$$r \geq -1-(P-2)r$$

$$\cancel{r \geq -1-r_1}$$

$$r+(P-2)r \geq -1$$

$$r[1+P-2] \geq -1$$

$$r \geq -\frac{1}{P-1} \quad \text{(proved)}$$

## Multivariate Probability Structure

Recall random experiment

↓

classical probability

↓

Statistical probability

↓

Axiomatic prob.

Let $(\Omega, \mathcal{S}, P)$ be a triplet

where $\Omega$ = sample space

$\Omega = \{1,2,3\}$

$\mathcal{S}$ : $\sigma$ field of subsets from $\Omega$

$\mathcal{S} = \{\{1\},\{2\},\{3\},\{1,2\},\dots,\phi,\Omega\}$

Random variable

$X : \Omega \to \mathbb{R}$

Suppose $\omega \in \Omega\,(S)$

X real value function $\to X(\omega) = x_1 \in \mathbb{R}$

Random vector

Imagine out P random experiments (different/ repetition of same experiment)

For each of random experiment sample space there exists sample space.

---

$\Omega_1 = \{H, T\}$

$\Omega_2 = \{1, 2, 3, 4, 5, 6\}$

$\Omega = \Omega_1 \times \Omega_2$

$= \{(H, 1), (H, 2), (H, 3), \ldots\}$

---

Let us name for $i^{th}$ random experiment sample space be

$\Omega_i$, $i = 1(1)P$

A real valued vector function $\underset{P\times 1}{X}$ is called a random vector if

it maps $\cancel{X(\omega) \in \mathbb{R}^P}$

$\omega \in \Omega_1 \times \Omega_2 \times \cdots \times \Omega_p$ to p-dimensional real line $\mathbb{R}^p$ i.e. $\underset{\sim}{X}(\omega) \in \mathbb{R}^p$

<u>Example</u> : $\Omega_1$ = Coin tossing, $\{H, T\}$

$\Omega_2$ = die throwing
$\{1, 2, 3, 4, 5, 6\}$

A : having odd number and H.

$X(\omega_1) = X\{1, H\}$

$X(\omega_2) = X\{3, H\}$

$X(\omega_3) = X\{5, H\}$

~~$f(x) = y$~~

$P(X(\omega_1)) = \frac{3}{6} \times \frac{1}{2} = \frac{1}{4}$

$$\underset{\sim}{X}(\omega) = \begin{pmatrix} X(1), X(H) \\ X(3), X(H) \\ X(5), X(H) \end{pmatrix}$$

$P(A) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

We write $\underset{\sim}{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}_{p \times 1}$

<u>Cumulative distribution function</u>

For a random vector,
$\underset{\sim}{X}_{p\times 1}$ c.d.f is defined as.

$F_{\underset{\sim}{X}}(\underset{\sim}{x})$ = joint probabilistic structure

$= Pr\left( X_1 \le x_1 \cap X_2 \le x_2 \cap X_3 \le x_3 \cap \dots \cap X_p \le x_p \right)$

In case the random vector $\underline{X}$ is discrete valued.

The prob. mass function is defined as

$$P_{\underline{X}}(\underline{x})$$

$$= P_r\left(X_1 = x_1 \cap X_2 = x_2 \cap X_3 = x_3 \cap \dots \cap X_p = x_p\right)$$

$$= Pr\left(X_1 = x_1, X_2 = x_2, \dots, X_p = x_p\right)$$

Def: Let $X_{p\times 1}$ be a continuous random vector. $f_{\underline{X}}(\underline{x})$ is called the prob. density function of $\underline{X}$ if $f_{\underline{X}}(\underline{x}) \cdot d\underline{\mathbf{x}}$

$$\doteq Pr\left(\underline{x} - \Delta\underline{x}/2 < \underline{X} < \underline{x} + \Delta\underline{x}/2\right)$$

where $\Delta\underline{x} \to \underline{0}$ (null vector)

$$= \int_{\underline{X} \in (\underline{x} - \frac{1}{2}\Delta\underline{x}\,;\, \underline{x} + \frac{1}{2}\Delta\underline{x}) \subset \mathbb{R}^p} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

Rough

$$d\underline{\mathbf{x}} = \begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_p \end{pmatrix}$$

$$\underline{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

Remark 1: For continuous r.v.

$$F_{\underline{X}}(\underline{x}) = Pr\left(\underline{X} \le \underline{x}\right)$$

$$= Pr\left(X_1 \le x_1 \cap X_2 \le x_2 \cap \dots \cap X_p \le x_p\right)$$

$$= \int_{x_1=-\infty}^{x_1} \int_{x_2=-\infty}^{x_2} \cdots \int_{x_p=-\infty}^{x_p} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_{p_1} \\ X_{p_1+1} \\ \vdots \\ X_p \end{pmatrix} \begin{matrix} \} X_{(1)} \\ \} X_{(2)} \end{matrix}$$

Remark 2 : If $\underset{p\times 1}{\underline{X}} = \begin{pmatrix} X_{(1)\, p_1 \times 1} \\ \cdots \\ \underline{X}_{(2)\, p-p_1 \times 1} \end{pmatrix}$

the conditional probability density of $\underline{X}_{(1)} | \underline{X}_{(2)}$ $\cancel{(\underline{X}_{(1)} | \underline{X})}$ is

$$f_{\underline{X}}(\underline{x}) \Big| g_{\underline{X}_{(2)}}(\underline{x}_{(2)})$$

[joint pdf $\underline{X}_{(1)}$ & $\underline{X}_{(2)}$ is $f_{\underline{X}}(\underline{x})$]

where $f_{\underline{X}_{(1)}}(\underline{x}_{(1)})$ is the marginal pdf of $\underline{X}_{(1)}$

$\cancel{f_{\underline{X}}(\underline{x})} =$

$$\underline{X} = \begin{pmatrix} \left.\begin{matrix} X_1 \\ X_2 \end{matrix}\right\} \underline{X}_{(1)} \\ \left.\begin{matrix} X_3 \\ X_4 \\ X_5 \end{matrix}\right\} \underline{X}_{(2)} \end{pmatrix}$$

or $(X_{(1)} | X)$ is $f_{\underline{X}_{(1)}}(\underline{x}_{(1)}) \Big| f_{\underline{X}}(\underline{x})$

$$f_{\underline{X}_{(1)}}(\underline{x}_{(1)}) = P_r \left( x_1 - \frac{dx_1}{2} < x_1 < x_1 + \frac{dx_1}{2} \cap x_2 - \frac{dx_2}{2} < x_2 < x_2 + \frac{dx_1}{2} \right)$$

$$f_{\underline{X}_{(2)}}(\underline{x}_{(2)}) = P_r \left( x_3 - \frac{dx_3}{2} < x_3 < x_3 + \frac{dx_3}{2} \cap \cdots \cap x_5 - \frac{dx_5}{2} < x_5 < x_5 + \frac{dx_5}{2} \right)$$

- Expectation and variance covariance matrix of $\underline{X}$

Suppose $\underline{X} \sim f_{\underline{X}}(\underline{x})$

then $E(\underline{X})$ is the termwise expectation

i.e.,
$$E(\underline{X})_{p\times 1} = \begin{pmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{pmatrix} = \underline{\mu}_{p\times 1} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$$

where $\mu_i = E(X_i)$

$$V(\underline{X}) = E\left[(\underline{X}-\underline{\mu})_{p\times 1} (\underline{X}-\underline{\mu})'_{1\times p}\right]$$

