T - test

- Recap Framework
- Recap Z - Test
- 1-sample T - Test
- 2- sample T - Test

Recap: Framework → You observe

Chance Vs Significant

Random Effect

↓ ↓ } Burden of proof

H_0 H_a

1. Null Vs Alternate
   H_0 H_a
2. Test Statistic, distribution
3. Left Vs Right Vs Two-tailed
4. p-value: prob of seeing data given H_0 is true
5. Significance level: $\underbrace{0.05}_{\alpha}$ → Compare p-value with α

p-value $< \alpha$ → reject H_0

A french cake shop claims that the average number of pastries they can produce in a day exceeds 500. The average number of pastries produced per day over a 70 day period was found to be 530. Assume that the population standard deviation for the pastries produced per day is 125. Test the claim using a z-test with the critical z-value = 1.64 at the alpha (significance level) = 0.05, and state your interpretation.

$H_0: \mu = 500 \quad (\mu \leq 500)$

$H_a: \mu > 500$

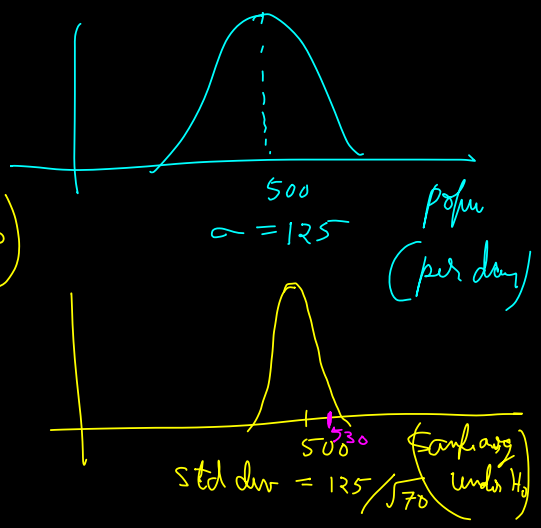
Test statistic : Sample avg of 70 days

Distribution : Gaussian (avg $= 500$ under H_0, stddev $= \frac{125}{\sqrt{70}}$)

Obs value of T-Stat

$p\text{-value} = P[T \geq 530 \mid H_0 \text{ is true}]$

$z\_stat = \frac{530 - 500}{125/\sqrt{70}}$



1 Sample T-Test

Improve IQ with a pill !!

Population avg IQ = 100

Try on a few people

110, 105, 98, 102 ----- Vs 100

8 people

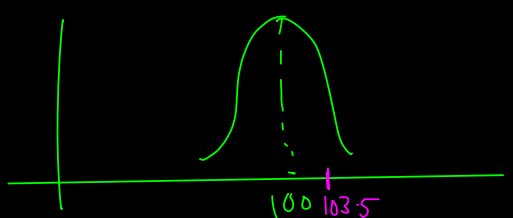
H_0 : $\mu = 100$ (pill has no effect) (μ → avg IQ of ppl taken pill)

H_a : $\mu > 100$ (pill has positive effect)

right tail

$\alpha = 0.01$ 99% Confident

Test statistic : sample avg of 8 people → 103.5



$$z\text{-stat} = \frac{103.5 - 100}{\sigma/\sqrt{8}}$$

σ missing

p-value from z-stat

This is not possible

$$t\text{-stat} = \frac{103.5 - 100}{s/\sqrt{8}}$$

s: Std dev of the samples (This is not Std error) Is also random

T-distribution

degree of freedom : 8 - 1 = 7

(In detail later in Chi Sq)

[If we know sample mean of 8 numbers, then only 7 values are enough to know the whole dataset]

Samples of IQ after taking pill vs 100

Scipy `ttest_1samp` → 1 sample T test

2 sample T-test ttest_ind (Independent)

IQ

Samples School 1 → 101.15

vs

Samples School 2 → 109.4

Sample drug

Drug recovery time drug 1 & drug 2

Which is better?

$x_1\ x_2\ x_3 \dots x_{20}$ vs $y_1\ y_2\ y_3 \dots y_{30}$

5 exclusive runs : Is scoring pattern similar or different in 1st inns Vs 2nd inns

1st inns → 46

2nd inns → 40

(A) ✓

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

"two-sided"

(B) ✓

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

"greater"

(C) ✗

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 < \mu_2$

"less"

win $\longrightarrow$ 51 $\longrightarrow \mu_1$

loss $\longrightarrow$ 35 $\longrightarrow \mu_2$

(A)

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 \neq \mu_2$

"two-sided"

$\downarrow$

H_a, low p-value

(B)

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

"greater"

$\downarrow$

H_a, low p-value

(C)

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 < \mu_2$

"less"

high p-value