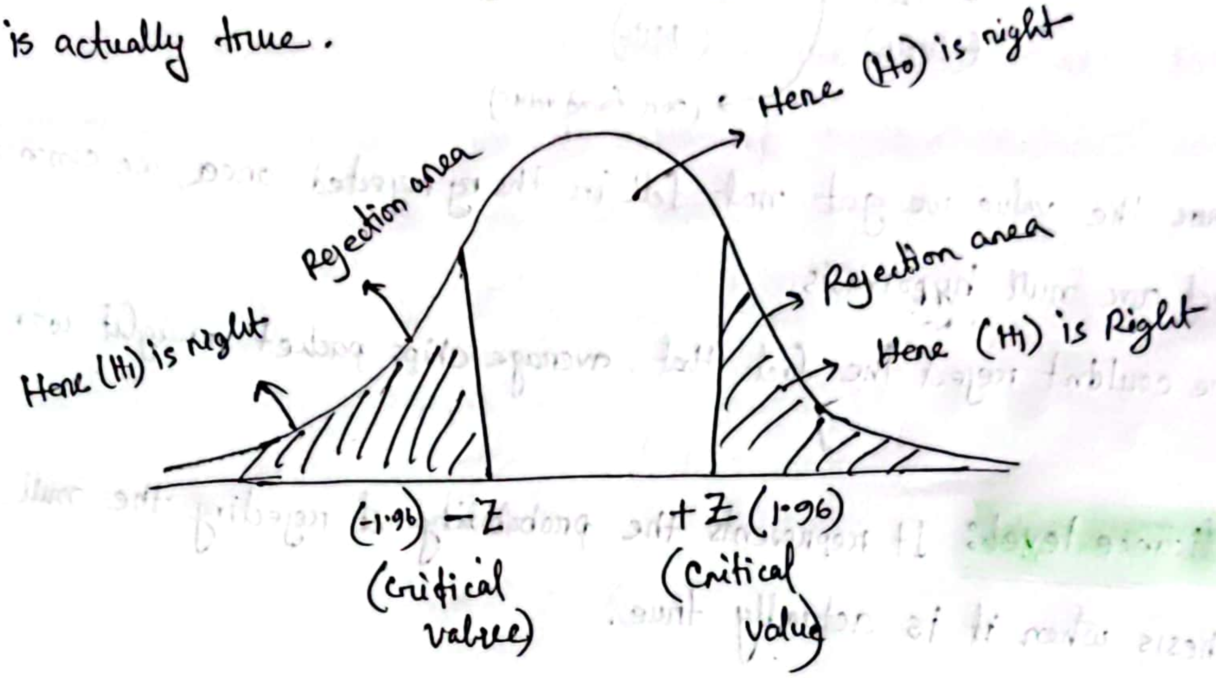
$\alpha = 10\%$, for 2 tailed test.



$\alpha = 30\%$, for 2 tailed test.



The more the α is, the greater the possibility of the rejection of the null hypothesis increase. That's why it is said that it represents the probability of rejecting the null hypothesis when it is actually true.



Problem with Rejection Region Approach:

First problem: Suppose, the value you found from your Z test is 1.95. From the above scenario case, it can be said that your Null hypothesis is true. But it is just happend for a 0.01 difference.

Suppose if after the Z test, if your value came -1.95, so you are rejecting the null hypothesis just for a 0.01 difference.

So critical point become very much important in rejection region approach.

Second problem: Suppose, you find your Z value = 2'00 in the previous scenario. Or suppose you get Z = 15. Both of ~~ca~~ the case, your Null hypothesis will be rejected. But the ~~evidence~~ strength of evidence can't be measured here. Z = 15 is more stronger evidence than Z = 2'00. Z = 15 means your data is lying very very far, even far from the rejection area. So it can't be detected in rejection region approach.

