# Lec. VII. P-VALUE & T-TEST

## ★ Type-I v/s Type-II Error:

In hypothesis testing, there are two types of errors that can occur when making a decision about the null hypothesis.

- **Type-I (false-positive):-** error occur when the sample results, lead to the rejection of the null hypothesis when it is in fact true.

  In other words, it's a mistake of finding a significant effect or relationship when there is none. The probability of comitting a Type-I error is denoted by $\alpha$ (alpha) which is also known as the significance level, researchers can control the risk of making a Type-I error.

- **Type-II (false-negative):** error occurs when based on the sample results, the null hypothesis is not rejected when it's in fact false.

  This means that the researcher fails to detect a significant effect or relationship when one actually exists. The probability of commiting a Type II error is denoted by $\beta$ (beta).

Truth about the population.

| | | Ho true | Ha true. |
|---|---|---|---|
| Decision based on sample | Reject Ho | Type-I error | Correct decision |
| | Accept Ho | Correct decision | Type-II error |

Type-I

The chance or probability that you will reject a null hypothesis that should not have Been rejected.

ie innocent declared as guilty.

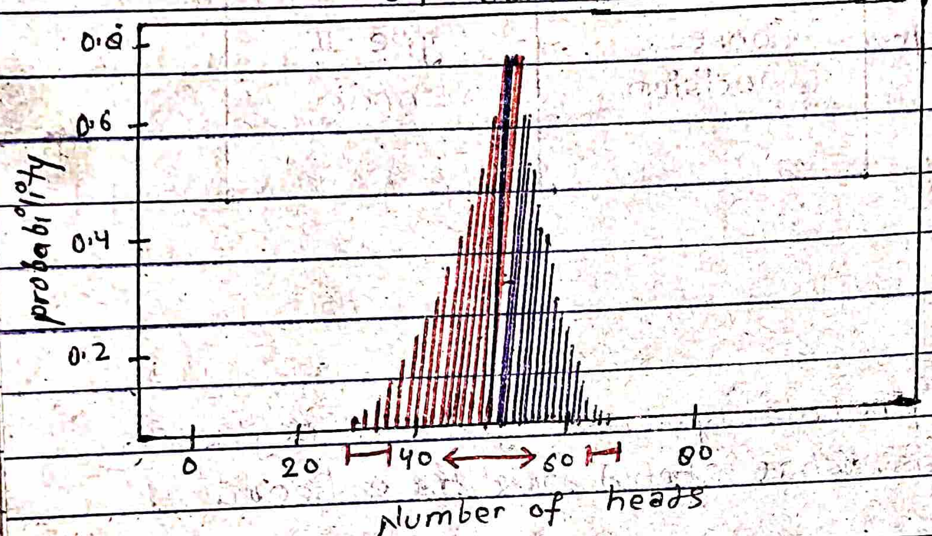
Type-II

The chance or probability that you willn't reject a null hypothesis when it should have Been rejected

ie guilty declared as innocent

p-value : p-value is the probability of getting a sample as or more extreme (having more evidence against $H_0$) than our own sample given the Null hypothesis ($H_0$) is true.

binomial distribution of coin toss




← pmf
ie 1 coin tossed 100 times

In simple words p-value is a measure of the strength of the evidence against null hypothesis that is provided by our sample data

ie tossing a coin 100 times

formulate → $H_0$ : $P(H) = P(T)$ (fair coin)
$H_a$ : $P(H) > P(T)$

let, say 53 head comes as o/p of the expt

prob corresponding to 53 H in plot ≈ 0.702

$P(H=53) = 0.7$

$P(H>53) = 1-0.7 = 0.3$ } → pvalues

↓

performing the expt 100 times we may get 53H → 30 times

Interpreting p-value :

- with significance level –

$\alpha = 0.05$ or $0.01$

p value $\leq \alpha \rightarrow$ reject $H_0$

- without significance level –

1. Very small p-values (ie, $p < 0.01$) indicate strong evidence against null hypothesis, suggesting the observed effect or difference is unlikely to have occured by chance alone.

2. Small p-values ($0.01 \leq p < 0.05$) indicate moderate evidence against the null hypothesis, suggesting that the observed effect or difference is less likely to have occured by chance alone.

3. Large p-values ($0.05 \leq p < 0.1$) indicate weak evidence against null hypothesis suggesting that the observed effect or difference might have occured by chance alone but there is still some level of uncertainity.

4. Very large p-values ($p \geq 0.1$) indicate weak or no evidence against the null hypothesis, suggesting that the observed effect or difference is likely to have occured by chance alone.

Q) Suppose 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 it's employees before implementing the training program. The avg. productivity was 50 units per day. After implementing the training program, the company measures a productivity of a random sample of 30 employees. The sample has an avg. productivity of 53 units per day and the pop. std is 4. The company wants to know if the new training program has significantly increased productivity?

