NAME: Vandana Chudasama

STD: MSC DIV: Previous. ROLL NO.:

YOUVA

SUBJECT: Multivariate Analysis

INDEX

SR. NO.	DATE	TITLE	PAGE NO.	TEACHER'S SIGN

Multivariate Analysis

In the study of statistics the variable under consideration so far where one or two or three and the objective of study was some characterstics of a group of individuals or objects. Then we have discussed descriptive and inferential statistics.

We may study more than three variables together so that it will generate a multivariate data in the form of a vector

$$X_{p\times 1} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}$$

, suppose we have collected the data on 'n' individuals / objects then we shall have 'n' 'p' components vectors as the observations and we have the observation matrix $X_{p\times n}$

The analysis of Multivariate data is known as the Multivariate Analysis.

The Variate

It is a linear combination of the variables with empirically determined weights. The variable are specified by the researcher where as the weights are determined by Multivariate techniques to meet the specific objectives. A variate of the 'n' variables $x_1, x_2, \ldots x_n$ can be stated Mathematically as variate value $= w_1x_1 + w_2x_2 + \ldots w_nx_n$ where x_n is the observed value and w_n is the weight determined by Multivariate techniques

* Analysis

$$x = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & \\ \vdots & & & \\ x_{2p} & \cdots & \cdots & x_{pn} \end{bmatrix}_{p \times n}$$

Define : $E(x) = \begin{bmatrix} E(x_1) \\ E(x_2) \\ \vdots \\ E(x_n) \end{bmatrix}_{n \times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{bmatrix}_{n \times 1} = \mu_{n \times 1}$

where $E(x_i) = \mu_i$

We have $var(x) = E[x_i - E(x)]^2$

and Multivariate case

$var(\underline{x}) = E[\underline{x} - E(\underline{x})](\underline{x} - E(\underline{x}))'$

Let $\Sigma_{p \times p}$ = be the variance - covariance matrix of $\underline{x}$ then (Dispersion matrix)

$$\Sigma = E \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \\ \vdots \\ x_p - \mu_p \end{bmatrix} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 & \cdots & x_p - \mu_p \end{bmatrix}$$

$$= E \begin{bmatrix} (x_1 - \mu_1)^2 & (x_1 - \mu_1)(x_2 - \mu_2) & \cdots & (x_1 - \mu_1)(x_p - \mu_p) \\ (x_2 - \mu_2)(x_1 - \mu_1) & (x - \mu_2)^2 & & \vdots \\ \vdots & \vdots & & \\ (x_p - \mu_p)(x_1 - \mu_1) & (x_p - \mu_p)(x_2 - \mu_p) & & (x_p - \mu_p)^2 \end{bmatrix}$$

OR

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{21} & \cdots & \sigma_{p1} \\ \sigma_{12} & \sigma_{22} & & \vdots \\ \vdots & & & \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_p^2 \end{bmatrix}$$

where $\sigma_i = E(x_i - \mu_i)^2 \quad i = 1, \dots p$

$$\sigma_{ij} = E(x_i - \mu_i)(x_j - \mu_j) \quad i \neq j$$

if $x_1, x_2, \dots x_p$ are independent then

$$\Sigma_{p \times p} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_p^2 \end{bmatrix}$$

Theorem: The variance - covariance matrix $\Sigma_{p \times p}$ is atleast positive semidefinite.

proof :- By definition we have $\Sigma = E(\underline{x} - \underline{\mu})(\underline{x} - \underline{\mu})'$
Define

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} \quad \text{and} \quad d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_p \end{bmatrix}$$

Also transform $\underline{Y} = (\underline{x} - \underline{\mu})$ then

$E(\underline{Y}) = E(\underline{x} - \underline{\mu}) = \underline{0}$.

then $d' \Sigma d = \underline{d}' E(\underline{Y}\,\underline{Y}') d = E(d' \underline{Y}\,\underline{Y}' d)$

Also

$$d' \underline{Y} = Z = d_1 Y_1 + d_2 Y_2 + \dots + d_p Y_p$$

$$E(Z) = E(d' \underline{Y}) = \underline{d}' E(\underline{Y}) = \underline{0}$$

$$Var(Z) = E(Z^2) - (E(Z))^2 = Var(Z) = E(Z^2) \geq 0$$

$\therefore \Sigma$ is atleast positive semidefinite

Note :- $\Sigma_{p \times p}$ is square symmetric and positive semidefinite Matrix

* Multivariate Normal distribution

For a random Variable X the univariate normal density function can be written as

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

which can be written in a general form

$$f(x;\mu,\sigma^2) = k e^{-\frac{1}{2}(x-B)'\alpha(x-B)} \quad \text{—(1)}$$

Where α is positive and the constant 'k' is determined such that the integral of eq (1) is unity.

The density function of Multivariate Normal Distribution (MVND) has an analogous form as that of a univariate Normal distribution.

Let $x_1, x_2, \ldots x_p$ are p- independent normal variate and $x_i \sim N(\mu_i, \sigma_i^2)$ then the p.d.f of $(x_1, x_2, \ldots x_p)$ is given by

$$f(x_1, x_2, \ldots x_p) = \prod_{i=1}^{p}\left(\frac{1}{\sqrt{2\pi}\,\sigma_i} e^{-\frac{1}{2}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2}\right) \quad \text{—— (*)}$$

$$= \frac{1}{(2\pi)^{p/2}} \frac{1}{\sigma_1 \cdot \sigma_2 \cdots \sigma_p} e^{-\frac{1}{2}\sum_{i=1}^{p}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2}$$

As x_i's are independent $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_p^2 \end{bmatrix}$

and $\Sigma^{-1} = \text{diag}\left(\frac{1}{\sigma_1^2} \ldots \frac{1}{\sigma_p^2}\right)$

as x is replaced by $\underline{x}$ consider the quantity

$= (\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})$

$$= (x_1-\mu_1, \; x_2-\mu_2 \ldots x_p-\mu_p)\begin{bmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2^2 & & \vdots \\ 0 & 0 & \cdots & 1/\sigma_p^2 \end{bmatrix}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ x_p-\mu_p \end{bmatrix}$$

$$= \sum_{i=1}^{p}\left(\frac{x_i-\mu_i}{\sigma_i}\right)^2$$

Also, $|\Sigma|^{1/2} = \sigma_1 \cdot \sigma_2 \cdots \sigma_p$

Substituting the above equation quantities in $(*)$ we get

$$f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{(\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})}$$

Notation: $X \sim N_p(\underline{\mu}, \Sigma)$

* <u>Derivation (General Form)</u>

In general the exponential family of distbution in a Multivariate form can be written as

$$f(x) = K e^{-\frac{1}{2}(\underline{x}-\underline{B})'A(\underline{x}-\underline{B})}$$

Where $\underline{B}$: Vector, A : +ve definite matrix (Symmetric)

as A is $(\underline{x}-\underline{B})'A(\underline{x}-\underline{B}) > 0$

$e^{-1/2(\underline{x}-\underline{B})'A(\underline{x}-\underline{B})} \leq 1 \Rightarrow f(x)$ will be bounded f^n

* Determination of K

Consider the transformation $(\underline{X}-\underline{B}) = C\underline{Y}$ where 'C' is a Non-Singular matrix (NSM) and is determined by using the result that: if 'A' is a positive definite matrix then there exists a NSM 'C' such that

$$C'AC = I \Rightarrow |C'||A||C| = |I|$$

Since $(\underline{X}-\underline{B}) = C\underline{Y}$

$\therefore$ Jackobian of the transformation is $|J| = |C|$

As $f(x) = k e^{-\frac{1}{2}(\underline{X}-\underline{B})' A (\underline{X}-\underline{B})}$

$$\therefore f(\underline{Y}) = k e^{-\frac{1}{2}\underline{Y}' C'AC\underline{Y}} \cdot |J|$$

or,

$$f(\underline{Y}) = k e^{-\frac{1}{2}\underline{Y}'\underline{Y}} \cdot |J| \quad \text{as } C'AC = I$$

or,

$$f(\underline{Y}) = k e^{-\frac{1}{2}\sum_{i=1}^{p} y_i^2} |C|$$

$$\therefore \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} k e^{-1/2 \sum_{i=1}^{p} y_i^2} |C| \, dy_p \cdots dy_1 = 1$$

$$\text{Or. } k|C| \prod_{i=1}^{p} \left\{ \int e^{-\frac{1}{2} y_i^2} \, dy_i \right\} = 1$$

$$\therefore k|C| (\sqrt{2\pi})^p = 1$$

as

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2 y_i^2} \, dy_i = 1$$

$$\Rightarrow \int_{-\infty}^{\infty} e^{-1/2 y_i^2} \, dy_i = \sqrt{2\pi}$$

$\therefore K = \frac{1}{(2\pi)^{p/2}|C|}$

Again $f(\underline{Y}) = \frac{1}{(2\pi)^{p/2}\cancel{|C|}} e^{-\frac{1}{2}\underline{Y}'\underline{Y}} \cdot \cancel{|C|}$

$= \frac{1}{(2\pi)^{p/2}} e^{-\frac{1}{2}\underline{Y}'\underline{Y}}$

as $C'AC = I \Rightarrow |C'|\,|A|\,|C| = 1$

$\Rightarrow |C|^2|A| = 1$ (as $|C'| = |C|$

$\Rightarrow |C| = \frac{1}{|A|^{1/2}}$

$\left(K = \frac{1}{(2\pi)^{p/2}} \frac{1}{\frac{1}{|A|^{1/2}}} = \frac{|A|^{1/2}}{(2\pi)^{p/2}} \right)$

* Determination of 'B'

we have $f(Y) = \left(\frac{1}{2\pi}\right)^{p/2} e^{-\frac{1}{2}\underline{Y}'\underline{Y}}$

Consider

$E(y_i) = \int_{-\infty}^{\infty} y_i \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y_i^2} dy \prod_{\substack{j=1 \\ j\neq i}}^{p} \left\{ \frac{1}{(2\pi)^{\frac{p-1}{2}}} e^{-\frac{1}{2}y_j^2} dy_j \right\}$

$= 0 \times \underline{1} = 0$

$\therefore E(y_i) = 0 \quad \forall i = 1, \dots p$

$\therefore E(\underline{Y}) = \underline{0}$ (Null vector)

Since $(\underline{X} - \underline{B}) = C\underline{Y} \Rightarrow \underline{X} = \underline{B} + C\underline{Y}$

$\therefore E(\underline{X}) = \underline{B} + C E(\underline{Y}) = \underline{B}$

But $E(\underline{X})$ is the mean vector $\underline{X}$ denoted by $\underline{\mu}$

$\boxed{\therefore \underline{B} = \underline{\mu}}$

* Determination of 'A'

By defn we have.

$\Sigma = E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'$

$= E(\underline{X} - \underline{B})(\underline{X} - \underline{B})'$

$= E(C\underline{Y}\underline{Y}'C')$ OR $\Sigma = C E(\underline{Y}\underline{Y}')C'$.

in order to obtain this consider the 2nd moment of $\underline{Y}$ and obtain $E(y_j, y_k)$ $j \neq k$

Now

$$E(y_j, y_k) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} y_j y_k \frac{1}{(2\pi)^{p/2}} e^{-\frac{1}{2}\underline{Y}'\underline{Y}} dy_p \ldots dy_1$$

$$= \left\{ \int_{-\infty}^{\infty} y_j \frac{1}{\sqrt{2\pi}} e^{-y_j^2/2} dy_j \right\} \left\{ \int_{-\infty}^{\infty} y_k \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y_k^2} dy_k \right\}$$

$$\prod_{\substack{l=1 \\ l \neq j \neq k}}^{p} \left\{ \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{p-1}{2}}} e^{-\frac{1}{2} y_l^2} dy_l \right\}$$

$\therefore$ or $E(y_j, y_k) = 0$ for $j \neq k$

if $j = k$,

$$E(y_j^2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \underline{y_j}^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} y_j^2/2} dy_j$$

$$\prod_{\substack{k=1 \\ k \neq j}}^{p} \left\{ \int_{-\infty}^{\infty} \frac{1}{(2\pi)^{\frac{p-1}{2}}} e^{-\frac{1}{2} y_k^2} dy_k \right\}$$

$$\therefore E(y_j, y_k) = 0, \quad j \neq k$$
$$= 1, \quad j = k$$

By Defn $\quad VCM(\underline{Y}) = E[\underline{Y} - E(\underline{Y})][\underline{Y} - E(\underline{Y})]'$

$$= E[\underline{Y} - 0][Y - 0]'$$
$$= E[Y'Y]$$

(VCM = Variance - covariance Matrix)

Again $(\underline{X} - \underline{B}) = C\underline{Y}$ or $(\underline{X} - \underline{\mu}) = C\underline{Y}$

$\Rightarrow \underline{Y} = C^{-1}(\underline{X} - \underline{\mu})$

So

$$E[C^{-1}(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'(C^{-1})'] = I_p$$

or $\quad C^{-1} E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'(C^{-1})' = I_p$

and $\quad E(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})' = \Sigma$

$$\therefore C^{-1} \Sigma (C^{1})^{-1} = I_p \Rightarrow \Sigma = CC' \rightarrow (1)$$

Also $\quad C'AC = I_p \Rightarrow A = (C'C)^{-1} I_p$

$$\Rightarrow A = C' I_p C^{-1} \quad — (2)$$

using (1) and (2) $\quad \boxed{A = \Sigma^{-1}}$

recall that $K = \frac{|A|^{1/2}}{(2\pi)^{p/2}}$ $\therefore K = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}$

Using the values of K, A, B in f(x) we obtain

$$F(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})}$$

is the P.d.f of random vector $X \sim N_p(\mu, \Sigma)$

the parameters are mean vector : $\mu_{p\times 1}$
Var-Covariance Matrix : $\Sigma_{p\times p}$

Note: If X is multivariate than each component x_i should be univariate.

Bivariate Normal Distribution

Univariate $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

Case 1: when X_1 and X_2 are independent

If X_1 and X_2 are independent

$F(x_1, x_2)$, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

($\because$ x_1 and x_2 are independent there covariance is zero)

Now,

$$f(x_1, x_2) = f(x_1) \cdot f(x_2)$$

$$= \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2} \times \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}$$

$$= \frac{1}{(2\pi)^{2/2}(\Sigma)^{1/2}} e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

$$= \frac{1}{(2\pi)^{2/2}(\Sigma)^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})'\Sigma(\underline{x}-\underline{\mu})}$$

→ Bivariate case

Now $\Sigma = \sigma_1^2 \sigma_2^2$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \underline{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

Now deriving exponential part

$$= \frac{-1}{2}\left[(x_1-\mu_1)\ (x_2-\mu_2)\right] \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}^{-1} \begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}$$

$$= -\frac{1}{2}\left[(x_1-\mu_1)\ (x_2-\mu_2)\right] \frac{1}{\sigma_1^2\sigma_2^2} \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix} \begin{bmatrix} (x_1-\mu_1) \\ (x_2-\mu_2) \end{bmatrix}$$

$$= -\frac{1}{2\sigma_1^2\sigma_2^2}\left[(x_1-\mu_1)\ (x_2-\mu_2)\right] \begin{bmatrix} \sigma_2^2\ (x_1-\mu_1) \\ \sigma_1^2\ (x_2-\mu_2) \end{bmatrix}$$

$$= \frac{-\sigma_1^2\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

(∵ Multiplying and dividing $\sigma_1^2\sigma_2^2$)

$$= -\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

$$F(x_1, x_2) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

Case 2:- At $p=2$ derive the Bivariate Normal Distribution using $N_p(\mu, \Sigma)$

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \underline{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad \rho_{\sigma_1\sigma_2} \ \operatorname{cov}(x_1, x_2)$$

Here x_1 and x_2 are dependent

$$\rightarrow \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}; \ \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

$$f(x_1, x_2) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}\left[(\underline{x}-\underline{\mu})'\Sigma^{-1}(\underline{x}-\underline{\mu})\right]} \quad \text{—(1)}$$

Now deriving exponential term

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

$$\begin{aligned} |\Sigma| &= \sigma_1^2\sigma_2^2 - \sigma_{12}^2 \\ &= \sigma_1^2\sigma_2^2 - (\rho_{12}\sigma_1\sigma_2)^2 \\ &= \sigma_1^2\sigma_2^2(1-\rho_{12}^2) \quad \text{—(2)} \end{aligned} \qquad \left(\because \rho = \frac{\sigma_{12}}{\sigma_1\sigma_2}\right)$$

$$\Sigma^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}\begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

Now, $-\frac{1}{2}\left[(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu})\right]$

Substituting value of Σ^{-1}

$$-\frac{1}{2}\left[(\underline{x}-\underline{\mu})'\right] \frac{1}{\sigma_1^2\sigma_2^2-\sigma_{12}^2}\begin{bmatrix}\sigma_2^2 & -\sigma_{12}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}\left[(\underline{x}-\underline{\mu})\right]$$

$$= \frac{1}{2\sigma_1^2\sigma_2^2-\sigma_{12}^2}\left[(x_1-\mu_1)\,(x_2-\mu_2)\right]\begin{bmatrix}\sigma_2^2 & -\sigma_{12}\\ -\sigma_{12} & \sigma_1^2\end{bmatrix}\begin{bmatrix}(x_1-\mu_1)\\ (x_2-\mu_2)\end{bmatrix}$$

$$= \frac{1}{2\sigma_1^2\sigma_2^2-\sigma_{12}^2}\left[(x_1-\mu_1)\,(x_2-\mu_2)\right]\begin{bmatrix}\sigma_2^2(x_1-\mu_1)-\sigma_{12}(x_2-\mu_2)\\ \sigma_1^2(x_2-\mu_2)-\sigma_{12}(x_1-\mu_1)\end{bmatrix}$$

$$= -\frac{1}{2\sigma_1^2\sigma_2^2-\sigma_{12}^2}\left[(x_1-\mu_1)^2\sigma_2^2-2\sigma_{12}(x_1-\mu_1)(x_2-\mu_2)+\sigma_1^2(x_2-\mu_2)^2\right]$$

$$= -\frac{1}{2}\,\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2\sigma_2^2-\sigma_{12}^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-\frac{2\sigma_{12}}{\sigma_1\sigma_2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]$$

[Multiplying $\sigma_1\sigma_2$ to numerator and denominator]

$$= -\frac{1}{2}\,\frac{\sigma_1^2\sigma_2^2}{\sigma_1^2\sigma_2^2-\rho_{12}\sigma_1^2\sigma_2^2}$$

$$\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

By using eqn(2)

$$= -\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho_{12}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]$$

Thus,

$$F(x_1, x_2) = \frac{1}{(2\pi)^{2/2}|\Sigma|^{1/2}}$$

$$e^{-\frac{1}{2(1-\rho_{12}^2)^2}\left[\left(\frac{x-\mu}{\sigma_1}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

★ Properties of MVND :- (17/8/22)

Scale & ORIGIN.

(1) Given that a random vector $\underline{x} \sim N_p(\mu, \Sigma)$
Consider $\underline{Y} = C\underline{X}$ where $C_{p\times p}$ NSM then
$\underline{Y} \sim N_p(C\mu, C\Sigma C')$

Proof: Since, $\underline{Y} = C\underline{X}$ is a non-singular transformation
$\therefore \underline{x} = C^{-1}Y$ and for such type of transformation $|J| = |C^{-1}|$

$$\therefore f(\underline{y}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(C^{-1}\underline{Y}-\underline{\mu})'\Sigma^{-1}(C^{-1}\underline{Y}-\underline{\mu})} \cdot |C^{-1}|$$

$$\therefore f(\underline{y}) = \frac{1}{(2\pi)^{p/2}|\underline{C}|^{1/2}|\Sigma|^{1/2}|\underline{C'}|^{1/2}}$$

$$\cdot\, e^{-1/2\{C^{-1}(\underline{Y}-C\underline{\mu})\}'\Sigma^{-1}\{\underline{C^{-1}C\underline{Y}} - C\underline{\mu})\}}$$

OR $$f(\underline{y}) = \frac{1}{(2\pi)^{p/2}|C\Sigma C'|^{1/2}} e^{\frac{1}{2}(Y-C\underline{\mu})'(C\Sigma C')^{-1}(Y-C\underline{\mu})}$$

$\therefore$ Comparing $f(\underline{y})$ with standard form of $f(\underline{X})$, we note that
$\underline{Y} \sim N_p(C\underline{\mu}, C\Sigma C')$

(2) Characteristic function :-

If $X \sim N_p(\underline{\mu}, \Sigma)$ then (show that)

$$\phi_X(t) = e^{i\underline{t}'\underline{\mu} - \frac{1}{2}\underline{t}'\Sigma\underline{t}}$$ where $t_{p\times 1}$ real vector

Proof: As Σ assumed to a NSM,
Σ^{-1} exists and hence there exists a NSM

'C' such that $C'\Sigma^{-1}C = I_p$

$\Rightarrow \Sigma^{-1} = (C')^{-1}C^{-1}$ or $\Sigma^{-1} = (CC')^{-1}$

$\Rightarrow \boxed{\Sigma = CC'} \rightarrow$ (Imp for 2marks) Prove that $\Sigma = CC'$

Again consider transformation

$(\underline{X} - \underline{\mu}) = C\underline{Y} \Rightarrow \underline{X} = \underline{\mu} + C\underline{Y}$

Also $E(\underline{X} - \underline{\mu}) = CE(\underline{Y}) = \underline{0}$

$\Rightarrow E(\underline{Y}) = \underline{0}$ as $E(\underline{X} - \underline{\mu}) = \underline{0}$

Now the var-cov matrix of $\underline{Y}$ is

$E(\underline{Y}\underline{Y}') = E[C^{-1}(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'(C^{-1})']$

or

$E(Y'Y) = C^{-1}E[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'](C')^{-1}$

$= C^{-1}\Sigma(C')^{-1}$

$= C^{-1}(CC')(C')^{-1}$

$= I_p$

$\therefore \underline{Y} \sim N_p(\underline{0}, I_p)$

as

$\underline{Y} \sim N_p(\underline{0}, I_p)$ it $\Rightarrow y_i \sim N(0,1)$

By def^n $\phi_{\underline{x}}(\underline{t}) = E[e^{it'x}]$

$$= E\left[e^{i\underline{t}'(\underline{\mu}+C\underline{Y})}\right]$$

$$= E\, e^{i\underline{t}'\underline{\mu}} \cdot e^{i\underline{t}'C\underline{y}}$$

<u>OR</u>

$$\phi_x(t) = e^{i\underline{t}'\mu} \cdot E\, e^{i\underline{t}'C\underline{y}}$$

OR

$$\phi_x(\underline{t}) = e^{i\underline{t}'\underline{\mu}} \cdot E\, e^{i\sum_{j=1}^{p} u_j y_j}$$

real variable

$$= e^{i\underline{t}'\underline{\mu}} \prod_{j=1}^{p} e^{iu_j y_j}$$

real constant

$$= e^{i\underline{t}'\mu} \prod_{j=1}^{p} E\, e^{iu_j y_j}$$

$$= e^{i\underline{t}'\mu} \prod_{j=1}^{p} e^{-\frac{1}{2}u_j^2}$$

$$= e^{i\underline{t}\mu}\, e^{-\frac{1}{2}\sum_j u_j^2}$$

$$= e^{i\underline{t}'\underline{\mu}} \cdot e^{-\frac{1}{2}\underline{u}\,\underline{u}'}$$

$$= e^{i\underline{t}\underline{\mu}} \cdot e^{-\frac{1}{2}\underline{t}'CC'\underline{t}}$$

$$= e^{i\underline{t}\mu - \frac{1}{2}\underline{t}'\Sigma\underline{t}}$$

$$= e^{-\frac{1}{2}t't} \longrightarrow N(0, I_p)$$

(Now $t'\underset{\sim}{c} = \underline{u}$ as $\underline{t}'_{1\times p} = C_{p\times p} \Rightarrow \underline{t}'c = u$ (say))

$\boxed{\text{INDEPENDENT}}$

(3) If $x_1, x_2, \ldots x_p$ have joint multivariate Normal Dist^n than Necessary and sufficient condition that one subset $X^{(1)}$ is independent of other subset $\underline{X}^{(2)}$ containing the remaining variables is that each co-variance of a variable from $\underline{X}^{(1)}$ and a variable from $\underline{X}^{(2)}$ be zero

Mathematically, let

$$\underline{X}_{p\times 1} = \left[\begin{array}{c} \underline{X}^{(1)} \\ \hline \underline{X}^{(2)} \end{array}\right] \begin{array}{l} \rightarrow q \\ \rightarrow (p-q) \end{array}$$

accordingly $\underline{\mu}_{p\times 1} = \left[\begin{array}{c} \underline{\mu}^{(1)} \\ \hline \underline{\mu}^{(2)} \end{array}\right] \begin{array}{l} \rightarrow q \\ \rightarrow (p-q) \end{array}$

$$\Sigma_{q\times q} = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$$

where $\Sigma_{11} \Rightarrow q\times q = E\left[(\underline{X}^{(1)} - \underline{\mu}^{(1)})(\underline{X}^{(1)} - \underline{\mu}^{(1)})'\right]$

$$\Sigma_{12\, q\times p-q} = E\left[(\underline{X}^{(1)} - \underline{\mu}^{(1)})(\underline{X}^{(2)} - \underline{\mu}^{(2)})'\right]$$

$$\Sigma_{22\,(p-q)(p-q)} = E\left[(\underline{X}^{(2)} - \underline{\mu}^{(2)})(\underline{X}^{(2)} - \underline{\mu}^{(2)})'\right]$$

Proof <u>Necessary Part</u> :-

Let $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ be <u>independent</u> and let X_i be a variable from the ~~ith~~ 1st subset $\underline{X}^{(1)}$ and X_j be any variable from $X^{(2)}$, then we have to show $\text{cov}(X_i, X_j) = 0$

By defn

$$\text{cov}(X_i, X_j) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} (x_i - \mu_i)(x_j - \mu_j)\, f(\underline{x})\, dx_1\, dx_2 \cdots dx_p$$

Since $X^{(1)}$ and $X^{(2)}$ are independent. So the <u>joint</u> density of $\underline{X}^{(1)}$ and $X^{(2)}$ will be the product of their marginals i.e $f(x) = f_1(x^{(1)}) \cdot f_2(x^{(2)})$ (say)

$$\therefore \operatorname{Cov}(x_i, x_j) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(x_i - \mu_i)\, f_1(\underline{x}^{(1)})\, d\underline{x}^{(1)}$$

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(x_j - \mu_j)\, f_2(\underline{x}^{(2)})\, d\underline{x}^{(2)}$$

Now

$\Rightarrow E(x_i - \mu_i) \cdot E(x_j - \mu_j) = 0$

i.e $\operatorname{cov}(x_i, x_j) = 0$ for $i = 1, 2, \ldots, q$
$j = q+1, \ldots, p$

this means all the elements of Σ_{12} are zero i.e $\Sigma_{12} = 0$

Sufficient Part

let $\Sigma_{12} = 0$ then we have to show that $\underline{x}^{(1)}$ and $\underline{x}^{(2)}$ are independent, Here

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right] = \begin{bmatrix} \Sigma_{11} & \underline{0} \\ \underline{0} & \Sigma_{22} \end{bmatrix}$$

$$\Rightarrow \Sigma^{-1} = \begin{bmatrix} \Sigma_{11}^{-1} & \underline{0} \\ \underline{0} & \Sigma_{22}^{-1} \end{bmatrix}$$

we have $f(\underline{x}) = \dfrac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x} - \underline{\mu})' \Sigma^{-1} (\underline{x} - \underline{\mu})}$

after partition consider the exponential terms:-

$$(\underline{x} - \underline{\mu})' \Sigma (\underline{x} - \underline{\mu})$$

$$\begin{bmatrix} \underline{x}^{(1)} - \underline{\mu}^{(1)} \\ \underline{x}^{(2)} - \underline{\mu}^{(2)} \end{bmatrix}' \begin{bmatrix} \Sigma_{11}^{-1} & \underline{0} \\ \underline{0} & \Sigma_{22}^{-1} \end{bmatrix} \begin{bmatrix} \underline{x}^{(1)} - \underline{\mu}^{(1)} \\ \underline{x}^{(2)} - \underline{\mu}^{(2)} \end{bmatrix}$$

$$(\underline{x}^{(1)} - \underline{\mu}^{(1)})' \Sigma_{11}^{-1} (\underline{x}^{(1)} - \underline{\mu}^{(1)}) + (\underline{x}^{(2)} - \underline{\mu}^{(2)})' \Sigma_{22}^{-1} (\underline{x}^{(2)} - \underline{\mu}^{(2)})$$

$$\therefore f_1(\underline{x}^{(1)}) = \frac{1}{(2\pi)^{q/2} |\Sigma_{11}|^{1/2}} e^{-\frac{1}{2}(\underline{x}^{(1)} - \underline{\mu}^{(1)})' \Sigma_{11}^{-1} (\underline{x}^{(1)} - \underline{\mu}^{(1)})}$$

$$f_2(\underline{x}^{(2)}) = \frac{1}{(2\pi)^{\frac{p-q}{2}} |\Sigma_{22}|^{1/2}} e^{-\frac{1}{2}(\underline{x}^{(2)} - \underline{\mu}^{(2)})' \Sigma_{22}^{-1} (\underline{x}^{(2)} - \underline{\mu}^{(2)})}$$

$$= N_q(\underline{\mu}^{(1)}, \Sigma_{11}) \times N_{p-q}(\underline{\mu}^{(2)}, \Sigma_{22})$$

$$\therefore f(\underline{x}) = f_1(\underline{x}^{(1)}) f_2(\underline{x}^{(2)})$$

$\Rightarrow$ $\underline{x}^{(1)}$ and $\underline{x}^{(2)}$ are independent

(4) Marginal Distribution

Let $\underline{X}_{p\times 1} \sim N_p(\mu, \Sigma)$ and $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ are two Subsets of $\underline{X}_{p\times 1}$ where $\underline{X}^{(1)}_{q\times 1}$ and $\underline{X}^{(2)}_{(p-q)\times 1}$ then the Marginal distn of any subset of $\underline{X}_{p\times 1}$ is also MVND with mean vector and Var-cov matrix obtained by taking proper components of μ and Σ.