Variance-covariance matrix of random vector $\underline{X}$

$$Var(\underline{X}) = E\left[\begin{pmatrix} X_1-\mu_1 \\ X_2-\mu_2 \\ \vdots \\ X_p-\mu_p \end{pmatrix} (X_1-\mu_1 \quad X_2-\mu_2 \quad \cdots \quad X_p-\mu_p)\right]$$

$$Var(\underline{X}) = \begin{pmatrix} E(X_1-\mu_1)^2 & E(X_1-\mu_1)(X_2-\mu_2) & \cdots & E(X_1-\mu_1)(X_p-\mu_p) \\ E(X_1-\mu_1)(X_2-\mu_2) & E(X_2-\mu_2)^2 & \cdots & E(X_2-\mu_2)(X_p-\mu_p) \\ \vdots & \vdots & & \vdots \\ E(X_1-\mu_1)(X_p-\mu_p) & E(X_2-\mu_2)(X_p-\mu_p) & \cdots & E(X_p-\mu_p)^2 \end{pmatrix}$$

If $\sigma_{ij} = Cov(X_i, X_j) = E(X_i - \mu_i)(X_j - \mu_j)$

and $\sigma_i^2 = E(X_i - \mu_i)^2 = Var(X_i)$, the above matrix

$$Var(\underline{X}) = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_p) \\ Cov(X_1, X_2) & Var(X_2) & \cdots & Cov(X_2, X_p) \\ \vdots & \vdots & & \vdots \\ Cov(X_1, X_p) & Cov(X_2, X_p) & \cdots & Var(X_p) \end{pmatrix}$$

$$Var(\underline{X}) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_2^2 & \cdots & \sigma_{2p} \\ \vdots & & & \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_p^2 \end{pmatrix}_{p \times p}$$

(1) $Var(\underline{X})$ is generally denoted by $\Sigma$

(2) $\Sigma$ is a symmetric matrix.

Exercise : Suppose $X_{3\times 1}$ be a random vector with mean vector $\underline{\mu} = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}$ and $Cov(X_i, X_j) = \frac{(i+j)}{2}$ $\forall (i, j)$

Construct $\Sigma$. Also find $E(2\underline{X})$ and $V(2\underline{X})$

$V(2\underline{X}) = E(2\underline{X} - \mu^*)(2\underline{X} - \mu^*)'$

$$\mu^* = E(2\underline{X})$$

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{pmatrix}$$

given, $\sigma_{ij} = cov(x_i, x_j) = \frac{i+j}{2}$

Therefore,

$$\Sigma = \begin{pmatrix} 1 & \frac{3}{2} & 2 \\ \frac{3}{2} & 2 & \frac{5}{2} \\ 2 & \frac{5}{2} & 3 \end{pmatrix}$$

Now, $E(2\underline{X}) = 2E(\underline{X}) = 2\underline{\mu} = \begin{pmatrix} 6 \\ -4 \\ 0 \end{pmatrix}$

$V(2\underline{X}) = E(2\underline{X} - \underline{\mu}^*)(2\underline{X} - \underline{\mu}^*)'$ where, $\mu^* = 2\underline{\mu}$

$= E(2\underline{X} - 2\underline{\mu})(2\underline{X} - 2\underline{\mu})'$

$$= A \cdot E\left[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'\right]$$

$$= 4\,V(\underline{X}) = 4\Sigma = \begin{pmatrix} 4 & 6 & 8 \\ 6 & 8 & 10 \\ 8 & 10 & 12 \end{pmatrix}$$

<u>Properties</u> : For a matrix $A_{m\times p}$ , ~~$V(A\underline{X})_{m\times p}$~~

$E(\underline{X}_{p\times1}) = \underline{\mu}$

$V(\underline{X}_{p\times1}) = \Sigma_{p\times p}$

$$V(A_{m\times p}\underline{X}_{p\times1}) = A_{m\times p}\, V(\underline{X}_{p\times1}) A'_{p\times m}$$

$$= A_{m\times p}\, \Sigma_{p\times p}\, A'_{p\times m}$$

<u>Pf</u> : $V(A\underline{X})$

$$= E\left[(A\underline{X} - E(A\underline{X}))(A\underline{X} - E(A\underline{X}))'\right]$$

$$= E\left[(A\underline{X} - A\underline{\mu})(A\underline{X} - A\underline{\mu})'\right]$$

$$= E\left[A(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'A'\right]$$

$$= A\left[E\left\{(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'\right\}\right]A'$$

$$= A\Sigma A'$$

Rough

$(A\underline{X} - A\underline{\mu})'$

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

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

$[(AB)' = B'A']$

In particular if $A$ is of order $1\times p$

$$V(A_{1\times p}\underline{X}_{p\times1}) = A_{1\times p}\,\Sigma_{p\times p}\,A'_{p\times1}$$

<u>Prop 2</u> : $V(X_{p\times1})$ [any variance covariance matrix] is always positive definite.

- $\underset{n\times n}{\underset{\sim}{A}}$ is called to be positive definite matrix if for a vector $\underset{n\times 1}{\underset{\sim}{x}} \neq \underset{\sim}{0}$

$$\underset{\sim}{x}' A \underset{\sim}{x} > 0$$

- If there ~~might~~ exist ~~be~~ some non-null vector for which $\underset{\sim}{x}' A \underset{\sim}{x}$ might be zero i.e. $\underset{\sim}{x}' A \underset{\sim}{x} \geq 0$ then $\underset{\sim}{A}$ is called positive semi definite matrix.

---

<u>proof</u> >> Consider any non null vector

$$\underset{\sim}{a}' \neq \underset{\sim}{0}$$

Then a Q.F (Quadratic form)

$$\underset{\sim}{a}'_{1\times p} V(\underset{\sim}{x}) \underset{\sim}{a}_{p\times 1} = \underset{\sim}{a}' \Sigma \underset{\sim}{a}$$

$$= \underset{\sim}{a}' E\left[(\underset{\sim}{x} - E(\underset{\sim}{x}))(\underset{\sim}{x} - E(\underset{\sim}{x}))'\right]\underset{\sim}{a}$$

$$= E\left[\underset{\sim}{a}'(\underset{\sim}{x} - E(\underset{\sim}{x}))(\underset{\sim}{x} - E(\underset{\sim}{x}))'\underset{\sim}{a}\right]$$

$$= E\left[(\underset{\sim}{a}'\underset{\sim}{x} - E(\underset{\sim}{a}'\underset{\sim}{x}))(\underset{\sim}{a}'\underset{\sim}{x} - E(\underset{\sim}{a}'\underset{\sim}{x}))'\right]$$

Let us define $\underset{1\times 1}{y} = \underset{1\times p}{\underset{\sim}{a}'} \underset{p\times 1}{\underset{\sim}{x}} - E(\underset{\sim}{a}'\underset{\sim}{x})$

Then the above is $E(y.y) = E(y^2)$

~~$y^2$ is~~ $y^2 > 0$ for any real $y$

$\therefore E(y^2) > 0$

$\Rightarrow \underline{a}' \Sigma \underline{a} > 0$

$\Rightarrow \Sigma$ is positive definite (p.d.)

If $\Sigma$ is positive semi definite i.e. $|\Sigma| = 0$ then the distribution of $\underline{X}$ is called <u>singular</u> distribution which is difficult to deal with. We deal with only non singular multivariate distribution.