In that ~~comes~~ where comes P-VALUE which can help to measure the strength of evidence

Type 1 vs Type 2 Error:

Type-1 (False Positive): This occurs when the sample results, lead to the rejection of the null hypothesis when it is in fact True

It is denoted by α (also known as the significance level)

Researchers can control the risk of making a Type 1 error.
(By reducing α)

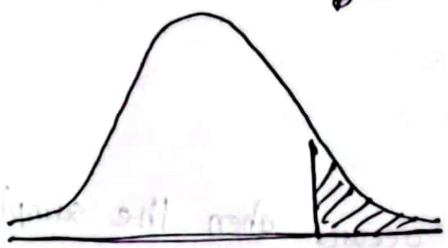
Type-II (False Negative): It occurs when based on the sample results, the null hypothesis is not reject when it is in fact false.

It is denote by β. This means, the researcher fails to detect a significant effect or relationship when one actually exists.

One tailed test: It is used when the researcher is interested in testing the effect in a ~~spefi~~ specific direction (either greater than or less than the value specified in the null hypothesis). The alternate hypothesis in one tailed test contains an ~~equality~~ inequallity (either '>' or '<')

Example: A researcher wants to test whether a new medication increases the average recovery rate compared to the existing medication.

Right tail test: → H_1 : μ > value for value < μ



Left tail test: H_1 : μ < value or value > μ

Two-tailed test: When the researcher is interested in testing the effect in both directions (i.e, whether the value specified in null hypothesis is different, either greater or smaller). The ~~a~~ alternate hypothesis in a two tailed test contains " ≠ " sign.

Example: A researcher wants to test whether a new medication has a different average recovery rate compared to the existing medication.

Advantage of performing two-tailed test:

1. Can detect effects in both side
2. Two tailed tests are more conservative because the significance level (α) is split in between the both tails of the distribution. This reduces the ~~ri~~ risk of Type I errors in cases where the direction of effect is uncertain.

Disadvantages:

1. Less powerful: Because the significance level (α) is getting divided into two parts and the area is getting reduced in both tail. So for rejecting null hypothesis, stronger evidence will be needed.
2. Not ideal for directional hypothesis ($>, <$ values).

Advantage of performing one-tailed test:

1. More powerful: As it is detecting effects in any one of the tail region of the distribution and significance level is not getting halved, so it is more stronger in terms of detecting because the tail area is bigger.

2. More appropriate to test for an effect in specific direction.

Disadvantages:

Missed effects: If some ~~test~~ defect can fell in the opposite direction, it can't detect.

Increased risk of Type I error: A significance level is not getting halved here, in of the tail region it would take more area, So null hypothesis probability will get shorten in that region. So that will Increase the type I error.

P Value: It is the probability of getting a sample as or more extreme having more evidence against ~~H~~ Null Hypothesis (H_0) than our own sample given the NULL Hypothesis is True.

(It's a bit confusing. Watch campusX session 46 if you forget)

Suppose you r tossing a fair coin. You tossed the coin 100 times. In this sample you saw that heads came 57 times.

100 times tossing a fair coin = 1 sample

Now, suppose your PValue is = 0·3

~~Fed~~ That means, If you take 100 samples of tossing a coin 100 times in every sample, 30 times in the samples, where Head will come 57 times after a 100 coin toss. This is what PValue says.

Relation between PValue & ~~α~~ significance level (α):

If P-Value $\leq \alpha$, you can reject your Null Hypothesis (H_0)

→ How to calculate P-Value?

Previously in some Z-test problems, we have used rejection region approach. But now, we will ignore Rejection region approach and will use P-Value instead. The problem is in the next page.

Problem: A company is evaluating the impact of a new training program on the productivity of it's employees. The company has data on the average productivity of its employees before implementing the training program. The average productivity was 50 units per day. After implementing the training program, the company measures the productivity of a random sample of 30 employees. The sample has an average productivity of 53 units per day, and the population ~~sample~~ std deviation is 4. The company wants to know if the new training program has significantly improved productivity.