Proof We have $\underline{X}_{p\times 1} = \begin{bmatrix} \underline{X}^{(1)} \\ \hline \underline{X}^{(2)} \end{bmatrix}$ acc. $\mu = \begin{bmatrix} \underline{\mu}^{(1)} \\ \hline \mu^{(2)} \end{bmatrix}$

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$$

consider the transformation
$\underline{Y}^{(1)} = \underline{X}^{(1)} + M\underline{X}^{(2)}$
& $\underline{Y}^{(2)} = \underline{X}^{(2)}$

i.e $\underline{Y} = \begin{bmatrix} I & M \\ 0 & I \end{bmatrix} \underline{X}$ where M is a NSM and is choosen such that

$\underline{Y}^{(1)}$ and $\underline{Y}^{(2)}$ are independent

$\therefore$ $cov[\underline{Y}^{(1)}, \underline{Y}^{(2)}] = 0$ determine 'M'

Now $cov[\underline{Y}^{(1)}, Y^{(2)}] = E[\underline{Y}^{(1)} - E(\underline{Y}^{(1)})][\underline{Y}^{(2)} - E(\underline{Y}^{(2)})]'$

Also, $E(\underline{Y}) = \begin{bmatrix} I & M \\ 0 & I \end{bmatrix} E(x)$

i.e $E\begin{bmatrix} \underline{Y}^{(1)} \\ \hline Y^{(2)} \end{bmatrix} = \begin{bmatrix} I & M \\ 0 & I \end{bmatrix} \begin{bmatrix} \mu^{(1)} \\ \hline \mu^{(2)} \end{bmatrix}$

$= \begin{bmatrix} \underline{\mu}^{(1)} + M\underline{\mu}^{(2)} \\ \hline \mu_{(2)} \end{bmatrix}$

So that

$cov[\underline{Y}^{(1)}, \underline{Y}^{(2)}] = E[\underline{X}^{(1)} + M\underline{X}^{(2)} - \underline{\mu}^{(1)} - M\underline{\mu}^{(2)}][\underline{X}^{(2)} - \underline{\mu}^{(2)}]'$

$= E[(\underline{X}^{(1)} - \underline{\mu}^{(1)})(\underline{X}^{(2)} - \underline{\mu}^{(2)})'] + ME[\underline{X}^{(2)} - \underline{\mu}^{(2)}][\underline{X}^{(2)} - \underline{\mu}^{(2)}]'$

$= \Sigma_{12} + M\Sigma_{22}$

Now, $cov(\underline{Y}^{(1)}, \underline{Y}^{(2)}) = 0$

$\Rightarrow \Sigma_{12} + M\Sigma_{22} = 0 \Rightarrow M = -\Sigma_{12}\Sigma_{22}^{-1}$

$\underline{Y}^{(1)} = \underline{X}^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\underline{X}^{(2)}$ OR $\underline{Y} = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix} \underline{X}$

$\underline{Y}^{(2)} = \underline{X}^{(2)}$

OR $\underline{Y} = C\underline{X}$ (say) 'C' ~~say~~ being a NSM

$\therefore$ Using property (1) if $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$

$\underline{Y} \sim N_p(C\underline{\mu}, C\Sigma C')$ we obtain

$$E(\underline{Y}) = \begin{bmatrix} \underline{\mu}^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\underline{\mu}^{(2)} \\ \hline \underline{\mu}^{(2)} \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \hline \gamma_2 \end{bmatrix} = \underline{\gamma} \text{ (say)}$$

and

$$C\Sigma C' = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} I^* & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix}'$$

$$= \left[\begin{array}{c|c} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ \hline 0 & \Sigma_{22} \end{array}\right]$$

$$= \begin{bmatrix} \Sigma_{11.2} & 0 \\ 0 & \Sigma_{22} \end{bmatrix} = \Omega \rightarrow \text{Partial Var-Covariance Matrix}$$

$$\therefore f(\underline{y}) = \frac{1}{(2\pi)^{p/2}|\Omega|^{1/2}} e^{-\frac{1}{2}(\underline{Y}-\underline{\gamma})'\Omega^{-1}(\underline{Y}-\underline{\gamma})}$$

$$= \frac{1}{(2\pi)^{q/2}|\Sigma_{11.2}|^{1/2}} e^{-\frac{1}{2}(\underline{Y}^{(1)}-\gamma^{(1)})'\Sigma_{11.2}^{-1}(\underline{Y}^{(1)}-\gamma^{(1)})}$$

$$\times \frac{1}{(2\pi)^{\frac{p-q}{2}}|\Sigma_{22}|^{1/2}} e^{-\frac{1}{2}(\underline{Y}^{(2)}-\gamma^{(2)})'\Sigma_{22}^{-1}(\underline{Y}^{(2)}-\gamma^{(2)})}$$

$= f(\underline{Y}^{(1)}) \times f(\underline{Y}^{(2)})$ as $Y^{(1)}$ and $Y^{(2)}$ are independent — (*)

As per the transformation $\underline{Y}^{(2)} = \underline{X}^{(2)}$
∴ The marginal density of $\underline{Y}^{(2)}$ will be obtained by integrating (*) over $\underline{Y}^{(1)}$
also it is to be noted that $f(\underline{Y}^{(1)})$ is a q component MVND.

So the marginal density of $\underline{Y}^{(2)}$ is given by

$$g(\underline{y})^{(2)} = \frac{1}{(2\pi)^{\frac{p-q}{2}} |\Sigma_{22}|^{1/2}} e^{-\frac{1}{2}(\underline{y}^{(2)}-\gamma^{(2)})' \Sigma_{22}^{-1} (\underline{y}^{(2)}-\gamma^{(2)})}$$

as $\underline{Y}^{(2)} = \underline{X}^{(2)}$

∴ the marginal density of $\underline{X}^{(2)}$ will be given by

$$F(\underline{x}^{(2)}) = \frac{1}{(2\pi)^{\frac{p-q}{2}} |\Sigma_{22}|^{1/2}} e^{-\frac{1}{2}(\underline{x}^{(2)}-\underline{\mu}^{(2)})' \Sigma_{22}^{-1} (\underline{x}^{(2)}-\underline{\mu}^{(2)})} \cdot |J|$$

(for transformation consider here $|J|=1$)

i.e $\underline{X}^{(2)} \sim N_{p-q}(\underline{\mu}^{(2)}, \Sigma_{22})$

(5) Conditional Distribution

Proof Let $X \sim N_p(\mu, \Sigma)$ and $\underline{X} = \begin{bmatrix} X^{(1)} \\ \hline X^{(2)} \end{bmatrix}$ accordingly

$$\underline{\mu} = \begin{bmatrix} \mu_{(1)} \\ \hline \mu_{(2)} \end{bmatrix}, \quad \Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$$

We wish to obtain the conditional density of $X^{(1)}$ given $X^{(2)} = x^{(2)} \to$ specified vector

Consider the transformation : $\underline{Y}^{(1)} = \underline{X}^{(1)} + M\underline{X}^{(2)}$
$\underline{Y}^{(2)} = \underline{X}^{(2)}$

where M is choosen such that $Y^{(1)}$ and $Y^{(2)}$ are independent (we know that)

$$M = -\Sigma_{12}\Sigma_{22}^{-1}. \quad \text{Also,}$$

$$\underline{Y} = \left[\begin{array}{c} \underline{Y}^{(1)} \\ \hline \underline{Y}^{(2)} \end{array}\right] \sim N_p\left[\left[\begin{array}{c} \gamma^{(1)} \\ \hline \gamma^{(2)} \end{array}\right], \left(\begin{array}{c|c} \Sigma_{11\cdot 2} & 0 \\ 0 & \Sigma_{22} \end{array}\right)\right]$$

$$\Rightarrow \underline{Y}^{(1)} \sim N_q\left[\underline{\mu}^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}, \Sigma_{11\cdot 2}\right]$$

and

$$\Rightarrow \underline{Y}^{(2)} \sim N_{p-q}\left[\underline{\mu}^{(2)}, \Sigma_{22}\right]$$

Now the conditional density of $\underline{X}^{(1)} / \underline{X}^{(2)} = \underline{x}^{(2)}$ is (say)

$$h\left[X^{(1)} \,\middle|\, \underset{\substack{\uparrow \\ \text{R.V}}}{X^{(2)}} = \underset{\substack{\uparrow \\ \text{value of R.V}}}{x^{(2)}}\right] = \frac{\phi\left(\underline{X}^{(1)}, \underline{X}^{(2)}\right)}{g\left(\underline{X}^{(2)} = x^{(2)}\right)}$$

We find the joint density $\phi(\underline{X}^{(1)}, \underline{X}^{(2)})$ from it by $|J|$ here

$$|J| = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{bmatrix} = 1$$

$$\therefore \phi(\underline{X}^{(1)}, \underline{X}^{(2)}) = \frac{1}{(2\pi)^{q/2}|\Sigma_{11\cdot 2}|}$$

$$e^{-\frac{1}{2}\left[\underline{X}^{(1)} - \Sigma_{12}\Sigma_{22}^{-1}\underline{X}^{(2)} - \underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}\underline{\mu}^{(2)}\right]' \Sigma_{11\cdot 2}^{-1}\left[\underline{X}^{(1)} - \Sigma_{11}\Sigma_{22}^{-1}\underline{X}^{(2)} - \underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}\underline{\mu}^{(2)}\right]}$$

$$\times \frac{1}{(2\pi)^{\frac{p-q}{2}}|\Sigma_{22}|^{1/2}} e^{-\frac{1}{2}(\underline{X}^{(2)} - \underline{\mu}^{(2)})'\Sigma_{22}^{-1}(\underline{X}^{(2)} - \underline{\mu}^{(2)})}$$

$\longrightarrow$ * *

$\therefore h = [\underline{X}^{(1)} \mid \underline{X}^{(2)} = x^{(2)}]$ = The first part of $(* *)$

or $h[\underline{X}^{(1)} \mid \underline{X}^{(2)}] \sim N_q[\underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}^{(2)} - \underline{\mu}^{(2)})', \Sigma_{11\cdot 2}]$

where $\Sigma_{11\cdot 2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

So the conditional mean will be
$E[\underline{X}^{(1)} \mid \underline{X}^{(2)} = x^{(2)}] = \underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}^{(2)} - \underline{\mu}^{(2)})$
is also known as Regression f^n.

The matrix $\Sigma_{12}\Sigma_{22}^{-1}$ is called matrix of Regr. Coeff and is denoted by $\Gamma\beta$

$$\Sigma_{12}\Sigma_{22}^{-1} = \Gamma\beta$$

$$\underline{X}^{(1)} = \underline{\mu}^{(1)} + \Sigma_{12}\Sigma_{22}^{-1}(\underline{X}^{(2)} - \underline{\mu}^{(2)})$$

$$(\underline{X})^{(1)} = \bar{X} + b_{xy}(Y - \bar{Y})$$

Note :- Regression is nothing but it is conditional mean when we find conditional dist^n → To get conditional mean.

SINGULAR DISTRIBUTION

Given that $\underline{X} \sim N_p(\mu, \Sigma)$ where $\Sigma_{p \times p} \rightarrow$ NSM. then the random vector X is said to have a Non-Singular or Proper distribution, on the other hand if the matrix Σ is a singular matrix i-e rank of Σ is 'r' (say) where $r < p$ then the random vector x is said to have a singular / degenerated distⁿ of rank 'r'.

→ In this case Σ^{-1} does not exists but it is possible to find 'g-inverse' of Σ and if $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the k eigen values of Σ in such a case the density fⁿ of $\underline{x}$ is given by

$$f(\underline{x}) = \frac{1}{(2\pi)^{k/2}(\lambda_1 \ldots \lambda_k)} e^{-\frac{1}{2}[(\underline{x}-\underline{\mu})'(\lambda_1, \lambda_2, \ldots, \lambda_k)^{-1}(\underline{x}-\underline{\mu})]}$$

Thm 1 Let $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ consider the transformation $Z = D\underline{X}$ where 'D' is a $t \times p$ matrix $(t \leq p)$ of rank 't' then $Z \sim N_t(D\underline{\mu}, D\Sigma D')$
(Marginal density)

Proof If $t < p$ we can find a matrix 'H' s.t $C_{p\times p} = \begin{bmatrix} D \\ H \end{bmatrix}$ is a NSM consider the transformation

$$Z^* = C\underline{X} = \begin{bmatrix} D \\ H \end{bmatrix} \underline{X} = \begin{bmatrix} D\underline{X} \\ H X \end{bmatrix} = \begin{bmatrix} Z \\ W \end{bmatrix} \text{ (say)}$$

Since Z^* is a Non-Singular transformation

$$\therefore Z^* \sim N_p(C\underline{\mu}, C\Sigma C') \text{ i.e}$$

$$Z^* = \begin{pmatrix} Z \\ W \end{pmatrix} \sim N_p\left[\begin{pmatrix} D\underline{\mu} \\ H\underline{\mu} \end{pmatrix}, \begin{pmatrix} D\Sigma D' & D\Sigma H' \\ H\Sigma D' & H\Sigma H' \end{pmatrix}\right]$$

$\therefore$ the marginal density of Z will be $N_t(D\underline{\mu}, D\Sigma D')$

Thm A random vector $\underline{X}_{p\times 1}$ with $E(\underline{X}) = \underline{\mu}$ and $var(\underline{X}) = \Sigma$ is said to have a MVND with mean vector $\underline{\mu}$ and Variance-Covariance Matrix Σ iff there is a linear transformation $\underline{X} = A\underline{Y} + \underline{\lambda}$ —(1), where $A_{p\times r}$ matrix $\underline{\lambda}_{p\times 1}$ Σ is of rank 'r' and

$$\underline{Y}_{r\times 1} \sim N_r(\underline{\gamma}, T)$$

NECESSARY PART

let (1) holds i.e $X = A\underline{Y} + \underline{\lambda}$ | as $\underline{Y} \sim N_r(\gamma, T)$

$\Rightarrow E(\underline{X}) = A\,E(\underline{Y}) + \underline{\lambda} = A\underline{\gamma} + \underline{\lambda}$

and

$E[X - E(\underline{X})][\underline{X} - E(\underline{X})'] = [A(\underline{Y} - \underline{\gamma})][(\underline{Y} - \gamma)'A']$

$A\,E(\underline{Y} - \gamma)(\underline{Y} - \underline{\gamma})'A' = ATA' = \Sigma$ (say). If $p > r$ then Σ is a singular matrix. Since, $\underline{X}$ is a linear combination of $\underline{Y}$, so $\underline{X}$ also has a MVND, now as the rank of Σ is 'r' which is less than 'p' so the distⁿ of $\underline{X}$ will be singular as Σ^{-1} does not exist, so we can not write the expression of p.d.f of $\underline{X}$ (Density free) approach.

SUFFICIANCY PART

let X has a singular MVND of rank 'r' then three exists a NSM 'B' s.t $B\Sigma B' = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$

now consider the transformation $B\underline{X} = \underline{V}$ (say) and consider the partioning $\underline{V} = \begin{bmatrix} V^{(1)} \\ V^{(2)} \end{bmatrix}$

So $\underline{V} = B\underline{X}$ defines it having the mean vector $E(\underline{V}) = BE(\underline{X}) = B\underline{\mu}$

Or,

$$E(\underline{V}) = \begin{pmatrix} \gamma^{(1)} \\ \hline \gamma^{(2)} \end{pmatrix} = \gamma \text{ (say) and } B\Sigma B' = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

$\Rightarrow$ VCM of $V^{(1)} = I_r$ and VCM of $V^{(2)} = 0$

But VCM $(V^{(2)}) = 0 \Rightarrow \underline{V}^{(2)} = \gamma^{(2)}$

$$\therefore BX = \begin{bmatrix} \underline{V}^{(1)} \\ \hline \underline{V}^{(2)} \end{bmatrix} = \begin{bmatrix} \underline{V}^{(1)} \\ \hline \gamma^{(2)} \end{bmatrix} \Rightarrow \underline{X} = \underset{[CD]}{B^{-1}} \begin{bmatrix} \underline{V}^{(1)} \\ \hline \gamma^{(2)} \end{bmatrix}$$

Now partition B^{-1} as $[CD]$

So that $\underline{X} = C\underline{V}^{(1)} + D\gamma^{(2)} = C\underline{V}^{(1)} + \lambda$ (say)

Or $\underline{X} = A\underline{Y} + \lambda$ form

$\therefore$ If $\underline{X}$ has a singular MVND it can be expressed as above.

07/09/2022

Result :- A p-component vector $\underline{U}$ is said to have a MVND iff every linear combination of $\underline{U}$ has a univariate normal distⁿ

NECCESSARY PART

Consider an arbitary vector $\underline{L}_{p\times1}$ and let $\underline{U}_{p\times1}$ has the MVND then since $\underline{L}$ is assumed to be NON-NULL, $\underline{L}'$ will be of rank 1 so that there exists a matrix B such that $B = \begin{pmatrix} L' \\ D \end{pmatrix}$ and B' is a NSM consider the ~~transform~~ transformation: $Z = BU$ or $Z = \begin{pmatrix} L'U \\ \hline DU \end{pmatrix}$

Also,

$$Z \sim N\left[\begin{pmatrix} L'U \\ \hline D\mu \end{pmatrix}, \left(\begin{array}{c|c} L'\Sigma L & L'\Sigma D' \\ \hline D\Sigma L & D\Sigma D' \end{array}\right)\right]$$

$\Rightarrow \underline{l}'\underline{U}_{(1\times1)} \sim N_1(\underline{l}'\underline{\mu}_{(1\times1)}, \underline{l}'\Sigma\underline{l}_{(1\times1)})$

i.e every linear combination of $\underline{U}$ is a univariate normal.

SUFFICIENT PART

let $\underline{l}_{p\times1}$ be any arbitary vector and $\underline{l}'\underline{U}$ has a univariate normal distⁿ with mean $\underline{l}'\underline{\mu}$ and variance $\underline{l}'\Sigma\underline{l}$ then the characterstic fn of $\underline{l}'\underline{U}$ will be

$$\phi_t(\underline{l}'\underline{U}) = E\left[e^{it\underline{l}'\underline{U}}\right] = e^{it\underline{l}'\underline{\mu} - \frac{1}{2}t^2\underline{l}'\Sigma\underline{l}}$$

now iff $t=1$ then

$$\phi_1(\underline{l}'\underline{U}) = e^{i\underline{l}'\underline{\mu} - \frac{1}{2}\underline{l}'\Sigma\underline{l}} \quad \text{(By defn)}$$

i.e $\phi_1[\underline{l}'\underline{U}] = E[e^{i\underline{l}'\underline{U}}]$ which is the characteristic fun

since $\underline{l}$ is any arbitrary vector so

$$\phi(\underline{l}'\underline{U}) = E\left[e^{i\underline{l}'\underline{U}}\right]$$ which is the ch fn of a MVND

$\therefore \underline{U}$ has a MVND using the uniqueness property of ch fn.

R

REPRODUCTIVE PROPERTY

Statement :- Let the i^{th} random vector $\underline{U}_i \sim N_p(\underline{\mu}_i, \Sigma_i)$ $i = 1, 2, \ldots n$ then for fixed constants $A_1, A_2, \ldots A_n$ Consider $Y = \sum_{i=1}^{n} A_i \underline{U}_i$ then

$$\underline{Y} \sim N_p\left[\sum_i A_i \underline{U}_i ,\ \sum_{i=1}^{n} A_i^2 \Sigma_i\right]$$

Proof We have $Y = A_1\underline{U}_1 + \ldots + A_n \underline{U}_n$. Consider an arbitrary linear combination $\underline{L}'\underline{Y} = A_1\underline{L}'\underline{U}_1 + \ldots + A_n\underline{L}'\underline{U}_n$ Since $\underline{U}_i \sim N_p(\underline{\mu}_i, \Sigma_i)$ $\underline{L}'\underline{U}_i$ is a univariate Normal. Again

$$\underline{L}'\underline{Y} = \sum_{i=1}^{n} \underline{L}' A_i \underline{U}_i \quad \text{or}$$

$$\underline{L}'\underline{Y} = \sum_{i=1}^{n} A_i Z_i \quad \text{where} \quad Z_i = \underline{L}'\underline{U}_i$$

Since it is a linear combination of univariate normal variable, so that $\underline{L}'\underline{Y}$ also has a univariate normal with MEAN $= E[\underline{L}'\underline{Y}]$

OR

$$\begin{aligned}
\text{MEAN} &= E[A_1\underline{L}'\underline{U}_1 + \ldots + A_n\underline{L}'\underline{U}_n] \\
&= \underline{L}'[A_1\underline{\mu}_1 + \ldots + A_n\underline{\mu}_n] \\
&= \underline{L}' \sum_{i=1}^{n} A_i \underline{\mu}_i
\end{aligned}$$

And

$$\begin{aligned}
\text{Variance} &= var[A_1\underline{L}'\underline{U}_1 + \ldots + A_n\underline{L}'\underline{U}_n] \\
&= A_1^2 var(\underline{L}'\underline{U}_1) + \ldots + A_n^2 var(\underline{L}'\underline{U}_n) \\
&= A_1^2 \underline{L}'\Sigma_1\underline{L} + \ldots + A_n^2 \underline{L}'\Sigma_n \underline{L} \\
&= \underline{L}' \sum_{i=1}^{n} A_i^2 \Sigma_i \underline{L} \quad \text{i.e}
\end{aligned}$$

$$\text{i.e} \quad \underline{L}'\underline{Y} \sim N_1\left(\underline{L} \sum_{i=1}^{n} A_i \underline{\mu}_i ,\ \sum_{i=1}^{n} A_i^2 \Sigma_i\right)$$

Since $\underline{L}$ is any arbitrary vector and $\underline{L}'\underline{Y}$ vector and $L'Y$ ~~combination~~ and $\underline{L}'\underline{Y}$ is a linear combination of $\underline{Y}$ which has a MVND

$$\therefore \cancel{A} Y \sim N_p\left(\sum_{i=1}^{n} A_i \mu_i,\ \sum_{i=1}^{n} A_i^2 \Sigma_i\right)$$

* MLE's of the parameters of $N_p(\underline{\mu}, \Sigma)$:

Rules of differentiation:-

(1) $\frac{\partial}{\partial \underline{X}} |X| = |X| X^{-1}$

(2) $\frac{\partial}{\partial \underline{X}} tr \cdot A\underline{X} = A$ [tr - trace]

(3) $\frac{\partial}{\partial \underline{X}} \underline{x}' A X = 2\underline{X}' A = 2A'\underline{X}$

$\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ where $\underline{\mu}$ is mean vector and Σ is var. cov Matrix of the Random Vector X

If $\underline{\mu}$ and Σ are known then the distribution of x is completely specified

However if they are not known then for specifying the distribution of x ~~are~~ we have to estimate $\underline{\mu}$ and Σ with the help of a random sample of vectors from $N_p(\underline{\mu}, \Sigma)$

In the estimation procedure we will assume that Σ is a non-singular matrix

Derivation :-

let $x_1, \ldots x_N$ be a r.s of N random vectors from $N_p(\underline{\mu}, \Sigma)$

Define :- the sample mean vector

$$\bar{X} = \frac{1}{N}\sum_{\alpha=1}^{N} X_\alpha = \begin{pmatrix} \bar{X}_1 \\ \bar{X}_2 \\ \vdots \\ \bar{X}_p \end{pmatrix} = \frac{1}{N}\begin{pmatrix} \sum_\alpha X_{1\alpha} \\ \sum_\alpha X_{2\alpha} \\ \vdots \\ \sum_\alpha X_{p\alpha} \end{pmatrix}$$

and let the sample VCM be,

$$S = \frac{1}{N}\sum_{\alpha=1}^{N}(\bar{X}_\alpha - \bar{X})(\bar{X}_\alpha - \bar{X})' \rightarrow (p\times p) \text{ matrix}$$

The likelihood function (LF) is given by

$$L = \prod_{i=1}^{N} f(\underline{x}_i, \underline{\mu}, \Sigma) = \frac{e^{-\frac{1}{2}\sum_{\alpha=1}^{N}(\underline{x}_\alpha - \underline{\mu})'\Sigma^{-1}(\underline{x}_\alpha - \underline{\mu})}}{(2\pi)^{\frac{Np}{2}} \, |\Sigma|^{N/2}}$$

Since, Σ is NSM, so Σ^{-1} will exist

$\therefore$ we find the MLE of Σ^{-1} and then using the invariance property of MLE we obtain the MLE of Σ.