For singular distribution, the entire multivariate probability stands on a p-dimensional ~~plane~~ line. $[ax_1 + bx_2 + cx_3 + \dots + dx_p = K]$

## Dependence and Independence

Two random vectors $\underset{p\times 1}{\underline{X}_{(1)}}$ and $\underline{X}_{(2)\,p\times 1}$ are said to be pairwise independent if

$$F_{\underline{X}_{(1)}, \underline{X}_{(2)}}(\underline{x}_{(1)}, \underline{x}_{(2)}) = F_{\underline{X}_{(1)}}(\underline{x}_{(1)}) \cdot F_{\underline{X}_{(2)}}(\underline{x}_{(2)})$$

$$\forall (\underline{x}_{(1)} \text{ and } \underline{x}_{(2)})$$

$$f_{\underline{X}_{(1)}\underline{X}_{(2)}}(\underline{x}_{(1)}, \underline{x}_{(2)}) = f_{\underline{X}_{(1)}}(\underline{x}_{(1)}) \cdot f_{\underline{X}_{(2)}}(\underline{x}_{(2)})$$

Jt. p.d.f. of $\underline{X}_{(1)}$ and $\underline{X}_{(2)}$

marginal pdf. of $\underline{X}_{(1)}$

marginal pdf of $\underline{X}_{(2)}$

- mutual independence of random variable

$$f_{X,Y,Z}(x,y,z) = f_X(x)\cdot f_Y(y)\cdot f_Z(z)$$

- mutual independence of ~~probability~~ n events

$$P(A_i \cap A_j) = P(A_i)\cdot P(A_j)$$

$$P(A_i \cap A_j \cap A_k) = P(A_i)\cdot P(A_j)\cdot P(A_k)$$

$$\vdots$$

total combination

$$\binom{n}{2} + \binom{n}{3} + \binom{n}{4} + \cdots + \binom{n}{n}$$

$$= 2^n - n - 1$$ no. of conditions satisfies

---

**Properties:** Three random vectors $\underline{X}_{(1)}, \underline{X}_{(2)}$ and $\underline{X}_{(3)}$ are ~~cally~~ called mutually independent if

$$f_{\underline{X}_{(1)}, \underline{X}_{(2)}, \underline{X}_{(3)}}\left(\underline{x}_{(1)}, \underline{x}_{(2)}, \underline{x}_{(3)}\right) = f(\underline{x}_{(1)})\cdot f(\underline{x}_{(2)})\cdot f(\underline{x}_{(3)})$$

$$\forall\ \underline{x}_{(1)}, \underline{x}_{(2)}, \underline{x}_{(3)}$$

pairwise independence $\not\Rightarrow$ mutual independence

EX: Consider a joint p.d.f

$$f_{\underline{X}=(X_1,X_2,X_3)}(x_1,x_2,x_3) = \frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2}(x_1^2+x_2^2+x_3^2)} \cdot \left(1 + x_1x_2x_3 e^{-\frac{1}{2}(x_1^2+x_2^2+x_3^2)}\right.$$

$$, -\infty < X_i < \infty$$

Show that $X_1, X_2, X_3$ are pairwise independent but not mutual ind.

$$\rightarrow f_{X_1, X_2}(x_1, x_2) = \int_{x_3=-\infty}^{\infty} f_{\underline{X}}(\underline{x})\, dx_3$$

$$= \frac{1}{(2\pi)^{3/2}} \int_{x_3=-\infty}^{\infty} \left[ e^{-\frac{1}{2}(x_1^2 + x_2^2 + x_3^2)} + x_1 x_2 x_3 e^{-(x_1^2 + x_2^2 + x_3^2)} \right] dx_3$$

$$= \frac{1}{(2\pi)^{3/2}} \left[ e^{-\frac{1}{2}(x_1^2 + x_2^2)} \cdot \sqrt{2\pi} + x_1 x_2 e^{-\frac{(x_1^2 + x_2^2)}{2}} \cdot \int_{x_3=-\infty}^{\infty} x_3 e^{-x_3^2/2}\, dx_3 \right]$$

$$\left[ \int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi} \quad \text{and} \quad \int_{x_3=-\infty}^{\infty} x_3 \cdot e^{-x_3^2/2}\, dx_3 = E(X_3) = 0 \right]$$

Standard normal

$$= \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}$$

$$\therefore f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} \qquad f_{X_3}(x_3) = \frac{1}{\sqrt{2\pi}} e^{-x_3^2/2}$$

$$f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi}} e^{-x_2^2/2}$$

So they are pairwise independent.

But, $f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot f_{X_3}(x_3) \neq f_{X_1,X_2,X_3}(x_1,x_2,x_3)$

not mutually independent.

## Multiple Regression theory (In probability version)

Suppose a random vector $\underset{\sim}{X}_{p\times 1}$ has mean vector as $\underset{\sim}{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$ and variance covariance matrix $V(\underline{X}) = \Sigma_{p\times p}$,

with corresponding correlation matrix $\rho$.

We define, $\sigma_{ij} = E\left[(X_i - \mu_i)\cdot(X_j - \mu_j)\right] = Cov(X_i, X_j)$

where $X_i$ and $X_j$ are the $i^{th}$ and $j^{th}$ component of $\underset{\sim}{X}$ and so is $\rho_{ij} = corr(X_i, X_j)$

$\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ X_j \\ \vdots \\ X_p \end{pmatrix}$

Also, $\sigma_i^2 = V(X_i)$, $\mu_i = E(X_i)$

Our objective is to study the association or dependence of $X_1$ on $X_2, X_3, \ldots, X_p$ i.e. the association of $X_1$ on $\underset{(p-1)\times 1}{\underline{X}_{(2)}} = \begin{pmatrix} X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix}$

This association or dependence is measured by conditional expectation of $X_1$ on $\underline{X}_{(2)}$

3) i.e. $E(X_1 / \underline{X}_{(2)})$

From regression theory this conditional expectation is called regression equation.

$$f_{X,Y}(x,y) = cxy$$

$y$ on $x$

$$E(Y|x) = \int_x y \cdot f_{Y|X}(y|x)\, dy$$

Suppose the regression form is linear i.e.

$$E(X_1 / \underline{X}_{(2)}) = \alpha + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_p$$

i) Thus the model is

$$X_1 = \alpha + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_p + \epsilon$$

$$= (\alpha \ \beta_2 \ \beta_3 \ \beta_4 \cdots \beta_p) \begin{pmatrix} 1 \\ X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix} + \epsilon$$

Using least square principle on ;

$E(X_1 - E(X_1 | \underline{X}_{(2)}))^2$ by minimizing to be

i.e. to minimize $E(X_1 - \alpha - \beta_2 X_2 - \cdots - \beta_p X_p)^2$ by parameters

$\alpha, \beta_2, \ldots, \beta_p$ resp.

