

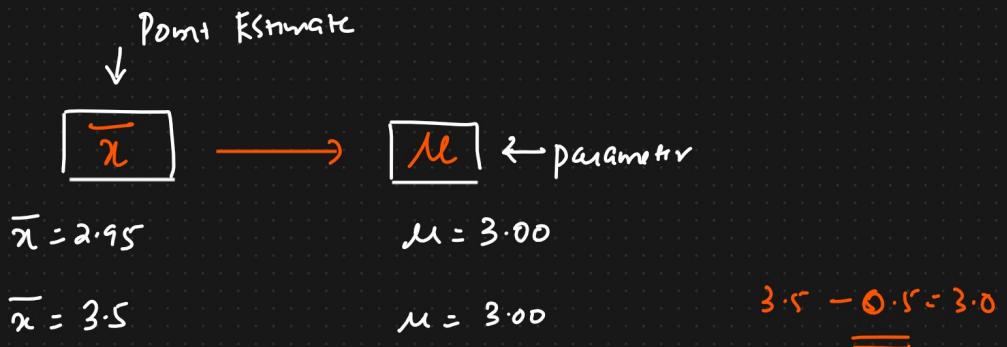
Agenda

- ① Point Estimate ✓
- ② Range of C.I ✓
- ③ Chi Square distribution ✓
- ④ F distribution. → F test ✓
- ⑤ [ANOVA] → Assignment \Rightarrow more than 2 groups

① Point Estimate

F Test

Dfn: The value of any statistic that estimates the value of a parameter is called Point Estimate



We rarely know if our point Estimate is correct because it is an estimation of the actual value

$$\bar{x}$$

We construct C.I to help estimate what the actual value of unknown population mean is.

Point Estimate \pm Margin of Error

Lower Range C.I = Point Estimate - Margin of Error

Higher Range C.I = Point Estimate + Margin of Error

- ① On the Verbal section of the CAT exam, a sample of 25 test takers has a mean of 520. With a standard deviation of 80. Construct a 95% C.I about the mean?

Ans) $\bar{x} = 520 \quad n=25 \quad [s=80] \quad C.I = 0.95 \quad \alpha = 0.05$

95%
↓
C.I

$$\begin{aligned} C.I &= \text{Point Estimate} \pm \text{Margin of Error} \\ &= \bar{x} \pm [t_{\alpha/2}] \left(\frac{s}{\sqrt{n}} \right) \Rightarrow \text{Standard Error} \\ &= 520 \pm 2.064 \left[\frac{80}{\sqrt{25}} \right] \quad [0.05] \end{aligned}$$

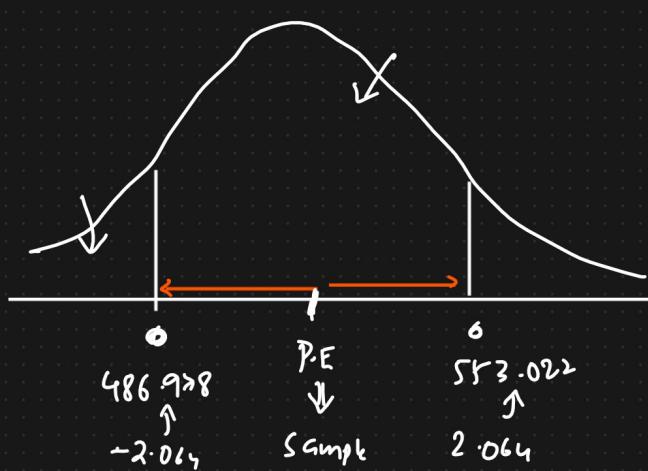
degree of freedom = $n-1 = 24$

= 24, /.

↓

$$\text{Lower C.I} = 520 - 2.064 \left[\frac{80}{\sqrt{5}} \right] = 486.978 \quad //.$$

$$\text{Higher C.I} = 520 + 2.064 \left[\frac{80}{\sqrt{5}} \right] = 553.022 \quad //$$



mean

② CHI SQUARE TEST

The Chi Square Test for Goodness of fit χ^2 tests claims about population proportion.
 [categorical variables]

It is a non parametric test that is performed on categorical data
 [ordinal, nominal data].

Eg: There is a population of Male who likes different color of Bike

	Theory	Sample	\Rightarrow Goodness of fit.
Yellow Bike	$\frac{1}{3}$	22	
Orange Bike	$\frac{1}{3}$	17	
Red Bike	$\frac{1}{3}$	59	\Rightarrow Observed categorical distribution

Theoretical
 (categorical distribution)

①

In 2010 Census of the city, the weight of the individuals in a small city were found to be the following

<50kg	50 - 75	>75
20%	30%	50%

In 2020, weight of $n=500$ individuals were sampled. Below are the results

<50	50 - 75	>75
140	160	200

Using $\alpha=0.05$, would you conclude the population difference of weights

has changed in last 10 years?

Ans)

In 2010

Expected

$< 50\text{kg}$	$50-75$	> 75
20%	30%	50%
0.2	0.3	0.5

In 2020

$n=500$

Observed

< 50	$50-75$	> 75
100	160	200



In 2010

Expected

$< 50\text{kg}$	$50-75$	> 75
500×0.2	500×0.3	500×0.5
= 100	= 150	= 250



①

Null hypothesis H_0 : The data meets the expectation

Alternate hypo H_1 : The data does not meet the expectation.