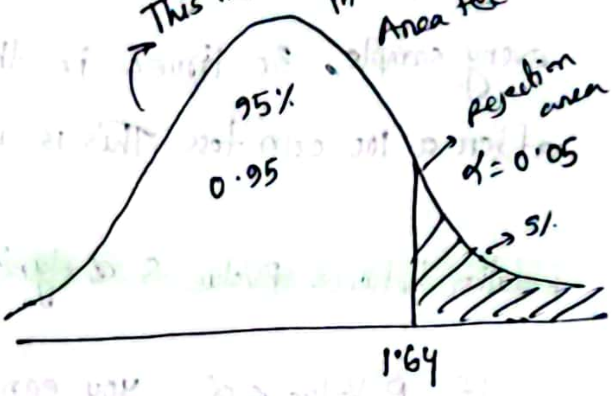
$\mu = 50, \; n = 30, \; \bar{x} = 53, \; \sigma = 4$

$\alpha = 0.05$ (Given)

Step 1: $H_0 = \mu_0 = 50$

$H_1 = \mu > 50$

This is what we would do in rejection Area technique

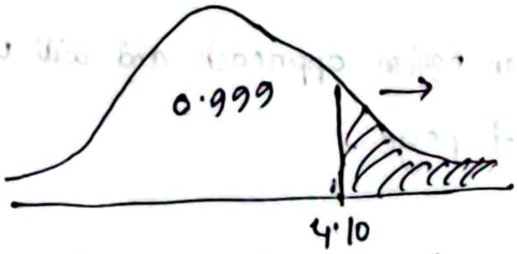


Z stat: $\dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{53 - 50}{4/\sqrt{30}}$

$= 4.10$

For using P value, we found our z stat which is 4.10.

We will calculate the area after 4.10



Now we will go to the z table to find the ~~area~~ value for 4.10.

The ~~da~~ value we will get from the Z-table is the left area of 4·10. So to get our desired outcome (value of right area after 4·10) we have to subtract it from 1.

value we get from z table for 4·10 = 0·999

$\therefore$ P-Value = 1 - 0·999 = 0·0001

as p-value $\leq \alpha$-value, we can reject our Null Hypothesis.

Problem 02: Suppose a chips company claims that their food weights of 50 grams per packate. To verify this claim, an org. decided to test a random sample of packets. The org. wants to determine whether the actual average weight differs significantly from the claimed 50 grams. The org. collects a random sample of 40 packets and measures their weights. They find that the sample has an average weight of 49 grams, with a population std of 5 grams.

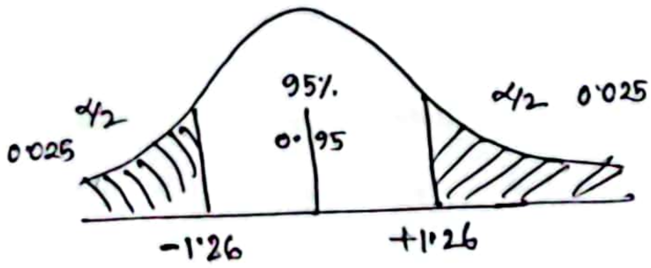
$\mu = 50$, $n = 40$, $\bar{x} = 49$, $\sigma = 5$, $\alpha = 0.05$

Here,

Null Hypothessis : Average weight of packet (μ) = 50

Alternate Hypothesis: Average weight $\neq$ 50 (Means weight can be > 50 or < 50)

Z test:
$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = -1.26$$



Here in this case where we have 2 areas, p value = area 1 + area 2

from z table we have found value for -1.26 which is = 0.206

As z table provide the left area value of a point so, we used (-1.26)

So that the left area value can be found. which is = 0.103

Now, as the data is normally distributed, the left area of -1.26 and the right area of +1.26 is same,

So, total area value = P-Value = 0.1038 + 0.103

= 0.206

Now, As P-Value > Significance level (0.05), so we can not reject the null hypothesis.

So the average packet weight is 50.

T-Test: T-test you perform when you don't have the value of population ~~samp~~ std deviation.

Three main types of T-test:

① One sample T-test: You just take one sample from the population and using it's mean ($\bar{x}$) you ~~tried~~ try to perform Hypothesis testing on the population.

② Independent Two Sample T-test: It is used to compare the mean of two independent samples. The null hypothesis states that there is no significant difference between the means of the two samples, While the alternate hypothesis say, there is a significant difference.

③ Paired t-test (dependent two-sample t-test) :

The paired t-test is used to compare the means of two samples that are dependent or paired, such as pre test and post test scores for the same group of subjects or measurements taken on the same subjects under two different conditions. The null hypothesis state that there is no significant difference between the means of the paired differences, while the ~~al~~ alternate hypothesis states that there is a significant difference.

Math problems for t-test are noted down in PW-Skills section