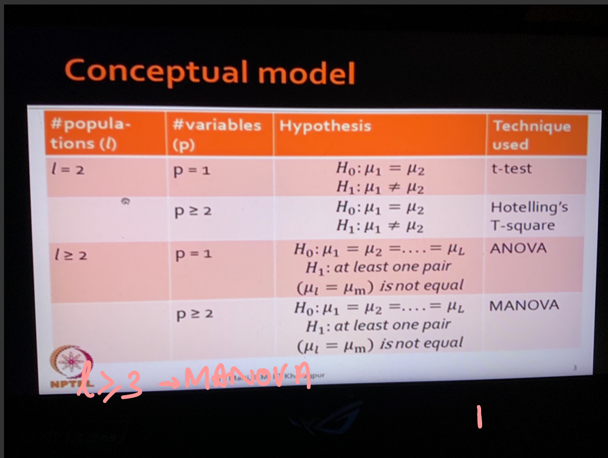
# MANOVA [Multivariate Analysis of Variance]

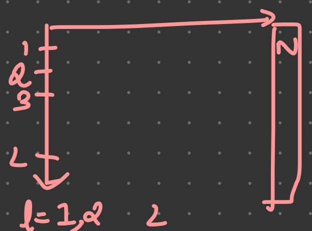


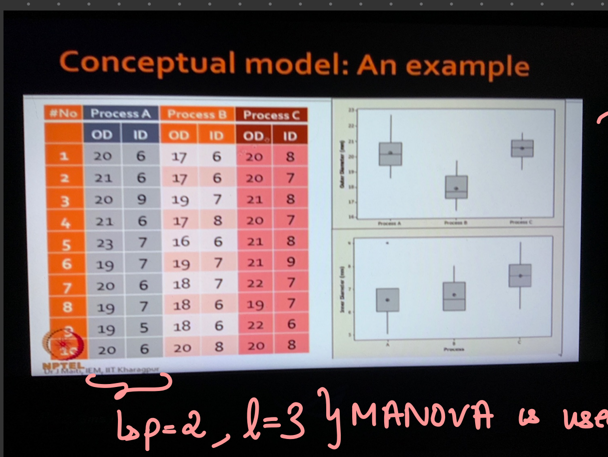

| #populations ($l$) | #variables (p) | Hypothesis | Technique used |
|---|---|---|---|
| $l = 2$ | p = 1 | $H_0: \mu_1 = \mu_2$<br>$H_1: \mu_1 \neq \mu_2$ | t-test |
| | p ≥ 2 | $H_0: \mu_1 = \mu_2$<br>$H_1: \mu_1 \neq \mu_2$ | Hotelling's T-square |
| $l \geq 2$ | p = 1 | $H_0: \mu_1 = \mu_2 = \ldots. = \mu_L$<br>$H_1$: at least one pair ($\mu_l = \mu_m$) is not equal | ANOVA |
| | p ≥ 2 | $H_0: \mu_1 = \mu_2 = \ldots. = \mu_L$<br>$H_1$: at least one pair ($\mu_l = \mu_m$) is not equal | MANOVA |