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

$H_0 : \mu = 50$ (new training program lead same productivity)

$H_a : \mu > 50$ (significant increase)

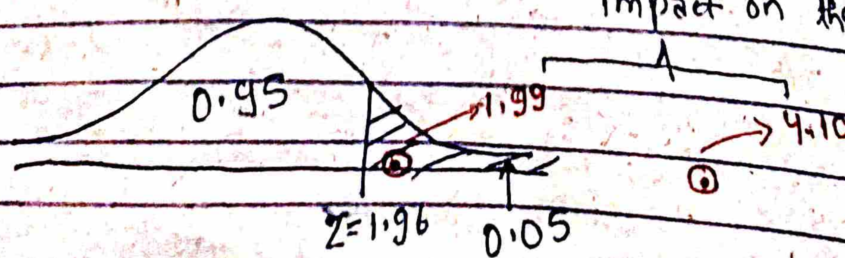
test-statistic $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \dfrac{53-50}{4/\sqrt{30}} \Rightarrow 4.10$

Let, 5% level of significance,

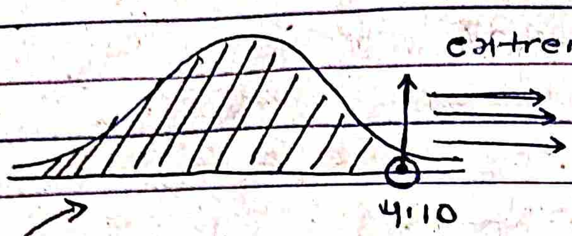
$1.96 < 4.10$ : reject NULL hypothesis (strongly)

what if $Z = 1.99$

couldn't differentiate b/w the impact on these two with z-test



p-value, $Z = 4.10$




on z-table leftmost area for 4.10

0.09
+4 $\boxed{0.9998}$

extremost area $\Rightarrow 1 - 0.9998$
$\Rightarrow 0.0002$

p-value $< 0.05$
reject null hypo. $\rightarrow$ event is Likely to have occured

Q). Suppose a snack food company claims that their Lays wafer packet contain an average weight of 50 gms per packet. To verify this claims, a consumer watchdog organization decides to test a random sample of Lays wafer packets. The organization wants to determine whether the actual avg. weight differs significantly from the claimed 50 gms. The organization collects a random sample of 40 Lays wafer packets and measure their weights. They find that the sample has an avg. weight of 49 gms, with a standard deviation of 5 gms ?.
(population)

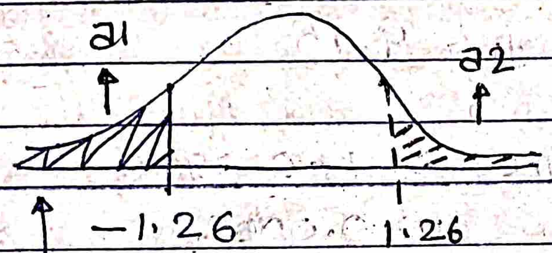
$\mu = 50$ ; $n = 40$ , $\bar{x} = 49$ , $\sigma = 5$

$H_0 : \mu = 50$
$H_a : \mu \neq 50$

$\alpha = 0.05$

p-value,

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \frac{49-50}{5/\sqrt{40}} \Rightarrow -1.26$$




p-value ≡ $a_1 + a_2$

→ $0.103 \times 2$

→ 0.206.

as not specified which extremost dirn.

from z-table

| | 0.06 |
|---|---|
| -1.2 | $\boxed{0.103}$ |

$0.206 > 0.05$

not reject null hypothesis

ways to calculate C.I

z-procedure → pop → std. dev (available)

t-procedure → p → std. dev (not available)

rule of thumb → for smaller sample sizes

★ **T-tests:**

- T-distribution is used while making assumptions about a mean when we don't know the std. dev.

- It's a bell shaped distribution whose mean is $\mu$ and the std. dev. is $\sigma$. The t-dist. is similar to normal dist. but flatter and shorter.

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

the estimate of standard error of difference b/w means.

- A **t-test** is a statistical test in hypothesis testing to compare the means of two samples or to compare a sample mean to a known population mean.

There are 3 main type of T-tests:

1. **Single Sample T-test:**

$$\text{1 sample} \xrightarrow{\text{work on}} \bar{x} \xrightarrow[\text{using}]{\text{to find}} \mu$$

Used to compare the mean of a single sample to a known population mean.

The null hypothesis states that there's no significant difference b/w sample mean & population mean, while alternative hypothesis state there's a significant difference.

## Assumptions for single sample T-test :

① Normality – population from which the sample is drawn is normally distributed.

② Independence – The observation in the sample must be independent, which means that the value of one observation shouldn't influence the value of another observation.

③ Random sampling – The sample must be random and representative set of population.

④ Unknown population standard deviation.

Q). Suppose a manufacturer claims that the avg. weight of their new choclate bars is 50 gms, we highly doubt & want to check this so we drew out a sample of 25 choclate bars and measured their weight, the sample mean came out to be 49.7 gms & the sample std. devi. was 1.2gms Consider the significance level to be 0.05.

$\mu = 50$, $n = 25$, $\bar{x} = 49.7$, $s = 1.2$

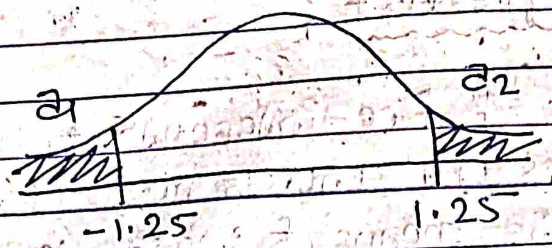
$H_0 : \mu = 50$
$H_a : \mu \neq 50$

$\alpha = 0.05$

assuming, it is normal

Test statistic – $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \Rightarrow \frac{49.7 - 50}{1.2/\sqrt{25}} \Rightarrow -1.25$

t-distribution.




t-value → a1 + a2

→ 0.11 × 2
↑ from table

→ 0.22

degree of freedom; df = n-1 → 24

0.22 > 0.05

couldn't reject null hypo.

★ Code to calculate t-value :

```
from scipy.stats import t

# set the t-values & dof
t_value = input()
dof = input()

# Calculate the CDF value
cdf_value = t.cdf(t_value, dof)
print(cdf_value).
```

The variance & the t-distribution is always greater than '1' and is limited only to 3 or more degrees of freedom. It means this distribution has a higher dispersion than the standard normal distribution.