Z - Test

- Recap Framework of Hypothesis testing
- Recap CLT
- Case study → Improve sales

Word cloud (from last class)

Null Hypothesis (H_0) & alternate (H_a)

Test statistic (from sample or data)

Distribution

p-value (Under assumption of H_0, how likely it is to see the data & all extreme values)

If p-value is low $\longrightarrow$ we reject H_0

Significance level (α) 0.05 or 0.01

5% significance 95% confidence)

1% significance 99% confidence

p-value $< \alpha \longrightarrow$ Reject H_0

Type I Type II error

Right Vs Left Vs Two Tailed

Recap of CLT

Avg height $\rightarrow$ 65 inch
std dev $\rightarrow$ 2.5

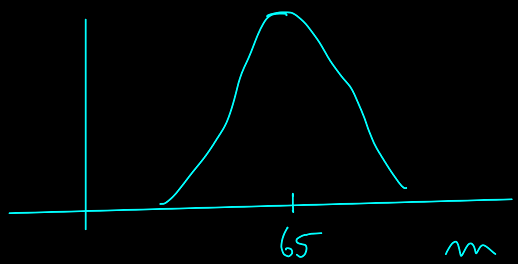
Take 50 samples

$m \rightarrow$ sample mean

$m \rightarrow$ Gaussian

avg or expected value of $m \rightarrow 65$

Stddev of $m \rightarrow \dfrac{2.5}{\sqrt{50}}$ "std error"



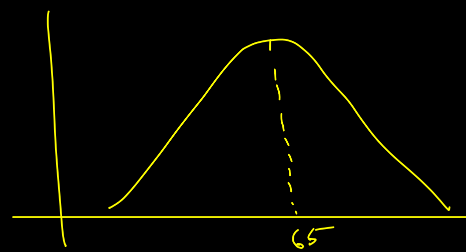
Std dev $\dfrac{2.5}{\sqrt{50}}$

Take 5 samples

$m \rightarrow$ Gaussian

avg or exp $\rightarrow 65$

Std dev $\rightarrow \dfrac{2.5}{\sqrt{5}}$



std dev $\dfrac{2.5}{\sqrt{5}}$

Sales case study : Shampoo bottle 2000 stores

Historical data : weekly sales $\mu = 1800$, $\sigma = 100$

Improve sales → market

Hire a team → expensive → test on a few stores

Before deploying to all 2000 stores, we test on 50 stores

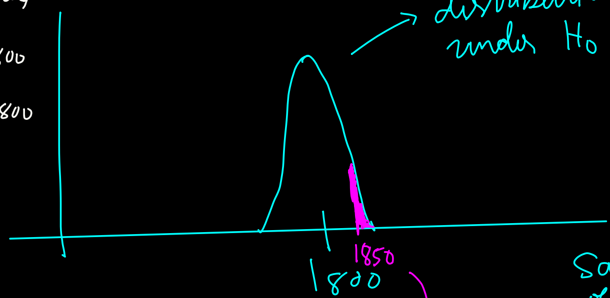
On those 50 stores, we see an average of 1850

1% significance, $\alpha = 0.01$

μ_m: sales with marketing

H_0 : marketing has no effect $\mu_m = 1800$

H_a : marketing has an effect $\mu_m > 1800$



Test statistic : sample mean

Distribution (under H_0) : Gaussian $1800, \frac{100}{\sqrt{50}}$

$$\text{Std dev} = \frac{100}{\sqrt{50}}$$

$$z = \frac{1850 - 1800}{(100/\sqrt{50})} = 3.53$$

p-value $P[T \geq 1850 \mid H_0] = 1 - \text{norm.cdf}(3.53)$

$= 0.0002$

p-value < 0.01 ? → Yes → Reject H_0

The marketing team had an effect

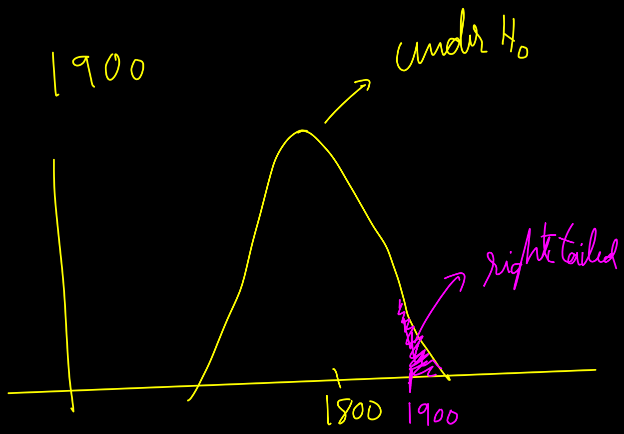
Second team :