Now

$$\text{Log } L = k - \frac{N}{2} \log |\Sigma| - \frac{1}{2}\sum_{\alpha=1}^{N}(\underline{X}_\alpha - \underline{\mu})'\Sigma^{-1}(\underline{X}_\alpha - \underline{\mu})$$

$$\log L = k + \frac{N}{2}\log|\Sigma^{-1}| - \frac{1}{2}\sum_{\alpha=1}^{N}(\underline{X}_\alpha - \underline{\mu})'\Sigma^{-1}(\underline{X}_\alpha - \underline{\mu})$$

Consider the exponent

$$\sum_{\alpha=1}^{N}(\underline{X}_\alpha - \underline{\mu})'\Sigma^{-1}(\underline{X}_\alpha - \underline{\mu})$$

$$\sum_{\alpha=1}^{N}(\underline{X}_\alpha - \bar{\underline{X}} + \bar{\underline{X}} - \underline{\mu})'\Sigma^{-1}(\underline{X}_\alpha - \bar{\underline{X}} + \bar{\underline{X}} - \underline{\mu})$$

$$= \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) + \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

$$+ \sum_{\alpha=1}^{N} (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) + \sum_{\alpha=1}^{N} (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

$$= \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) + \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu})' \Sigma^{-1} (\underline{x} - \underline{\mu})$$

$$= \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) + N (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

Again,

$$\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu})' \Sigma^{-1} (\underline{x}_\alpha - \underline{\mu})$$

$$= tr \left[\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) \right] + N (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

$$= \sum_{\alpha=1}^{\alpha} tr\, \Sigma^{-1} (\underline{x}_\alpha - \bar{\underline{x}}) (\underline{x}_\alpha - \bar{\underline{x}})' + N (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

let $A = \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}}) (\underline{x}_\alpha - \bar{\underline{x}})'$ then,

$$\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu})' \Sigma^{-1} (\underline{x}_\alpha - \underline{\mu})$$

$$= tr (A \Sigma^{-1}) + N (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

$$\log L = \cancel{k}\ \cancel{k}\ k + \frac{N}{2} \log |\Sigma^{-1}| - \frac{1}{2} \left[tr (\Sigma^{-1} A) + N (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu}) \right]$$

$$= k + \frac{N}{2} \log |\Sigma^{-1}| - \frac{1}{2} tr \cdot \Sigma^{-1} A - \frac{N}{2} (\bar{\underline{x}} - \underline{\mu})' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu})$$

— $(*)$

Differentiating (*) w.r.t $\underline{\mu}$ and Σ^{-1} respectively are obtain:-

$$\frac{\partial \log L}{\partial \underline{\mu}} = \frac{N}{\cancel{\Sigma}}\,\cancel{\Sigma}\,\Sigma^{-1}(\bar{x}-\underline{\mu}) = 0 \qquad \text{---(1)}$$

$$\Rightarrow (\bar{x}-\hat{\underline{\mu}}) = \underline{0} \Rightarrow \boxed{\hat{\underline{\mu}} = \bar{x}}$$

Also, $$\frac{\partial \log L}{\partial \Sigma^{-1}} = \frac{N}{2}\,\frac{\cancel{|\Sigma^{-1}|}\,|\Sigma^{-1}|^{-1}}{\cancel{|\Sigma^{-1}|}} - \frac{A}{2} - \frac{N}{2}(\bar{x}-\underline{\mu})(\bar{x}-\underline{\mu})' = 0$$

$$\Rightarrow \Sigma^{-1} = \left(\frac{A}{N}\right)^{-1} = NA^{-1}$$

$$\therefore \hat{\Sigma} = \frac{1}{N}\sum_{\alpha=1}^{N}(\underline{x}_\alpha-\bar{x})(x_\alpha-\bar{x})'$$

<u>Note</u> It has been shown that,

$$\sum_{\alpha=1}^{N}(\underline{x}_\alpha-\underline{\mu})(\underline{x}_\alpha-\underline{\mu})' = A \,\&\, N(\bar{x}-\mu)(\bar{x}-\underline{\mu})'$$

and $$\sum_{\alpha=1}^{N}(\underline{x}_\alpha-\underline{\mu})'\Sigma^{-1}(\underline{x}_\alpha-\underline{\mu})$$

$$= tr(\Sigma^{-1}A) + N(\bar{x}-\underline{\mu})'\Sigma^{-1}(\bar{x}-\underline{\mu})$$

Thus the j p.d.f of $x_1, \ldots x_N$ can be written as:

$$K\, e^{-1/2\, N(\bar{x}-\underline{\mu})'\Sigma^{-1}(\bar{x}-\underline{\mu})} \cdot \frac{1}{2} e^{-\frac{1}{2} tr\, \Sigma^{-1}A}$$

Thus $\bar{x}$ and A form a set of jointly sufficient statistic for μ and Σ. If Σ is known then $\bar{x}$ is sufficient for μ however if μ is known then A/N is Not sufficient for Σ but $\frac{1}{N}\sum_{\alpha=1}^{N}(\underline{x}_\alpha-\underline{\mu})(\underline{x}_\alpha-\underline{\mu})'$ is sufficient for Σ

* Elliptical Distribution

Elliptical distⁿs provide a generalized for a multivariate normal distⁿ. In the 2 and 3 dimention cases the joint distⁿ function of the variables for Ellips or Ellipsoids.

It is to be noted that the multivariate normal distⁿ is constant on ellipsoids i.e

$(\underline{x}-\underline{\mu})' \Sigma^{-1} (\underline{x}-\underline{\mu}) = C$ — (1)

For every positive value of 'c' in p-dimensional space, the center of each ellipsoid is at the point μ. The shape and orientation is determined by Σ.

because (1) is sphere if $\Sigma = \sigma^2 I$ then $N_p(\mu, \sigma^2 I)$ is known as spherical normal density.
It is noted that MVND with mean μ and VCM Σ is constant on concentric ellipsoids

A general class of distⁿ with this property is the class of Elliptically contoured distⁿ with the density.

$|\Lambda|^{-1/2} g[(\underline{x}-\delta)' \Lambda (x-\delta)]$ where Λ is a +ve definite matrix

If C is a NSM s.t $C'\Lambda^{-1}C = I$ then the transformation $(\underline{x}-\delta) = C\underline{Y}$ carries the general density to Multivariate normal density then the contours of constants density then are the spheres centered at origin. the class of such density is known as spherically ~~cond~~ centered distn.

Applications :-

Elliptical distn are important in the study of economics and financials situation where various factors of portfolio analysis including mutual fan separation theorem and capital assets pricing model hold good for elliptical distn.

UNIT-II

Wishart Distribution (1928)

another name is multivariate χ^2 or generalization of χ^2

The wishart distⁿ is the distⁿ of the sample variance ~~co~~ co variance matrix it is also called as distⁿ of the moment matrix

Wishart distⁿ is a multivariate generalization of χ^2 - distⁿ

The sample VCM is given by

$$S = \frac{1}{(N-1)} \sum_{\alpha=1}^{N} (\underline{X}_\alpha - \bar{\underline{X}})(\underline{X}_\alpha - \bar{\underline{X}})'$$

OR

$$(N-1)S = \sum_{\alpha=1}^{N} (\underline{X}_\alpha - \bar{\underline{X}})(\bar{\underline{X}}_\alpha - \bar{X})$$

let $Z_\alpha = (\underline{X}_\alpha - \bar{\underline{X}})$ then:

$Z_\alpha \sim N_p(\underline{0}, \Sigma)$ Now the p.df of Z_α is given by

$$\frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{\frac{1}{2} Z_\alpha' \Sigma^{-1} Z_\alpha}$$

$\therefore$ The pdf of $Z_1, \ldots Z_N$. (say) $f(z)$ is given by

$$f(z) = \frac{e^{-\frac{1}{2}\sum_{\alpha=1}^{N} Z_\alpha' \Sigma^{-1} Z_\alpha}}{(2\pi)^{\frac{(N-1)p}{2}} |\Sigma|^{N-1/2}}$$

Exponent is $-\frac{1}{2} \sum_{\alpha=1}^{N-1} Z_\alpha' \Sigma^{-1} Z_\alpha$

which is a scalar quantity $\forall \alpha$ since $\underline{Z}_\alpha' \Sigma^{-1} \underline{Z}_\alpha$ is scalar.

$$\therefore \sum_{\alpha=1}^{N-1} \underline{Z}_\alpha' \Sigma^{-1} \underline{Z}_\alpha = tr\, \underline{Z}_\alpha' \Sigma^{-1} \underline{Z}_\alpha$$

$$= \sum_{\alpha=1}^{N-1} tr\, \underline{Z}_\alpha' \Sigma^{-1} \underline{Z}_\alpha = \sum_{\alpha=1}^{N-1} tr\, \Sigma^{-1} \underline{Z}_\alpha \underline{Z}_\alpha'$$

$$= tr\, \Sigma^{-1} \sum_{\alpha=1}^{N-1} \underline{Z}_\alpha \underline{Z}_\alpha'$$

$$= tr.(\Sigma^{-1} A) \text{ as } A = \sum_\alpha (\underline{X}_\alpha - \bar{X})(\underline{X}_\alpha - \bar{X})$$

$$f(\underline{z}) = \frac{1}{(2\pi)^{\frac{(N-1)p}{2}} |\Sigma|^{(N-1)/2}} e^{-\frac{1}{2} tr.\, \Sigma^{-1} A}$$

As $f(\underline{z})$ is p.d.f $\int_R f(\underline{z})\, d\underline{z} = 1$

i.e $\int e^{-1/2\, tr.\, \Sigma^{-1} A} = (2\pi)^{(N-1)p/2} |\Sigma|^{\left(\frac{N-1}{2}\right)}$ —— (1)

Since the Matrix A is a Symmetric matrix the no of distinct elements of the matrix will be

$\frac{p(p+1)}{2}$ and let these are $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ \vdots & \cdots & \cdots & a_{pp} \end{bmatrix}$

Now the ch. f^n of A is given by

$\phi(A) = E\left[e^{tr.\, iUA}\right]$ where U is a matrix such that

$$U_{ij} = U_{ji} = \frac{t_{ij}}{2}$$

i.e

$$U = \begin{bmatrix} t_{11} & t_{12}/2 & \cdots & t_{1p}/2 \\ t_{21}/2 & t_{22} & \cdots & t_{2p}/2 \\ \vdots & \vdots & & \vdots \\ t_{p1}/2 & t_{p2}/2 & \cdots & t_{pp} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & & \vdots \\ a_{p1} & \cdots & \cdots & a_{pp} \end{bmatrix}$$

$$\therefore tr(UA) = t_{11}a_{11} + t_{12}a_{12} + \cdots + t_{1p}a_{1p}$$
$$t_{2p}a_{2p} + \cdots + t_{2p}a_{2p}$$
$$+ \cdots + $$
$$+ \cdots + t_{pp}a_{pp} \quad \text{--- } (**)$$

$$\therefore \phi_A(U) = E\left[\exp\{i(t_{11}a_{11} + t_{12}a_{12} + \cdots + t_{pp}a_{pp})\}\right]$$

Again

$$\phi_A(U) = E\left[e^{tr.iUA}\right] = \int e^{tr.iUA} f(z)\,dz$$

$$= \frac{1}{(2\pi)^{\frac{(N-1)p}{2}} |\Sigma|^{\left(\frac{N-1}{2}\right)}} \int e^{tr.iUA} \cdot e^{-\frac{1}{2} tr.\Sigma^{-1}A}\,dz$$

$$= \frac{1}{(2\pi)^{\frac{(N-1)p}{2}} |\Sigma|^{\left(\frac{N-1}{2}\right)}} \int e^{-\frac{1}{2}\{tr.(\Sigma^{-1} - 2iU)A\}}\,dz \quad \text{--- } (2)$$

But from the analogy of (1) we have

$$\int e^{-\frac{1}{2} tr.(\Sigma^{-1} - 2iU)A}\,dz = |\Sigma^{-1} - 2iU|^{-\left(\frac{N-1}{2}\right)} (2\pi)^{\frac{(N-1)p}{2}}$$

$$\phi_A(U) = \frac{|\Sigma^{-1} - 2iU|^{-\left(\frac{N-1}{2}\right)}}{(2\pi)^{\frac{(N-1)p}{2}} |\Sigma^{-1}|^{\left(\frac{N-1}{2}\right)}} \times (2\pi)^{\frac{(N-1)p}{2}}$$

using $(*)$

$$\phi_A(U) = \frac{1}{|\Sigma^{-1} - 2iU|^{\frac{(N-1)}{2}} \, |\Sigma|^{(N-1)/2}}$$

OR

$$\phi_A(U) = \frac{1}{|I - 2iU|^{\frac{(N-1)}{2}}}$$

Now Suppose that the pdf of A is given by

$$P(A) = C|\Sigma|^{-\frac{(N-1)}{2}} |A|^{\frac{(N-p-2)}{2}} \exp\{-\tfrac{1}{2} tr.\Sigma^{-1}A\} \quad \text{— (\#)}$$

where

$$C = \frac{1}{(2)^{\frac{(N-1)p}{2}} \pi^{\frac{p(p-1)}{2}} \prod_{j=1}^{p} \Gamma\frac{N-j}{2}}$$

Again $\int P(A)\, dA = 1$ gives

$$\int |A|^{\frac{(N-p-2)}{2}} c|\Sigma|^{-\frac{(N-1)}{2}} \exp\{-\tfrac{1}{2} tr - \Sigma^{-1}A\}\, dA = 1$$

$$\Rightarrow \int |A|^{\frac{(N-p-2)}{2}} \exp\{-\tfrac{1}{2} tr.\Sigma^{-1}A\}\, dA = \frac{|\Sigma|^{\frac{(N-1)}{2}}}{C} \quad \text{— (3)}$$

and $\phi_A(U) = E\left[e^{tr\, iUA}\right]$

$$= C|\Sigma|^{-\frac{(N-1)}{2}} \int |A|^{\frac{(N-p-2)}{2}} e^{-\frac{1}{2} tr(\Sigma^{-1} - 2iU)A}\, dA$$

$$= \cancel{C}\, |\Sigma|^{-\frac{(N-1)}{2}} \, \frac{|\Sigma^{-1} - 2iU|^{-\frac{(N-1)}{2}}}{\cancel{C}}$$

$$= |I - 2iU\cancel{\Sigma}|^{-\frac{(N-1)}{2}} = \frac{1}{|I - 2iU\cancel{\Sigma}|^{-\frac{(N-1)}{2}}} \quad \text{— (4)}$$

Hence, it is observed that the ch.fn 'A' is same as that given by (4) which is obtained by using the proposed density $P(A)$ $\therefore$ using the uniqueness theorem ch.fn we conclude that the pdf of A is the proposed one given by (#). This dist^n of 'A' is called the wisharts dist^n and is denoted by $W_p(n, \Sigma)$ where $n = (N-1)$

— x —

Remark :- Wishart dist^n is a Multivariate analogue of χ^2 density i.e for $p=1$ we obtain χ^2_n from $W_p(n, \Sigma)$

Proof For $p=1$, & $\Sigma_{1\times 1} = \sigma^2$ (say)

$$A = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})^2 = (N-1)s^2 = ns^2$$

For $p=1$,
$$C = \frac{1}{2^{\frac{(N-1)}{2}} \pi^0 \Gamma\frac{N-1}{2}} = \frac{1}{2^{n/2}\Gamma\frac{n}{2}}$$

$$P(A) = C|\sigma^2|^{-\frac{(N-1)}{2}} (ns^2)^{\frac{(N-1-2)}{2}} \exp\{-1/2\, tr.\, \Sigma^{-1}A\}$$

$$= C \cdot \frac{1}{\sigma^2} \left(\frac{ns^2}{\sigma^2}\right)^{\frac{n}{2}-1} e^{-\frac{ns^2}{2\sigma^2}}$$

OR

$$P(A) = \frac{1}{2^{n/2}\Gamma\frac{n}{2}\,\sigma^2} \left(\frac{ns^2}{\sigma^2}\right)^{n/2-1} e^{-\frac{ns^2}{2\sigma^2}} \cdot dA$$

let $A = ns^2 \Rightarrow dA = n\,ds^2$

putting $\frac{ns^2}{\sigma^2} = u \Rightarrow \frac{n\,ds^2}{\sigma^2} = du$

$$\boxed{P(U)\,dU = \frac{1}{2^{n/2}\,\Gamma\frac{n}{2}} U^{\frac{n}{2}-1} e^{-\frac{U}{2}}\,dU} \; \chi^2_n$$

Properties of wishart distribution

(1) The chfn of A (wishart distn) is

$$\phi_A = |I - 2iU\Sigma|^{-\frac{(N-1)}{2}}$$

(2) Reproductive property

If $A_i \sim W_p(n_i, \Sigma)$ $i=1,2$. then
$A_1 + A_2 \sim W_p(n_1+n_2, \Sigma)$

Proof By defn $\phi_A(U) = E[e^{tr.iUY}]$

$$= E[e^{tr\,iU(A_1+A_2)}]$$
$$= E[e^{tr\,iUA_1}] \cdot E[e^{tr\,iUA_2}]$$
$$= \phi_{A_1}(U) \cdot \phi_{A_2}(U)$$
$$= |I - 2iU\Sigma|^{-\frac{n_1}{2}} \cdot |I - 2iU\Sigma|^{-\frac{n_2}{2}}$$
$$= |I - 2iU\Sigma|^{-\frac{(n_1+n_2)}{2}}$$

$\Rightarrow Y \sim W_p(n_1+n_2, \Sigma)$

In general, if A_i $(i=1,\ldots k)$ are independently distributed as $W_p(n_i, \Sigma)$ then

$$\sum_{i=1}^{k} A_i \sim W_p\left(\sum_{i=1}^{k} n_i, \Sigma\right)$$

(3) If $A \sim W_p(n, \Sigma)$ consider CA, where C is a scalar constant then $CA \sim W_p(n, C\Sigma)$

proof: let $Y = CA$ then $\phi_Y(U) = E[e^{tr\, iUY}]$
$= E[e^{tr\, iUCA}]$,

let $U^* = UC$ so $\phi_Y(U) = E[e^{tr\, iU^*A}]$
$= \phi_A(U^*)$

or $\phi_Y(U) = |I - 2iUC\Sigma|^{-n/2}$

$\Rightarrow Y \sim W_p(n, C\Sigma)$

(4) <u>Marginal Distribution</u>:-

let A and Σ be partitioned as:

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \text{ and } \Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{22} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right]$$

where $A_{11} = q \times q$ $\quad A_{21} = (p-q) \times q$
$A_{12} = q \times (p-q)$ $\quad A_{22} = (p-q) \times (p-q)$

$\Sigma_{11} = q \times q$ $\quad \Sigma_{21} = (p-q) \times q$
$\Sigma_{12} = q \times (p-q)$ $\quad \Sigma_{22} = (p-q) \times (p-q)$

If $A \sim W_p(n, \Sigma)$ then $A_{11} \sim W_q(n, \Sigma_{11})$

proof: Given that $A \sim W_p(n, \Sigma)$ we also write:
$A = \sum_{\alpha=1}^{N-1} Z_\alpha Z_\alpha'$ where $Z_\alpha \sim N_p(0, \Sigma)$ and are iid.

Consider the partitioning of $Z_\alpha = \begin{bmatrix} Z_\alpha^{(1)} \\ Z_\alpha^{(2)} \end{bmatrix} \begin{array}{l} \to (q \times 1) \\ \to (p-q) \times 1 \end{array}$

such that

$\underline{Z}_\alpha^{(1)} \sim N_q(\underline{0}, \Sigma_{11})$

Thus, $A = \sum_{\alpha=1}^{N-1} \underline{Z}_\alpha \underline{Z}_\alpha' = \sum_{\alpha=1}^{n} \begin{bmatrix} \underline{Z}_\alpha^{(1)} \\ \underline{Z}_\alpha^{(2)} \end{bmatrix} \begin{bmatrix} \underline{Z}_\alpha^{(1)'} & \underline{Z}_\alpha^{(2)'} \end{bmatrix}$

Or,

$$A = \begin{bmatrix} \sum_{\alpha=1}^{n} \underline{Z}_\alpha^{(1)} \underline{Z}_\alpha^{(1)'} & \sum_{\alpha=1}^{n} \underline{Z}_\alpha^{(1)} \underline{Z}_\alpha^{(2)'} \\ \sum_{\alpha=1}^{n} \underline{Z}_\alpha^{(2)} \underline{Z}_\alpha^{(1)'} & \sum_{\alpha=1}^{n} \underline{Z}_\alpha^{(2)} \underline{Z}_\alpha^{(2)'} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$\therefore A_{11} = \sum_{\alpha=1}^{n} \underline{Z}_\alpha^{(1)} \underline{Z}_\alpha^{(1)'}$, $\underline{Z}_\alpha \sim N_q(0, \Sigma_{11})$ and iid

$\therefore A_{11} \sim W_q(n, \Sigma_{11})$

5. If $A \sim W_p(n, \Sigma)$ and $\underline{L}_{p\times 1}$ is any arbitrary fixed vector then $\underline{L}' A \underline{L} \sim \chi^2_n \cdot \sigma^2$

Proof: Given that $A \sim W_p(n, \Sigma)$ and $A = \sum_{\alpha=1}^{n} Z_\alpha Z_\alpha'$

$Z_\alpha \sim N_p(0, \Sigma)$ and are iid

$\therefore \underline{L}' A \underline{L} = \underline{L}' \sum_{\alpha=1}^{n} Z_\alpha \underline{Z}_\alpha' \underline{L} = \sum_{\alpha=1}^{n} (\underline{L}' Z_\alpha)(\underline{L}' Z_\alpha)'$

Since $\underline{L}' Z_\alpha$ is a linear combination of $\underline{Z}_\alpha$ also it is scalar $\therefore \underline{L}'A\underline{L} = \sum_{\alpha=1}^{n} (\underline{L}' \underline{Z}_\alpha)^2$

Again, $Z_\alpha \sim N_p(0, \Sigma)$ $\therefore \underline{L}' Z_\alpha \sim N_1(0, \underline{L}' A \underline{L})$

Or,

$\underline{L}' Z_\alpha \sim N_1(0, \sigma^2)$ where $\underline{L}' \Sigma \underline{L} = \sigma^2$ (say)

$\therefore \underline{L}' A \underline{L} = \sum_{\alpha=1}^{n} (\underline{L}' Z_\alpha)^2$ is the sum of square of SNV's

$\therefore \frac{\underline{L}'A\underline{L}}{\sigma^2} \sim \chi^2_n$

Or, $\boxed{\underline{L}'A\underline{L} \sim \sigma^2 \chi^2_n}$

6 If $A \sim W_p(n, \Sigma)$ then $B'AB \sim W_q(n, B'\Sigma B)$ where B is a $(p \times p)$ matrix

Proof Given that $A \sim W_p(n, \Sigma)$ where $A = \sum_{\alpha=1}^{n} Z_\alpha Z_\alpha'$,

$Z_\alpha \sim N_p(0, I)$

now $B'AB = B' \sum_{\alpha=1}^{n} Z_\alpha Z_\alpha' B = \sum_{\alpha=1}^{n} B' Z_\alpha Z_\alpha' B = \sum_{\alpha=1}^{n} W_\alpha W_\alpha'$ (say)

where

$W_\alpha = B'Z_\alpha$ and are indep. for $\alpha = 1, 2, \ldots n$

$\therefore B'AB \sim W_q(n, B'\Sigma B)$

7 CONDITIONAL DISTRIBUTION

let $A \sim W_p(n, \Sigma)$ and A and Σ are partitioned as: $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$

then $A_{11.2}$ is distributed as $W_p[n-(p-q), \Sigma_{11.2}]$

where $A_{11.2} = A_{11} - A_{12} A_{22}^{-1} A_{21}$

and $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$

* DISTRIBUTION OF SAMPLE MEAN VECTOR AND SAMPLE VCM :-

let $x \sim N_p(\mu, \Sigma)$ then L.f of $\underline{x}_1, \underline{x}_2, \ldots \underline{x}_N$ is:

$$L = \frac{1}{(2\pi)^{\frac{Np}{2}} |\Sigma|^{N/2}} e^{-\frac{1}{2} tr. \Sigma^{-1} A} \cdot e^{-1/2 (\bar{\underline{x}} - \underline{\mu})' (\bar{\underline{x}} - \underline{\mu})} \quad \text{--- (1)}$$

The MLE^s are:

$$\hat{\underline{\mu}} = \bar{x} \quad , \quad \hat{\Sigma} = A/N \quad \text{or,} \quad \hat{\Sigma} = S : \text{Sample VCM}$$

Consider the transformation $Y = P X$
where P is an orthogonal matrix
and its last row is taken as $\frac{1}{\sqrt{n}}$

in this case $|J| = 1$ and consider the partitioning of $Y = \begin{bmatrix} Z \\ \sqrt{n}\,\bar{X} \end{bmatrix}$ where $Z_{(n-1) \times p}$

We can write $(\underline{x} - \underline{\mu})'(\underline{x} - \underline{\mu}) = (x - \underline{\mu})' P P' (\underline{x} - \underline{\mu})$
as $PP' = I$

$$= \begin{bmatrix} Z \\ \sqrt{n}\,(\bar{x} - \mu) \end{bmatrix}' \begin{bmatrix} Z \\ \sqrt{n}\,(\bar{x} - \mu) \end{bmatrix}$$

$$= Z Z' + N (\underline{\bar{x}} - \underline{\mu})' (\bar{\underline{x}} - \underline{\mu})$$

$$= \frac{1}{(2\pi)^{Np/2} |\Sigma|^{N/2}} e^{-\frac{1}{2} tr. \Sigma^{-1} [ZZ' + N(\bar{x} - \mu)'(\bar{x} - \mu)]}$$

which can be partitioned as ZZ' and $N(\bar{x} - \mu)'(\bar{x} - \mu)$ Again Z is a matrix of $(N-1)$ obser from $(p-1)$ variables.

$\therefore$ $Z'Z \sim W_{p-1}(n-1, \Sigma)$ as $Z'Z = Y'Y - N \bar{X} \bar{X}'$

$\therefore Z'Z' \sim W_{p-1}(n-1, \Sigma)$

as $Z'Z = Y'Y - n\bar{X}\bar{X}'$
$= X'P'PX - n\bar{X}\bar{X}'$
$= nS = n\hat{\Sigma}$.