$l \geq 3 \rightarrow$ MANOVA

ANOVA

$i = 1, 2, 3$    n observations



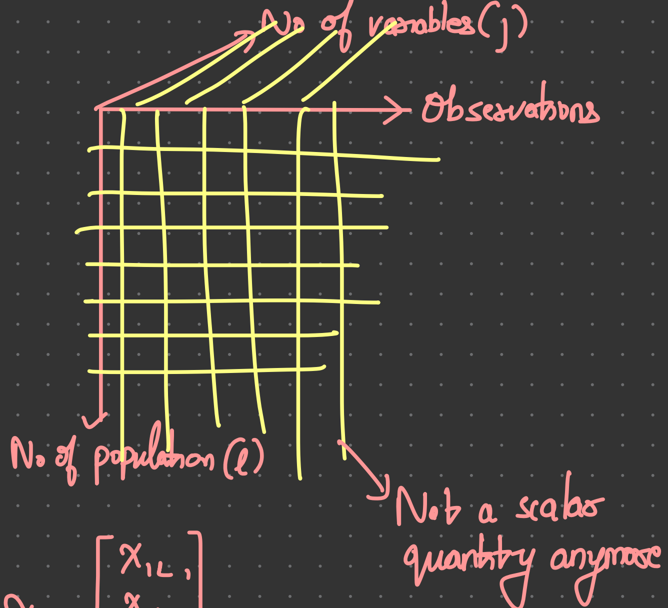




| #No | Process A OD | Process A ID | Process B OD | Process B ID | Process C OD | Process C ID |
|---|---|---|---|---|---|---|
| 1 | 20 | 6 | 17 | 6 | 20 | 8 |
| 2 | 21 | 6 | 17 | 6 | 20 | 7 |
| 3 | 20 | 9 | 19 | 7 | 21 | 8 |
| 4 | 21 | 6 | 17 | 8 | 20 | 7 |
| 5 | 23 | 7 | 16 | 6 | 21 | 8 |
| 6 | 19 | 7 | 19 | 7 | 21 | 9 |
| 7 | 20 | 6 | 18 | 7 | 22 | 7 |
| 8 | 19 | 7 | 18 | 6 | 19 | 7 |
| 9 | 19 | 5 | 18 | 6 | 22 | 6 |
| 10 | 20 | 6 | 20 | 8 | 20 | 8 |

} will test collectively

$\hookrightarrow p = 2,\ l = 3$ } MANOVA is used

$p \geq 2$ for MANOVA

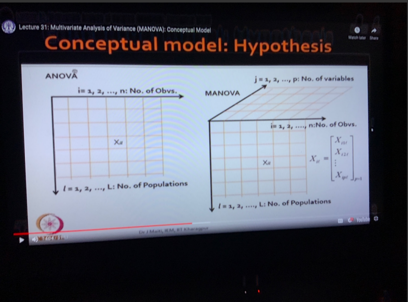

$$X_{il} = \begin{bmatrix} x_{il1} \\ x_{il2} \\ \\ \\ x_{ilp} \end{bmatrix}_{p\times 1}$$

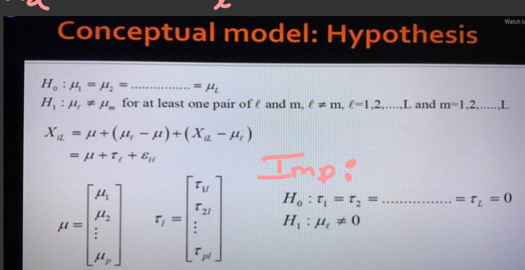
General Observations

$$\mu_l = \begin{bmatrix} \mu_{l1} \\ \mu_{l2} \\ \\ \\ \mu_{lp} \end{bmatrix}_{p\times 1} \qquad \Sigma_l = \begin{bmatrix} & \\ & \end{bmatrix}_{p\times p} \qquad l = 1, 2 \quad l$$

$$\Sigma_1, \Sigma_2 - - \Sigma_l$$

$$\mu_1, \mu_2 \quad - - \mu_l$$



# ANOVA

$$x_{1l} = \mu + (\mu_l - \mu) + (x_{1l} - \mu_l)$$

$$= \mu + \tau_l + \varepsilon_{11}$$

Grand Mean

Pop Effect

Random Error

# MANOVA

$$x_{1l} = \boxed{\mu} + \boxed{\tau_l} + \boxed{\varepsilon_{1l}}$$

$$x_{1l} = \boxed{\mu} + \boxed{(\mu_l - \mu)} + \boxed{(x_{1l} - \mu_l)}$$

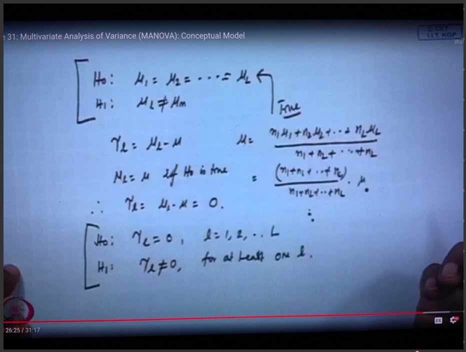
$$\begin{bmatrix} x_{1l1} \\ \vdots \\ x_{1lp} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} + \begin{bmatrix} \mu_{l1} - \mu_1 \\ \mu_{l2} - \mu_2 \\ \vdots \\ \mu_{lp} - \mu_p \end{bmatrix} + \begin{bmatrix} x_{1l1} - \mu_{l1} \\ \vdots \\ x_{1lp} - \mu_{lp} \end{bmatrix}$$