1st normal equation

$$\frac{\partial}{\partial \alpha} E(X_1 - \alpha - \beta_2 X_2 - \cdots - \beta_p X_p)^2 = 0$$

$$\Rightarrow \mu_1 = \alpha + \beta_2\mu_2 + \cdots + \beta_p \mu_p = (1\ \beta_2 \cdots \beta_p)\begin{pmatrix}\alpha \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_p\end{pmatrix} = (1\ \underline{\beta}')\begin{pmatrix}\alpha \\ \underline{\mu}_{(2)}\end{pmatrix}$$

where, $\underline{\beta}_{p-1\times1}$ and $\underline{\mu}_{(2)} = \begin{pmatrix}\mu_2 \\ \mu_3 \\ \vdots \\ \mu_p\end{pmatrix}_{p-1\times1}$

The other normal equations

2nd Nor. Equ$^n$.

$$(-2)\, E\left(X_2\left(X_1 - \alpha - \beta_2 X_2 - \cdots - \beta_p X_p\right)\right) = 0$$

$$\Rightarrow E\Big(X_2\big(X_1 - (\mu_1 - \beta_2\mu_2 - \cdots - \beta_p\mu_p) - \beta_2 X_2 - \cdots - \beta_p X_p\big)\Big) = 0$$

$\Rightarrow$ ~~$E(X_1 X_2) - \mu_1 E(X_2)$~~

$$\Rightarrow E\Big[X_2\big(X_1 - \mu_1 - \beta_2(X_2 - \mu_2) - \beta_3(X_3 - \mu_3) - \cdots - \beta_p(X_p - \mu_p)\big)\Big] = 0$$

Rough

$$E(X_2(X_1 - \mu_1))$$

$$= E\big((X_2 - \mu_2)(X_1 - \mu_1) + \mu_2(X_1 - \mu_1)\big)$$

$$= E(X_2 - \mu_2)(X_1 - \mu_1) + \mu_2 \underbrace{E(X_1 - \mu_1)}_{0}$$

$$= E(X_2 - \mu_2)(X_1 - \mu_1)$$

$$= \sigma_{12}$$

$$\Rightarrow E(X_2(X_1 - \mu_1)) - \beta_2 E(X_2(X_2 - \mu_2)) - \cdots - \beta_p E(X_p(X_p - \mu_p)) = 0$$

$$\Rightarrow \sigma_{12} = \beta_2 \sigma_2^2 + \beta_3 \sigma_{23} + \cdots + \beta_p \sigma_{2p}$$

The other normal equations:

$$\sigma_{12} = \beta_2 \sigma_2^2 + \beta_3 \sigma_{23} + \cdots + \beta_p \sigma_{2p}$$

$$\sigma_{13} = \beta_2 \sigma_{23} + \beta_3 \sigma_3^2 + \cdots + \beta_p \sigma_{3p}$$

$$\vdots$$

$$\sigma_{1p} = \beta_2 \sigma_{2p} + \beta_3 \sigma_{3p} + \cdots + \beta_p \sigma_p^2$$

$$\Rightarrow \begin{pmatrix} \sigma_{12} \\ \sigma_{13} \\ \vdots \\ \sigma_{1p} \end{pmatrix} = \begin{pmatrix} \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2p} \\ \sigma_{23} & \sigma_3^2 & \cdots & \sigma_{3p} \\ \vdots & & & \\ \sigma_{2p} & \sigma_{3p} & \cdots & \sigma_p^2 \end{pmatrix} \begin{pmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\text{If } \Sigma = \left(\begin{array}{c|cccc} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \hline \sigma_{21} & \sigma_{22} & \sigma_{23} & \cdots & \sigma_{2p} \\ \vdots & & \ddots & & \\ & \vdots & & & \\ \sigma_{p1} & \sigma_{p2} & \sigma_{p3} & \cdots & \sigma_{pp} \end{array}\right)$$

$$= \begin{pmatrix} \sigma_{11\ 1\times 1} & \underline{\sigma}'_{1\times p-1} \\ \underline{\sigma}_{p-1\times 1} & \Sigma_{2\ p-1\times p-1} \end{pmatrix}$$

Then the above is

$$\boxed{\underline{\sigma} = \Sigma_2 \underline{\beta}}$$

From the above $\underline{\beta}_{p-1\times 1} = \Sigma_2^{-1} \underline{\sigma}$

provided, $\Sigma_2$ is a non singular matrix.

$$\text{Thus } \beta_j = \frac{\begin{vmatrix} \sigma_{22} & \sigma_{23} & \cdots & \sigma_{2,j-1} & \sigma_{21} & \sigma_{2,j+1} & \cdots & \sigma_{2p} \\ \vdots & & & \vdots & \sigma_{31} & & & \vdots \\ \vdots & & & \vdots & \vdots & & & \\ \sigma_{p2} & \sigma_{p2} & & \sigma_{p,j-1} & \sigma_{p1} & \sigma_{p,j+1} & & \sigma_{pp} \end{vmatrix}}{\begin{vmatrix} \sigma_{22} & \sigma_{23} & \cdots & \sigma_{2p} \\ \vdots & & & \\ \sigma_{p2} & \cdots & & \sigma_{pp} \end{vmatrix}}$$

(Cramer's Rule)

$$= \frac{(-1)^{j-2}\begin{vmatrix} \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2,j-1} & \sigma_{2,j+1} & \cdots & \sigma_{2p} \\ \sigma_{13} & \sigma_{23} & \cdots & \sigma_{3,j-1} & \vdots & & \vdots \\ \vdots & \vdots & & \sigma_{p,j-1} & \sigma_{p,j+1} & & \sigma_{pp} \\ \sigma_{1p} & \sigma_{2p} & & & & & \end{vmatrix}}{\begin{vmatrix} & & '' & & \end{vmatrix}}$$

$$= (-1)^{j-2}\frac{\sigma_1}{\sigma_j}\frac{\rho_{1,j}}{\rho_{1,1}} = -\frac{\sigma_1}{\sigma_j}\cdot\frac{\rho_{1,j}}{\rho_{1,1}}$$

where, $\rho_{i,j}$ is the cofactor of $(i,j)$ element from correlation matrix $\rho$.

Multiple regression line or conditional expectation

$$E\left(X_1 \mid \underline{X}^{(2)}\right) = \mu_1 - \sum_{j=2}^{p}\frac{\sigma_1}{\sigma_j}(X_j - \mu_j)\frac{\rho_{1,j}}{\rho_{1,1}}$$

$$X_1 = \mu_1 - \sum_{j=2}^{p}\frac{\sigma_1}{\sigma_j}(X_j - \mu_j)\frac{\rho_{1,j}}{\rho_{1,1}}$$

Remark

Remember $\rho = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix}^{-1} \Sigma \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix}$

$$= \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p)^{-1} \Sigma \, \text{diag}(\sigma_1 \sigma_2, \ldots \sigma_p)^{-1}$$

---

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & & & \\ \sigma_{p1} & \cdots & & \sigma_{pp} \end{pmatrix}$$

$$= \begin{pmatrix} \rho_{11}\sigma_1\sigma_1 & \rho_{12}\sigma_1\sigma_2 & \cdots & \rho_{1p}\sigma_1\sigma_p \\ \vdots & & & \\ \vdots & & & \rho_{pp}\sigma_p\sigma_p \\ \rho_{p1}\sigma_p\sigma_1 & \cdots & & \end{pmatrix}$$

$$\cancel{\sigma_1 \sigma_2} \cdots$$

$$\rho = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \vdots & & & & \\ \vdots & & & & \\ \rho_{p1} & \cdots & & & 1 \end{pmatrix}$$

Rough

---

## Multiple Correlation :

Def >> Measure of association between $X_1$ and the set of independent variables $(X_2, \ldots, X_p)$

It is denoted by,

$$\rho_{1 \cdot 23 \ldots p} = \frac{Cov(X_1, X_{1 \cdot 23 \ldots p})}{\sqrt{V(X_1) V(X_{1 \cdot 23 \ldots p})}}$$

where the model is

$$X_1 = X_{1.23\cdots p} + \varepsilon_{1.23\cdots p}$$

We assume $cov\left(X_{1.23\cdots p}, \varepsilon_{1.23\cdots p}\right) = 0$

and $X_{1.23\cdots p} = \alpha + \beta_2 X_2 + \cdots + \beta_p X_p$

where $\alpha, \beta_2, \cdots, \beta_p$ are estimated by least square principle.