Further $\bar{X} \sim N_p(\underline{\mu}, \frac{\Sigma}{n})$

UNIT-3

* Problem of classification and discrimination :-

INTRODUCTION :-

In many real life situations Sometimes we have a group of individuals or objects or some conditions and we wish to classify them into different categories.

This may be done by taking into account certain observations on a given set of individuals or objects the condition and on the basis of their charachtersstics we wish to decide to which group a perticular individual to which group belongs. This is the case where all the individual/of objects/condition to be classified.

In the problem of classification we do not know the number of categories into which individuals/objects/conditions are to be classified, here we have to make our own categories according to our satisfaction. However in case where the no. of categories are known and we have to only separate the observation into thes known groups is known as the problem of discrimination.

For the both purposes we measure p-charachterstice on each individuals/object/condition and thus we have a vector $\underline{X}_{p\times 1}$ as the observation Corresponding to k-categories, we can assume k- underlying pop^{ns} and when we say a vector belongs to a category it is equivalent to saying test the individual (objects/ conditions belongs to corresponding pop^{ns}

* AREAS OF APPLICATIONS:

1. Anthropological/ Archeological studies (when the information is lost)

2. Medical diagnosis of rare disease (like:- brain cancer etc)

3. Financial sector

4. In general where the measurement of identification variable is difficult or expensive.

* Mathematical formulation:

Let $\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ for the sake of simplicity we consider only two categories into which $\underline{X}_{p\times 1}$ is to be classified

Then any random vector $\underline{X}_{p\times 1}$ may come from either the pop^n π_1 or π_2

For taking our decision we divide (?) our p-dimensional space into two region R_1 and R_2

and the decision is made as: if the observed value of $\underline{X} \in R_1$, we say that the individual belongs to the Popn π_1, otherwise it belongs to π_2. R_1 and R_2 must be mutually exclusive

Decision Taken	Actual Population	
	π_1	π_2
π_1	No Error	Second type of misclassification (M.C)
π_2	firsty type of m.c	No Error

In any problem while choosing R_1 and R_2 there is same associated loss, so our purpose is to choose R_1 and R_2 in such a way that the average loss or cost is minimum

For this we define:-

$C(2|1)$: cost of 1st type of m.c i.e cost of misclassifying an order as coming from π_2 when actually it belongs to π_1

$C(1|2)$: Cost of 2nd type of m.c i.e cost of missclassifying and order as coming from π_1, when actually it belongs to π_2

we also assume that $C(2|1) \geq 0$, $C(1,2) \geq 0$ and $C(1|1) = 0$ & $C(2|2) = 0$

let $f_i(\underline{x})$ be the p.d.f of the i^{th} population (π_i)

define: $P(j|i)$ = Prob. of misclassifying an abses into j^{th} population when actually it belong to i^{th} pop^n

$= Pr.[\underline{x} \in \pi_j \mid \pi_i$ is actual $pop^n]$

Some Definitions

1. Averaget Cost (A.C), it is defined as.

$$\sum_{j=1}^{k} P(j|i)\, c(j|i) = \underline{\gamma_i}(R) \quad (\text{say})$$

now, if we have different rules of classification (say) : R', R', R'' etc. we will have different A.C like $\gamma_1(R)$, $\gamma_2(R')$, $\gamma_3(R'')$ etc.

2. Among the rules of classifications (say) R_1 and R_1^*, R_1 is said to be better than

R^* if $\gamma_i(R_1) \leq \gamma_i(R^*)$ $i = 1, \ldots k$

with atleast one inequality

R_1 is said to be ADMISSIBLE. if there does not exist any R^* which is better than R_1

3 BAYES RULE

let q_i : be the prob. that an individual belongs to π_i (i.e aprior prob). then the A.C. of m.c is defined as $\sum_{i=1}^{k} q_i \gamma_i(R)$ w.r.t apriori prob. $(q_1, q_2 \ldots, q_k)$

A Bayes rule is defined as:
Given $q:(q_1, q_2, \ldots q_k)$ a classification rule which minimizes the a.c. of m.c is called the Bayes rule w.r.t. q

<u>Note</u> This may result is large prob. of m.c though it is optimum.

★ Problem of classification into Two population (General Theory)

let q_1 be the probability that an observation selected at random belongs to π_1 and q_2 be the probability that it belongs to π_2. Again let us suppose that $p_1(x)$ be the pdf of π_1 and $p_2(x)$ be the pdf of π_2 define

$P(2|1)$: Prob of misclassificationing an observation π_2 when actually it belongs to π_1

$= Pr[x \in R_2 \mid \text{when } P_1(x) \text{ is the actual pdf}]$

i.e

$P(2|1) = \int_{R_2} P_1(\underline{x})\,d\underline{x}$, similarly $P(1|2) = \int_{R_1} P_2(\underline{x})\,d\underline{x}$

$\therefore$ the prob of 1^{st} type of M.C $= q_1 P(2|1)$
and the prob of 2^{nd} type of M.C $= q_2 P(1|2)$

So the Average cost (A.C)

$$= C(1/2) \cdot q_2 P(1|2) + C(2|1)\, q_1 P(2|1)$$

$$= C(1|2) \cdot q_2 \int_{R_1} P_2(\underline{x})\,d\underline{x} + C(2|1) \int_{R_2} P_1(x)\,d\underline{x}$$

now we wish to obtain R_1 and R_2 for a good classification in such a way that $A.C\,(R_1, R_2)$ is minimum.

i.e If R_1^* & R_2^* be any other regions of classification than $A.C\,(R_1, R_2) \leq A.C\,(R_1^*, R_2^*)$

i.e

$$q_1 C(2|1) \int_{R_2} P_1(\underline{x})\,dx + q_2 C(1|2) \int_{R_1} P_2(\underline{x})\,dx \leq$$

$$q_1 \cdot C(2|1) \cdot \int_{R_2^*} P_1(\underline{x})\,d\underline{x} + q_2 C(1|2) \int_{R_1^*} P_2(\underline{x})\,d\underline{x} \quad — (*)$$

as $P_1(\underline{x})$ and $P_2(\underline{x})$ are the pdf

$$\int_{R_1 \cup R_2} P_i(\underline{x})\,d\underline{x} = 1 \quad \text{and} \quad \int_{R_1^* \cup R_2^*} P_i(\underline{x})\,d\underline{x} = 1$$

i.e

$$\int_{R_2} P_i(\underline{x})\,d\underline{x} = 1 - \int_{R_1} P_i(\underline{x})\,d\underline{x} \quad \text{and}$$

$$\int_{R_2^*} P_i(\underline{x})\,d\underline{x} = 1 - \int_{R_1^*} P_i(\underline{x})\,d\underline{x} \quad \text{for } i = 1, 2$$

Substituting all the above values in the Expression $(*)$ we get

$$q_1 C(2|1) \left[1 - \int_{R_1} P_1(\underline{x})\,d\underline{x} \right] + q_2 C(1|2) \int_{R_1} P_2(\underline{x})\,d\underline{x}$$

$$\leq q_1 C(2|1) \left[1 - \int_{R_1^*} P_1(\underline{x})\,d\underline{x} \right] + q_2 C(1|2) \int_{R_1^*} P_2(\underline{x})\,d\underline{x}$$

i.e.

$$\int_{R_1} \left[q_2 C(1|2) P_2(\underline{x}) - q_1 C(2|1) P_1(\underline{x})\,d\underline{x} \right]$$

$$- \int_{R_1^*} \left[q_2 C(1|2) P_2(\underline{x}) - q_1 C(2|1) P_1(\underline{x})\,d\underline{x} \right] \leq 0$$

this condition will be satisfied, if we choose R_1 s.t

$q_2 C(2|2) p_2(\underline{x}) - q_1 C(2|1) p_1(\underline{x}) \leq 0$ inside R_1

and $q_2 C(1|2) p_2(\underline{x}) - q_1 C(2|1) p_1(\underline{x}) \geq 0$ outside R_1

So, R_1 is the region where

$$q_1 C(2|1) p_1(\underline{x}) \geq q_2 C(1/2) p_2(\underline{x})$$

or,
$$\boxed{R_1 : \frac{p_1(\underline{x})}{p_2(\underline{x})} \geq \frac{q_2}{q_1} \cdot \frac{C(1|2)}{C(2|1)}}$$

and
$$\boxed{R_2 : \frac{p_1(\underline{x})}{p_2(\underline{x})} < \frac{q_2}{q_1} \cdot \frac{C(1|2)}{C(2|1)}}$$

If q_1, q_2 are known and $C(1|2), C(2|1)$ are also known we can obtain R_1 and R_2 if $p_1(\underline{x})$ and $p_2(\underline{x})$ are completely specified.

SPECIAL CASE : If $q_1 = q_2 = \frac{1}{2}$ and

$C(1|2) = C(2|1)$ then RHS $= 1$

(*) Problem of classification (In to two MVND) :-

Let $\underline{X}_{p\times 1}$ be an observations vector and suppose that we have to classify it into two categories i.e we assume that an observation can come either from the popn $\pi_1 : N_p(\mu^{(1)}, \Sigma)$ or from the popn $\pi_2 : N_p(\mu^{(2)}, \Sigma)$ equality of Σ

$$f_1(\underline{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\mu^{(1)})' \Sigma^{-1} (\underline{x}-\mu^{(1)})}$$

$$f_2(\underline{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}^{(2)})'\Sigma^{-1}(\underline{x}-\underline{\mu}^{(2)})}$$

$$\frac{f_1(\underline{x})}{f_2(\underline{x})} = \frac{e^{-\frac{1}{2}(\underline{x}-\underline{\mu}^{(1)})\Sigma^{-1}(\underline{x}-\underline{\mu}^{(1)})}}{e^{-\frac{1}{2}(\underline{x}-\underline{\mu}^{(2)})\Sigma^{-1}(\underline{x}-\underline{\mu}^{(2)})}}$$

$$= e^{-\frac{1}{2}(\underline{x}-\underline{\mu}^{(1)})\Sigma^{-1}(\underline{x}-\underline{\mu}^{(1)}) + \frac{1}{2}(\underline{x}-\underline{\mu}^{(2)})\Sigma^{-1}(\underline{x}-\underline{\mu}^{(2)})}$$

the Region of classification are given by

$$R_1 : \frac{f_1(x)}{f_2(x)} \geq k \quad \text{and} \quad R_2 : \frac{f_1(x)}{f_2(x)} < k$$

so, here

$$R_1 = e^{\frac{1}{2}[(\underline{x}-\underline{\mu}^{(2)})'\Sigma^{-1}(\underline{x}-\underline{\mu}^{(2)}) - (\underline{x}-\underline{\mu}^{(1)})'\Sigma^{-1}(\underline{x}-\underline{\mu}^{(1)})]} \geq k$$

OR

$$R_1 = \frac{1}{2}[(\underline{x}-\underline{\mu}^{(2)})'\Sigma^{-1}(\underline{x}-\underline{\mu}^{(2)}) - (\underline{x}-\underline{\mu}^{(1)})'\Sigma^{-1}(\underline{x}-\underline{\mu}^{(1)})] \geq \log k$$

OR

$$R_1 = \frac{1}{2}\Big[\{\underline{x}'\Sigma^{-1}\underline{x} - \underline{x}'\Sigma^{-1}\underline{\mu}^{(2)} - \underline{\mu}^{(2)'}\Sigma^{-1}\underline{x} + \underline{\mu}^{(2)'}\Sigma^{-1}\underline{\mu}^{(2)}\}$$
$$- \{\underline{x}'\Sigma^{-1}\underline{x} - \underline{x}'\Sigma^{-1}\underline{\mu}^{(1)} - \underline{\mu}^{(1)'}\Sigma^{-1}\underline{x} + \underline{\mu}^{(1)'}\Sigma^{-1}\underline{\mu}^{(1)}\}\Big]$$
$$\geq \log k$$

OR $R_1 = \frac{1}{2}$ Since $\underline{x}'\Sigma^{-1}\underline{\mu}^{(2)} = \underline{\mu}^{(2)'}\Sigma^{-1}\underline{x}$

and $\underline{x}'\Sigma^{-1}\underline{\mu}^{(1)} = \underline{\mu}^{(1)'}\Sigma^{-1}\underline{x}$

$$\therefore R_1 = \frac{1}{2}\left[2\underline{x}'\Sigma^{-1}\underline{\mu}^{(1)} - 2\underline{x}'\Sigma^{-1}\underline{\mu}^{(2)} - \underline{\mu}^{(1)'}\Sigma^{-1}\underline{\mu}^{(1)} + \underline{\mu}^{(2)'}\Sigma^{-1}\underline{\mu}^{(2)}\right]$$
$$\geq \log k$$

OR

$$R_1 = \underline{X}'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) - \frac{1}{2}\left\{(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})\right\} \geq \log k$$

OR

$$R_1 = \underline{X}'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) \geq \frac{1}{2}\left\{(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})\right\} + \log k$$

$$R_2 = X'\boxed{\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})} < \frac{1}{2}\left\{(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})\right\} + \log k$$

Under the assumption of the two population equally likely i-e $q_1 = q_2 = \frac{1}{2}$ and if the two costs of misclassification are SAME i-e $C(1|2) = C(2|1)$, $\log k = 0$. ~~Also,~~

Also, let $\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) = \underline{d}_{p\times 1}$ then

$$R_1 : \underline{X}'\underline{d} \geq \frac{1}{2}(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})$$

$$R_2 : \underline{X}'\underline{d} < \frac{1}{2}(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})$$

$$C(1|2)\,P(1|2)\,q_2 = C(2|1)\,P(2|1)\,q_1 = C \text{ (say)}$$

let $U = X'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) - \frac{1}{2}\left\{(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})'\Sigma^{-1}(\underline{\mu}^{(1)} - \underline{\mu}^{(2)})\right\}$

then $R_1 : U \geq c$ and $R_2 : U < c$,

In order to find out the value of 'c' we must know the distⁿ of U. If $\underline{X}$ comes from Π_1 then $E_1(U)$ can be obtained.

We have

$P(1|2) = Pr\{U \geq c \mid X \in \Pi_2\} = Pr\{U \geq c \mid X \sim N_p(\mu^{(2)}, \Sigma)\}$

and $P(2|1) = Pr\{U < c \mid X \in \Pi_1\} = Pr\{U < c \mid X \sim N_p(\mu^{(1)}, \Sigma)\}$

it is a ~~observation~~ observed that U is a linear combination of X which is normally distributed So U is NORMALLY distributed

1. when $X \sim N_p(\mu^{(1)}, \Sigma)$,

$$E_1(U) = \mu^{(1)'}\Sigma^{-1}(\mu^{(1)}-\mu^{(2)}) - \frac{1}{2}\{(\mu^{(1)}+\mu^{(2)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})\}$$

OR

$$E_2(U) = \frac{1}{2}(\mu^{(1)}-\mu^{(2)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})$$

Also, Mahalanobis $-D^2$ (Distance) is given by

$$(\mu^{(1)}-\mu^{(2)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)}) = \alpha \text{ (say)}$$

i.e $\boxed{E_1(U) = \alpha/2}$

Again, $Var_1(U) = E[U - E_1(U)]^2$

OR

$$Var_1(U) = \left[X'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)}) - \frac{1}{2}\{(\mu^{(1)}+\mu^{(2)})\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})\right]^2$$

$$= E\left[\{(X-\mu^{(1)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})\}'\{(X-\mu^{(1)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})\}\right]$$

$$= (\mu^{(1)}-\mu^{(2)})'\Sigma^{-1}E(X-\mu^{(1)})(X-\mu^{(1)})'\Sigma^{-1}(\mu^{(1)}-\mu^{(2)})$$

$$= (\mu^{(1)}-\mu^{(2)})'\Sigma(\mu^{(1)}-\mu^{(2)})$$

$$= \alpha$$

i.e $\boxed{V_1(U) = \alpha}$

2. If $\underline{X} \sim N_p(\underline{\mu}^{(2)}, \Sigma)$ then

$$\boxed{E_2(U) = -\alpha/2} \quad \text{and} \quad \boxed{V_2(U) = \alpha}$$

So,

If $\underline{X} \in \pi_1$, $U \sim N(\alpha/2, \alpha)$ and
If $\underline{X} \in \pi_2$, $U \sim N(-\alpha/2, \alpha)$

$$x \in \pi_1 \quad P_1(U) = \frac{1}{\sqrt{2\pi}\,\alpha} e^{-\frac{1}{2\alpha}\left(U - \frac{\alpha}{2}\right)^2}$$

$$x \in \pi_2 \quad P_2(U) = \frac{1}{\sqrt{2\pi}\,\alpha} e^{-\frac{1}{2\alpha}\left(U + \frac{\alpha}{2}\right)^2}$$

$$\therefore C(2|1) = \int_{-\infty}^{c} \frac{1}{\sqrt{2\pi}\,\alpha} e^{-\frac{1}{2\alpha}\left(U - \frac{\alpha}{2}\right)^2} du$$

and

$$\therefore C(1|2) = \int_{c}^{\infty} \frac{1}{\sqrt{2\pi}\,\alpha} e^{-\frac{1}{2\alpha}\left(U + \frac{\alpha}{2}\right)^2} du$$

let us take $\frac{U - \alpha/2}{\sqrt{\alpha}} = z$ and $\frac{U + \alpha/2}{\sqrt{\alpha}} = t$

$$\text{then } C(2|1)\, q_1 \int_{-\infty}^{\frac{c}{\sqrt{\alpha}} - \frac{1}{2}\sqrt{\alpha}} \phi(z)\, dz$$

$$= C(1/2)\, q_2 \int_{\frac{c}{\sqrt{\alpha}} + \frac{1}{2}\sqrt{\alpha}}^{\infty} \phi(t)\, dt \quad \text{——} \quad (*)$$

$\therefore$ knowing the values of q_1, q_2 $C(1|2)$ and $C(2|1)$ $(*)$ can be evalued using the tables of Standard Normal distn and we can obtain the value of 'c'.

Hence

$R_1 : \underline{x}' \Sigma^{-1} (\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) \geq \frac{1}{2} \left[(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})' \Sigma^{-1} (\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) \right] + C$

$R_2 : \underline{x}' \Sigma^{-1} (\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) < \frac{1}{2} \left[(\underline{\mu}^{(1)} + \underline{\mu}^{(2)})' \Sigma^{-1} (\underline{\mu}^{(1)} - \underline{\mu}^{(2)}) \right] + C$

* Classification when the Population Parameters are Estimated :-

(Fisher's Criterion 1936)

Let $\underline{x}_{p\times q}$ be a vector which it has to be descriminated into two population $N_p(\underline{\mu}^{(1)}, \Sigma)$ and $N_p(\underline{\mu}^{(2)}, \Sigma)$ consider another vector $\underline{l}_{p\times 1}$ then fisher's suggested a linear combination $\underline{l}'\underline{x}$ for the discrimation where $\underline{l}$ is to be choosen such that $[E_1(\underline{l}'\underline{x}) - E_2(\underline{l}'\underline{x})]^2$ is maximum such that ~~consider~~ $var(\underline{l}'\underline{x}) = k$ (constant) generally taken to be unity. Now, in vector to discriminate $\underline{x}$ we have to maximize $(\underline{l}'\underline{\mu}^{(1)} - \underline{l}'\underline{\mu}^{(2)})$ such that ~~constant~~ consider $var(\underline{l}'\underline{x}) = \underline{l}'\Sigma\underline{l} = 1$

Again, $[\underline{l}'\underline{\mu}^{(1)} - \underline{l}'\underline{\mu}^{(2)}]^2 = (\underline{l}'\underline{d})^2 = \underline{l}'\underline{d}\,\underline{d}'\underline{l}$

where $\underline{d} = (\underline{\mu}^{(1)} - \underline{\mu}^{(2)})$

$\therefore$ inorder to maximize we consider a f^n ϕ using lagrange's multiplier 'λ' i.e

$$\phi = \underline{l}'\underline{d}\,\underline{d}'\underline{l} - \lambda[\underline{l}'\Sigma\underline{l} - k]$$
$$= \underline{l}'\underline{d}\,\underline{d}'\underline{l} - \lambda\underline{l}'\Sigma\underline{l} - \lambda k$$

differentiating ϕ w.r.t l, we have:

$$\frac{\partial \phi}{\partial l} = 2\, d\, d' l - 2\lambda \Sigma l = 0$$

$\Rightarrow d' d\, l = \lambda \Sigma l$, premultiplying it by l'

we have

$$l' d\, d' l = \lambda l' \Sigma l = \lambda k \quad \Big| \quad \text{Eg } l' \Sigma l = k$$

i.e

$$(l'd)^2 = \lambda k \Rightarrow l'd = \sqrt{\lambda k} = d'l$$

we have

$$d\sqrt{\lambda k} = \lambda \Sigma l \Rightarrow l = \frac{\sqrt{\lambda k}}{\lambda} \Sigma^{-1} d = \sqrt{\frac{k}{\lambda}}\, \Sigma^{-1} d$$

i.e

$$l = \sqrt{\frac{k}{\lambda}}\, \Sigma^{-1} (\mu^{(1)} - \mu^{(2)})$$

So, the best linear discriminatat function is $l'x$ where $l \propto \Sigma^{-1} (\mu^{(1)} - \mu^{(2)})$

If the parameter $\mu^{(1)}$, $\mu^{(2)}$ and Σ are unknown

$$\bar{x}^{(1)} = \frac{1}{N_1} \sum_{\alpha=1}^{N_1} x_\alpha^{(1)}, \alpha \qquad \bar{x}^{(2)} = \frac{1}{N_2} \sum_{\alpha=1}^{N_2} x_\alpha^{(2)}$$

and

$$S = \frac{(N_1 - 1) S_1 + (N_2 - 1) S_2}{(N_1 + N_2 - 2)}$$

$$\text{where } S_i = \frac{1}{(N_i - \lambda)} \sum_{\alpha=1}^{N_i} (x_\alpha^{(i)} - \bar{x}_\alpha^{(i)})(x_\alpha^{(i)} - \bar{x}_\alpha^{(i)})'; \quad i = 1, 2$$

$$R_1 = \underline{X}' \Sigma^{-1} (\bar{\underline{X}}^{(1)} - \bar{\underline{X}}^{(2)})' \geq \frac{1}{2} (\bar{\underline{X}}^{(1)} + \bar{\underline{X}}^{(2)})' \Sigma^{-1} (\bar{\underline{X}}^{(1)} - \bar{\underline{X}}^{(2)})$$

$$R_2 = \underline{X}' \Sigma^{-1} (\bar{\underline{X}}^{(1)} - \bar{\underline{X}}^{(2)})' \ngtr \frac{1}{2} (\bar{\underline{X}}^{(1)} + \bar{\underline{X}}^{(2)})' \Sigma^{-1} (\bar{\underline{X}}^{(1)} - \bar{\underline{X}}^{(2)})$$

→ **Remark :-** when $\Sigma_1 \neq \Sigma_2$ then FISHER suggested to use $(\underline{X} - \mu^{(1)})' \Sigma_1^{-1} (\underline{X} - \mu^{(1)}) - (X - \mu^{(2)})' \Sigma_2^{-1} (\underline{X} - \mu^{(2)})$ as the discriminant f^n it is of a Q.F.
Quadratic ~~form~~ distn function

obiviously it is not a linear discriminant but it is of a quadratic form therefore we shall called it a quadratic discriminant function

* DIFFICULTIES :-

The two region of classification are given by R_1 and R_2 but if in practical situations we get an equality then what to do?

Then it becomes subjective and a difficult problem to classify the random vector X either in π_1 or π_2

Hence instead of two regions we must have a third region also in which the observation may fall. Otherwise we must have some more observation to classify properly the random vector X.

* Canonical / Correlations Analysis (CCA)

Introduction (CCA)

The most commonly used dependence analysis in multivariate is the multiple regression analysis. The technique deals with the study of one variable on a set of predictor variables. But MRA doesn't provide any information about the correlation of the two sets of variables taking into the consideration of the interrelationship among the variables in the ~~wide set~~. Y-Set.

CCA is an alternative multivariate method use to study the correlation of the variables in the Y-set and the X-set taking consideration of the interdependence of the variables in the Y-set.

For Example :- In biological experiments we may be interested in studying the impact of the amount of insecticide and its concentration on the number of dead insects and on the no. of days required to kill the insects.

Criterion sets include the no. of dead insects and the no. of days required to kill the insect. The predictor set includes the amount of insecticide and its concentration. The frame set is usually denoted by Y-set and the later set is denoted by X-set

The MCA deals with one variable at a time from Y set.

However CCA is appropriate when the criterion variable themselves are correlated. The main purpose of CCA is to study the structural relationship of two sets of variable simultaneously.

Canonical Variables and Canonical Correlation

Let $\underline{X}_{p\times 1}$ be a random vector with $\underline{\mu}$ and Σ as the parameter of the pop^n under consideration. Let $\underline{X}$ be partitioned into two subset $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ having p_1 and p_2 components respecting $(p_1 < p_2) \rightarrow (?)$
the Σ is also partitioned as:

$$\Sigma = \left[\begin{array}{c|c} \Sigma_{11} & \Sigma_{12} \\ \hline \Sigma_{21} & \Sigma_{22} \end{array}\right].$$

Let $\underline{\alpha}$ and $\underline{\gamma}$ be any two $(p_1 \times 1)$ and $(p_2 \times 1)$ vectors and let $U = \underline{\alpha}' X^{(1)}$ and $V = \underline{\gamma}' X^{(2)}$ be two linear combination of $\underline{X}^{(1)}$ and $\underline{X}^{(2)}$ respectively. We have

$$\mathrm{Var}(U) = \mathrm{Var}(\underline{\alpha}' \underline{X}^{(1)}) = \underline{\alpha}' \Sigma_{11} \underline{\alpha} \text{ and}$$
$$\mathrm{Var}(V) = \mathrm{Var}(\underline{\gamma}' \underline{X}^{(2)}) = \gamma' \Sigma_{22} \gamma \text{ and}$$
$$\mathrm{cov}(U, V) = \underline{\alpha}' \Sigma_{12} \gamma.$$

We wish to choose $\underline{\alpha}$ and $\underline{\gamma}$ in such a way that the correlation between U and V is maximum such that consider that their variance are constant.

