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.

(H_0) Null Hypothesis: It is basically choosing a statement. It is noted down in PW-skills pocket.

(H_1) 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

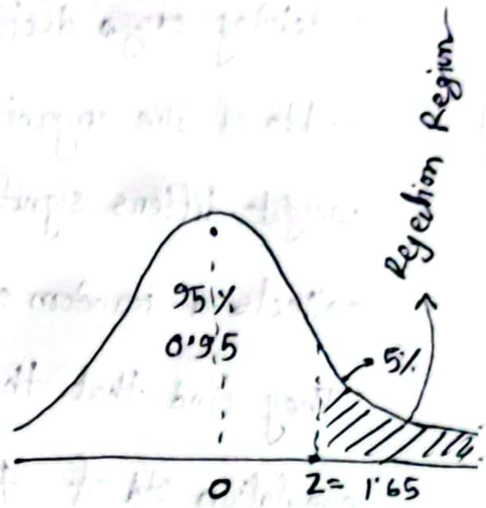
$H_1 \rightarrow \mu > 50$

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: $Z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{53-50}{\frac{5}{\sqrt{30}}} = \dfrac{3}{5/\sqrt{30}} = 3{\cdot}28$



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)

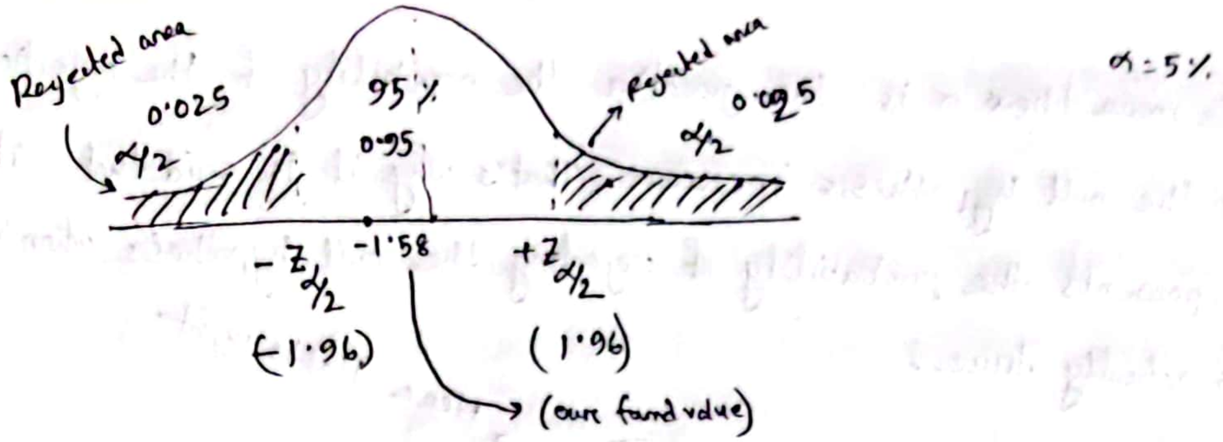
Step: $(H_1) \rightarrow \mu \,!= 50$ (Average weight is not 50)

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

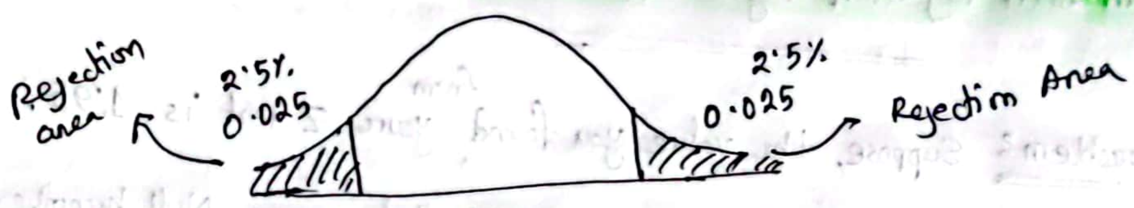


As ~~our~~ the value we got not fell in the ~~rej~~ rejected area, we cannot reject our null hypothesis.

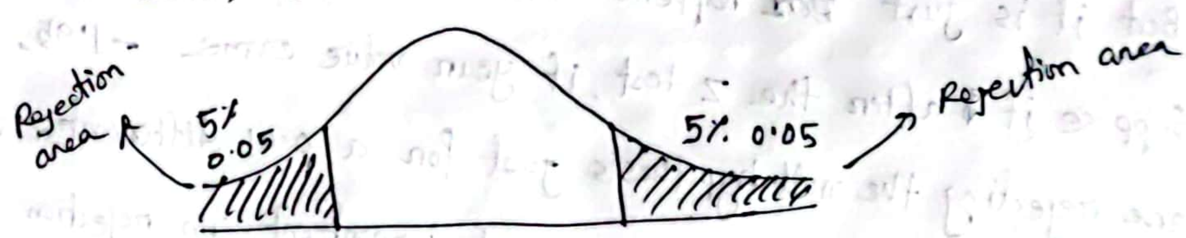
So, we couldn't reject the fact that, average chips packet weight was 50 gm.