$$\therefore E\left(X_{1.23\cdots p}\right) \cancel{=}$$

$$= E\left(\mu_1 - \Sigma \frac{\sigma_1}{\sigma_j} \cdot (x_j - \mu_j) \frac{P_{1,j}}{P_{1,1}}\right)$$

$$= \mu_1$$

$$cov\left(X_1, X_{1.23\cdots p}\right) = Cov\left(X_{1.23\cdots p} + \varepsilon_{1.23\cdots p}, X_{1.23\cdots p}\right)$$

$$= V\left(X_{1.23\cdots p}\right) + 0$$

$$= V\left(X_{1.23\cdots p}\right)$$

$$\varepsilon_{1.23\cdots p} = X_1 - X_{1.23\cdots p}$$

$$\cancel{= X_1}$$

$$= X_1 - \left(\alpha + \beta_2 X_2 + \cdots + \beta_p X_p\right)$$

$$= X_1 - \left(\mu_1 - \beta_2 \mu_2 - \beta_3 \mu_3 - \cdots - \beta_p \mu_p + \beta_2 X_2 + \cdots + \beta_p X_p\right)$$

$$= (1, -\beta_2, -\beta_3, \ldots, -\beta_p) \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{pmatrix}$$

$$= (1, -\underline{\beta}') \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{pmatrix}$$

$V(\varepsilon_{1.23\ldots p})$

$$= V\left[ (1, -\underline{\beta}') \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{pmatrix} \right]$$

$$= (1, -\underline{\beta}') \, V \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{pmatrix} (1, -\underline{\beta}')'$$

$$= (1, -\underline{\beta}') \, \Sigma \begin{pmatrix} 1 \\ -\underline{\beta} \end{pmatrix}$$

$$= (1, -\underline{\beta}') \begin{pmatrix} \sigma_{11} & \underline{\sigma}' \\ \underline{\sigma} & \Sigma_2 \end{pmatrix} \begin{pmatrix} 1 \\ -\underline{\beta} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11} - \underline{\beta}'\underline{\sigma} & \underline{\sigma}' - \underline{\beta}'\Sigma_2 \end{pmatrix} \begin{pmatrix} 1 \\ -\underline{\beta} \end{pmatrix}$$

$$= \sigma_{11} - \underline{\beta}'\underline{\sigma} - \underline{\sigma}'\underline{\beta} + \underline{\beta}'\Sigma_2\underline{\beta}$$

$$= \cancel{\sigma_{11} - 2\underline{\beta}'\underline{\sigma}} = \sigma_{11} - 2\underline{\beta}'\underline{\sigma} + \underline{\beta}'\underline{\sigma}$$ [As from normal equ$^n$ $\underline{\sigma} = \Sigma_2 \underline{\beta}$]

$$= \sigma_{11} - \underline{\beta}'\underline{\sigma}$$

Also, $V(X_{1.23\cdots p})$

$$= \cancel{V(X_1)} = V(X_1 - \varepsilon_{1.23\cdots p})$$

$$= V(X_1) + V(\varepsilon_{1.23\cdots p}) - 2\,Cov(X_1, \varepsilon_{1.23\cdots p})$$

$$\boxed{\begin{aligned} &Cov(X_{1.23\cdots p} + \varepsilon_{1.23\cdots p}, \varepsilon_{1.23\cdots p}) \\ &= V(\varepsilon_{1.23\cdots p}) \end{aligned}}$$

$$= V(X_1) - V(\varepsilon_{1.23\cdots p})$$

$V(X_{1.23\cdots p})$

$$= \sigma_{11} - \sigma_{11} + \underline{\beta}'\underline{\sigma} = \underline{\beta}'\underline{\sigma}$$

$$\therefore \rho_{1.23\cdots p} = \frac{Cov(X_1, X_{1.23\cdots p})}{\sqrt{V(X_1)\cdot V(X_{1.23\cdots p})}}$$

$$= \sqrt{\frac{V(X_{1\cdot 23\cdots p})}{V(X_1)}} = \sqrt{\frac{\underline{\beta}'\underline{\sigma}}{\sigma_{11}}}$$

$$= \sqrt{\frac{\underline{\sigma}'\Sigma_2^{-1}\underline{\sigma}}{\sigma_{11}}}$$

$\underline{\beta} = \Sigma_2^{-1}\underline{\sigma}$

$\underline{\beta}' = \underline{\sigma}'(\Sigma_2^{-1})'$

$= \underline{\sigma}'\Sigma_2^{-1}$

(as $\Sigma_2$ sym)

---

$$\rho = \left(\begin{array}{c|cccc} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} \cdots & \rho_{1p} \\ \hline \rho_{21} & \rho_{22} & \rho_{23} & \cdots & \rho_{2p} \\ \rho_{31} & & & & \\ \vdots & & & & \\ \rho_{p1} & \rho_{p2} & \cdots & & \rho_{pp} \end{array}\right)_{p\times p} = \left(\begin{array}{c|c} 1 & \underline{\rho}'_{1\times p-1} \\ \hline \underline{\rho}_{p-1\times 1} & \rho_{2\cdot p-1\times p-1} \end{array}\right)$$

Splitting the correlation matrix

$\rho$ in $\begin{pmatrix} \rho_{11} & \underline{\rho}'_{1\times p-1} \\ \underline{\rho} & \rho_2 \end{pmatrix}$ using the conversion rule

$$\Sigma_{p\times p} = \text{diag}(\sigma_1\ \sigma_2 \cdots \sigma_p)\ \rho_{p\times p}\ \text{diag}(\sigma_1 \cdot \sigma_2 \cdots \sigma_p)$$

we have

$$\Sigma_2 = \text{diag}(\sigma_2, \ldots, \sigma_p)\ \rho_2\ \text{diag}(\sigma_2, \sigma_3 \ldots \sigma_p)$$

$$\Sigma = \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix} \rho \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p \end{pmatrix}$$

$$\Sigma_2^{-1} = \operatorname{diag}\left(\frac{1}{\sigma_2}\ \frac{1}{\sigma_3} \cdots \frac{1}{\sigma_p}\right) \rho_2^{-1} \operatorname{diag}\left(\frac{1}{\sigma_2}\ \frac{1}{\sigma_3} \cdots \frac{1}{\sigma_p}\right)$$

$$\rho_{1.23\cdots p} = \sqrt{\frac{\sigma_1^2 (\sigma_2\rho_{12}\ \sigma_3\rho_{13} \cdots \sigma_p\rho_{1p}) \operatorname{diag}\left(\frac{1}{\sigma_2} \cdots \frac{1}{\sigma_p}\right) \rho_2^{-1} \operatorname{diag}\left(\frac{1}{\sigma_2} \cdots \frac{1}{\sigma_p}\right)}{\sigma_1^2}}$$