General Obs vector

Grand Mean Vector

Population effect vector

Random Error Vector




$$x_{1l} = \mu + \tau_l + \varepsilon_{1l}$$

$$\sum_{l=0}^{L} \tau_l = 0 \rightarrow \text{Equal sample size}$$

$$\sum_{l=1}^{L} n_l \tau_l = 0 \rightarrow \text{Unequal Sample size}$$

$$\sum_{l=1}^{L} n_l \tau_l = 0$$

$$LHS = \sum n_l(\mu_l - \mu)$$

$$LHS = \sum n_l \mu_l - \sum n_l \mu$$

$$LHS = (n_1\mu_1 + n_2\mu_2 \ldots) - (n_1 + n_2 \ldots)\mu$$

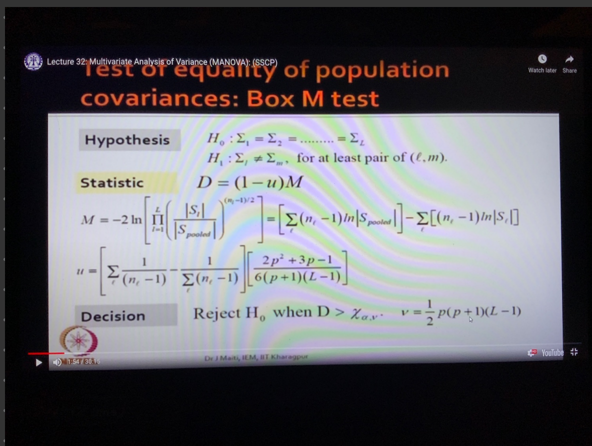
$$LHS = 0$$

$$\mu = \boxed{\text{Grand Mean} = \frac{n_1\mu_1 + n_2\mu_2 \ldots + n_L\mu_L}{n_1 + n_2 \ldots + n_L}}$$

# Lecture-2

Assumptions:

1. Population covariances are equal
2. Errors are normally distributed
3. Errors are IID




$$S_{pooled} = \frac{(n_1-1)S_1 + (n_2-1)S_2 \;--\; + (n_L-1)S_L}{[n_1+n_2 \;-\;-\; + n_L] - L}$$