On 5 stores $\longrightarrow$ avg 1900

H_0 : marketing has no effect

H_a : marketing has positive effect

T : sample mean



std dev = $\frac{100}{\sqrt{5}}$

$z = \frac{1900-1800}{100/\sqrt{5}} = 2.23$

p-value : $P[T \geq 1900 \mid H_0]$

$= 1 - \text{norm.cdf}(2.23)$

$= 0.012$

$\alpha = 0.01$

if p-value $< \alpha \longrightarrow$ no ! $\rightarrow$ do not reject H_0

Marketing has no effect

Supply chain example

$\alpha = 0.01$

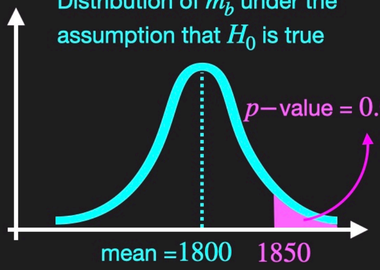
$\mu = 1800$
$\sigma = 100$

50 stores with average of 1850

$H_0 : \mu_b = 1800$

$H_a : \mu_b > 1800$

Distribution of m_b under the assumption that H_0 is true



std dev = $\dfrac{100}{\sqrt{50}}$

$$z = \frac{1850 - 1800}{100/\sqrt{50}} = 3.53$$

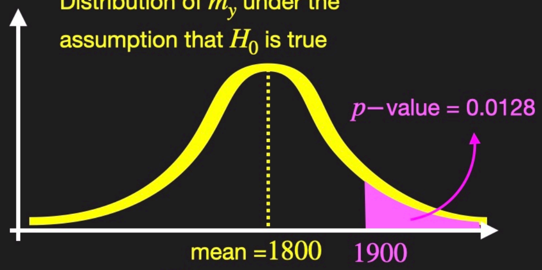
Reject H_0

5 stores with average of 1900

$H_0 : \mu_y = 1800$

$H_a : \mu_y > 1800$

Distribution of m_y under the assumption that H_0 is true

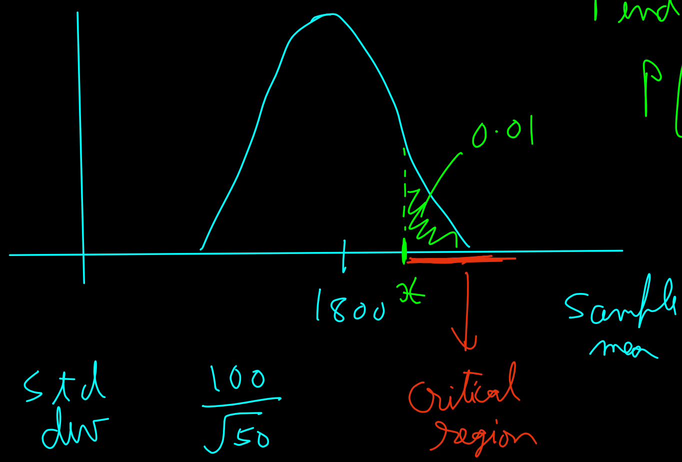


std dev = $\dfrac{100}{\sqrt{5}}$

$$z = \frac{1900 - 1800}{100/\sqrt{5}} = 2.23$$

Fail to reject H_0

$\mu = 1800$, $\sigma = 100$, $n = 50$ stores $\alpha = 0.01$



std dev $\frac{100}{\sqrt{50}}$

Find x such that

$$P[T \geq x \mid H_0] = 0.01$$

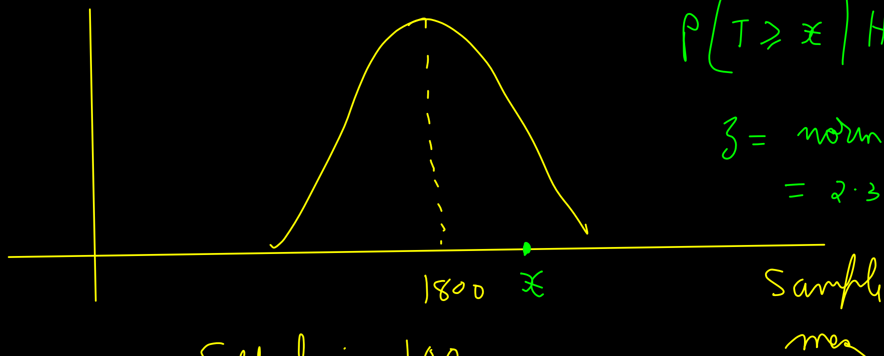
$$z = \text{norm.ppf}(0.99) = 2.32$$

$$x = 1800 + \frac{(2.32)(100)}{\sqrt{50}}$$

$$\boxed{x = 1832.8}$$

Critical value

$\mu = 1800$, $\sigma = 100$, 5 story, $\alpha = 0.01$



$P[T \geq x \mid H_0] = 0.01$

$z = \text{norm.ppf}(0.99) = 2.32$

Std dev: $\dfrac{100}{\sqrt{5}}$

$$x = 1800 + (2.32)\frac{100}{\sqrt{5}}$$

critical value $\boxed{x = 1903.7}$

