HYPOTHESIS TESTING - 3

Z Test

Burger Company

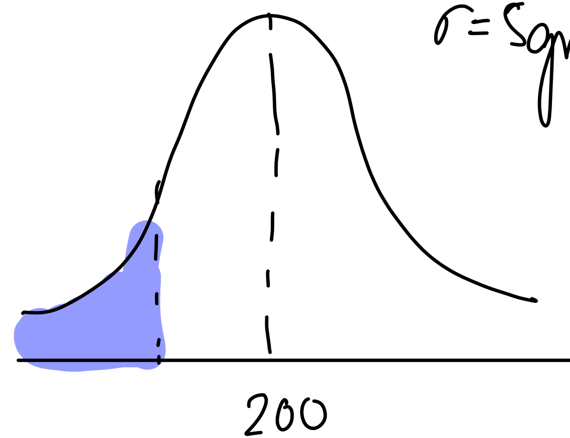
A company selling burgers claims that its burgers weigh 200 gms on an average, with std. of 5 gms.

An unsatisfied hungry customer wants to disprove this claim.

$H_0 : \mu = 200$

$H_a : \boxed{\mu < 200}$ $\boxed{\mu > 200}$ $\boxed{\mu \neq 200}$

$\sigma = 5$ gms



left tailed test

AI chip company

Google → TPU

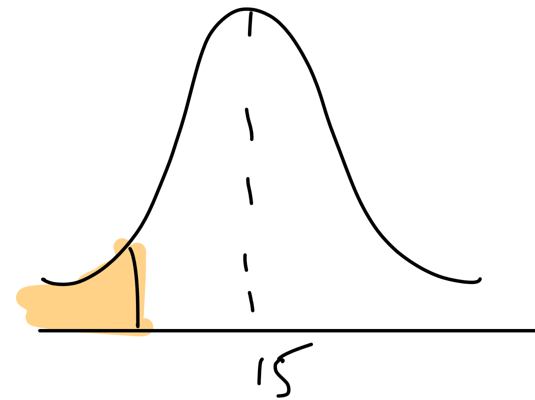
The company wants to claim that it is better than GPU.

The training time for ResNet is 15 minutes on the GPU.

H_0 : training time $\mu = 15$

H_a : ✓ $\mu < 15$ $\qquad$ $\mu > 15$

$\mu \,!= 15$



1000 images in 15 mins
GPU

TPU

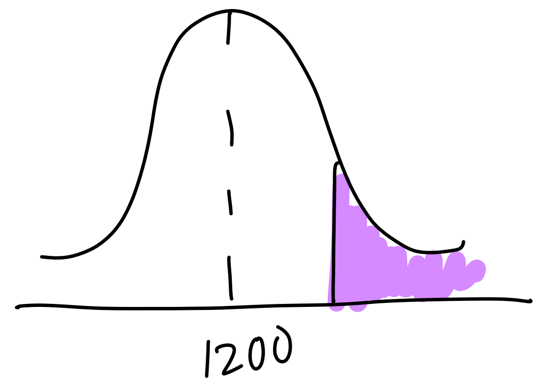
Machine Learning Deployment

A ML model in deployment processes 1200 images/sec on average with Std. of 10 images/sec.

We want to show that the new model can process images at a higher rate.

H_0 : Rate of processing $\mu = 1200$

H_a : $\boxed{\mu > 1200}$



Right tailed test.

Height from your state.

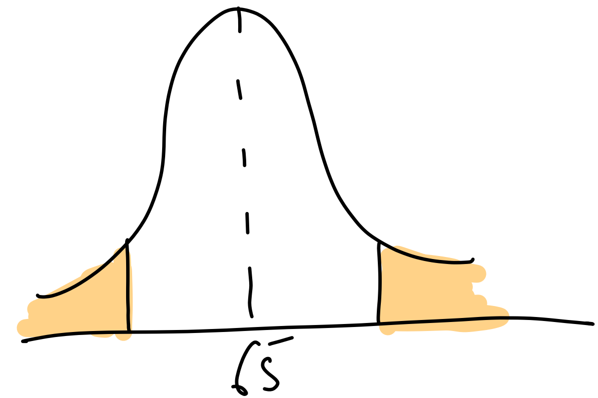
The average height of Indians is 65 inches, with std, of 2.5 inches.