We may normalize the linear combination so that $var(U) = var(V) = 1$

i.e $\underline{\alpha}' \Sigma_{11} \underline{\alpha} = \underline{\gamma}' \Sigma_{22} \underline{\gamma} = 1$ —— (1)

Now, the correlation between U and V is

$$\rho_{UV} = \frac{Cov(U, V)}{\sqrt{var(U)} \cdot \sqrt{var(V)}} = \frac{\underline{\alpha}' \Sigma_{12} \underline{\gamma}}{\sqrt{\underline{\alpha}' \Sigma_{11} \underline{\alpha}} \sqrt{\underline{\gamma}' \Sigma_{22} \underline{\gamma}}}$$

$$= \underline{\alpha}' \Sigma_{12} \underline{\gamma} \quad \text{if (1) holds}$$

i.e $\rho_{UV} = \underline{\alpha}' \Sigma_{12} \underline{\gamma}$ —— (2)

Obviously maximizing ρ_{UV} is equivalent to maximizing $\underline{\alpha}' \Sigma_{22} \underline{\gamma}$ such that consider (1)
For this consider the function

$$\phi = \underline{\alpha}' \Sigma_{12} \underline{\gamma} - \frac{\lambda_1}{2} (\underline{\alpha}' \Sigma_{11} \underline{\alpha} - 1) - \frac{\lambda_2}{2} (\underline{\gamma}' \Sigma_{22} \underline{\gamma} - 1)$$

where λ_1 and λ_2 are the ~~log~~ lagrange's multiplier. Now, differentiating ϕ w.r.t $\underline{\alpha}$ and $\underline{\gamma}$ respectively.
we have

$$\frac{\partial \phi}{\partial \underline{\alpha}} = \Sigma_{12} \underline{\gamma} - \frac{\lambda_1}{2} 2 \Sigma_{11} \alpha = 0$$

$$\Rightarrow \Sigma_{12} \underline{\gamma} - \lambda_1 \Sigma_{11} \underline{\alpha} \quad \text{—— (3)}$$

$$\frac{d\phi}{d\underline{\gamma}} = \underline{\alpha}'\Sigma_{12} - \frac{\lambda_2}{2}\cdot 2\,\underline{\gamma}'\Sigma_{22} = 0$$

$\Rightarrow \underline{\alpha}'\Sigma_{12} - \lambda_2 \underline{\gamma}'\Sigma_{22} = 0$ or.

$\Rightarrow \Sigma_{21}\underline{\alpha} - \lambda_2 \Sigma_{22}\underline{\gamma} = 0$ —— (4)

(transforming the equation) (transposing the eqn)

premultiplying (3) by $\underline{\alpha}$ and (4) by $\underline{\gamma}'$ we have.

$\underline{\alpha}'\Sigma_{12}\underline{\gamma} - \lambda_1 \underline{\alpha}'\Sigma_{11}\underline{\alpha} = 0$ i.e $\underline{\alpha}'\Sigma_{12}\underline{\gamma} = \lambda_1$ as $\underline{\alpha}'\Sigma_{11}\underline{\alpha} = 1$

and $\underline{\gamma}'\Sigma_{21}\underline{\alpha} - \lambda_2 \underline{\gamma}'\Sigma_{22}\underline{\gamma} = 0$ i.e $\underline{\gamma}'\Sigma_{21}\underline{\alpha} = \lambda_2$ $\quad \underline{\gamma}'\Sigma_{22}\underline{\gamma} = 1$

as $\underline{\alpha}'\Sigma_{12}\underline{\gamma}$ and $\underline{\gamma}'\Sigma_{21}\underline{\alpha}$ are same being the correlation between U & V

$\therefore \lambda_1 = \lambda_2 = \rho_{UV} = \rho$ (say)

Again from (3) and (4) we have.

$$\Sigma_{12}\underline{\gamma} - \rho\Sigma_{11}\underline{\alpha} = 0 \quad \text{—— (5)}$$
$$\Sigma_{21}\underline{\alpha} - \rho\Sigma_{22}\underline{\gamma} = 0 \quad \text{—— (6)}$$

Or, $-\rho\Sigma_{11}\underline{\alpha} + \Sigma_{12}\underline{\gamma} = 0$

$\Sigma_{21}\underline{\alpha} - \rho\Sigma_{22}\underline{\gamma} = 0$

which can be written in matrix form as:

$$\begin{bmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{bmatrix} \begin{bmatrix} \underline{\alpha}' \\ \underline{\gamma} \end{bmatrix} = 0 \quad \text{—— (7)}$$

Since, $\underline{\alpha}$ and $\underline{\gamma}$ are NON-zero vectors

$\therefore$ for a non-zero solution of (7) will exist only if the coefficient matrix is singular i.e

$$\begin{vmatrix} -\rho\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & -\rho\Sigma_{22} \end{vmatrix} \quad \text{—— (8)}$$

Since, the LHS of (8) is a polynomial of order 'p' so there will be 'p' roots of ρ.

Again, premultiplying (5) by $\Sigma_{21}\Sigma_{11}^{-1}$ and (6) by ρ we get

$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\underline{\gamma} - \rho\,\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{11}\underline{\alpha} = 0$ OR
$\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\underline{\gamma} - \rho\,\Sigma_{21}\underline{\alpha} = 0$

$\rho\,\Sigma_{21}\underline{\alpha} - \rho^2\Sigma_{22}\underline{\gamma} = 0$ OR $\rho\Sigma_{21}\underline{\alpha} - \rho^2\Sigma_{22}\underline{\gamma} = 0$

Adding both we get
$(\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\underline{\gamma} - \rho^2\Sigma_{22})\,\underline{\gamma} = 0$ —— (9)

Again, $\underline{\gamma}$ is a non-zero vector, so we must have $|\Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} - \rho^2\Sigma_{22}| = 0$ —— (10)

Since, the LSH of (10) is a polynomial of order p_2 in ρ^2, so it gives p_2 roots of ρ^2 (say) $\rho_1^2, \rho_2^2, \ldots \rho_{p_2}^2$.

Now premultiplying (5) by ρ and (6) by $\Sigma_{12}\Sigma_{22}^{-1}$ we have (get)

$$\left.\begin{aligned}\rho\,\Sigma_{12}\underline{\gamma} - \rho^2\Sigma_{11}\underline{\alpha} &= 0\\ \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\underline{\alpha} - \rho\,\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}\underline{\gamma} &= 0\end{aligned}\right\}\text{ Adding these}$$

$(\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} - \rho^2\Sigma_{11})\,\underline{\alpha} = 0$ —— (11) Again, $\alpha \neq 0$ $\therefore$
$|\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22}| - \rho^2|\Sigma_{11}| = 0$ —— (12)
↳ which is a polynomial of order P_1 in ρ^2 which gives P_1 roots as : $\rho_1^2, \rho_2^2 \ldots \rho_{p_1}^2$.

~~Again~~

As the equations (9) and (11) have been obtained from (7) hence there non-zero roots should be equal and there must be $(p_2 - p_1)$ zero roots, So the number of non-zero roots of ρ is:

$$p - (p_2 - p_1) = (p_1 + p_2) - (p_2 - p_1) = 2p_1$$

$\therefore$ ρ^2 has p_1 non-zero roots and ρ has $\boxed{2p_1}$ non-zero roots of the form $= \rho = \pm \rho_i \; (i = 1, 2, .., p_1)$

now, the positive roots are arranged as:

$\rho(1) \geq \rho(2) \geq \cdots \geq \rho(p_1)$.

Here $\rho_{(1)}$ is called the FIRST CANONICAL CORR and substituting $\rho = \rho_{(1)}$ in (9) and (11) and solving them for α and γ

Let $\alpha^{(1)}$ and $\gamma^{(1)}$ be the solution then

$$\boxed{U_1 = \alpha^{(1)'} x^{(1)}} \quad \text{and} \quad \boxed{V_1 = \gamma^{(1)'} x^{(2)}}$$

are the 1^{st} Canonical variables

Similarly $\rho_{(2)}$ representing the SECOND CANONICAL CORRELATION and

$$\boxed{U_2 = \alpha^{(2)'} x^{(1)}} \quad \text{and} \quad \boxed{V_2 = \gamma^{(2)'} x^{(2)}}$$

are the 2^{nd} Canonical variables

(*) CCA Properties :-

(1) For different values of cannonical correlation we can obtain the vector $\underline{\alpha}^{(1)} \ldots \alpha^{(P_1)}$ and $\underline{\gamma}^{(1)} \gamma^{(2)} \ldots \gamma^{(P_2)}$ etc. define the matrices

$$A = [\underline{\alpha}^{(1)} \ldots \alpha^{(P_1)}] \quad \text{and} \quad \Gamma = [\gamma^{(1)} \gamma^{(2)} \ldots \gamma^{(P_2)}]$$

then we have

$$A' \Sigma_{11} A = I_{p_1},$$

$$A' \Sigma_{12} \Sigma_{22}^{-1} \cdot \Sigma_{21} A = R^* = \log(\rho_1^2, \rho_2^2, \ldots, \rho_{P_1}^2)$$

and

$$\Gamma' \Sigma_{22} \Gamma = I_{p_2}$$

$$\Gamma' \Sigma_{12} \Sigma_{11}^{-1} \Sigma_{21} \Gamma = \text{diag}(\rho_1^2, \rho_2^2, \ldots \rho_{P_2}^2)$$

(2) The number of non-zero roots of

$$|\Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} - \rho^2 \Sigma_{11}| = 0 \quad \text{or.}$$

$|\Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12} - \rho^2 \Sigma_{22}| = 0$ is equal to the roots of Σ_{12}

(*) Method of TEST :-

The principle of CCA is to find out ~~to~~ two linear combination $\alpha' X^{(1)}$ of the variables in $X^{(1)}$ and $X^{(2)}$ so that the linear combinations are highly correlated.

this is possible only if the variables in $X^{(1)}$ and $X^{(2)}$.

However, the CCA will not be useful if the variables in the two subsets are not related. Therefore, before performing CCA we should test

$$H_0 : \Sigma_{12} = 0$$
$$\text{ag.} \quad H_1 : \Sigma_{12} \neq 0$$

Bartlett (1951) proposed a criterian which was further developed by WILK so it is known as WILK'S criterion, and the proposed test statistics is

$$\Delta = \frac{|S_{22}^{-1} S_{21} S_{11}^{-1} S_{12}|}{|S_{22}|}$$

it is present that $\Delta \sim \chi^2_{p_1 p_2, \alpha}$

where $\chi^2 = -[(n-1) - \frac{1}{2}(p_1 + p_2 + 1)] \log \Delta$

If $\chi^2_{cal} > \chi^2_{p_1 * p_2, \alpha}$, Ho is rejected
$\therefore$ Proceed for CCA.

⊛ BEST REGRESSION EQUATION :-

We have $\underset{p\times 1}{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}$

Assume $E(\underline{x}) = 0$

Consider the partitioning of $\underline{X}_{p\times 1}$ as

$\underline{x}_{p\times 1} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$. accordingly

$$\Sigma = \begin{bmatrix} Var(x_1) & cov(x_1, \underline{x}^{(2)}) \\ cov(\underline{x}^{(2)}, x_1) & var(\underline{x}^{(2)}) \end{bmatrix}$$

where $\underline{X}^{(2)}_{(p-1)\times 1}$

denote $Var(x_1) = \sigma_{11}$ (say)

$cov(x_1, \underline{x}^{(2)}) = \underline{\sigma}$

$Var(\underline{x}^{(2)}) = \Sigma_{22}$

Now, let 'predicted' value of x_1 be $\underline{\beta}' \underline{x}^{(2)}$

where $\underline{\beta} = \begin{bmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{bmatrix}$ is the $(p-1)\times 1$ vector.

we wish to obtain $\underline{\beta}$ in such a way that it minimizes the residual sum of square let it be denoted by Q. Then

$$Q = E[x_1 - \underline{\beta}' \underline{x}^{(2)}]^2$$

$$= E[(x_1 - \underline{\beta}' \underline{x}^{(2)})][(x_1 - \underline{\beta}' \underline{x}^{(2)})]'$$

Or

$$Q = E[x_1^2 - x_1 \underline{x}^{(2)} \beta - \beta' \underline{x}^{(2)} x_1 + \beta' \underline{x}^{(2)} x^{(2)'} \beta]$$

$$= \sigma_{11} - 2\beta' \sigma + \beta' \Sigma_{22} \beta$$

Now differentiating Q with respect to β

$$\frac{\partial Q}{\partial \underline{\beta}} = -2\underline{\sigma} + 2\beta \Sigma_{22} = 0$$

$$\Rightarrow \underline{\beta} \Sigma_{22} = \underline{\sigma} \quad \Rightarrow \boxed{\beta = \Sigma_{22}^{-1} \underline{\sigma}}$$

So the predicted value of x_1 called Regression is given by

$$\boxed{\hat{x}_1 = \underline{\beta}\, \underline{x}^{(2)} = \underline{\sigma}' \Sigma_{22}^{-1} \underline{x}^{(2)}}$$

↳ Regr of x_1 on $(x_2, x_3, \ldots x_p)$

$\therefore [x_1 - \underline{\sigma}' \Sigma_{22}^{-1} x^{(2)}]$ is the residual denoted by $x_{1.23\cdots p}$.

Note that $\operatorname{cov}[x_{1.23\cdots p}, x^{(2)}] = 0$

i.e residual an $x^{(2)}$ are uncorrelated

STA 2203 C03

M T W T F S S

Page No.: 1

Date: 21/12/22

YOUVA

Multivariate Analysis - II

Theorem If an m-component vector $Y \sim N_m(0, T)$ where 'T' is a Non Singular Matrix then $Y' T^{-1} Y \sim \chi^2_m$ d.f

Proof Since T is NSM we can find another Matrix C such that: $CTC' = I$

Now consider the Transformation

$$\underline{Z} = C\underline{Y}$$

then

$$Z \sim N_m(0, CTC') \quad \text{or} \quad Z \sim N_m(0, I)$$

the components of Z are iid $N_1(0, 1)$

Now

$$Y' T^{-1} Y = (C^{-1}Z)' T^{-1} (C^{-1}Z)$$

$$= Z' (C' T C)^{-1} Z = Z' I Z$$

$$= \underline{Z' Z} = \sum_{i=1}^{m} Z_i^2$$

Since Z_i^s are iid $N(0,1)$

$$\underline{Y' T^{-1} Y} = \sum_{i=1}^{m} Z_i^s \sim \chi^2_m$$

<u>Application</u>:- Them is used in testing $H_0: \mu = \mu_0$ when Σ is <u>known</u>. As we know that $\bar{X} \sim N_p\left(\mu, \frac{\Sigma}{N}\right) \Rightarrow (\bar{x} - \mu) \sim N_p\left(0, \frac{\Sigma}{N}\right)$

$$\Rightarrow \sqrt{N}(\bar{x} - \mu) \sim N_p(0, \Sigma)$$

Or.

$$\underbrace{N(\bar{x} - \mu)' \Sigma^{-1} (\bar{x} - \mu)}_{\text{calculated}} \sim \underbrace{\chi^2_{p;\alpha\%}}_{\text{tabulated}}$$

So reject H_0 if $\chi^2_{cal} > \chi^2_{tab}$

In univariate case for testing the hypothesis $H_0: \mu = \mu_0$ when σ^2 is unknown we use

$$t = \frac{(\bar{x} - \mu)}{s/\sqrt{n}} \sim t_{n-1}$$

Now $t^2 = n(\bar{x}-\mu)'(s^{-1})(\bar{x}-\mu)' \sim f_{1,n-1}$
on the same analogy Hotelling derived a test for one sample / two sample as:

let $x_1, \ldots x_N \sim N_p(\mu, \Sigma)$, Σ unknown,
if we wish to test $H_0: \mu = \mu_0$ (specified)
he proposed : $T^2 = N(\bar{x}-\mu_0)' S^{-1} (\bar{x}-\mu_0)$
where $S = \frac{1}{(N-1)} \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})'$

* **Likelihood Ratio test for the Mean** (Σ unknown)

We wish to test $H_0: \mu = \mu_0$, Σ being unknown
let $X_1, X_2, \ldots X_N$ be a r.s of size N from $N_p(\mu, \Sigma)$ then its likelihood f^n will be:

$$L = \frac{1}{(2\pi)^{\frac{Np}{2}} |\Sigma|^{N/2}} e^{-\frac{1}{2}\sum_{\alpha=1}^{N}(x_\alpha-\mu)'\Sigma^{-1}(x_\alpha-\mu)} \quad \text{---(1)}$$

we define the LRT $\lambda = \dfrac{\max_{H_0} L}{\max_{H_1} L}$

The L.F will be max if the pmf are replaced by their MLE^s, we have $\hat{\mu} = \bar{x}$, $\hat{\Sigma} = \hat{\Sigma}_1 = A/N$
where
$$A = \sum_{\alpha=1}^{N} (x_\alpha - \bar{x})(x_\alpha - \bar{x})' \longrightarrow \text{under } H_1$$
and
$$\mu = \mu_0, \hat{\Sigma} = \hat{\Sigma}_0 = \frac{1}{N}\sum_{\alpha=1}^{N}(x_\alpha - \mu_0)(x_\alpha - \mu_0)'$$
$\hookrightarrow$ under H_0

Under H_0 we have

$$\max_{H_0} L = \frac{1}{(2\pi)^{\frac{Np}{2}} |\hat{\Sigma}_0|^{N/2}} e^{-\frac{1}{2}\sum_{\alpha=1}^{N}(x_\alpha-\mu_0)\hat{\Sigma}_0^{-1}(x_\alpha-\mu_0)}$$

Consider the Exponent : $\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)' \hat{\Sigma}_0^{-1} (\underline{x}_\alpha - \underline{\mu}_0)$ $\rightarrow$ 1×1 scalar

$$\therefore \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)' \hat{\Sigma}_0^{-1} (\underline{x}_\alpha - \underline{\mu}_0) = tr \cdot \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)' \hat{\Sigma}_0^{-1} (\underline{x}_\alpha - \underline{\mu}_0)$$

$$= \sum_{\alpha=1}^{N} tr\, (\underline{x}_\alpha - \underline{\mu}_0)' \hat{\Sigma}_0^{-1} (\underline{x}_\alpha - \underline{\mu}_0)$$

$$= \sum_{\alpha=1}^{N} tr\, \hat{\Sigma}_0^{-1} (\underline{x}_\alpha - \underline{\mu}_0)' (\underline{x}_\alpha - \underline{\mu}_0)$$

$$= tr\, \hat{\Sigma}_0^{-1} \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)' (\underline{x}_\alpha - \underline{\mu}_0)$$

$$= tr \cdot \hat{\Sigma}_0^{-1} (N) \hat{\Sigma}_0$$

$$= N\, tr\, I_p$$

$$= Np$$

$$\therefore \max_{H_0} L = \frac{1}{(2\pi)^{\frac{Np}{2}} |\hat{\Sigma}_0|^{N/2}} e^{-\frac{1}{2} Np} \quad \text{--- (2)}$$

Now

$$\max_{H_1} L = \frac{1}{(2\pi)^{\frac{Np}{2}} |\hat{\Sigma}_1|^{N/2}} e^{\frac{1}{2} \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}})' \hat{\Sigma}_1^{-1} (\underline{x}_\alpha - \bar{\underline{x}})} \quad \text{--- (3)}$$

Again it can be proved that the exponent of (3) is $-\frac{Np}{2}$

$\therefore$ Using (2) and (3)

$$\lambda = \frac{|\hat{\Sigma}_1|^{N/2}}{|\hat{\Sigma}_0|^{N/2}} = \frac{|A|^{\frac{N}{2}}}{\left|\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)(\underline{x}_\alpha - \underline{\mu}_0)'\right|^{N/2}}$$

Now $\sum_{\alpha=1}^{N} (\underline{x}_\alpha - \underline{\mu}_0)(\underline{x}_\alpha - \underline{\mu}_0)'$

$$= \sum_{\alpha=1}^{N} (\underline{x}_\alpha - \bar{\underline{x}} + \bar{\underline{x}} - \underline{\mu}_0)(\underline{x}_\alpha - \bar{\underline{x}} + \bar{\underline{x}} - \underline{\mu}_0)'$$

$$= \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})' + \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(\bar{X} - \mu_0)' +$$

$$\sum_{\alpha=1}^{N} (\bar{X} - \mu_0)(\bar{X}_\alpha - \bar{X})' + \sum_{\alpha=1}^{N} (\bar{X} - \mu_0)(\bar{X} - \mu_0)'$$

$$= \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})'(X_\alpha - \bar{X}) + \sum_{\alpha=1}^{N} (\bar{X} - \mu_0)(\bar{X} - \mu_0)'$$

$$= A + N(\bar{X} - \mu_0)(\bar{X} - \mu_0)'$$

$$\therefore \lambda^{2/N} = \frac{|A|}{|A + N(\bar{X} - \mu_0)(\bar{X} - \mu_0)'|}$$

Consider the denominator

$$|A + N(\bar{X} - \mu_0)(\bar{X} - \mu_0)| = \begin{vmatrix} A & \sqrt{N}(\bar{X} - \mu_0) \\ -\sqrt{N}(\bar{X} - \mu_0)' & 1 \end{vmatrix} \quad \text{---} (*)$$

Now using the result: if $|B| \neq 0$ and E is a Square matrix then $\begin{vmatrix} B & C \\ D & E \end{vmatrix} = |B|\,|E - DB^{-1}C|$

applying it to $(*)$ we have

$$\begin{vmatrix} A & \sqrt{N}(\bar{X} - \mu_0) \\ -\sqrt{N}(\bar{X} - \mu_0)' & 1 \end{vmatrix} = |A|\,|1 + N(\bar{X} - \mu_0)' A^{-1} (\bar{X} - \mu_0)|$$

$$\lambda^{2/N} = \frac{|A|}{|A|\,|1 + N(\bar{X} - \mu_0)' A^{-1} (\bar{X} - \mu_0)} = \frac{1}{|1 + N(\bar{X} - \mu_0)' A^{-1} (\bar{X} - \mu_0)|}$$

Again $A = (N-1)S \Rightarrow A^{-1} = \dfrac{S^{-1}}{(N-1)}$

Let $T^2 = N(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0)$
then

$$\lambda^{2/N} = \frac{1}{1 + N(\bar{X} - \mu_0) \frac{S^{-1}}{(N-1)} (\bar{X} - \mu_0)} = \frac{1}{1 + \frac{T^2}{(N-1)}}$$

$$\frac{1}{1+\frac{T^2}{(N-1)}} \leq \lambda_0^{2/N} \Rightarrow 1+\frac{T^2}{(N-1)} \geq \lambda_0^{-2/N}$$

The LRT is to Reject H_0 if $\lambda \leq \lambda_0$ (specified)

OR

$$\lambda^{2/N} \leq \lambda_0^{2/N} \quad \text{OR} \quad \frac{1}{1+\frac{T^2}{(N-1)}} \leq \lambda_0^{2/N}$$

i.e $T^2 \geq [\lambda_0^{-2/N} - 1](N-1) = T_0^2$ (say) is a another constant to be determined in such a way that

$$P_r[T^2 \geq T_0 \mid H_0] = \alpha \text{ (size of the test)}$$

Hence, for a desired size α test to obtain T_0^2 we must know that the dist^n of T^2

* **Distribution of T^2**

In general Hotling T^2 is defined as

$$T^2 = \underline{Y}'S^{-1}\underline{Y} \quad \text{where} \quad \underline{Y} \sim N_p(\underline{\gamma}, \Sigma)$$

and

$$S = \frac{1}{(N-1)} \sum_{\alpha=1}^{N} \underline{Z_\alpha}\underline{Z_\alpha}' \; ; \; Z_\alpha \sim N_p(\underline{0}, \Sigma) \quad \text{and}$$

are independent for $\alpha = 1, 2, \ldots (N-1)$ i.e

$$(N-1)\, S = \sum_{\alpha=1}^{N} Z_\alpha \underline{Z_\alpha}'$$

Again,

$$T^2 = N(\overline{\underline{X}} - \mu_0)'S^{-1}(\overline{X} - \mu_0) \;\Big|\; \text{under } H_0$$

let D be a NSM such that $D\Sigma D = I$ —(1)

Define:

$$\left.\begin{aligned} Y^* &= DY \\ V^* &= D\underline{\gamma} \\ S^* &= DSD' \end{aligned}\right\} \text{—(2)}$$

then $Y^* \sim N_p(D\underline{\gamma}, D\Sigma D')$ or $\underline{Y^*} \sim N_p(V^*, I)$ —(3)

Again in general : $T^2 = \underline{Y}' S^{-1} \underline{Y}$

$$= (D^{-1}Y^*)' S^{-1} (D^{-1}Y^*)$$
$$= \underline{Y}^* (DSD')^{-1} \underline{Y}^*$$

OR $$T^2 = \underline{Y}^* (S^*)^{-1} \underline{Y}^* \quad - (4)$$

Also $\underline{\delta}' \Sigma^{-1} \underline{\delta} = (D^{-1}\underline{V}^*)' \Sigma^{-1} (D^{-1}V^*)$

$$= (\underline{V}^*)' (D\Sigma D')^{-1} \underline{V}^*$$
$$= \underline{V}^{*\prime} I \underline{V}^* = \underline{V}^{*\prime} \underline{V}^* \quad - (5)$$

We see that $(N-1) S^* \sim \sum_{\alpha=1}^{N-1} \underline{Z_\alpha^*} \underline{Z_\alpha^*}'$

where $Z_\alpha^* \sim D Z_\alpha \sim N_p(0, I)$ as $D \Sigma D' = I$
and each $\underline{Z_\alpha^*}$ is indep for $\alpha = 1, 2, \ldots N$

Consider an orthogonal matrix $Q_{p \times p}$ s.t its 1st row is given by the element

$$q_{1j} = \frac{Y_j^*}{\sqrt{\underline{Y}^{*\prime} \underline{Y}^*}}$$

for $j = 1, 2, \ldots p$ where Y_j^* is j^{th} component of $\underline{Y}^* = \begin{pmatrix} Y_1^* \\ \vdots \\ Y_p^* \end{pmatrix}$

now if we can show that $\sum q_{1j}^2 = 1$ then we shall be able to define such a matrix we have

$$\sum_{j=1}^{p} q_{1j}^2 = \frac{\sum_{j=1}^{p} Y_j^{*2}}{\underline{Y}^{*\prime}\underline{Y}^*} = \frac{\sum_{j=1}^{p} Y_j^{*\prime} Y_j^*}{\underline{Y}^{*\prime}\underline{Y}^*} = \frac{\underline{Y}^{*\prime}\underline{Y}^*}{\underline{Y}^{*\prime}\underline{Y}^*} = 1$$

Since $\sum_{j=1}^{p} q_{1j}^2 = 1$

Hence the choice ~~of the matrix~~ of the orthogonal matrix with q_{1j} as the elements in the 1st row is permissible

Now, we consider another transformation

$$U_{p\times 1} = Q\underline{Y}^* = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_p \end{pmatrix}$$

where $U_\alpha = \sum_{j=1}^{p} q_{\alpha j}\, Y_j^*$

Again, Since Q is an orthogonal matrix so

$$\sum_{j=1}^{p} q_{1j}\, q_{\alpha j} = 0 \quad \text{if } \alpha \neq 1$$
$$= 1 \quad \text{if } \alpha = 1$$

Thus, $\dfrac{\sum_{j=1}^{p} Y_j^*}{\sqrt{Y^{*\prime} Y^*}} \cdot q_{\alpha j} = 0 \quad$ if $\alpha \neq 1$

$\Rightarrow \sum_{j=1}^{p} Y_j^* q_{\alpha j} = 0 \quad$ if $\alpha \neq 1$

$\Rightarrow U_\alpha = 0 \quad$ if $\alpha \neq 1$ —— (6)

$U_{p\times 1} = \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_p \end{pmatrix}$ Again we have

$$U_1 = \sum_{j=1}^{p} q_{1j}\, Y_j^*$$

$$= \frac{\sum_{j=1}^{p} Y_j^{*\prime} Y_j^*}{\sqrt{Y^{*\prime} \underline{Y}^*}} = \frac{Y^{*\prime} Y^*}{\sqrt{Y^{*\prime} Y^*}}$$