---

Rough

$$\sigma_2\rho_{12}\ \sigma_3\rho_{13} \cdots \sigma_p\rho_{1p} \begin{pmatrix} \frac{1}{\sigma_2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_3} & 0 \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \frac{1}{\sigma_p} & 0 \end{pmatrix}$$

$$= (\rho_{12}\ \rho_{13} \cdots \rho_{1p})$$

---

$$= \sqrt{(\rho_{12}\ \rho_{13} \cdots \rho_{1p}) \rho_2^{-1} \begin{pmatrix} \rho_{12} \\ \rho_{13} \\ \vdots \\ \rho_{1p} \end{pmatrix}}$$

$$= \sqrt{\underline{\rho}' \rho_2^{-1} \underline{\rho}}$$

$$\rho^2_{1.23\cdots p} = \frac{\underline{\sigma}' \Sigma_2^{-1} \underline{\sigma}}{\sigma_{11}}$$

$$\begin{pmatrix} \sigma_2 \rho_{12} \\ \vdots \\ \sigma_p \rho_{1p} \end{pmatrix}$$

$$1 - \rho^2_{1\cdot 23\ldots p} = \frac{\sigma_{11} - \underline{\sigma}' \Sigma_2^{-1} \underline{\sigma}}{\sigma_{11}}$$

where $\Sigma = \left(\begin{array}{c|c} \sigma_{11} & \underline{\sigma}' \\ \hline \underline{\sigma} & \Sigma_2 \end{array}\right)$

$$|\Sigma| = |\Sigma_2| \left(\sigma_{11} - \underline{\sigma}' \Sigma_2^{-1} \underline{\sigma}\right)$$

$$1 - \rho^2_{1\cdot 23\ldots p} = \frac{|\Sigma_2| \left(\sigma_{11} - \underline{\sigma}' \Sigma_2^{-1} \underline{\sigma}\right)}{|\Sigma_2|\, \sigma_{11}}$$

$$\boxed{\Sigma^{-1} = \frac{\text{Adj } \Sigma}{|\Sigma|} \quad \text{Cofactor of } \sigma_{11} \text{ is } \Sigma_2}$$

$$= \frac{|\Sigma|}{|\Sigma_2|\, \sigma_{11}}$$

$$= \frac{1}{\frac{|\Sigma_2|}{|\Sigma|} \sigma_{11}} = \frac{1}{\sigma^{11} \sigma_{11}}$$

where $\sigma^{11}$ being the $(1,1)$ element in $\Sigma^{-1}$

$$\therefore \rho^2_{1\cdot 23\ldots p} = \left(1 - \frac{1}{\sigma^{11} \sigma_{11}}\right)$$

$$\boxed{\rho_{1\cdot 23\ldots p} = \sqrt{1 - \frac{1}{\sigma^{11} \sigma_{11}}}}$$

Ex: Let $X$ and $Y$ be such that, pdf

$$f_{XY}(x,y) = cx^2y \,, \quad 0 < Y \le X \le 1$$

(i) Find $c$

(ii) Find $P\left(Y \le \frac{X}{2}\right)$

(iii) Find $P\left(Y \le \frac{X}{4} \,\middle|\, Y \le \frac{X}{2}\right)$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y) = 1$$

(i) $$\int_{x=0}^{1}\int_{y=0}^{x} cx^2y \, dy \, dx = 1$$

$$\Rightarrow \int_{x=0}^{1} cx^2 \cdot \frac{y^2}{2}\Big|_0^x \, dx = 1$$

$$\Rightarrow \int_{x=0}^{1} cx^2 \cdot \frac{x^2}{2} \, dx = 1$$

$$\Rightarrow \frac{c}{2} \frac{x^5}{5}\Big|_0^1 = 1$$

$$\Rightarrow \frac{c}{10} = 1 \rightarrow \boxed{c = 10}$$

(ii) $P(Y \le \frac{X}{2})$

$$= \int_{x=0}^{1}\int_{y=0}^{x/2} 10x^2 y \, dy \, dx$$

$$= \int_{x=0}^{1} 10x^2 \int_{y=0}^{x/2} y \, dy \, dx$$

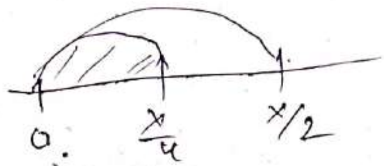
$$= \int_{x=0}^{1} 10x^2 \left.\frac{y^2}{2}\right|_{0}^{x/2} dx$$

$$= \int_{x=0}^{1} 10x^2 \cdot \frac{1}{2}\left[\frac{x^2}{4}\right] dx$$

$$= \frac{10}{8}\int_{x=0}^{1} x^4 \, dx$$

$$= \frac{\cancel{10}^{2}}{\underset{4}{\cancel{8}}} \cdot \frac{1}{\cancel{5}}$$

$$= \frac{1}{4}$$




(iii) $P\left(Y \le \frac{X}{4} \mid Y \le \frac{X}{2}\right)$

$$= \frac{P\left(Y \le \frac{X}{4} \cap Y \le \frac{X}{2}\right)}{P\left(Y \le \frac{X}{2}\right)} = \frac{P\left(Y \le \frac{X}{4}\right)}{P\left(Y \le \frac{X}{2}\right)}$$

$$P\left(Y \le \frac{X}{4}\right) = \int_{x=0}^{1} 10x^2 \cdot \frac{1}{2}\left[\frac{x^2}{16}\right] dx$$

$$= \frac{\cancel{10}^{2}}{32} \cdot \frac{1}{\cancel{5}} = \frac{1}{16}$$

Find $P\left(Y \le \frac{X}{4} \mid Y \le \frac{X}{2}\right)$ if $-1 < Y \le X \le 1$

if $-1 < x < 0$: $P\left(Y \le \frac{x}{4} \mid Y \le \frac{x}{2}\right) = P\left(Y \le \frac{x}{2}\right)$

$$= \begin{cases} 1 & \text{if } -1 < x < 0 \\ 1/16 & \text{if } 0 < x < 1 \end{cases}$$

Ex : Find the regression curve of Y on X for (X, Y) whose j.t pdf

$$f(x,y) = Ke^{-x(y+1)}, \quad x \geq 0, \; y \geq 0$$

$$E(Y|X) = \int_{y=0}^{\infty} y f(y|x)\, dy$$

(dependent range)

$$f(y|x) = \frac{f(x,y)}{g(x)}$$ $\longrightarrow$ marginal of $x$

$$g(x) = k\int_{y=0}^{\infty} e^{-x(y+1)}\, dy$$

$$\int_0^{\infty}\int_0^{\infty} Ke^{-x(y+1)}\, dy\, dx$$

$$= \int_0^{\infty}\int_0^{\infty} K(e^{-xy} \cdot e^{-x})\, dy\, dx.$$

$\ln x = z$

$\Rightarrow \frac{1}{x} dx = dz$

$$= \int_0^{\infty} Ke^{-x}\left[-\frac{e^{-xy}}{x}\right]_0^{\infty} dx$$

$$= \int_0^{\infty} Ke^{-x} \cdot \left(\frac{1}{x}\right) dx$$

$$= k\int_0^{\infty} x^{-1} e^{-x}\, dx \Rightarrow \underline{\text{undefined}}.$$

Ex: $f(x,y) = \begin{cases} 1 & \text{if } |Y| < X, \quad 0<X<1 \\ 0 & \end{cases}$

Show that regression function of Y on X is linear but regression function of X on Y is nonlinear.

To ~~show~~ find

$\rightarrow E(Y|X)$

$= \int_{y=-x}^{x} y f(y|x)\, dy$

$|Y| < X$

$-X < Y < X$

to find $f(y|x)$

$f(y|x)$

$= \dfrac{f(x,y)}{g(x)}$

$g(x) = $ p.d.f. of X

$= \int_{|y|=-x}^{x} 1\, dy$

$= 2\int_0^x dy = 2x$

(even func.)