You want to show that the height of people in your state is not 65

$H_0 : \mu = 65$

$H_a :$ $\boxed{\mu < 65}$ $\boxed{\mu > 65}$

$\boxed{\mu \,!= 65}$



Two tailed test.

Summary

BurgerCompany

H_0 : $\mu = 200$

H_a : $\mu < 200$

} left tailed Test

ML Model

H_0 : $\mu = 1200$

H_a : $\mu > 1200$

} Right tailed test

Height

H_0 : $\mu = 65$

H_a : $\mu != 65$

} Two tailed test.

Hypothesis Testing Framework

1) Setup the Null and Alternate Hypothesis

2) Choose the right test statistic distribution

3) Left tailed vs Right tailed vs Two-Tailed

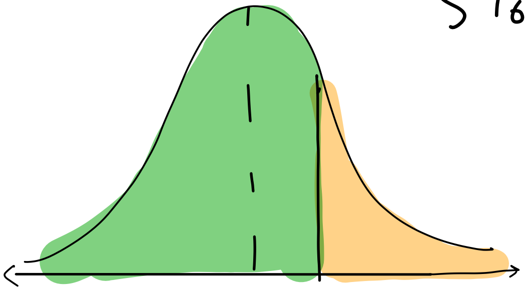
4) Compute P-value

5) If P- value is less than α , then reject the null hypothesis.

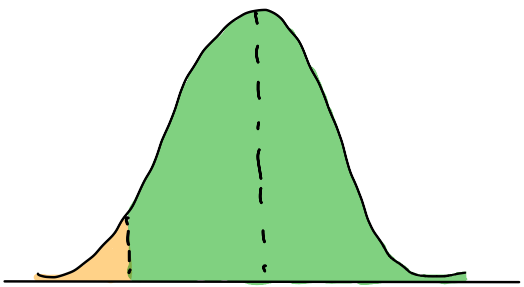
Sign in

Sign in

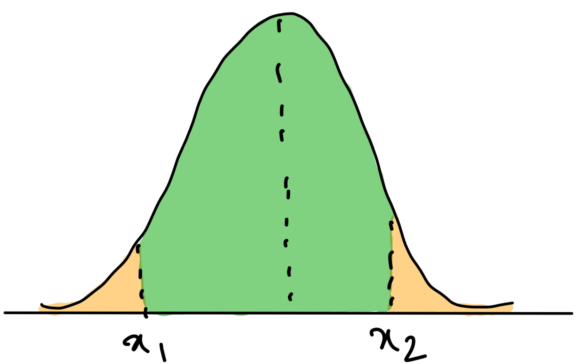
5%



Right tailed $1 - cdf(\quad)$



Left tailed $cdf(\quad)$



Two tailed

$$\underbrace{cdf(x_1)}_{\text{left}} + \underbrace{(1 - cdf(x_2))}_{\text{Right}}$$