If $\Sigma_1 = \Sigma_2 \quad \Sigma_L$, then $|S_l| = |S_{pooled}|$

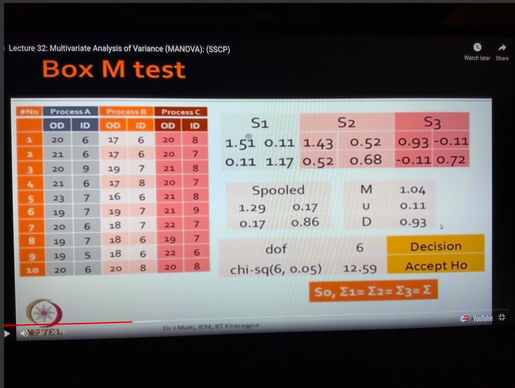
↳ Ratio will become one

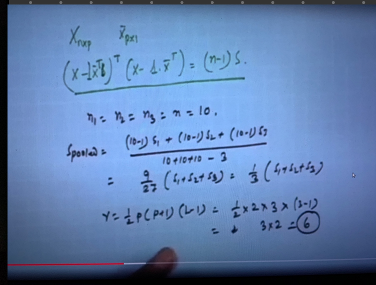
ln is taken for linearisation

$v = \frac{1}{2}p(p+1)(L-1)$ — DOF for D

If $D > \chi^2_{\alpha,v}$ we reject $H_0$ — variances aren't equal

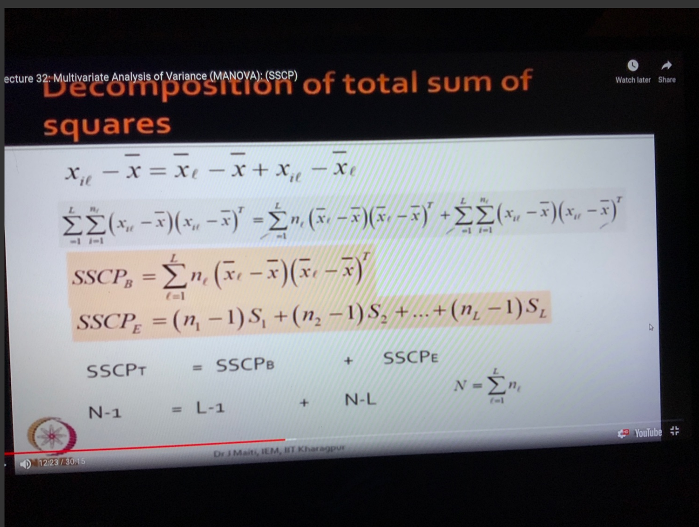
Example:




$X_{n\times p}$ $\quad \bar{x}_{p\times 1}$

$$(X-1\cdot\bar{x}^T)^T(X-1\bar{x}^T) = (n-1)S$$ → Same formula is used here



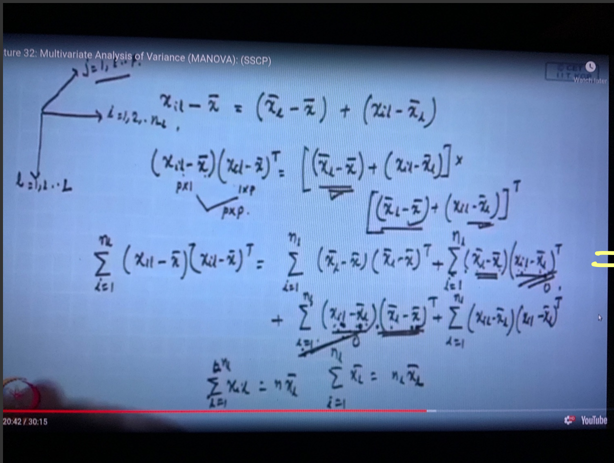

For ANOVA,

$$x_{il} = u + (u_l - u) + (x_{il} - u_l)$$

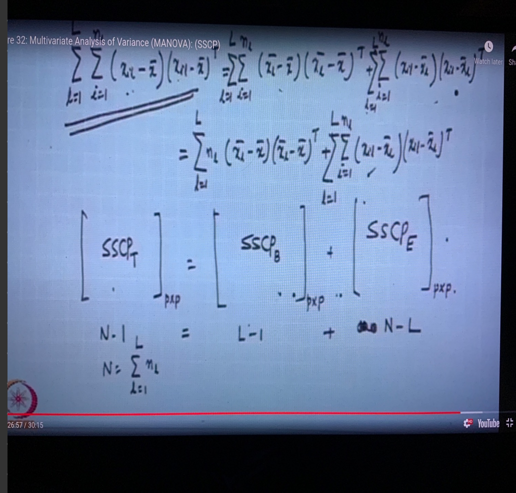
$$\hat{u} = \bar{x}, \quad \hat{u}_l = \bar{x}_l$$

$$\underset{(p\times 1)}{x_{il}} = \underset{(p\times 1)}{\bar{x}} + \underset{(p\times 1)}{(\bar{x}_l - \bar{x})} + \underset{(p\times 1)}{(x_{il} - \bar{x}_l)}$$