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

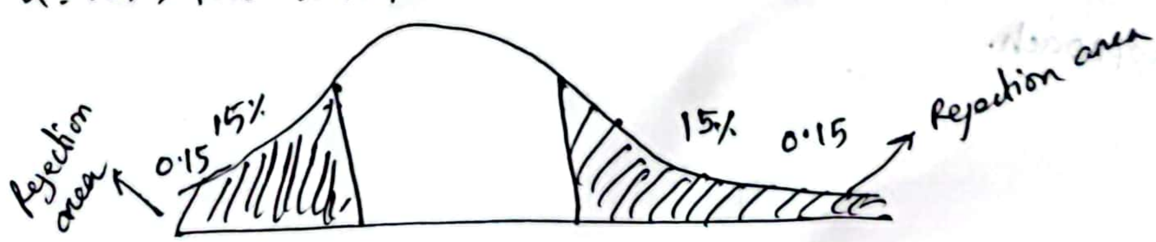
Means, Suppose, $\alpha = 5\%$, for 2 tailed test



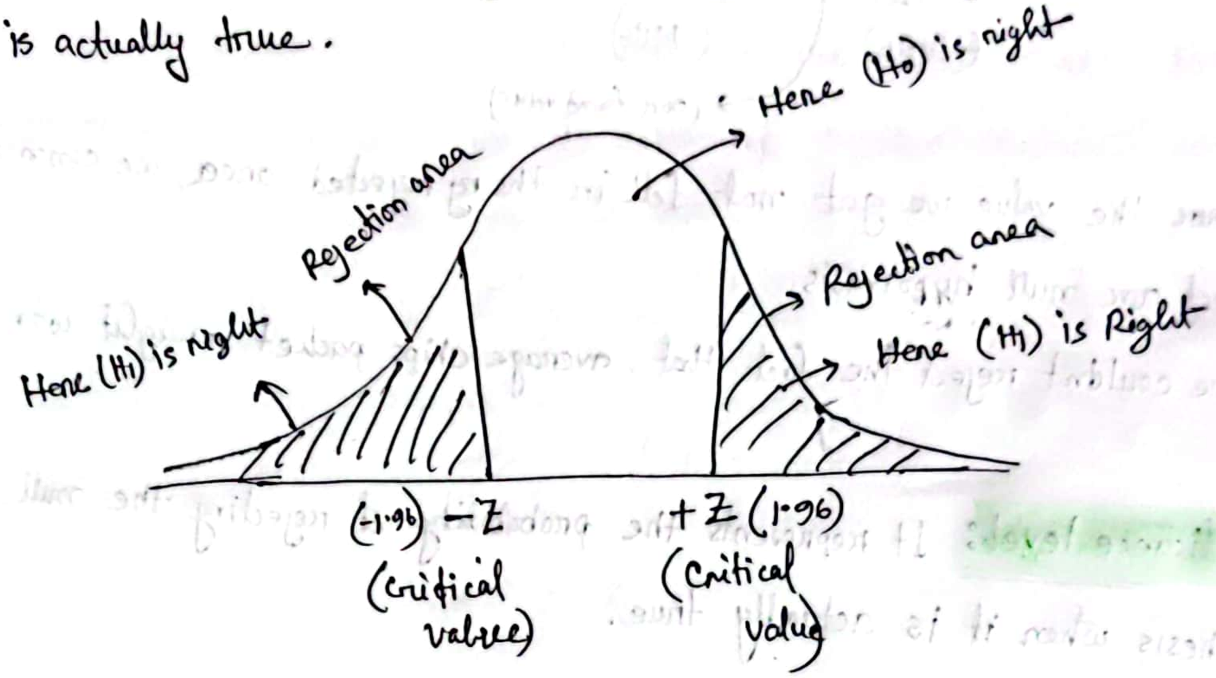
$\alpha = 10\%$, for 2 tailed test.



$\alpha = 30\%$, 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 scenrio 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)

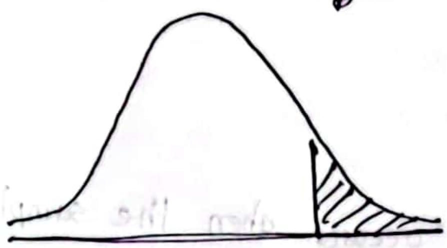
Type-II (False Negative): It occurs when based on the sample results, the null hypothesis is not reject when it is in fact false.

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~~ inequality (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: → $H_1: \mu >$ value or value $< \mu$



Left tail test: $H_1: \mu <$ value or value $> \mu$

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 alternate hypothesis in a two tailed test contains "$\neq$" 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 ~~it~~ 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.

CamScanner

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 shorter in that region. So that will increase the type I error.