Recap

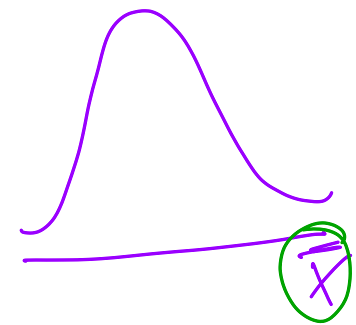
Avg height of people = 65 inches

$\sigma = 2.5$ inches

2

* 50 individuals $\mu = 65$, $SE = \frac{2.5}{\sqrt{50}}$

* 5 individuals $\mu = 65$, $SE = \frac{2.5}{\sqrt{5}}$



Let "m" be my sample mean

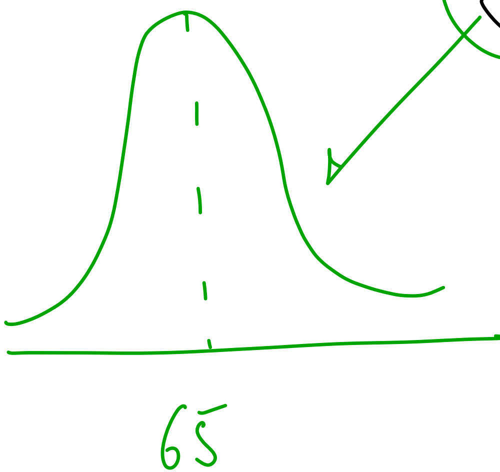
Is 'm' a random variable → ✓

Distribution → Gaussian Distribution

$E(m) = 65$

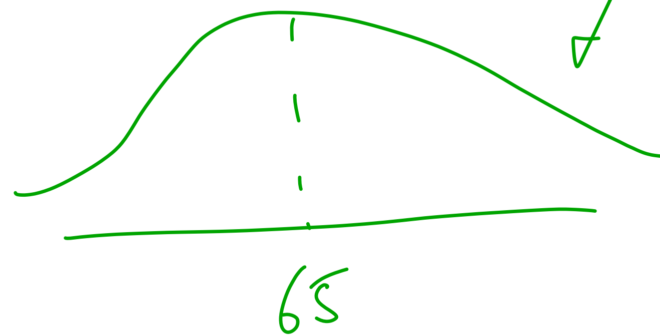
① 50 $\mu = 65$

$$std_1 = \frac{2.5}{\sqrt{50}}$$



② 5 $\mu = 65$

$$std_2 = \frac{2.5}{\sqrt{5}}$$



Retail Flipkart

$\mu = 1800$

$\sigma = 100$

Shampoo

① M1

Apply 50 Stores

Avg Sales = 1850

② M2

Apply 5 Stores

Avg Sales = 1900

$H_0 : \mu = 1800$ (marketing had no effect)

$H_a : \mu > 1800$ (marketing had an affect)

① MJ → 50 stores

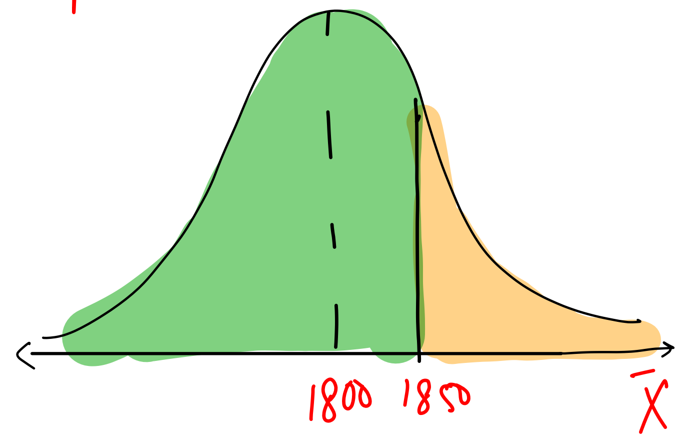
Right tailed Test.

test statistic : "m" sample mean

Distribution : Gaussian

$E[m] = 1800$

$SE = \frac{100}{\sqrt{50}}$



$p = P[m \geq 1850 \mid H_0 \text{ is true}]$

$p = 0.0002$ $\alpha = 0.05$

"$p < \alpha$"