Similarly done for MANOVA




$\Longrightarrow$



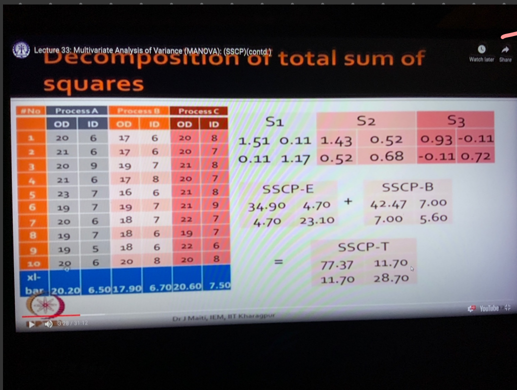

$$SSCP_B = \sum_{l=1}^{L} n_l (\bar{x}_l - \bar{x})(\bar{x}_l - \bar{x})^T$$

$$SSCP_E = (n_1 - 1)S_1 + (n_2 - 1)S_2 \; \cdots \; + (n_L - 1)S_L$$

$$\boxed{SSCP_T = SSCP_B + SSCP_E}$$

→ Decomposition of SSCP Matrix

# Lecture 3



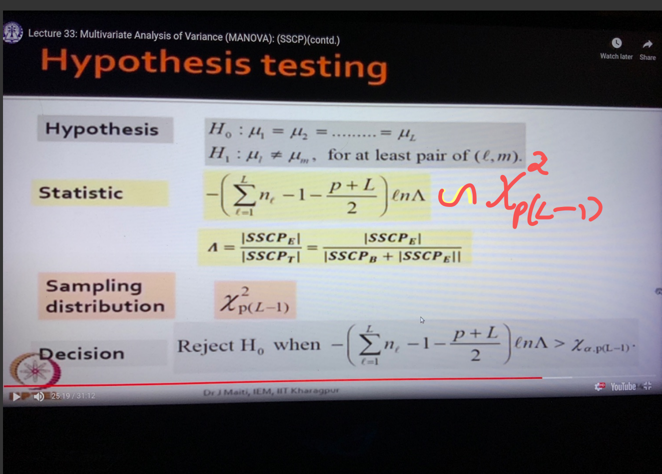

→ Example solving at start

## One way MANOVA [Population of one kind]

| Source of Variation | SSCP | DOF | Wilk's $\Lambda$ |
|---|---|---|---|
| population | $SSCP_B$ | $L-1$ | $\Lambda = \frac{\lvert SSCP_E \rvert}{\lvert SSCP_T \rvert}$ |
| error | $SSCP_E$ | $N-L$ | $\Lambda = \frac{\lvert SSCP_E \rvert}{\lvert SSCP_E \rvert + \lvert SSCP_B \rvert}$ |
| Total | $SSCP_T$ | $N-1$ | |

should be also low as possible

$$N = \sum_{l=1}^{L} n_l$$

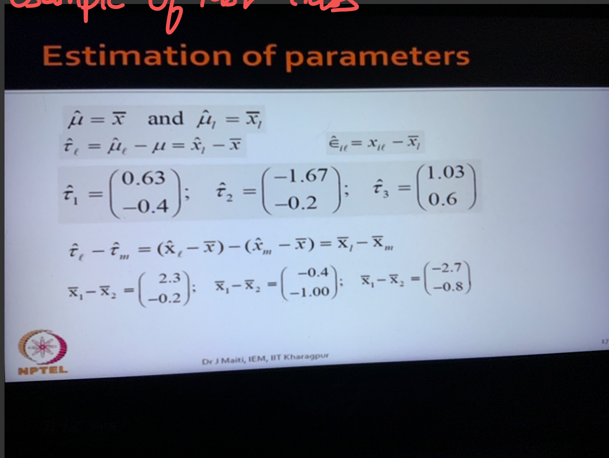



$\sim \chi^2_{p(L-1)}$

Example in end

# Lecture 4

Using example of last class




$$E(\hat{\tau}_\ell - \hat{\tau}_m) = E[\bar{x}_\ell - \bar{x}_m]$$

$$= E[\bar{x}_\ell] - E[\bar{x}_m]$$ [As independent]