$= \dfrac{1}{2x}, |y| < X$

~~$f(x,y) =$~~

$$E(Y|X) = \int_{y=-x}^{x} y \cdot \frac{1}{2x}\, dy$$

$$= \frac{1}{2x} \cancel{2\int_{y=0}^{x} y\, dy} \cdot 0$$

$$= 0 \quad (y \text{ odd fun})$$

$E(Y|X) = 0$

Model is $Y=0$ is nothing but the X axis which is a straight line.

$E(X|Y)$

$$= \int_{|y|}^{1} x \cdot f(x|y)\, dx$$

$$f(x|y) = \frac{f(x,y)}{g(y)}$$

$g(y)$ = pdf of $Y$

$$= \int_{x=|y|}^{1} 1\, dx = 1-|y|$$

$$f(x|y) = \frac{1}{1-|y|}$$

$$E(X|Y) = \int_{|y|}^{1} \frac{x}{1-|y|}\, dx = \frac{1}{1-|y|} \cdot \frac{x^2}{2}\Big|_{|y|}^{1}$$

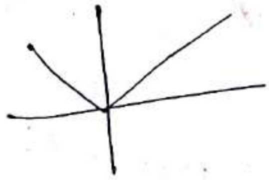
$$= \frac{1}{2} \frac{(1-|y|^2)}{1-|y|}$$

Model is $X = \frac{1}{2}(1+|y|)$

$\therefore$

$X = \frac{1}{2}(1-y)$

$X = \frac{1}{2}(1+y)$



. Regression function of $X$ on $Y$ is nonlinear.

## Partial Correlation :

For a multiple regression set up where

$X_1 = \alpha + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_p X_p$ is the linear relation between $X_1$ and $\underline{X}_{(2)} = \begin{pmatrix} X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix}$

We are interested to calculate the partial correlation between $X_1$ and $X_2$, i.e., $\rho_{12.34\dots p} = \text{Corr}(\varepsilon_{1.34\dots p}, \varepsilon_{2.34\dots p})$

where $\varepsilon_{1.34\dots p}$ being the error for the model

$$X_1 = X_{1.34\dots p} + \varepsilon_{1.34\dots p}$$

and $\varepsilon_{2.34\dots p}$ being the error for the model

$$X_2 = X_{2.34\dots p} + \varepsilon_{2.34\dots p}$$

For the submodel,

$$X_{1.34\ldots p} = \alpha + \beta_3 X_3 + \cdots + \beta_p X_p$$

the normal equations are

$$\mu_1 = \alpha + \beta_3 \mu_3 + \cdots + \beta_p \mu_p$$

$$\underline{\sigma}_{(1)}' = \Sigma_3 \underline{\beta} \Rightarrow \underline{\beta} = \Sigma_3^{-1} \underline{\sigma}_{(1)}'$$

$$\text{where } \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \underline{\sigma}_{(1)}' \\ \sigma_{21} & \sigma_{22} & \underline{\sigma}_{(2)}' \\ \underline{\sigma}_{(1)} & \underline{\sigma}_{(2)} & \Sigma_3 \end{pmatrix}$$

$$\underline{\beta} = \begin{pmatrix} \beta_3 \\ \beta_4 \\ \vdots \\ \beta_p \end{pmatrix}$$

$$\underline{\sigma}_{(1)}' = (\sigma_{13}\ \sigma_{14} \cdots \sigma_{1p})$$

$$\underline{\sigma}_{(2)}' = (\sigma_{23}\ \sigma_{24} \cdots \sigma_{2p})$$

Similarly for the second submodel

$$X_{2.34\ldots p} = \alpha^0 + \beta_3^0 X_3 + \cdots + \beta_p^0 X_p$$

In the performance of least square estimation process the normal equations are -

$$\mu_2 = \alpha^0 + \beta_3^0 \mu_3 + \cdots + \beta_p^0 \mu_p$$

$$\underline{\sigma}_2' = \Sigma_3 \underline{\beta}^0, \quad \underline{\beta}^0 = \begin{pmatrix} \beta_3^0 \\ \beta_4^0 \\ \vdots \\ \beta_p^0 \end{pmatrix}$$

$$\varepsilon_{1.34\ldots p} = X_1 - X_{1.34\ldots p}$$

$$= (X_1 - \mu_1) - (\beta_3\ \beta_4 \cdots \beta_p) \begin{pmatrix} X_3 - \mu_3 \\ X_4 - \mu_4 \\ \vdots \\ X_p - \mu_p \end{pmatrix}$$

$$= (x_1 - \mu_1) - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix}$$

$$\left(\Sigma_3^{-1} \underline{\sigma}_{(1)}\right)' = \underline{\sigma}_{(1)}' \Sigma_3^{-1}$$

$$\varepsilon_{2.34\cdots p} = X_2 - X_{2.34\cdots p}$$

$$= (x_2 - \mu_2) - (\beta_3^\circ \ \beta_4^\circ \cdots \beta_p^\circ) \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix}$$

$$= (x_2 - \mu_2) - \underline{\sigma}_{(2)}' \Sigma_3^{-1} \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix}$$

$$\operatorname{cov}\left(A_{m\times n} X_{n\times 1}, B_{m\times n} Y_{n\times 1}\right) = A \operatorname{cov}(X,Y) B'$$

$$\operatorname{cov}(\varepsilon_{1.34\cdots p}, \varepsilon_{2.34\cdots p})$$

$$= \operatorname{cov}\left[ (x_1 - \mu_1) - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix}, \ (x_2 - \mu_2) - \underline{\sigma}_{(2)}' \Sigma_3^{-1} \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix} \right]$$

$$= \operatorname{cov}\left[ (x_1 - \mu_1), (x_2 - \mu_2) \right] - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \operatorname{cov}\left[ \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_p \\ \vdots \\ x_p - \mu_p \end{pmatrix}, (x_2 - \mu_2) \right]$$

$$- \operatorname{cov}\left[ (x_1 - \mu_1), \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_p \\ \vdots \\ x_p - \mu_p \end{pmatrix} \right] \Sigma_3^{-1} \underline{\sigma}_2$$

P.T.O

$$+\underline{\sigma}_{(1)}'\Sigma_3^{-1}\, Cov\left[\begin{pmatrix}X_3-\mu_3\\ X_4-\mu_4\\ \vdots\\ X_p-\mu_p\end{pmatrix}, \begin{pmatrix}X_3-\mu_3\\ X_4-\mu_4\\ \vdots\\ X_p-\mu_p\end{pmatrix}\right]\Sigma_3^{-1}\underline{\sigma}_{(2)}$$

$$=\sigma_{12}-\underline{\sigma}_{(1)}'\Sigma_3^{-1}\underline{\sigma}_{(2)}-\underline{\sigma}_{(1)}'\Sigma_3^{-1}\underline{\sigma}_{(2)}+\underline{\sigma}_{(1)}'\Sigma_3^{-1}\Sigma_3\Sigma_3^{-1}\underline{\sigma}_{(2)}$$

---

$$\Sigma_3=\begin{pmatrix}\sigma_{33} & \sigma_{34} & \cdots & \sigma_{3p}\\ \sigma_{43} & \sigma_{44} & \cdots & \sigma_{4p}\\ \vdots & & & \\ \sigma_{p3} & \sigma_{p4} & \cdots & \sigma_{pp}\end{pmatrix}$$

$$Cov\left[\begin{pmatrix}X_3-\mu_3\\ X_4-\mu_4\\ \vdots\\ X_p-\mu_p\end{pmatrix}, \begin{pmatrix}X_3-\mu_3\\ X_4-\mu_4\\ \vdots\\ X_p-\mu_p\end{pmatrix}\right]$$

$$=\begin{pmatrix}Cov(X_3-\mu_3, X_3-\mu_3) & Cov(X_3-\mu_3, X_4-\mu_4) & \cdots\\ \vdots & & \\ Cov(X_p-\mu_p, X_3-\mu_3) & \cdots & \end{pmatrix}$$