So $U_1 = \sqrt{\underline{Y}^{*\prime} \underline{Y}^*}$ —— (7)

$$\therefore U = Q\underline{Y}^* = \begin{pmatrix} \sqrt{Y^{*\prime} Y^*} \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Again we have $T^2 = Y^{*\prime} S^{*\prime} \underline{Y}^*$

$$= (Q^{-1}\underline{U})' S^{*\prime} (Q^{-1}\underline{U})$$
$$= \underline{U}'(Q^{-1})' S^* (Q^{-1}) \underline{U}$$

OR $\quad T^2 = \underline{U}'(QS^*Q')^{-1}\underline{U}$

Also $(N-1) S^* = A^* \Rightarrow S^* = \dfrac{A^*}{(N-1)}$

$$T^2 = \underline{U}' \left\{ Q \frac{A^*}{(N-1)} Q' \right\}^{-1} \underline{U}$$

OR $\dfrac{T^2}{(N-1)} = \underline{U}' B^{-1} \underline{U}$ — (8)

where $B = Q A^* Q'$

$$\therefore \frac{T^2}{(N-1)} = \underline{U}' B^{-1} \underline{U}$$

Now let $B = \begin{bmatrix} b_{11} & \underline{b}_{(1)} \\ b_{(1)}' & B_{22} \end{bmatrix}$, $B^{-1} = \begin{bmatrix} b^{11} & \underline{b}^{(1)} \\ b^{(1)'} & B^{22} \end{bmatrix}$

As $BB^{-1} = I$ we have

$$B = \begin{array}{c} \qquad\qquad b^{(1)} \\ \left[\begin{array}{c|ccc} b_{11} & b_{12} & \cdots & b_{1p} \\ \hline \vdots & & & \\ b_{(1)}' & & & \\ b_{p1} & b_{p2} & \cdots & b_{pp} \end{array}\right] \end{array} B_{22}\ (p-1)\times(p-1)$$

$$I = \begin{array}{c} (\text{Only } 1) \qquad \text{Null} \\ \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 1 & 0 & \cdots & \\ \vdots & 0 & 1 & \cdots & \\ 0 & \vdots & \vdots & \cdots & 1 \end{array}\right] \\ \text{Matrix} \end{array}$$

$$b_{11} b^{11} + \underline{b}_{(1)} \underline{b}^{(1)} = 1 \rightarrow (A)$$
$$b_{11} \underline{b}^{(1)} + \underline{b}_{(1)} B^{22} = 0 \rightarrow (B)$$
$$\underline{b}_{(1)}' b^{11} + B_{22} \underline{b}^{(1)'} = 0 \rightarrow (C)$$
$$\underline{b}_{(1)}' \underline{b}^{(1)} + B_{22} B^{22} = I_{p-1} \rightarrow (D)$$

Using C we have

$$\cancel{\underline{b}^{(1)'} = -B_{22}^{-1} \underline{b}_{(1)} b^{11}}$$

$$\underline{b}^{(1)'} = -B_{22}^{-1} \underline{b}_{(1)}' b^{11}$$

Subsitituting this in (A) we have

$$b_{11} b^{11} - b_{(1)} B_{22}^{-1} b'_{(1)} b^{11} = 1$$

OR $$[b_{11} - b_{(1)} B_{22}^{-1} b'_{(1)}] b^{11} = 1$$

$$\Rightarrow \boxed{b_{11.2} = \frac{1}{b^{11}}}$$

where $b_{11.2} = b_{11} - b_{(1)} B_{22}^{-1} b'_{(1)}$

$$\frac{T^2}{(N-1)} = (U_1\ 0 \cdots 0) \begin{bmatrix} b^{11} & b^{(1)} \\ b^{(1)'} & B^{22} \end{bmatrix} \begin{pmatrix} U_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= U_1 b^{11} U_1$$
$$= U_1^2 b^{11}$$

Again $$\frac{T^2}{(N-1)} = \underline{U}' B^{-1} \underline{U}$$

$$= \begin{pmatrix} U_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} B^{-1} (U_1\ 0 \cdots 0)$$

OR $$\frac{T^2}{N-1} = \frac{Y^{*'} Y^*}{b_{11.2}} \quad -(9)$$

as $b_{11.2} = 1/b^{11}$

As $Y^* \sim N_p(\gamma', I)$ and it's components are independent ~~$\alpha = 1, 2$~~ also if $\gamma^* \neq 0$ then
$\underline{Y}^{*'} Y^* \sim \chi^2_{(\delta, p)}$

where δ is Non central Parameter and
$\delta = \underline{\gamma}^{*'} \underline{\gamma}^*$

Thus the numerator of (9) has $\chi^2_{(\delta, p)}$ dist^n
Now we wish to obtain dist^n of denominator

for that we have $A^* = \sum_{\alpha=1}^{N-1} Z_\alpha^{**} Z_\alpha^{*'}$

where each $Z_\alpha^* \sim N_p(0, I)$ and are independent for $\alpha = 1, \ldots (N-1)$

matrix Q containing elements of Y^* is a random matrix

$\therefore B = QA^*Q'$ also become random

Hence we consider the conditional of B for given Q (i.e Y^*) conditionally.

$$B \sim \sum_{\alpha=1}^{N-1} Q Z_\alpha^* Z_\alpha^{*'} Q' = \sum_{\alpha=1}^{N-1} U_\alpha U_\alpha' \text{ (say)}$$

where $U_\alpha = Q Z_\alpha^*$ and $U_\alpha \sim N_p(0, I)$
as $QQ' = I$

and are independent

Consider $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \sum_{\alpha=1}^{N-1} Z_\alpha Z_\alpha'$

$Z_\alpha \sim N_p(0, \Sigma)$

then $A_{11.2} = A_{11} - A_{12} A_{22}^{-1} A_{21}$

$$= \sum_{\alpha=1}^{(N-1)-(p-q)} W_\alpha W_\alpha'$$

CONDITIONAL DISTN OR WISHART

$W_\alpha \sim N_p(0, \Sigma_{11.2})$ are independent $\alpha = 1, 2 \ldots (N-1)$

applying this result the conditional dist^n of $W_{11.2}$ is $\sum_{\alpha=1}^{(N-1)(p-q)} W_\alpha W_\alpha'$ $\quad W_\alpha \sim N_1(0, \Sigma)$

and are independent for $\alpha = 1, 2 \ldots (N-)$

we have

$$I_p = \begin{bmatrix} 1 & 0 \\ 0 & I_{p-1} \end{bmatrix}$$

$$\therefore I_{11.2} = 1 - I_{(1)} I_{p-1}^{-1} I_{(1)}' = 1$$

$\therefore W_\alpha \sim N(0,1)$

So the conditional distribution of

$$b_{11.2} = \sum_{\alpha=1}^{(N-P)} W_\alpha^2$$

where W_α^2 are Standard Normal variables and are independent $\alpha = 1, 2, \ldots (N-1)$ $\chi^2_{(N-1)}$ df which implies the conditional distribution of $\frac{T^2}{(N-1)}$ given Q (or Y^*) is same as that of

$\frac{\chi^2_{(\delta, p)}}{\chi^2_{(N-P)}}$ as this does not involve Y^* or Q matrix, so this become the unconditional dist^n of $\frac{T^2}{(N-1)}$ also

Thus $$\frac{T^2}{(N-1)} \sim \frac{\chi^2_{(\delta, p)}}{\chi^2_{(N-P)}}$$

OR

$$\boxed{\frac{T^2}{(N-1)} \cdot \frac{(N-P)}{P} \sim F_{p, (N-P), \delta}}$$

$\uparrow$ T^2 - distribution

Here $\delta = \gamma' \Sigma^{-1} \gamma$

if $\delta = 0$ T^2 becomes central distribution

Applications Of T^2

(1) **One Sample Problem:-**

Suppose we are interested in testing the Hypothesis $H_0: \mu = \mu_0$; Σ being unknown

Let $x_1, x_2, \ldots x_N$ be a random sample of size N from $N_p(\mu, \Sigma)$

Then the proposed Test statistics is
$$T^2 = N(\bar{X} - \mu_0)' S^{-1} (\bar{X} - \mu_0)$$

where $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$ and $S = \frac{1}{(N-1)} \sum_{\alpha=1}^{N-1} (X_\alpha - \bar{X})(X_\alpha - \bar{X})'$

So,
$$\frac{T^2}{(N-1)} \cdot \frac{(N-p)}{p} \sim F_{p, N-p; \alpha\%}$$

and ∴ the critical region is : $F_{cal} > F_{p,(N-p); \alpha\%}$

(2) <u>Two Sample Problem</u> $(H_0 : \mu^{(1)} = \mu^{(2)})$

Suppose that we wish to test $\mu_1^{(1)} = \mu_2^{(2)}$ from 2 populations $N_p(\mu^{(1)}, \Sigma_1)$ & $N_p(\mu^{(2)}, \Sigma_2)$ as $\Sigma_1 = \Sigma_2 = \Sigma$ say

let $X_\alpha^{(1)}$ $(\alpha = 1, 2, \ldots N_1)$ and $X_\beta^{(2)}$ $(\beta = 1, 2, \ldots N_2)$ be two independent random sample from these pop^{ns} respectively. We calculate

$$\bar{X}^{(1)} = \frac{1}{N_1}\sum_{\alpha=1}^{N_1} X_\alpha^{(1)} \quad \text{and} \quad \bar{X}^{(2)} = \frac{1}{N_2}\sum_{\beta=1}^{N_2} X_\beta^{(2)}$$

and
$$S_i = \frac{1}{(N_i - 1)} \sum_{\alpha=1}^{N_i} (X_\alpha^{(i)} - \bar{X}^{(i)})(X_\alpha^{(i)} - \bar{X}^{(i)})', \quad i = 1, 2.$$

also calculated the pooled variance covariance matrix given by :-

$$S = \frac{(N_1 - 1)S_1 + (N_2 - 1)S_2}{(N_1 + N_2 - 2)}$$

We know that $\bar{X} \sim N_p(\mu, \frac{\Sigma}{N})$ using this we have

$$\bar{X}^{(1)} \sim N_p\left(\mu^{(1)}, \frac{\Sigma}{N}\right), \quad \bar{X}^{(2)} \sim N_p\left(\mu^{(2)}, \frac{\Sigma}{N}\right)$$

$$\Rightarrow \sqrt{\frac{N_1 N_2}{N_1 + N_2}}\left(\bar{X}^{(1)} - \bar{X}^{(2)}\right) \sim N_p(0, \Sigma) \quad \text{under } H_0$$

$\therefore$ the proposed test statistics is:

$$T^2 = \frac{N_1 N_2}{(N_1 + N_2)}\left(\bar{X}^{(1)} - \bar{X}^{(2)}\right)' S^{-1}\left(\bar{X}^{(1)} - \bar{X}^{(2)}\right)$$

Or,

$$\frac{T^2}{(N_1 + N_2 - 2)} \cdot \frac{N_1 + N_2 - p + 1}{p} \sim F_{p, N_1 + N_2 - p - 1; \alpha\%}$$

③ q - Sample problem

Let $X_\alpha^{(i)}$ $(\alpha = 1, 2, \dots N_i)$, $i = 1, \dots q$ be independent random sample from $N_p(\mu^{(i)}, \Sigma)$ equality of Σ^s is assumed and we wish to test

$H_0 : \sum_{i=1}^{q} \beta_i \mu^{(i)} = \mu_0$ (specified) where β_i's are scalars and VCM as

$$\bar{X}^{(i)} = \frac{1}{N_i}\sum_{\alpha=1}^{N_i} X_\alpha^{(i)} \quad \text{and} \quad S_i = \frac{1}{(N_i - 1)}\sum_{\alpha=1}^{N_i}\left(X_\alpha^{(i)} - \bar{X}^{(i)}\right)\left(X_\alpha^{(i)} - \bar{X}^{(i)}\right)'$$

and the pooled VCM

$$S = \frac{1}{\sum_{i=1}^{q} N_i - q}\sum_{i=1}^{q}\sum_{\alpha=1}^{N_i}\left(X_\alpha^{(i)} - \bar{X}^{(i)}\right)\left(X_\alpha^{(i)} - \bar{X}^{(i)}\right)'$$

then for testing H_0 the proposed test statistic is:

$$T^2 = C\left[\sum_{i=1}^{q} \beta_i \bar{X}^{(i)} - \mu_0\right]' S^{-1}\left[\sum_{i=1}^{q} \beta_i \bar{X}^{(i)} - \mu_0\right] \sim T^2_{\left(\sum_{i=1}^{q} N_i - q\right) d.f}$$

So under H_0 :

$$\frac{T^2}{\sum_{i=1}^{q} N_i - q} \cdot \frac{\sum_{i=1}^{q} N_i - q - (p-1)}{p} \sim F_{p, \sum_{i=1}^{q} N_i - q - (p-1); \alpha\%}$$

$\mu = \mu_0$, Σ, σ^2 }

(4) Problem of Symmetry

Let $X_{p\times1} \sim N_p(\mu, \Sigma)$, Σ assumed to any known. We may be interested in testing $H_0: \mu_1 = \mu_2 = \dots \mu_p$ (ie all the components of the mean vectors are same) $H_0': (\mu_1 - \mu_p) = (\mu_2 - \mu_p) = \dots = (\mu_{p-1} - \mu_p) = 0$ which can be written ~~in~~ matrix is from as:

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & -1 \\ 0 & 1 & 0 & \cdots & -1 \\ \vdots \\ 0 & 0 & 0 & \cdots & 1 \; -1 \end{bmatrix} \mu = 0$$

In order to test this H_0 we choose a matrix C such that it is of order $(p-1)\times p$ and it is of ~~order~~ rank $(p-1)$ and $C\underline{1} = 0$ where $\underline{1}_{p\times1}$ is a unit vector.

$$\therefore C\underline{1} = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1p} \\ C_{21} & C_{22} & \cdots & C_{2p} \\ \vdots & & & \vdots \\ C_{p-1,1} & \cdots & \cdots & C_{p-1,p} \end{bmatrix}_{(p-1)\times p} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} C_{11} + C_{12} + \cdots + C_{1p} \\ C_{21} + C_{22} + \cdots + C_{2p} \\ \vdots \\ C_{p-1,1} + \cdots + C_{p-1,p} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \underline{0}$$

Consider the transformation $\underline{Y} = C\underline{x}$ then

$\underline{Y} \sim N_p(C\mu, C\Sigma C')$ or

$Y \sim N_p(\eta, \phi)$ where $\eta = C\mu$

and $\phi = C\Sigma C'$

If H_0 is true all the components of μ will be equal and then

$\eta = C\mu = C\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = 0$ is $H_0 \to H_0' : \eta = 0$ in this case the proposed test statistics is

$T^2 = N\bar{Y}' S_y^{-1} \bar{Y}$ and then we can proceed with as usual testing as one sample problem.

Ex: let $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$

$x_1 \to$ Score in Physics
$x_2 \to$ Maths
$x_3 \to$ English
$x_4 \to$ Chemistry
$x_5 \to$ Social Science

Suppose we wish to test the hypothesis that the student is good in all the subjects then we can have $E(\underline{X}) = \mu$ and we can propose H_0 as above and again we can test the hypothesis.

(*) Properties of T^2.

T^2-test is (i) UMP (ii) UMPU - Unbiased
(iii) UMPS - Similar (iv) UMPI - Invariance

T^2-test has a monotonic power function.

(*) Optimum property

T^2 is unaffected by change of origin and scale of origin and scale of the response variable i.e It is Invariant all the transformation of type:

$$Y = C\underline{X} + \underline{d}$$

where C is NSM and $\underline{d}$ is non-random vector of component

proof We have $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$.

Consider $\underline{Y} = C\underline{X} + \underline{d}$ and suppose we wish to test $H_0 : \underline{\mu} = \underline{\mu}_0$ then $E(\underline{Y}) = C\underline{\mu} + \underline{d} = C\underline{\mu}_0 + \underline{d}$ and suppose we wish to test

$$H_0 : \underline{\mu} = \underline{\mu}_0 \text{ then } E(\underline{Y}) = C\underline{\mu} + \underline{d}$$
$$= C\underline{\mu}_0 + \underline{d} \quad (\text{under } H_0)$$

for this we have $T_x^2 = N(\bar{\underline{X}} - \underline{\mu}_0)' S_x^{-1} (\bar{\underline{X}} - \underline{\mu}_0)$

Now $T_y^2 = N(\bar{Y} - \underline{\mu}_0)' S_y^{-1} (\bar{Y} - \underline{\mu}_0)$

OR

$$T_y^2 = N[(C\bar{\underline{X}} + \underline{d} - C\underline{\mu}_0 - \underline{d})' S_y^{-1} (C\bar{\underline{X}} + d - C\underline{\mu}_0 - d)]$$
$$= N[(C\bar{\underline{X}} + \underline{d} - C\underline{\mu}_0 - \underline{d})' (C S_x C')^{-1} (C\bar{\underline{X}} + d - C\underline{\mu}_0 - d)]$$
$$= N[(\bar{x} - \underline{\mu}_0)' C' (C S_x C')^{-1} C (\bar{\underline{X}} - \underline{\mu}_0)]$$
$$= N[(\bar{\underline{X}} - \underline{\mu}_0)' C' (C')^{-1} S_x^{-1} (C^{-1}) C (\bar{\underline{X}} - \underline{\mu}_0)]$$
$$= N[(\bar{\underline{X}} - \underline{\mu}_0)' S_x^{-1} (\bar{\underline{X}} - \underline{\mu}_0)]$$
$$= T_x^2$$

$\therefore$ ~~T_x~~ T^2 is invariant under a group of transformation

Repeated Measurement (Repeated)

Frequently the experimental observation are collected at different times but on the same sample units the simplest example of this Repeated Measurement design consists of a controlled treatment (a placebo, standard drug or anyother null treatment) on one ocassion and the treatment of interest on other occassions.

It is reasonable to assume that the treatment has an additive effect than we may write the responses of the i^{th} subject under control and treatment conditions as:-

$$\left.\begin{aligned} x_{i1} &= \mu + e_{i1} \\ x_{i2} &= \mu + \tau + e_{i2} \end{aligned}\right\} \quad \text{— (1)}$$

where μ is the general mean effect common to all the subjects and τ is the treatment effect of the experimental condition.
And e_{i1}, e_{i2}, are the random disturbance term

Let us extend the paired observation case to the general setup of 'K' (say) responses collected on the same experimental unit at successive times or under a variety of experimental condition.

We shall assume that these 'K' responses constitute a fixed set of observation. Let the observations on N- Independent units under 'K' conditions are arranged as:-

Sample units	Condition
1	1 2 3 . . . K
1	x_{11} x_{12} . . . x_{K1}
2	x_{21} x_{22} . . . x_{K2}
⋮	⋮ ⋮
N	x_{N1} x_{N2} . . . x_{NK}

Observations from a repeated Measurements Experiment.

The mathematical model for the j^{th} response in the i^{th} sampling unit may be written as:

$$x_{ij} = \mu + \mu_j + e_{ij} \quad , \quad i = 1,2,\ldots N \quad j = 1,2,\ldots k$$

— (1)

where μ : general mean effect (or general level of the p.m. common to all observations)

μ_j : effect of j^{th} specification and

e_{ij} : random disturbance

e_{ij} reflects both the interaction of the i^{th} unit with the j^{th} response and the experimental error.

To obtain the test of hypothesis for μ_j it will be necessary to assume that the error vector $(e_{i1}, e_{i2}, \ldots e_{iK}) = (0, 0, \ldots 0) = \underline{0}$ — (2) should be a null vector and the covariance matrix

$$\Sigma = E\left[\begin{pmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{iK} \end{pmatrix} (e_{i1}\ e_{i2} \ldots e_{iK})\right] \quad — (3)$$

write :-

$$\mu = [\mu + \mu_1, \mu + \mu_2, \ldots \mu + \mu_K] \quad — (4)$$

then the null hypothesis of equal responses effect is :

$$H_0 : \mu_1 = \mu_2 = \ldots = \mu_K$$

$$\text{Vs } H_1 : \mu_i \neq \mu_j \quad \text{for atleast one pair of } i \text{ and } j$$

the vector statements of H_0 and H_1 are :

$$H_0 : \begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \\ \vdots \\ \mu_{K-1} - \mu_K \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \underline{0} \quad \text{vs} \quad H_1 : \begin{pmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \\ \vdots \\ \mu_{K-1} - \mu_K \end{pmatrix} \neq 0 \quad — (5)$$

The test of H_0 stated above can be now carried out with the help of Hotlines T^2 statistics computed from the mean vector and VCM of the differences

$$y_{ij} = [x_{ij} - x_{i,j+1}] \quad \text{—} (6)$$

of the observations on adjacent responses. If we ~~element~~ denote $\bar{Y}'$ the $(k-1)$ component vector of differences of successive response means:-

$$\bar{Y}' = [\bar{y}_1, \bar{y}_2, \ldots \bar{y}_{k-1}] = [\bar{x}_1 - \bar{x}_2, \ldots \bar{x}_{k-1} - \bar{x}_k] \quad \text{—} (7)$$

and let $S_{y(k-1)(k-1)}$ denotes the covariance matrix of the y_{ij} differences. Then the test statistics for testing H_0 of equal response effect is:

$$T^2 = N\,\bar{Y}' S_y^{-1} \bar{Y} \quad \text{and under } H_0$$

$$\frac{T^2 (N-k+1)}{(N-1)(k-1)} \sim F_{(k-1),(N-k+1)};\alpha\%.$$

then H_0 may be rejected if $F_{cal} > F_{tab,\alpha\%}$. O.W H_0 may be accepted.

3 (a) Multiple Comparisons:-

If the hypothesis of equal treatment effects is rejected consider any linear combination

$$\sum_{j=1}^{k} b_j \mu_j \quad -(1) \quad \text{of the treat effects}$$

Where the coefficients b_j's sum to zero

We have to choose a matrix c so that the element generate the contrast of adjacent treatments. Now let $\underline{a}'$ be any $(k-1)$ component row vector then by the UNION-INTERSECTION principle it follows that all the inequalities like:

$$[\underline{a}' c (\bar{\underline{x}} - \underline{\mu})]^2 \leq \frac{1}{N}\, \underline{a}' c S c' \underline{a}\; T^2_{(k-1)(N-k+1),\alpha\%}$$

$$(*)$$

formed for the different choices of the elements of $\underline{a}$ hold simultaneously with prob $(1-\alpha)$ But since the row sums of C of each row is zero, so must the sum of elements of $a'C$ be zero $b'\underline{\mu} = \underline{a}'C\mu$ is $\therefore$ a contrast of the treatment effects

So,

$$b'\mu = \sum_{j=1}^{k} b_j (\mu + \mu_j) = \sum_{j=1}^{k} b_j \mu_j$$

expansion of (*) leads to a set of $100(1-\alpha)\%$ simultaneously C.I.s whose general member is :

$$\underline{b}'\bar{\underline{X}} - \sqrt{\frac{b'Sb}{N} T^2} \leq \underline{b}'\mu \leq \underline{b}'\bar{x} + \sqrt{\frac{b'S\underline{b}}{N}} T^2_{(k-1)(N-k+1);\alpha}$$

if the interval includes zero, the hyp. $H_0: \sum_{j=1}^{k} b_j \mu_j = 0$ is acceptable in the sense of multiple comparison at $\alpha\%$ level.

☆ Profile Analysis

Profile analysis is an equivalent of RM MANOVA mainly concern with the test scores obtained by a group of persons from educational and pschological assesment.

Profile also refers to an outline or a short description of someone's life or achievements

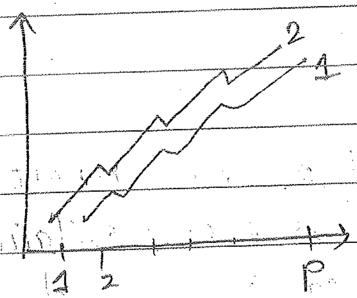
In a battery of say 'p' psychological tests administered at to a group. Let $\mu_1, \mu_2, \ldots \mu_p$ are the expected scores in these 'p' tests, then the PROFILE of the group is defined as the GRAPH obtained by joining the points $(1, \mu_1)$ $(2, \mu_2)$ ... (p, μ_p) successively.

In practice the μ_i's value are estimated from the average score in a sample.

If we have two possibly different groups we will have two profiles, one determined by joining the points (i, μ_i) $i=1,2,\dots p$ and the other by (i, γ_i) $i=1,2,\dots p$ where $\gamma_1, \gamma_2, \dots \gamma_p$ are the true mean scores of the other group.

The profiles are set to be similar if the segment of line joining (i, μ_i) to $(i+1, \mu_{i+1})$ is PARALLEL to the corresponding segment of the line joining (i, γ_i) to $(i+1, \gamma_{i+1})$ of the other profiles each i.

Obviously this will be so if the hypothesis
$H_0 : (\mu_i - \mu_{i-1}) = (\gamma_i - \gamma_{i-1})$ for $i = 1, 2, \dots p$ is TRUE



Let us assume that the test scores $\underline{x}_{p\times 1}$ are normally distributed and their VCM is same for both the group.

let X be $(p \times n_1)$ matrix of scores of n_1 individual from the first group and similarly Y be $(p \times n_2)$ matrix of scores of n_2 individual from the 2^{nd} group. Then the matrices X and Y represent the samples of the sizes n_1 and n_2 from ~~NCp~~ $N_p(\underline{\mu}, \Sigma)$ and $N_p(\underline{\gamma}, \Sigma)$ resp.

The Hypothesis H_0 can be rewritten as
H_0 : $E(C_1\bar{X})$ is same in both the groups
where C_1 is the contrast matrix given by

$$C_1 = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & -1 & 1 \end{bmatrix}$$

Note that $C_1\bar{X}$ has the distribution $N_{p-1}(C_1\underline{\mu}, C_1\Sigma C_1')$ in the 1st group and $N_{p-1}(C_1\underline{\gamma}, C_1\Sigma C_1')$ in the 2nd group and $C_1\underline{X}, C_1\underline{Y}$ have the samples these distn.