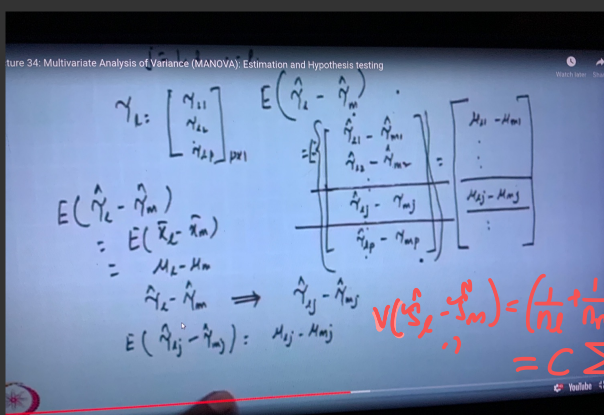
$V(\hat{\tau}_\ell - \hat{\tau}_m)$

$$V(\bar{x}_\ell - \bar{x}_m) = V(\bar{x}_\ell) + V(\bar{x}_m)$$

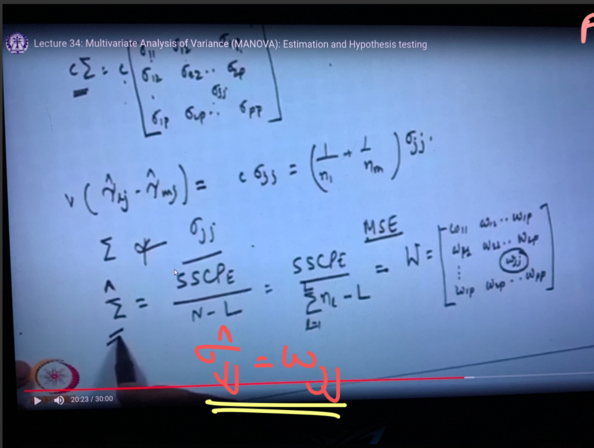
Carably

$$= \frac{\Sigma_\ell}{n_\ell} + \frac{\Sigma_m}{n_m} = \boxed{\left(\frac{1}{n_\ell} + \frac{1}{n_m}\right)\Sigma}$$

Assum $\Sigma_1 = \Sigma_2 - \Sigma_L$




$$V(\hat{\tau}_\ell - \hat{\tau}_m) = \left(\frac{1}{n_\ell} + \frac{1}{n_m}\right)\Sigma$$
$$= c\,\Sigma$$



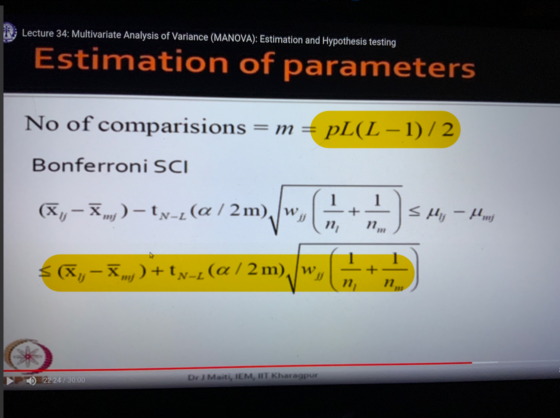

$\hat{\sigma}_{jj} = w_{jj}$

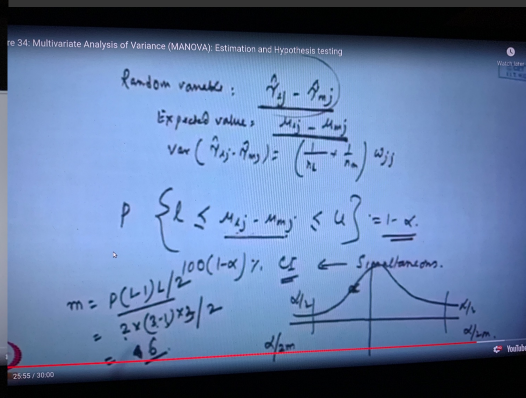
Find interval estimate

Random Variable: $\hat{\tau}_{\ell j} - \hat{\tau}_{mj}$

Expectation: $\mu_{\ell j} - \mu_{mj}$

$$Var(\hat{\tau}_{\ell j} - \hat{\tau}_{mj}) = \left(\frac{1}{n_\ell} + \frac{1}{n_m}\right) w_{jj}$$




If any interval contains zero, then that variable is not differenciating.

Ex $-0.25 \le$ ___ $\le 1.30$

Contains 0 [Not creating differenciating variable]