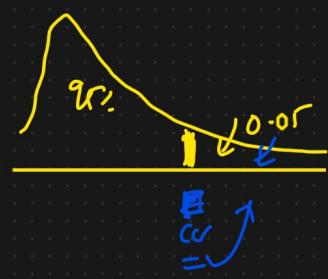
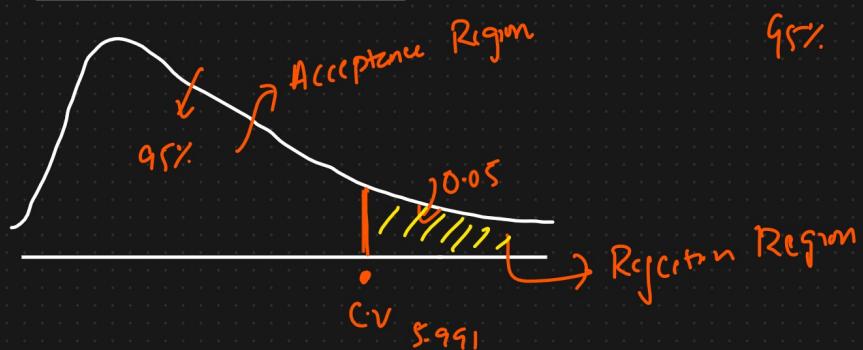
②

$$\alpha = 0.05 \quad (\cdot I = 95\%)$$

③ Degree of freedom

$$dof = k - 1 = 3 - 1 = 2$$

④ Decision Boundary



Chi square Test $\chi^2 > 5.991$ { Reject the Null Hypothesis }.

⑤ Calculate Chi Square Test Statistics

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$$= \frac{(140 - 110)^2}{100} + \frac{(160 - 150)^2}{150} + \frac{(200 - 250)^2}{250}$$

$$\chi^2 = 26.67$$

$\chi^2 > 5.991$ { Reject the Null Hypothesis }.

③ F-distribution

The F-distribution with d_1 and d_2 degrees of freedom is the distribution of

$$F = \frac{S_1/d_1}{S_2/d_2}$$

$S_1 \rightarrow$ Independent Random Variables } Chi square
 $S_2 \rightarrow$ Independent Random Variables } distribution

$d_1 \rightarrow$ Degree of freedom (S_1)

$d_2 \rightarrow$ Degree of freedom (S_2)

F test \rightarrow Variance Ratio Test { Comparing the variance between 2 groups }.

F Test [Variance Ratio Test].

- ① The following data shows the no. of bulbs produced daily for some days by 2 workers A and B

A	B	Can we consider based on the data or not
40	39	Worker B is more stable and efficient
30	38	$\alpha = 0.05$
38	41	$\underline{\underline{=}}$ 95% C.I.
41	33	
38	32	
35	39	
40	40	
34		

Ans) Null Hypothesis $H_0 : \sigma_1^2 = \sigma_2^2$

Alternate Hypothesis $H_1 : \sigma_1^2 \neq \sigma_2^2$

A			B		
X_i	\bar{x}	$(x_i - \bar{x})^2$	X_2	\bar{x}_2	$(x_2 - \bar{x})^2$
40	37	9	39	37	4
30	37	49	38	37	1
38	37	1	41	37	16
41	37	16	33	37	16
38	37	1	32	37	25
35	37	4	39	37	4
<hr/>			40	37	9
$\bar{X}_1 = 37$		$\sum (x_i - \bar{x})^2 = 80$	$\bar{X}_2 = 37$		$\frac{9}{7} = 84$

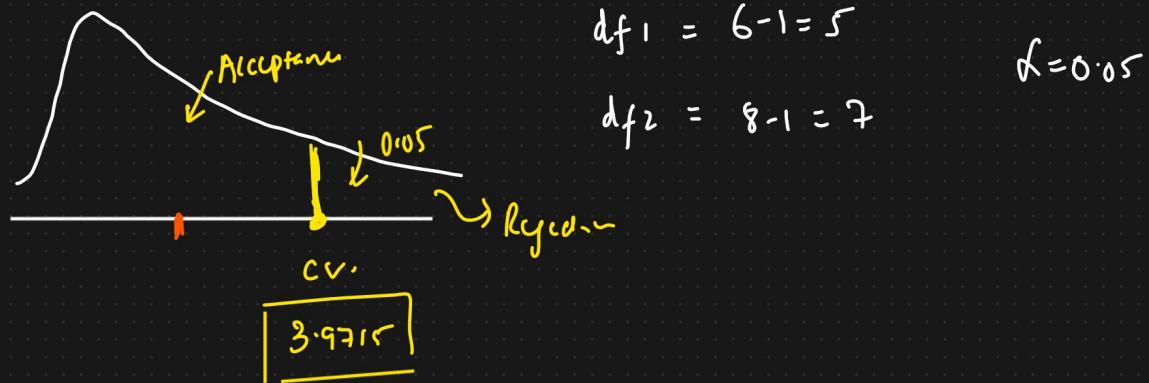
$$S_1^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{80}{5} = 16.$$

$$S_2^2 = \frac{84}{7} = 12.$$

* Variance Ratio: $[F\text{-test}]$

$$F = \frac{S_1^2}{S_2^2} = \frac{16}{12} = 1.33$$

* Decision Rule $[F \text{ distribution}]$



$F\text{-test} > 3.9715 \quad \{ \text{Reject } H_0 \}$.

Is $1.33 > 3.9715 \Rightarrow \text{False}$

We fail to Reject the Null Hypothesis



Worker A \approx Worker B