The problem of testing of H_0 is same as the test of equality of two mean vectors from two multivariate Normal distn with the same unknown Σ VCM, then the test is provided by the following test statistics:

TEST STATISTICS:-

$$\frac{(n_1+n_2)-(p-1)-1}{(p-1)} \cdot \frac{n_1 n_2}{(n_1+n_2)} \cdot \frac{D^2}{(n_1+n_2-2)} \sim F_{p-1,\, n_1+n_2-(p-1)-1;\ \alpha\%} \quad \text{—}(*)$$

Where $D^2 = (n_1+n_2-2)(C_1\underline{d})'(C_1SC_1')^{-1}(C_1\underline{d})$

$d = (\bar{X}-\bar{Y})$, $\bar{X}$: Sample mean vector of population 1

$\bar{Y}$: Sample mean vector of population 2

$$S = X\left[I - \frac{1}{n_1}E_{n_1,n_1}\right]X' + Y\left[I - \frac{1}{n_2}E_{n_2,n_2}\right]Y'$$

Where : $E_{p,q}$ $(p\times q)$ matrix of all the elements being unit

If we decide to accept the hypothesis of similarity of two profiles on the basis of the test given by the (*), the next logical question will be whether the average true scores of the two group is same i.e whether

H_1 : $\frac{1}{p}(\mu_1+\mu_2+\dots\mu_p) = \frac{1}{p}(\gamma_1+\gamma_2+\dots\gamma_p)$ OR $E_{1p}\underline{\mu} = E_{1p}\underline{\gamma}$

Observe that $E_{1p}\underline{X}$ has a univariate normal dist^n with variance $E_{1p}\Sigma E_{1p}$ and the means $E_{1p}\underline{\mu}$, $E_{1p}\underline{\gamma}$ in the two groups and hence the hypo H_1 can be tested by the usual t-test for testing the equality of the mean of two normal populations with the same variance.

If we square t for testing H_1 we will get the F distribution or F-test statistics.

If Ho is true and if in addition H_1 is also true then we would like to verify that $\mu_1=\gamma_1$, $\mu_2=\gamma_2$, $\dots$ $\mu_p=\gamma_p$. i.e BOTH THE PROFILES ARE EXACTLY SAME.

If H_0, H_1 and H_2 ALL hold then the distribution of X is same in both the groups and then X and Y are samples from the same group. So now we can POOL these giving the GRAND MEAN VECTOR :

$$\underline{g} = \frac{n_1\bar{X} + n_2\bar{X}}{(n_1+n_2)}$$ as an estimate

of the common value of μ and γ and it's dist^n is $N_p(\mu, \Sigma/N)$ where $N=(n_1+n_2)$

The hypo H_2 is equivalent to $E(C\underline{g})=0$, where 'C' is the contrast matrix and so the test reduce to the test of sign of contrasts which can be accomplished by

$$\frac{(n_1+n_2-2)-(p-1)-1}{(p-1)} \cdot \frac{T^2}{(n_1+n_2-2)} \sim F_{p-1,\, n_1+n_2-p-1}$$

$$T^2 = N(Cg)'(CSC')^{-1}(Cg)$$

Note: If H_2 is accepted the common profile of two group consist of parallel line to the x-axis at a distance equal to the common value of $\underline{y}$ and $\underline{\gamma}$

Notations :-

→ $\bar{X} = X E_{n_1,1}/n_1$ → Vector of sample means for 1st Sample

→ $\bar{Y} = Y E_{n_2,1}/n_2$ → vector of sample means for 2nd Sample

→ $S_X = X\left(I - \left(\frac{1}{n}\right) E_{n,n_2}\right) X'$: Matrix of corrected S.S and of SP of the observation from the 1st sample

→ $S_Y = Y\left[I - \left(\frac{1}{n}\right) E_{n_1,n_2}\right] Y'$: Matrix of corrected SS and of SP of the observation from the 2nd Sample

Also

$\bar{X} \sim N_p(\underline{y}, n_1^{-1} \Sigma)$ $\quad S_X \sim W_p(n_1-1, \Sigma_1)$

$\bar{Y} \sim N_p(\underline{\gamma}, n_2^{-1} \Sigma)$ $\quad S_Y \sim W_p(n_2-1, \Sigma_2)$

✱ Fishers Behren's Problem:-

Suppose we wish the test $H_0: \underline{y} = \underline{\gamma}$ —(1)
where $\underline{y}$ is the mean vector of $N_p(\underline{y}, \Sigma_1)$ and γ is the mean vector of $N_p(\underline{\gamma}, \Sigma_2)$, Σ_1 and Σ_2 are unknown and unequal

Consider that we have random samples of sizes n_1 and n_2 from the two multivariate normal population and let X and Y be $(p \times n_1)$ and $(p \times n_2)$ matrices of sample observations. We also assume that $n_1 < n_2$ and now partition the Y matrix as

$$Y = [\underset{n_1}{Y_1} \mid \underset{m}{Y_2}] \quad \text{—(2)} \quad \text{where } m = (n_2 - n_1)$$

Note that $E(X) = \underline{\mu} E_{1n_1}$, $Var(X) = \Sigma_1 \otimes I_{n_1}$ —(3)

$A \otimes B$: represent the Kronecker product of A & B

Similarly,

$$E(Y_1) = \underline{\gamma} E_{1n_1}, \quad E(Y_2) = \underline{\gamma} E_{1m} \quad \text{—(4)}$$

$$var(Y_1) = \Sigma_2 \otimes I_{n_1}, \quad var(Y_2) = \Sigma_2 \otimes I_m \quad \text{—(5)}$$

Further, $Cov(Y_1, Y_2) = 0$ —(6)

As all the cols of Y are indep define a $(p \times n_1)$ matrix Z as:

$$Z = X - \left(\frac{n_1}{n_2}\right)^{1/2} Y_1 \left(I - \frac{1}{n_1} E_{n_1, n_1}\right) - \frac{1}{n_2} Y E_{n_2, n_1} \quad \text{—(7)}$$

It can be observed from (3) & (4) that

$$E(Z) = (\underline{\mu} - \underline{\gamma}) E_1 n_1 \quad \text{—(8)}$$

Or,

$$Z = X - Y_1 P - \frac{1}{n_2} Y_2 E_{m n_1} \quad \text{—(9)}$$

where

$$P = \left(\frac{n_1}{n_2}\right)^{1/2} \left(I - \frac{1}{n_1} E_{n_1, n_1}\right) + \frac{1}{n_2} E_{n_1, n_1} \quad \text{—(10)}$$

using (2) Z can be written as

$$Z = X - \left(\frac{n_1}{n_2}\right)^{1/2} Y_1 \left(I - \frac{1}{n_1} E_{n_1, n_1}\right) - \frac{1}{n_2} \left(Y_1 E_{n_1 n_1} + Y_2 E_{m n_1}\right)$$

then using (3), (4) and (5) we ~~have~~ obtain :-

$$Var(z) = \Sigma_1 \otimes I_{n_1} + \Sigma_2 \otimes (P' I_{n_1} P) + \frac{1}{n_2}(E_{n_1 m} I_{m n_1})$$

$$= \Sigma_1 \otimes I_{n_1} + \Sigma_2 \otimes \left(P'P + \frac{m}{n_2^2} E_{n_1 n_1} \right)$$

$$= \left(\Sigma_1 + \frac{n_1}{n_2}\Sigma_2\right) \otimes I_{n_1} \quad \text{—— (11)}$$

OR

$$Var(z) = \Sigma \otimes I_{n_1} \quad \text{where } \Sigma = \Sigma_1 + \left(\frac{n_1}{n_2}\right)\Sigma_2 \quad \text{—— (12)}$$

Σ is non singular also z is a linear combination of the variables from X and Y, it is observed that each column of z will have a multivariate normal distribution with mean $(\underline{\mu} - \gamma)$ and a var-cov matrix Σ given by eq (12)

Further, as all the cols of z are independent then if the hypothesis $H_0 : \underline{\mu} = \gamma$ of the z will be testing that $z = \underline{0}$ (null vector)

Which can be again tested by using hotline's T^2 ~~t-so~~ given by

$$\frac{T^2}{(n_1-1)} = n_1 \bar{Z}' S_z^{-1} \bar{Z} \quad \text{where } \bar{Z} = \frac{1}{n_1} Z E_{n_1 1} \quad \text{—— (13)}$$

mean vector of obs in z and $S_z = Z\left(I - \frac{1}{n_1} E_{n_1, n_1}\right) Z'$ —— (15)

∴ the proposed test statistics is

$$\frac{(n_1-P)}{P} \cdot \frac{T^2}{(n_1-1)} \sim F_{p, n-p}$$

Note that $\bar{Z} = (\bar{X} - \bar{Y})$ and

$$S_U = U\left(I - \frac{1}{n_1} E_{n_1, n_1}\right) U' \quad \text{where } U = \left(\frac{n_1}{n_2}\right)^{1/2} Y_1 \quad \text{—— (15)}$$

it is $\therefore$ easier to form the matrix U by substracting from every col of X, $\left(\frac{n_1}{n_2}\right)^{1/2}$ times the corresponding col. of Y. Then we may write:

$$\frac{T^2}{(n_1-1)} = n_1 (\bar{X}-\bar{Y})' S_0^{-1} (\bar{X}-\bar{Y}) \quad \text{--- (16)}$$

it is to be noted that all the n_2 observations are utillizes ~~that~~ for finding out $\bar{Y}$ but only n_1 observations are utillizes in S_0 so which of the (n_2-n_1) columns of Y are ignored is arbitrary in this test and therefore this is the weakness of this test.

* Test of Symmetry of ORGANS:-

In many biological and anthropological problem let us consider the mesurements characterstics on the LEFT side are $x_1, x_2, \ldots x_p$ and $x_{p+1}, \ldots x_{2p}$ represent the same mesurements on the right sides. or $x_i\ (i=1,\ldots p)$ represents the characterstics of a twin pair and $x_{p+i}\ (i=1,\ldots p)$ are those for the other members and we are interested in testing the hypothesis of symmetry of the left and right sides or the equality of the measurement of the twins.

Mathematically the hypothesis to be tested can be expressed as

$$H_0: \mu_i = \mu_{p+i} \quad (i=1,\ldots p) \quad \text{--- (1)}$$

where $\mu_1, \mu_2, \ldots \mu_{2p}$ are the true means of $x_1, \ldots x_{2p}$.

let $\bar{X}$ and S represent the mean vector and the matrix of corrected SS and S.P of 'm' samples observation. from this 2p-variate normal popn.

Now, consider the partitioning of $\bar{X}$ and S as

$$\bar{X} = \begin{bmatrix} \bar{X}^{(1)} \\ \bar{X}^{(2)} \end{bmatrix} \begin{matrix} \to p \\ \to p \end{matrix} \quad \text{and} \quad S = \begin{bmatrix} S_1 & S_2 \\ S_2' & S_3 \end{bmatrix} \begin{matrix} \to p \\ \to p \end{matrix} \qquad —(2)$$

then we know that

$$\sqrt{n}\,\bar{X} \sim N_p(\sqrt{n}\,\mu, \Sigma) \text{ and } S \sim W_{2p}(n-1, \Sigma)$$

where $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{2p} \end{bmatrix}$ and Σ is $(2p \times 2p)$ VCM of 2p variate

$$\therefore Y = \sqrt{n}\,(\bar{X}^{(1)} - \bar{X}^{(2)}) = \sqrt{n}\,A\bar{X} \quad —(3)$$

has $N_p[\sqrt{n}(\mu^{(1)} - \mu^{(2)}), A\Sigma A']$ where $\mu = \begin{bmatrix} \mu^{(1)} \\ \mu^{(2)} \end{bmatrix} \begin{matrix} \to p \\ \to p \end{matrix}$

then the distn of $ASA' = (S_1 - S_2' - S_2 + S_3)$ is $W_p(ASA', n-1)$ —— (4)
and it is independent of Y.

Hence, if H_0 is true i.e $\mu^{(1)} = \mu^{(2)}$, then

$$\frac{(n-1)-(p-1)}{p} \cdot \frac{T^2}{(n-1)} \sim F_{p,(n-p);\alpha\%}$$

where $\frac{T^2}{(n-1)} = Y'(ASA')^{-1}Y$

OR, $\frac{T^2}{(n-1)} = n(\bar{X}^{(1)} - \bar{X}^{(2)})'(S_1 - S_2' - S_2 + S_3)^{-1}(\bar{X}^{(1)} - \bar{X}^{(2)})$

* The POWER OF TESTS ON MEAN VECTOR(S) :-

So far we have been concerned with the size of the test, however in a properly designed experiment the prob of type-2 error should also be controlled.

We ~~kn~~ use $T^2 = Y' S^{-1} Y$ — (1) is calculated from the p dimentional which has a multivariate normal distⁿ $N_p(\mu, \Sigma)$ and we have also estimated the parameter μ and σ based on n degrees of freedom then we know that quantity

$$F = \frac{T^2 (n-p+1)}{p}$$, has a NON-CENTRAL F-distⁿ with

d.f $p, (n-p+1)$ and ncp $\delta^2 = \mu' \Sigma \mu$

in particular if we are testing

$H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$

by the T^2-statistics, then ~~ncp $\delta^2 = N(\mu$~~

$$ncp\ \delta^2 = N(\mu - \mu_0)' \Sigma^{-1} (\mu - \mu_0)$$

Again, for the two sample problem, Testing

$H_0: \mu^{(1)} = \mu^{(2)}$ vs $H_1: \mu^{(1)} \neq \mu^{(2)}$

the test statistics has a NON-CENTRAL F-dist with p.m

$$\frac{N_1 N_2}{(N_1 + N_2)} (\mu^{(1)} - \mu^{(2)})' \Sigma^{-1} (\mu^{(1)} - \mu^{(2)})$$

and the df $p, N_1 + N_2 - (p-1)$

In either case the POWER FUNCTION of the test is

$$1 - \beta(\delta^2) = P[F' > F_{p, n-p+1; \alpha\%}]$$

where F' has non-central F-distribution with prescribed parameters

PEARSON and HEARTLY have prepared the power charts of the power function for $\alpha = 1\%$ & 5%.

Refrences :-

1. Page - 149 (MORRISON BOOK)
2. Investigate relevance profiling and Serial crime (3rd edition) Dr. CLAIRE FERGUSON (2014)

UNIT - II

PRINCIPLE COMPONENT ANALYSIS

INTRODUCTION

In Multivariate studies if the number of variables on which the data is collected is too large and we are interested in only few of them, then a way of reducing a number of variable is to discarded those linear combination of the variables which have similar or smaller variance.

PCA is a technique to help in this direction as it trys to find few linear combinations of the original variables which can be used to sumarize the data without loosing much information.

It is a techique to transform the original variables in such a way that most of the variation is explained by these linear combination. The linear combination is so selected are called the principle components and the major objective of PCA is to reduce the dimension of the data into few components.

A PCA is concerned with explaining the variance covariance structure of the given multivariate data.

It is important to note that the total variation remains the same even after the transformation of original variables into principale component.

Mathematics of PCA

Let $\underline{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ be a random vector of p-component with its dispersion matrix Σ, Since we are in the var-cov structure, we assume that $E(\underline{X}) = 0$. Let β be a $(p\times1)$ comp. vector s.t $\beta'\beta = 1$ and let $\underline{U} = \beta'\underline{X}$ be a linear combination of the components of X we wish to find out β in such a way that $var(\beta'\underline{X})$ is maximum. i.e

$$Var(\underline{U}) = var(\beta'X) = \underline{\beta}' var(\underline{X})\underline{\beta} = \beta'\Sigma\underline{\beta} \quad -(1)$$

where β is a Normal vector define a f^n:

$$\Phi = \underline{\beta}'\Sigma\underline{\beta} - \lambda(\underline{\beta}'\underline{\beta} - 1);$$

λ be the language's multiplier

now,

$$\frac{\partial\Phi}{\partial\underline{\beta}} = 2\Sigma\underline{\beta} - 2\lambda\underline{\beta} = 0 \quad \text{(equating it to zero)}$$

$$\Rightarrow \Sigma\beta - \lambda\underline{\beta} = 0 \quad \text{i.e } (\Sigma - \lambda\underline{I})\underline{\beta} = 0$$

Since $\beta \neq 0$

for a non-zero solution of the above equations we must have $|\Sigma - \lambda I| = 0 \quad -(2)$

Thus, $\underline{\lambda}$ is the ch. root of Σ, as Σ is a $(p\times p)$ matrix, $|\Sigma - \lambda I| = 0$, will be a polynomial of order 'p' which will give 'p' ch. roots for Σ (say): $\lambda_1, \lambda_2, \ldots \lambda_p$.

Now, let $\underline{\beta}_i$ be the ch. vector of Σ corresponding to the ch. root λ_i.

$\therefore$ it must satisfy the original equations and we must have

$$\Sigma \beta_i - \lambda_i \beta_i = 0,$$

now premultiplying it by β_i, we get

$\beta_i' \Sigma \beta_i - \beta_i' \lambda_i \beta_i = 0$ or $\beta_i' \Sigma \beta_i - \lambda_i = 0$

as $\underline{\beta}'\underline{\beta} = 1 \Rightarrow \underline{\beta}_i'\beta_i = 1$

$\therefore$ $\beta_i' \Sigma \beta_i = \lambda_i$ Or, $Var(\underline{\beta}_i' \underline{x}) = \lambda_i$,

We wish to maximize $var(\underline{\beta}_i' x)$
so, we should choose λ_i to be Maximum

now, arranging the 'p' roots as : $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p$
If $\underline{\beta}_1$ is the ch. vector corresponding to λ_1 (largest ch. root), then $(\underline{\beta}_1' x)$ is the NORMALIZED linear combination of $\underline{x}$ with this max. vari $= \lambda_1$

So, $\boxed{U_1 = \underline{\beta}_1' x} \rightarrow 1^{st}$ PC (Principle component)

For second principle component again we find out in such a way that $(\underline{\beta}' \underline{x})$. In such a way that

$$Var(\underline{\beta}' x) = \underline{\beta}' \Sigma \underline{\beta} \text{ is maximum}$$

Subjected to the conditions that
(i) $\underline{\beta}'\underline{\beta} = 1$ and
(ii) $\underline{\beta}' \underline{x}$ is uncorrected to U_1 i.e
$cov(U_1, \beta' \underline{x}) = 0$ or $\beta_1' \Sigma \underline{\beta} \Rightarrow \underline{\beta}' \Sigma \beta_1 = 0$

As β_1 is vector corresponding to λ_1

$\therefore \Sigma \underline{\beta}_1 = \lambda_1 \underline{\beta}_1$, now premultiplying it by $\underline{\beta}'$, we have

under (ii)

$$\underline{\beta}' \Sigma \underline{\beta}_1 = \underline{\beta}' \lambda_1 \underline{\beta}_1 \quad \text{i.e } \lambda_1 \underline{\beta}' \underline{\beta}_1 = \underline{\beta}' \Sigma \underline{\beta}_1 = 0$$

i.e $\lambda_1 \underline{\beta}' \underline{\beta}_1 = 0$ but $\lambda_1 \neq 0$

$\therefore \underline{\beta}' \underline{\beta}_1 = 0$,

Again define $\phi = \underline{\beta}' \Sigma \underline{\beta} - \lambda(\underline{\beta}'\underline{\beta} - 1) - \mu(\underline{\beta}' \Sigma \underline{\beta}_1 - 0)$

Where λ and μ are the lagrange's multipliers.

Now,

$$\frac{\partial \phi}{\partial \underline{\beta}} = 2 \Sigma \underline{\beta} - 2\lambda \underline{\beta} - \mu \Sigma \underline{\beta}_1 = 0,$$

premultiplying it by $\underline{\beta}_1'$, we have

$$2 \underline{\beta}_1' \Sigma \underline{\beta} - 2 \underline{\beta}_1' \lambda \underline{\beta} - \mu \underline{\beta}_1' \Sigma \underline{\beta}_1 = 0$$

Now, $\text{cov}(\underline{\beta}_1' \underline{X}, \underline{\beta}' \underline{X}) = \underline{\beta}' \Sigma \underline{\beta}_1 = \underline{\beta}' \Sigma \underline{\beta} = 0$

$$\therefore \frac{\partial \phi}{\partial \beta} = 2 \Sigma \underline{\beta} - 2\lambda \underline{\beta} = 0 \quad \text{or,} \quad (\Sigma - \lambda I)\underline{\beta} = 0$$

again, $\underline{\beta}'\underline{\beta} = 1 \quad \therefore \underline{\beta} \neq 0 \quad \therefore |\Sigma - \lambda I| = 0$

$\therefore \lambda$ is the ch. root of Σ and $\underline{\beta}$ is the corresponding ch. vector, Now, to maximize var $(\underline{\beta}' \underline{X})$ we choose 2^{nd} largest ch. root of Σ which is the λ_2 and let β_2 be the corresponding ch. vector then

$\boxed{U_2 = \underline{\beta}_2' \underline{X}}$ is the SECOND PRINCIPLE COMPONENT which is uncorrelated with U_1

Proof. It is to be noted that $\beta_1 = \beta_2$ satisfied $\beta_2' \beta_1 = 0$ and $\beta_2' \Sigma \beta_1 = 0$ which $\Rightarrow \beta_2$ and β_1 are orthogonal. Now, let $\mathbb{B} = [\beta_1, \beta_2, \dots \beta_p]$ then $\mathbb{B}$ represent a $(p \times p)$ orthogonal matrix

i.e $\mathbb{B}' \Sigma \mathbb{B} = \text{diag}(\lambda_1, \lambda_2, \dots \lambda_p)$ proceeding like this at the $(i+1)^{th}$ steps, we have to determine $\underline{\beta}$ s.t. $\text{var}(\underline{\beta}'\underline{x}) = \beta' \Sigma \underline{\beta}$ is maximum subjected to conditions

(i) $\underline{\beta}' \beta = 1$ and

(ii) $\beta' x$ is uncorrelated with $1^{st}, 2^{nd}, \dots i^{th}$ PC's

Now, let the i^{th} P.C $U_i = \underline{\beta}_i' \underline{x}$ and β_i the ch vector corresponding to λ_i and

(ii) implies that $\text{cov}(\underline{\beta}' x, \underline{\beta}_r' \underline{x}) = 0$ for $r = 1, 2, \dots i$

i.e $\underline{\beta}' \Sigma \beta_r = 0$ for $r = 1, 2, \dots i$ also for $r \le i$ the original equation must be satisfied

i.e $\Sigma \underline{\beta}_r = \lambda_r \beta_r$,

now, premultiplying it by $\underline{\beta}'$, we get

$$\underline{\beta}' \Sigma \beta_r = \underline{\beta}' \lambda_r \underline{\beta}_r = \lambda_r \underline{\beta}' \underline{\beta}_r \Rightarrow \underline{\beta}' \beta_r = 0, \; r = 1, 2, \dots i$$

$$\text{as } \underline{\beta}' \Sigma \underline{\beta}_r = 0 \quad \text{for } r \le i$$

Now define a function

$$\phi = \underline{\beta}' \Sigma \underline{\beta} - \lambda(\beta' \underline{\beta} - 1) - \sum_{r=1}^{i} \gamma_r \underline{\beta}' \Sigma \beta_r$$

$$\therefore \frac{\partial \phi}{\partial \underline{\beta}} = 2 \Sigma \underline{\beta} - 2 \lambda \underline{\beta} - \sum_{r=1}^{i} \gamma_r \Sigma \underline{\beta}_r$$

again premultiplying it by β_j $(j \le i)$, we get

$$2 \beta_j \Sigma \underline{\beta} - 2 \lambda \underline{\beta}_j \underline{\beta} - \sum_{r=1}^{i} \gamma_r \underline{\beta}_j \Sigma \underline{\beta}_r$$

$$\Rightarrow \gamma_j = 0 \quad \text{for } j \le i$$

∴ the equation reduces to the form

$\Sigma \underline{\beta} - \lambda \underline{\beta} = 0$ or $(\Sigma - \lambda I)\beta = 0 \Rightarrow |\Sigma - \lambda I| = 0$

then, we choose λ_{i+1} th ch. roots and let β_{i+1} be the $(i+1)^{th}$ corresponding ch. vector then,

$\boxed{U_{i+1} = \underline{\beta}'_{i+1} X}$ $(i+1)^{th}$ P.Cs with which is independent of all the other i P.Cs

★ MLE'S of Principle Components

For a given multivariate data involving the var-cov matrix Σ, it may happen that Σ is unknown, hence we proceed with an estimate of Σ.

let $x_1, x_2, \ldots x_N$ $(N > p)$ be random sample of size N and we find the MLE of Σ as $\hat{\Sigma} = S$ then we have to find out the characteristics roots of S which are given by $|S - \lambda I| = 0$ eqn.

In general, the matrix S is non-singular and let $l_1, l_2, \ldots l_p$ be the ch. roots of S such that $l_1 \geq l_2 \geq \ldots \geq l_p$. The ch. vector $\underline{b_i}$ corresponding to l_i will be obtaining by solving $|S - l_i I| = 0$ for $\underline{b_i}$ clearly $\underline{b_i}$ are such that $\underline{b}_i' \underline{b_i} = 1$

let $\underline{b_1}, \underline{b_2}, \ldots \underline{b_p}$ be the ch vector then the estimate of 'p' PCA^s are $\underline{b}_1' \underline{x}, \underline{b_2}' \underline{x}, \ldots \underline{b}_p' \underline{x}$

or $\hat{U} = \begin{pmatrix} \hat{\underline{b}}_1' \\ \hat{\underline{b}}_2' \\ \vdots \\ \hat{\underline{b}}_p \end{pmatrix} \underline{X}$ we may write

$B = [\underline{b_1}, \underline{b_2}, \ldots \underline{b_p}]$

then 'B' will be an orthogonal matrix

then, $B'SB = diag(l_1, l_2, \ldots, l_p) = D$ (say)

then,

$$S = (B')^{-1} D B^{-1} = B D B' \text{ (as } B \text{ is orthogonal)}$$

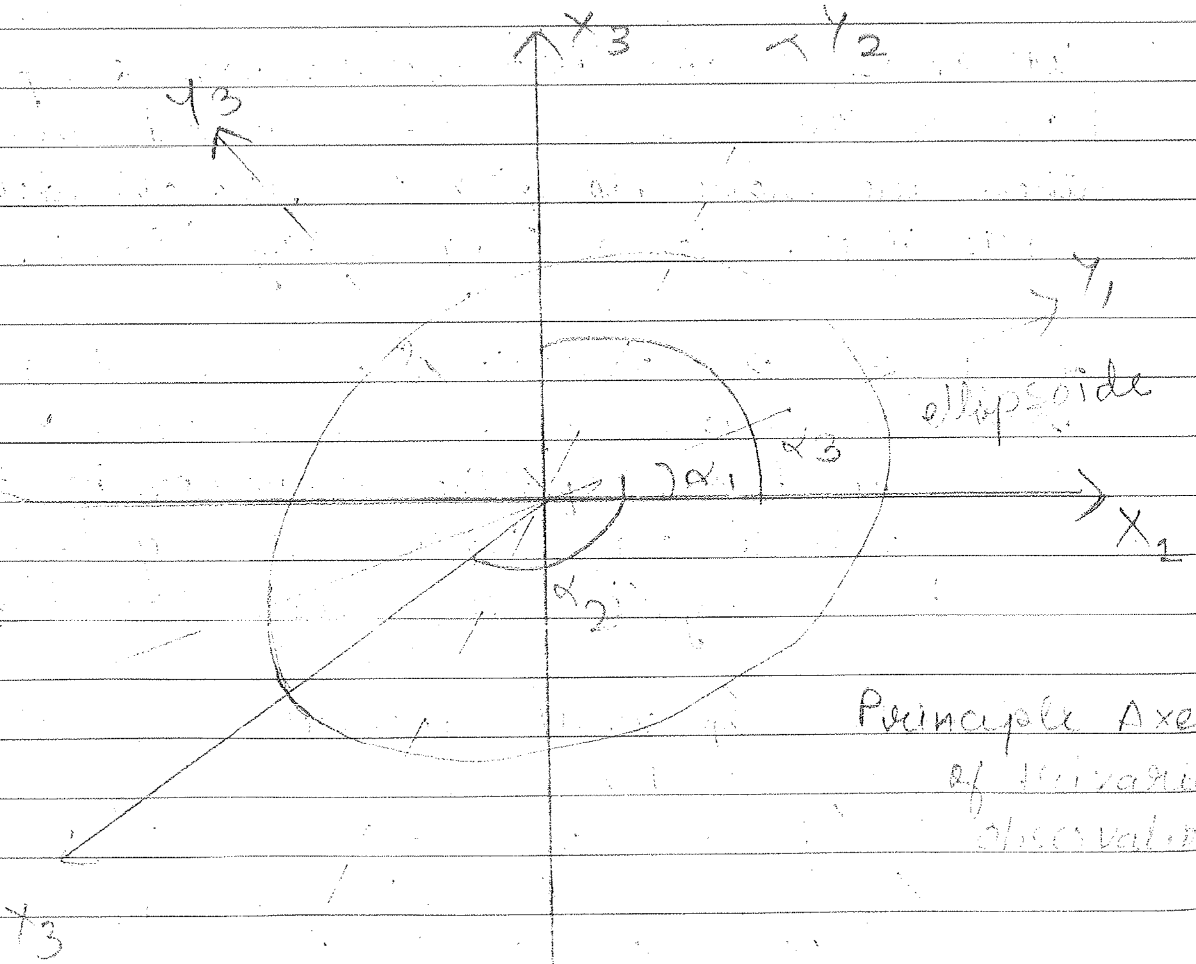
i.e. $S = (\underline{b}_1, \underline{b}_2, \ldots, \underline{b}_p) \begin{pmatrix} l_1 & 0 & \cdots & 0 \\ 0 & l_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & l_p \end{pmatrix} \begin{pmatrix} \underline{b}_1' \\ \underline{b}_2' \\ \vdots \\ \underline{b}_p' \end{pmatrix}$

i.e. $\boxed{S = l_1 \underline{b}_1 \underline{b}_1' + l_2 \underline{b}_2 \underline{b}_2' + \cdots + l_p \underline{b}_p \underline{b}_p'}$

if the rank of $S = r$ ($r < p$) then

$$\boxed{S = l_1 \underline{b}_1 \underline{b}_1' + \cdots + l_r \underline{b}_r \underline{b}_r'}$$

Geometrical Meaning Of Principle Component (Trivariate Set Of observation)



Principle Axes of trivariate observation

Imagine that a sample of N trivariate obs have the scatter plot shown as above where the origin of the response axes has been taken at the sample means.