$\boxed{\text{Reject } H_0}$

$Z = \frac{1850 - 1800}{100/\sqrt{50}}$ $SE = \frac{100}{\sqrt{50}}$

Z → Z_{stat} Z statistic

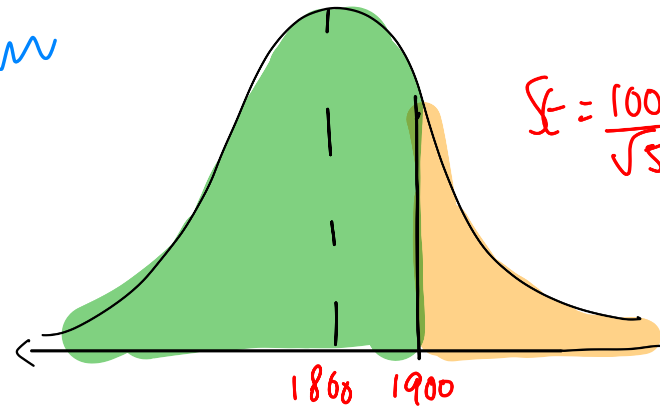
② M2 → 5 Stores

test statistic : m Sample mean

Distribution : Gaussian

$E[m] = 1800$

$SE = 100/\sqrt{5}$



$SE = \frac{100}{\sqrt{5}}$

$P[m \geq 1900 \mid H_0 \text{ is true}]$

$p = 0.012$ $\alpha = 0.05$

$\boxed{\text{Reject } H_0}$

$z = \dfrac{1900 - 1800}{100/\sqrt{5}}$

z statistics

* 95% Confidence $\quad 0.05 = \alpha$ Significance

$\alpha = 0.05$

1. Reject H_0 $\quad p < \alpha \quad 0.0002$
2. Reject H_0 $\quad p < \alpha \quad 0.012$

* 99% Confidence

$\alpha = 0.01$

Marketing leader effect

M1
1. $p = 0.0002 \quad p < \alpha$ Reject H_0

M2
2. $p = 0.012 \quad p > \alpha$ Fail to Reject H_0

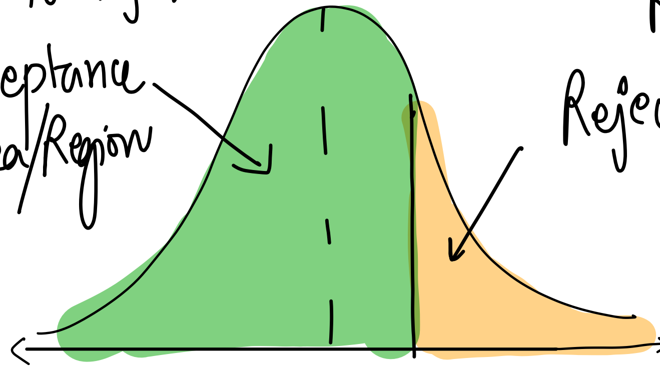
Confidence = 99% → 1% {chances}

95% → 5% {chances}

1/20 → Wrong Judgement

95%

Right Tailed Test

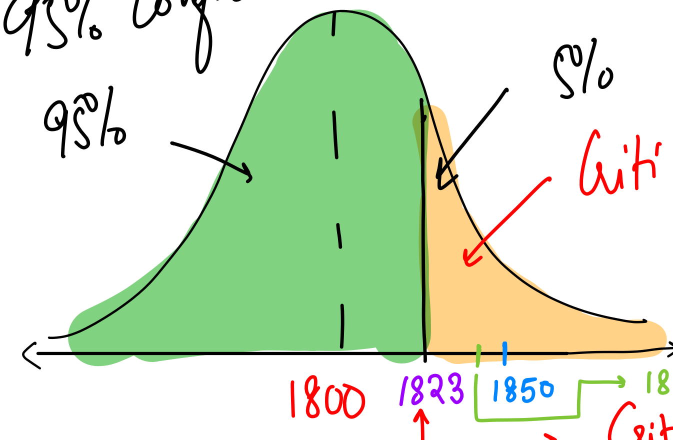


① MI

$\mu = 1800$

std $= 100/\sqrt{50}$

1850



① 95%

$Z_{95} = 1.64$

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

$x = 1800 + 1.64 \times \frac{100}{\sqrt{50}} = 1823$

$\boxed{x = \mu + Z \frac{\sigma}{\sqrt{n}}}$

99%

$Z_{99} = 2.32$

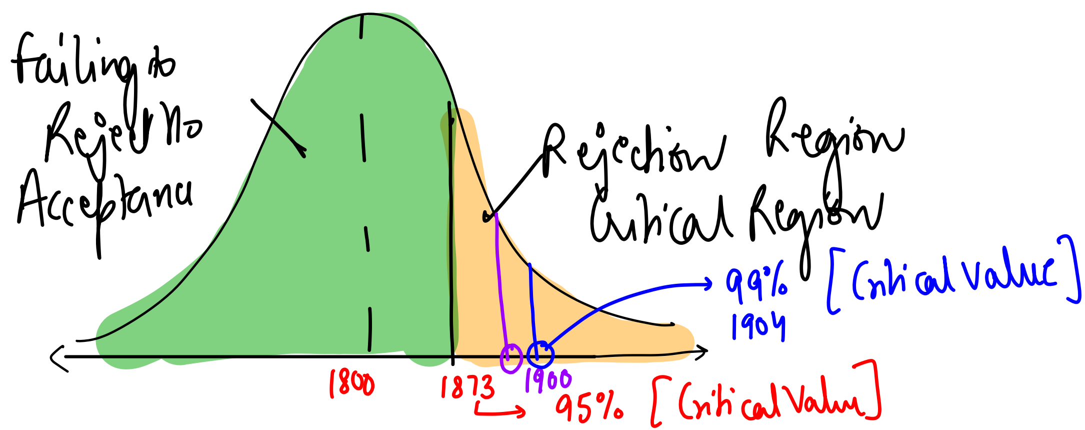
$x = 1800 + 2.32 \times \frac{100}{\sqrt{50}} = 1832.89$