Supply chain example

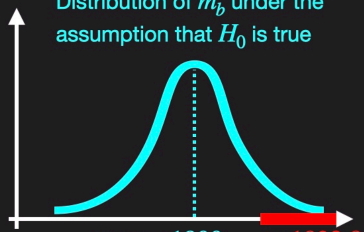
$\alpha = 0.01$

$\mu = 1800$
$\sigma = 100$

50 stores

$H_0 : \mu_b = 1800$
$H_a : \mu_b > 1800$

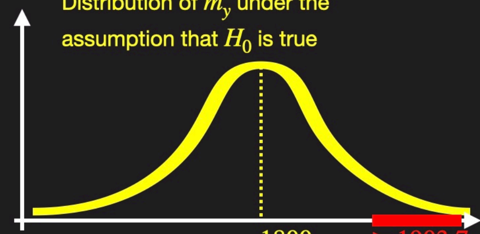
Distribution of m_b under the assumption that H_0 is true



5 stores

$H_0 : \mu_y = 1800$
$H_a : \mu_y > 1800$

Distribution of m_y under the assumption that H_0 is true



Note: For right-tailed test, the critical region is on the right
The probability associated with critical region is α

The rule to reject is very simple: If the observed test statistic is in the critical region, then reject the null hypothesis

Premature Babies : Once they grow up, how will their IQ be?

lower, normal, higher

Pop IQ : avg is 100, std dev 15

$\mu \rightarrow$ avg IQ of premature

$H_0 : \mu = 100$ (premature IQ = population IQ)

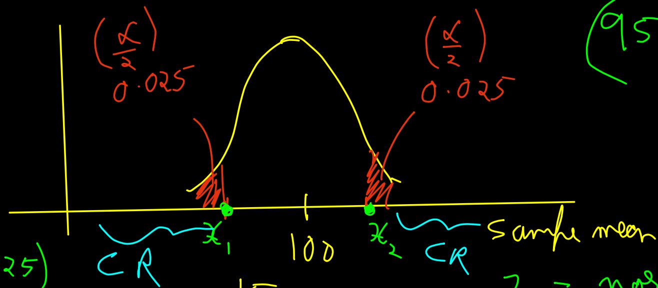
$H_a : \mu \neq 100 \longrightarrow$ 2-tailed

50 premature children

95% confidence (5% significance)

$\alpha = 0.05$

distribution of sample mean (n=50) under H_0



$(95.84, 104.15) \rightarrow$ normal IQ

$\underbrace{}_{H_0}$

$\frac{15}{\sqrt{50}}$

$z_1 = \text{norm.ppf}(0.025)$
$= -1.96$

$x_1 = 100 - 1.96\left(\frac{15}{\sqrt{50}}\right) = 95.84$

$\downarrow$
lower critical value

$z_2 = \text{norm.ppf}(0.975)$
$= 1.96$

$x_2 = 100 + 1.96\left(\frac{15}{\sqrt{50}}\right) = 104.15$

$\downarrow$
upper critical value

Framework

1. Setup Null & Alternate
2. Choose test statistic
3. Left Vs Right Vs 2-tailed
4. Compute p-values (computer critical region)
5. Compare p-value with α (test stat is in critical region)