The swarm of points seems to have generally an ellipsoid shape with MAJOR axis Y_1 and less well defined minor axis Y_2 and Y_3

Let us denote the angles with the original response axis as : $\alpha_1, \alpha_2, \alpha_3$, Y_1 passes through the sample mean points, its ~~obs~~ orientation is completely determined by the direction cosines.

$$\left.\begin{aligned} a_{11} &= \cos\alpha_1 \\ a_{21} &= \cos\alpha_2 \\ a_{31} &= \cos\alpha_3 \end{aligned}\right\} \quad -(1)$$

where $a_{11}^2 + a_{12}^2 + a_{13}^2 = 1$

It is known from the analytics geometry that, the value of the obs $[x_{l1}, x_{l2}, x_{l3}]$ on the new coordinate axis Y_1 will be given by

$$Y_{l1} = a_{11}(x_{l1} - \bar{x}_1) + a_{21}(x_{l2} - \bar{x}_2) + a_{31}(x_{l3} - \bar{x}_3) \quad -(2)$$

Now the mean of Y_1 is

$$\bar{Y}_1 = \frac{1}{N}\sum_{l=1}^{N} Y_{l1} = \frac{1}{N}\sum_{l=1}^{N}\sum_{j=1}^{3}(x_{lj} - \bar{x}_j)\, a_{j1} \quad -(3)$$

let us propose the notion of major axis as by defining that axis which passes through the direction of maximum variance in the points. In the present case of three responses that variance is :-

$$\frac{1}{(N-1)}\sum_{l=1}^{N} Y_{l1}^2 = \frac{1}{(N-1)}\sum_{l=1}^{N}\left[\sum_{j=1}^{3} a_{j1}(x_{lj}-\bar{x}_j)\right]^2 \quad -(4)$$

the angles of Y_1 could be found by differentiating this expression w.r.t a_{j1} and solving for the a_{j1} which makes the derivatives zero. The solution will be the characteristic vector of the greatest root of the sample VCM and Y_1 would be the continua of the continuum of the 1^{st} P.C of the system!
(continuum — 3^{rd} Image which contain 1^{st} P.C)

Definition:- The P.C of the sample of N, p-dimentional obs^n are the new variates specified by the axes of a rigid rotation of the original response coordinate system into an orientation corresponding to the directions of maximum variance in the sample scatter configuration.

The direction cosine of the new axes are normalised characteristic vectors corresponding to the successively a smaller characteristics roots of the sample covariance matrix.
If in a given situation two or more roots are equal the directions of the associated axes are not unique and it

may be therefore choosen in several positions

Therefore, it was suggested that if the components are computed using the correlation matrix instead of VCM the same geometrical interpretation holds and it would be easy to find out the Principle Component without overlapping.

Some Patterned Matrices And their P.C :-

(1) Equi - Correlation Matrix :-

In this case the $(p \times p)$ covariance matrix as

$$\Sigma = \sigma^2 = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & & & & \\ \rho & \rho & \cdots & \cdots & 1 \end{bmatrix} \quad \text{--- (1)}$$

we shall require $0 < \rho \leq 1$

The greatest ch. root of the matrix is

$$\lambda_1 = \sigma^2 [1 + (p-1)\rho] \quad \text{--- (2)}$$

and its Normalized ch. vector is -

$$\underline{a}_1' = \left[\frac{1}{\sqrt{p}}, \ldots \frac{1}{\sqrt{p}}\right] \quad \text{as} \quad \underline{a}_1 \underline{a}_1' = \begin{bmatrix} 1/\sqrt{p} \\ \vdots \\ 1/\sqrt{p} \end{bmatrix} \left[\frac{1}{\sqrt{p}} \cdots \frac{1}{\sqrt{p}}\right]$$

$$= \frac{1}{p} + \ldots + \frac{1}{p} = 1 \quad \text{--- (3)}$$

i.e. the 1^{st} P.C will be:

$$Y_1 = \frac{1}{\sqrt{p}} \sum_{j=1}^{p} X_j \quad \text{--- (4)}$$

i.e. it is merely proportional to the MEAN of the original p- responses and it accounts for

$$\frac{100 [1 + (p-1)\rho]}{p} \text{ of the}$$

total variance in the set

The remaining $(p-1)$ Ch. roots are all equal to : ~~σ^2 $1+\rho$~~ $\sigma^2 (1-\rho)$ --- (5)

and their corresponding vector are any of the $(p-1)$ linearly independent sol^n of the eq^n $\sigma^2 \rho (a_{12} + a_{22} + \dots + a_{p2}) = 0$ --- (6)

Note that this eq^n amounts to the requirement that each of the $(p-1)$ Ch. vector be orthogonal to the 1^{st} vector i.e

$$\underline{a_1}' \underline{a_2} = \underline{0} \quad \text{--- (7)}$$

\# Note: In many real life situation equi correlation matrix corresponding to the biological variables such as size of the living things

Eg consider the matrix $\rho = \begin{bmatrix} 1 & 0.6 & 0.6 & 0.6 \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$

assume that

$\sigma^2 = 1$

Here $\lambda_1 = [1 + (p-1)\delta] = 1 + (4-1) \times 0.6 = 2.8$

$\frac{100 \times 2.8}{4} = 70\%$ of the variation will be explained by 1st P.C since the ch. ~~root~~ vector is $\underline{a}_1' = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2})$

$\therefore Y_1 = \frac{1}{2}(X_1 + X_2 + X_3 + X_4)$.

The remaining 30% of the variance is EQUALLY attributable to the three new variates symmetrically distributed about Y_1 axis.

Principal Component Regression Analysis (PCRA)

Consider the regression model $Y = X\underline{\beta} + \underline{U}$ — (1)
where $\underline{Y}_{n\times1} \to$ dependent variable
X (n×p) matrix of p-explanatory variables,
β is a vector of p.m.s and
$\underline{U}$ is a vector of random errors. It is assumed that $E(\underline{U}) = 0$ and $E(UU') = \sigma^2 I_n$ then the OLS estimate of $\underline{\beta}$ is $\hat{\beta} = (X'X)^{-1}X'\underline{Y}$

This estimator exists subjected to the condition that the rank of X matrix is 'p' but this is violated if the explainatory variable are correlated. The perfect correlation among the predictors gives the $|X| = 0$ and in this case $(X'X)^{-1}$ does not exist offcourse g-inverse will exist and in such situation the OLS method of estimation break down.

Situation where the predictors are correlated is known as Multicolinearity problem where the variables $x_1, x_2, \ldots x_p$ may be not perfectly colinear but there may be some kind of relationship exists among them.

The problem of MC can be tackled if the columns of X matrix are 'orthogonalized' or the variables which cause the MC problem are droped from the analysis but droping of a variable may not be a wise decision because the variable being dropped may be important to study the variation in the dependent variable.

The principal component technique is used to orthogonolized the X matrix, in PCA the component are so selected that the selected component can explain maximum variation in Data. But in case of Regresion analysis the variables are selected in such a way that the correlation between y and selected variable(s) or component is maximum Mittle Shanmer and Barritelle (1977) have discussed the method of selection of principal component in regession analysis.

METHOD:

Without loss of generality let us consider that $\bar{y} = \bar{x}_j = 0$ and $var(\bar{y}) = var(\bar{x}_j) = 1$ $\quad j = 1, 2, \ldots p.$

and let $\underline{U} \sim N_n(0, \sigma^2 H)$ where $H = I - \frac{1}{n}\underline{1}\,\underline{1}'$ where $\underline{1} = (1, 1, \ldots 1)$ and

let $\Gamma = (\gamma_1, \gamma_2, \ldots \gamma_p)$ be an orthogonal matrix where γ_j is the j^{th} eigen vector corresponding to the eigen value λ_j. Let γ_{ij} be the i^{th} element of γ_j vector. If $\lambda_1 > \lambda_2 > \cdots > \lambda_p$ consider the transformation of X matrix to Z matrix s.t.

$Z = \Gamma' X$ where the elements of Z are $Z_1, Z_2, \ldots Z_p$. Now if we replace X by Z in (1) we have

$$Y = Z\Gamma'\beta + U = Z B + U \quad \text{—— (2)}$$

where $\beta = B = \Gamma'\beta$

$\therefore$ the LSE of B is: $\hat{B} = (Z'Z)^{-1} Z'Y$

This $\hat{B}$ is unbiased as $E(\hat{B}) = E\,(Z'Z)^{-1}Z'(ZB+U)$

$$= B + (Z'Z)^{-1}Z'\,E(U)$$

as $E(U) = 0$ $\qquad = B$

$$\boxed{E(\hat{B}) = B}$$

Also,

$$\text{Var}(\hat{B}) = E(\hat{B}-B)(\hat{B}-B)'$$

$$= E\,(Z'Z)^{-1}Z'UU'Z(Z'Z)^{-1}$$

$$= \Delta^{-1}\sigma^2 \text{ (say)}$$

$\therefore$ $\hat{B} \sim N_p(B, \Delta^{-1}\sigma^2)$ we can write

$$\hat{B}_i = \frac{1}{n\lambda_i} Z_i'Y \text{ and}$$

$$\boxed{\text{Var}(B_i) = \frac{\sigma^2}{n\lambda_i}} \quad i = 1, 2, \ldots p$$

10|03|2023

From this PCRA we obtain the sum of square due to regression

$$SS(\text{Regr}) = \hat{B}'Z'Y$$

and $$SS(\text{error}) = Y'Y - \hat{B}'Z'Y$$

an estimator of σ^2 is: $\dfrac{1}{(n-p-1)}\left[Y'Y - \hat{B}'Z'Y\right]$

Now, the test statistics for testing the sign of $H_0: B_i = 0$ is $t = \frac{\hat{B}_i}{S.E(\hat{B}_i)}$ we have

estimated $B \rightarrow$ the vector of pm's from this the original parameter vector β is estimated β, where $\boxed{\hat{\beta} = \Gamma \hat{B}}$

thus, $E(\hat{\beta}_i) = \sum_{j=1}^{p} \gamma_{ij} B_j$ with

$$var(\hat{\beta}_i) = \sigma^2 \sum_{j=1}^{p} \frac{\gamma_{ij}^2}{n\lambda_j}$$

from sample the estimator $\hat{\beta}_i$ is obtained by

$$\hat{\beta}_i = \sum_{j=1}^{p} g_{ij} B_j \quad \text{and} \quad var(\hat{\beta}_i) = \sum_{j=1}^{p} g_{ij}^2 / n l_j$$

where l_j is the j^{th} eigen value of $\frac{1}{n}$ ($X'X$ matrix)
g_{ij} : i^{th} element of eigen values corresponding to l_j ($j = 1, 2, .. p$)

UNIT 3

(One-way) MULTIVARIATE ANALYSIS OF VARIANCE : (MANOVA)

In many real life situation often more than than two population need to be compared, for this purpose random samples are calculated from each of the 'g' population (say) and these are arranged as:-

$$\left.\begin{array}{l} pop^n - 1 : x_{11}, x_{12}, \ldots x_{1n_1} \\ pop^n - 2 : x_{21}, x_{22}, \ldots x_{2n_2} \\ \vdots \\ pop^n - g : x_{g1}, x_{g2}, \ldots x_{gn_g} \end{array}\right\} \quad \text{---} \quad (1)$$

Assumptions :-

1. $X_{l1}, X_{l2}, \ldots X_{ln_l}$ is a random sample of size n_l from a popⁿ having mean μ_l, $l = 1, 2, \ldots g$ the r.s from different popⁿ are independent.

2. All the popⁿ have common covariance matrix Σ

3. Each population is multivariate normal then the hypothesis of interest is :-

$$H_0 : \begin{bmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{p1} \end{bmatrix} = \begin{bmatrix} \mu_{12} \\ \mu_{22} \\ \vdots \\ \mu_{p2} \end{bmatrix} = \cdots = \begin{bmatrix} \mu_{1g} \\ \mu_{2g} \\ \vdots \\ \mu_{pg} \end{bmatrix}$$

OR

$$H_0 : \underline{\mu}_1 = \underline{\mu}_2 = \cdots = \underline{\mu}_g$$

THE MODEL:

Analogy to univariate case, the model for comparing g-population mean vector is :

$$X_{lj} = \underline{\mu} + \tau_l + \underline{e}_{lj} \; ; \; j = 1, 2, \ldots n_l, \; l = 1, \ldots g$$

where $e_{lj} \sim N_p(0, \Sigma)$

μ = Overall mean vector

τ_l = lth treatment effect with $\sum_{l=1}^{g} n_l \tau_l = 0$

According to model (?), each component of the obs vector X_{lj} satisfies the assumption of univariate model. Now, a vector of observations may be decomposed (as suggested by the model)

Thus $\underline{x}_{lj} = \bar{\underline{x}} + (\bar{\underline{x}}_l - \bar{\underline{x}}) + (\underline{x}_{lj} - \bar{\underline{x}}_l)$ — (3)

(observ) overall sample mean $\downarrow \hat{\mu}$; estimated treatment effect $\rightarrow \hat{\tau}_l$; Residual $\downarrow \hat{e}_{lj}$

the decomposition in (3) leads to the Multivariate analogue of the univariate SS. break-up

Note that the cross-product (C.P) $(\underline{x}_{lj} - \bar{\underline{x}})(\underline{x}_{lj} - \bar{\underline{x}})'$ can be written as:-

$$(\underline{x}_{lj} - \bar{\underline{x}})(\underline{x}_{lj} - \bar{\underline{x}})' = [(\underline{x}_{lj} - \bar{\underline{x}}_l) + (\bar{\underline{x}}_l - \bar{\underline{x}})][(\underline{x}_{lj} - \bar{\underline{x}}_l) + (\bar{\underline{x}}_l - \bar{\underline{x}})]'$$

$$= (\underline{x}_{lj} - \bar{\underline{x}}_l)(\underline{x}_{lj} - \bar{\underline{x}}_l)' + (\underline{x}_{lj} - \bar{\underline{x}}_l)(\bar{\underline{x}}_l - \bar{\underline{x}})' + (\bar{\underline{x}}_l - \bar{\underline{x}})(\underline{x}_{lj} - \bar{\underline{x}}_l)' + (\bar{\underline{x}}_l - \bar{\underline{x}})(\bar{\underline{x}}_l - \bar{\underline{x}})'$$

$\sum_{j=1}^{n_l} (\underline{x}_{lj} - \bar{\underline{x}}_l) = 0 \rightarrow$ matrix

Next, $\sum_{l=1}^{g} n_l (\bar{\underline{x}}_l - \bar{\underline{x}}) = 0$ | as $\sum_l n_l \tau_l = 0$

∴ Summing the C.P over l and j gives:

$$\sum_{l=1}^{g} \sum_{j=1}^{n_l} (\underline{x}_{lj} - \bar{\underline{x}})(\underline{x}_{lj} - \bar{\underline{x}})'$$

Total SS

$$= \underbrace{\sum_{l=1}^{g} n_l (\bar{\underline{x}}_l - \bar{\underline{x}})(\bar{\underline{x}}_l - \bar{\underline{x}})'}_{\text{treat SS (Between)}} + \underbrace{\sum_{l=1}^{g} \sum_{j=1}^{n_l} (\underline{x}_{lj} - \bar{\underline{x}}_l)(\underline{x}_{lj} - \bar{\underline{x}}_l)'}_{\text{Residual S.S (within)}}$$

now, the with SS and C.P matrix can be expressed as:

$$W = \sum_{l=1}^{g} \sum_{j=1}^{n_l} (\underline{x}_{lj} - \bar{\underline{x}}_l)(\underline{x}_{lj} - \bar{\underline{x}}_l)'$$

$= (n_1-1)S_1 + (n_2-1)S_2 + \dots + (n_g-1)S_g$

where S_l is the sample cov. matrix for lth sample W matrix is the generalization of $(n_1+n_2-2)S_p$ matrix. Analogas to univariate result the hypothesis of No treatment effect is $H_0: \tau_1 = \tau_2 = \dots = \tau_g = 0$

MANOVA Table

Source of Variation	Degree of freedom.	Matrix of SS. & CP
Treatments	$(g-1)$	$B = \sum_{l=1}^{g} n_l(\bar{x}_l-\bar{x})(\bar{x}_l-\bar{x})'$
Residual (Error)	$(\sum_{l=1}^{g} n_l - g)$	$W = \sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj}-\bar{x}_l)(x_{lj}-\bar{x}_l)'$
Total	$(\sum_{l=1}^{g} n_l - 1)$	$B + W = \sum_{l=1}^{g}\sum_{j=1}^{n_l}(x_{lj}-\bar{x})(x_{lj}-\bar{x})'$

One of the tests to ~~the~~ test H_0 involves GENERALISED VARIANCES.

We reject H_0 is the ratio of the generalized variances $\Lambda^* = \frac{|W|}{|W+B|}$ is to small. Proposed by WILKS

Wilk Λ^* has relation with likelihood ratio criterian and the exact distribution of Λ^* can be derived for some special cases. However for large sample a modification of Λ^* is due to

Bartlett ~~He has shown that~~ BARTLET and can be used to test H_0.

Page No.: $\sim \chi^2_{p(g-1)}$ df

He has shown that $\left[(n-1) - \frac{(p+g)}{2}\right] \log_e \Lambda^* \sim \chi^2_{p(g-1)}$ df

MANOVA Two-way fixed effect Model (with Interaction)

→ (Univariate) → Denoting the r^{th} obser at the level 'l' of factor 1 and level 'k' of factor 2 by x_{lkr}, we may write a two-way univariate model as: $x_{lkr} = \mu + \tau_l + \beta_k + \gamma_{lk} + e_{lkr}$

$l = 1, \ldots, g$ and $k = 1, \ldots, b$, $r = 1, 2, \ldots n$

proceeding by analogy, we specify two-way fixed effect for a vector response consisting of p-component as:-

(fixed effect)

$$\underline{X}_{lkr} = \underline{\mu} + \underline{\tau}_l + \underline{\beta}_k + \underline{\gamma}_{lk} + \underline{e}_{lkr} \quad \text{---(1)}$$

$$\text{where } \sum_{l=1}^{g} \tau_l = \sum_{k=1}^{b} \beta_k = \sum_{l=1}^{g} \gamma_{lk} = \sum_{k=1}^{b} \gamma_{lk} = \underline{0}$$

the vectors are all of order $(p \times 1)$ and $\underline{e}_{lkr} \sim N_p(\underline{0}, \Sigma)$. thus, the responses consists of p-measurements repeated 'n' times at each possible combinations of level of factor - 1 and factor - 2. we can decompose the observation vector $\underline{X}_{lkr}$ as:

$$\underline{X}_{lkr} = \bar{\underline{X}} + (\bar{\underline{X}}_{l.} - \bar{\underline{X}}) + (\bar{\underline{X}}_{.k} - \bar{\underline{X}}) + (\bar{\underline{X}}_{lk} - \bar{\underline{X}}_{l.} - \bar{\underline{X}}_{.k} + \bar{\underline{X}}) + (\underline{X}_{lkr} - \bar{\underline{X}}_{lk}) \quad \text{--- (2)}$$

where $\bar{X}$: overall average of the obser. vectors

$\bar{X}_{\ell}$: overall average of the obser vectors at the ℓth level of factor 1

$\bar{X}_{k}$: Overall average of the observ vectors at ~~th~~ kth level of factor 2

$\bar{X}_{\ell k}$: Overall average of the observ vector at ℓth level of factor 1 and kth level of factor 2

We have :

$$\sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(X_{\ell kr}-\bar{X})(X_{\ell kr}-\bar{X})' =$$

$$\sum_{\ell=1}^{g} bn(\bar{X}_{\ell\cdot}-\bar{X})(\bar{X}_{\ell\cdot}-\bar{X})' + \sum_{k=1}^{b} gn(\bar{X}_{\cdot k}-\bar{X})(\bar{X}_{\cdot k}-\bar{X})'$$

$$+\sum_{\ell=1}^{g}\sum_{k=1}^{b} n(\bar{X}_{\ell k}-\bar{X}_{\ell\cdot}-\bar{X}_{\cdot k}+\bar{X})(\bar{X}_{\ell k}-\bar{X}_{\ell\cdot}-\bar{X}_{\cdot k}+\bar{X})'$$

$$+\sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(X_{\ell kr}-\bar{X}_{\ell k})(X_{\ell kr}-\bar{X}_{\ell k})' \quad \text{--- (3)}$$

degree of freedom

$$(gbn-1) = (g-1)+(b-1)+(g-1)(b-1)+gb(n-1) \quad \text{--- (4)}$$

MANOVA Table

Source of Variation	Degree of freedom	Matrix of SS and C.P
Factor - 1	$(g-1)$	S.S fact 1 = $\sum_{l=1}^{g} bn(\bar{X}_{l.}-\bar{X})(\bar{X}_{l.}-\bar{X})'$
Factor - 2	$(b-1)$	S.S fact 2 = $\sum_{k=1}^{b} gn(\bar{X}_{.k}-\bar{X})(\bar{X}_{.k}-\bar{X})'$
Interaction	$(g-1)(b-1)$	S.S. Interaction = $\sum_{l=1}^{g}\sum_{k=1}^{b} n(\bar{X}_{lk}-\bar{X}_{l.}-\bar{X}_{.k}+\bar{X})(\bar{X}_{lk}-\bar{X}_{l.}-\bar{X}_{.k}+\bar{X})'$
Residual	$gb(n-1)$	S.S. Residual = $\sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{l=1}^{n}(X_{lkr}-\bar{X}_{lk})(X_{lkr}-\bar{X}_{lk})'$
Total	$(gbn-1)$	SS Total = $\sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{l=1}^{n} = (X_{lkr}-\bar{X})(X_{lkr}-\bar{X})'$

Hypothesis :-

A test of $H_0: \gamma_{11} = \gamma_{12} = \cdots = \gamma_{gb} = \underline{0}$
(No interaction effect)

ag

H_1: atleast one $\gamma_{lk} \neq 0$

is conducted we reject H_0 for smaller values of

$$\Lambda^* = \frac{|S.S.P\ res|}{|SS\ Int + SS\ Pres|} \quad \text{--- (5)}$$

Note: The LRT procedure requires that $p \leq - gb(n-1)$ So that SS residual will be a positive Definite matrix

For large sample Λ^* can be compared with χ^2-distribution using Bartlett's correction to improve χ^2- approximation. Therefore we rej H_0 if

$$-\left[gb(n-1) - \frac{(p+1) - (g-1)(b-1)}{2}\right] \log \Lambda^* \sim \chi^2_{(g-1)(b-1); \alpha\%} \quad \text{--- (6)}$$

Remark

→ In two way manova generally the test for ~~interested~~ interaction is carried out before testing the main factor effects

→ As if the interaction effect exists then the factor effects do not have clear interpretation Therefore from a practical interpretation. ~~Therefrom~~ Therefore, from a practical Standpoint it is not advisable to proceed with additional multivariate test

→ In multivariate MANOVA model we proceed for testing of factor-1 and factor 2 as the main effects if no interaction is present

First we consider the testing of

$H_0: \tau_1 = \tau_2 = \cdots = \tau_g = 0$

vs H_1: atleast one $\tau_l \neq 0$

This hypo specifies No factor - 1 effects and Some factors - 1 effects respectively.

The prop test stat is

$$\Delta^* = \frac{|S.S.P\ res|}{|S.S.P\ fac\ 1 + S.S.P\ res|} \quad \text{--- (7)}$$

→ Again using Bartlett's correction, the L.R for testing of No factor - 1 effects is conducted by rej H_0 at level α if

$$-\left[gb(n-1) - \frac{(p+1)(g-1)}{2}\right] \log \Delta^* > \chi^2_{(g-1)} \; \alpha\%$$

--- (8)

In similar manner i factor - 2 effects are tested by considering

$H_0: \beta_1 = \beta_2 = \cdots = \beta_k = 0$

vs H_1: atleast on $\beta_k \neq 0$,

then the test stat is:

$$\Delta = \frac{|S.S.P\ Res|}{|S.S.P\ fac\ 2 + S.S.P\ res|} \quad \text{--- (9)}$$

Again using Bartlett's correction the L.R for testing of No factor-2 effects is conducted by rej H_0 at level α if

$$-\left[gb(n-1) - \frac{(p+1)(b-1)}{2}\right] \log \Delta^* > \chi^2_{(b-1);\,\alpha\%}. \quad -(10)$$

—x—