M2

$\mu = 1800$

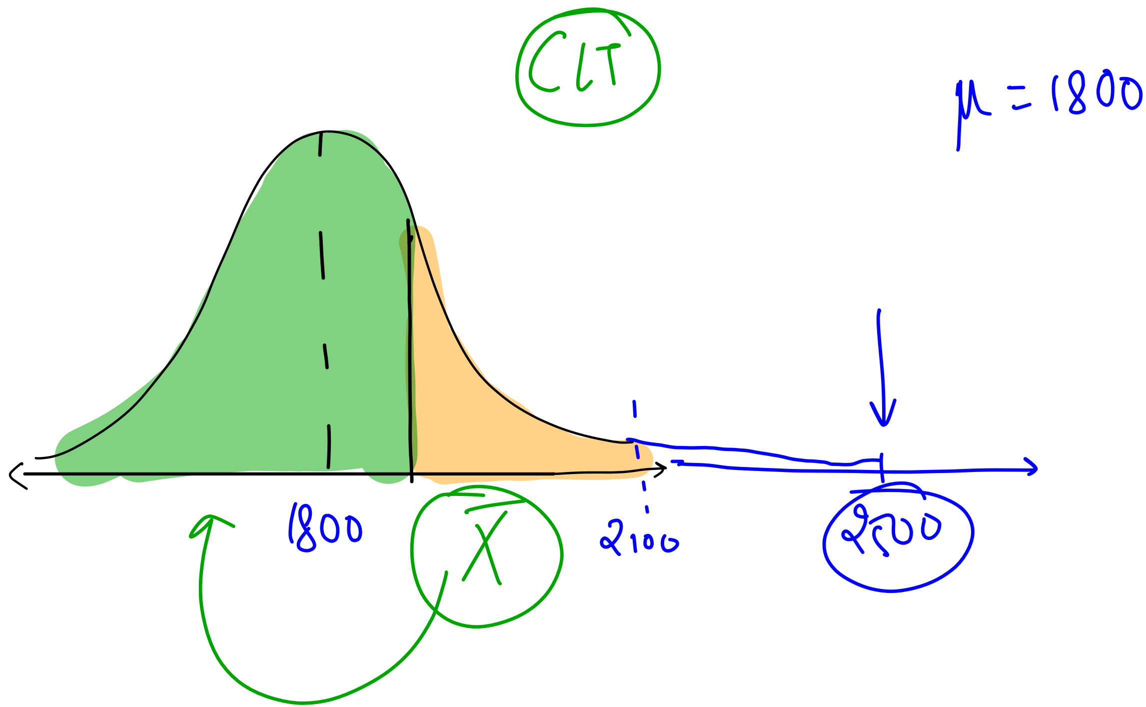
$SE = \frac{100}{\sqrt{5}}$

obs = 1900



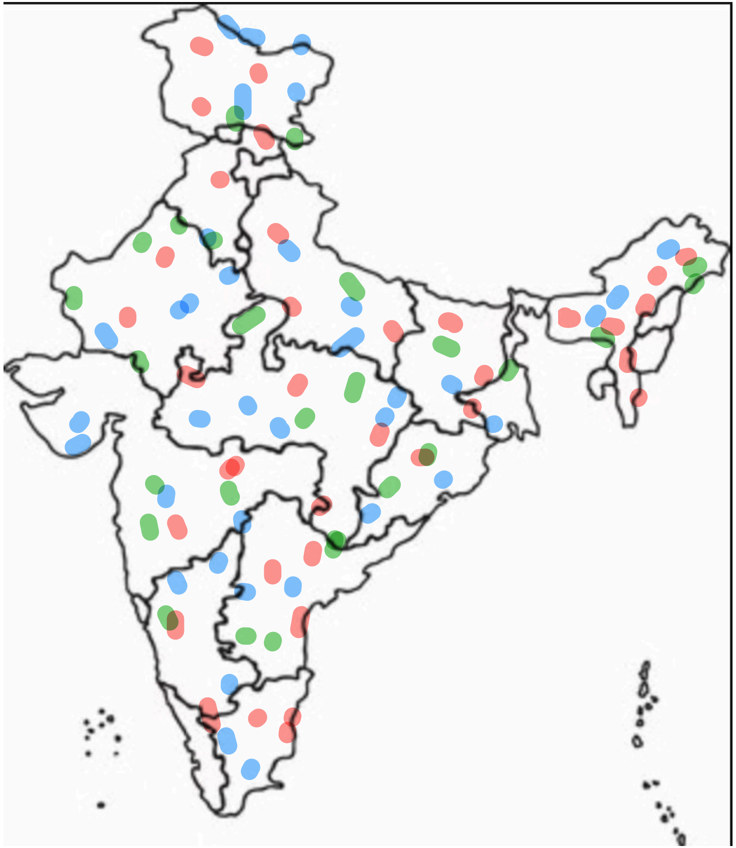
95% → Reject H_0

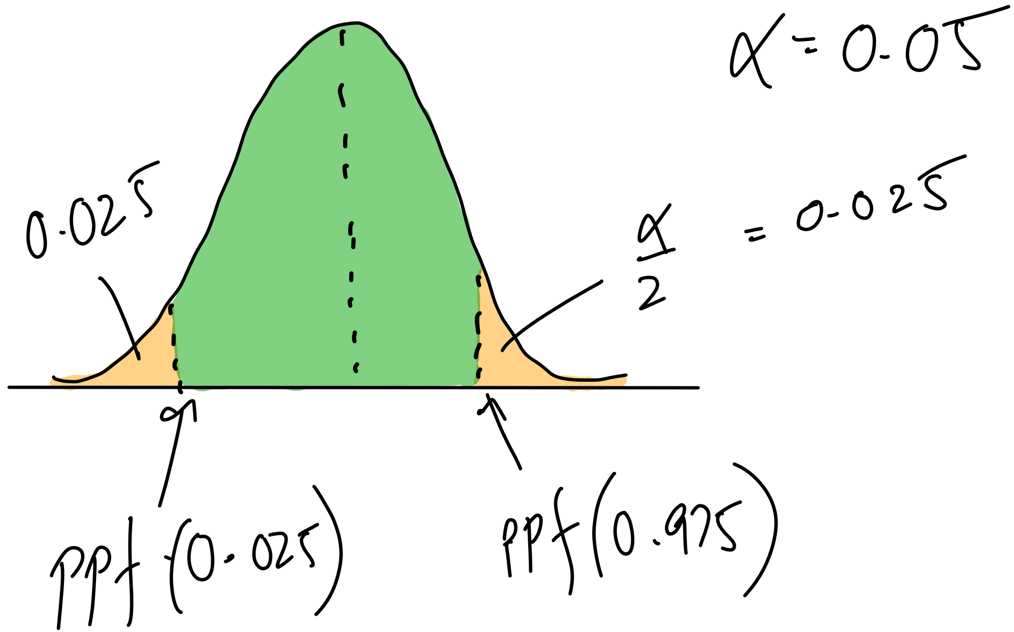
99% → Fail to Reject H_0.



$E(n)$

1. Binomial
2. Normal
3. Poisson
4. Exponential





α = 0.05
α/2 = 0.025
0.025
ppf(0.025)
ppf(0.975)