---

$$=\sigma_{12}-\underline{\sigma}_{(1)}'\Sigma_3^{-1}\underline{\sigma}_{(2)}$$

$$V\left(\varepsilon_{1.34\cdots p}\right)$$

$$= V\left[(x_1 - \mu_1) - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \begin{pmatrix} x_3 - \mu_3 \\ x_4 - \mu_4 \\ \vdots \\ x_p - \mu_p \end{pmatrix}\right]$$

$$\boxed{\begin{aligned} &V(AX) \\ &= AV(x)A' \end{aligned}}$$

$$= \sigma_{11} + \underline{\sigma}_{(1)}' \Sigma_3^{-1} \Sigma_3 \Sigma_3^{-1} \underline{\sigma}_{(1)}$$

$$- 2\, \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(1)}$$

$$= \sigma_{11} - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(1)}$$

$$V\left(\varepsilon_{2.34\cdots p}\right) = \sigma_{22} - \underline{\sigma}_{(2)}' \Sigma_3^{-1} \underline{\sigma}_{(2)}$$

$$\rho_{12.34\cdots p} = \frac{\operatorname{Cov}\left(\varepsilon_{1.34\cdots p}, \varepsilon_{2.34\cdots p}\right)}{\sqrt{V\left(\varepsilon_{1.34\cdots p}\right) V\left(\varepsilon_{2.34\cdots p}\right)}}$$

$$= \frac{\sigma_{12} - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(2)}}{\sqrt{\left(\sigma_{11} - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(1)}\right) \cdot \left(\sigma_{22} - \underline{\sigma}_{(2)}' \Sigma_3^{-1} \underline{\sigma}_{(2)}\right)}}$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \underline{\sigma}_{(1)}' \\ \sigma_{21} & \sigma_{22} & \underline{\sigma}_{(2)}' \\ \underline{\sigma}_{(1)} & \underline{\sigma}_{(2)} & \Sigma_3 \end{pmatrix}$$

rough

$$= \frac{|\Sigma_3| \left(\sigma_{12} - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(2)}\right)}{\sqrt{|\Sigma_3| \left(\sigma_{11} - \underline{\sigma}_{(1)}' \Sigma_3^{-1} \underline{\sigma}_{(1)}\right) . |\Sigma_3| \left(\sigma_{22} - \underline{\sigma}_{(2)}' \Sigma_3^{-1} \underline{\sigma}_{(2)}\right)}}$$

$$= \frac{\begin{vmatrix} \sigma_{12} & \underline{\sigma}_{(1)}' \\ \underline{\sigma}_{(2)} & \Sigma_3 \end{vmatrix}}{\sqrt{\begin{vmatrix} \sigma_{11} & \underline{\sigma}_{(1)}' \\ \underline{\sigma}_{(1)} & \Sigma_3 \end{vmatrix} \begin{vmatrix} \sigma_{22} & \underline{\sigma}_{(2)}' \\ \underline{\sigma}_{(2)} & \Sigma_3 \end{vmatrix}}}$$

$$= \frac{-\Sigma_{21}}{\sqrt{\Sigma_{22}\,\Sigma_{11}}}$$

where $\Sigma_{ij}$ being the cofactor of $(i, j)$ element in $\Sigma$.

$$\rho_{12\cdot 34\cdots p} = \frac{-\Sigma_{21}/|\Sigma|}{\sqrt{\Sigma_{22}/|\Sigma| \cdot \Sigma_{11}/|\Sigma|}}$$

$$= \frac{\sigma^{21}}{\sqrt{\sigma^{11}\,\sigma^{22}}}$$

$\sigma^{ij}$ being the $(i, j)$ element in $\Sigma^{-1}$.

- Prove that $\rho_{12\cdot 34\cdots p} = \dfrac{-P_{2,1}}{\sqrt{P_{22}\,P_{11}}}$

$\boxed{\sigma_{12} = \sigma_1 \cdot \sigma_2 \cdot \rho_{12}}$

$$\Sigma = \text{diag}\begin{pmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_p\end{pmatrix} P \,\text{diag}\left(\sigma_1\ \sigma_2 \ldots \sigma_p\right)$$

$$P = \begin{pmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} & \cdots & \rho_{1p} \\ \rho_{21} & \rho_{22} & \rho_{23} & & \cdots & \rho_{2p} \\ \vdots & \vdots & & & & \\ & & & & & \rho_{pp} \\ \rho_{p1} & \rho_{p2} & \rho & \cdots & & \end{pmatrix}_{p\times p}$$

$$P_{2,1} = \begin{pmatrix} \rho_{12} & \rho_{13} & \rho_{14} & \cdots & \rho_{1p} \\ \rho_{32} & \rho_{33} & & \cdots & \rho_{3p} \\ \vdots & & & & \\ \rho_{p2} & \rho_{p3} & & \cdots & \rho_{pp} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \cdots & \sigma_{2p} \\ \vdots & & & & \\ \sigma_{p1} & \sigma_{p2} & \sigma_{p3} & \cdots & \sigma_{pp} \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_1 \sigma_1 \rho_{11} & \sigma_1 \sigma_2 \rho_{12} & \sigma_1 \sigma_3 \rho_{13} & \cdots & \sigma_1 \sigma_p \rho_{1p} \\ \sigma_1 \sigma_2 \rho_{21} & \sigma_2^2 \rho_{22} & & \cdots & \sigma_2 \sigma_p \rho_{2p} \\ \vdots & & & & \\ \sigma_1 \sigma_p \rho_{p1} & \sigma_2 \sigma_p \rho_{p2} & & \cdots & \sigma_p^2 \rho_{pp} \end{pmatrix}$$

$$\Sigma_{2,1} = \begin{vmatrix} \sigma_1 \sigma_2 \rho_{12} & \sigma_1 \sigma_3 \rho_{13} & \cdots & \sigma_1 \sigma_p \rho_{1p} \\ \sigma_2 \sigma_3 \rho_{23} & \sigma_3 \sigma_3 \rho_{33} & \cdots & \sigma_3 \sigma_p \sigma_{3p} \\ \vdots & & & \\ \sigma_2 \sigma_p \rho_{p2} & \sigma_3 \sigma_p \rho_{p3} & \cdots & \sigma_p^2 \rho_{pp} \end{vmatrix}$$

$$= \sigma_1 \cancel{\sigma_2} \sigma_3 \cdots \sigma_p \cdot \sigma_2 \sigma_3 \cdots \sigma_p \; | \; \rho_{2,1}$$

$$= \sigma_1 \sigma_2 \sigma_3^2 \cdots \sigma_p^2 \, P_{2,1}$$

$$\Sigma_{1,1} = \begin{vmatrix} \sigma_2^2 \rho_{22} & \sigma_2 \sigma_3 \rho_{23} & \cdots & \sigma_2 \sigma_p \sigma_{2p} \\ \sigma_2 \sigma_3 \rho_{32} & \sigma_3^2 \rho_{33} & \cdots & \sigma_3 \sigma_p \sigma_{3p} \\ \vdots & & & \vdots \\ \sigma_2 \sigma_p \sigma_{p2} & \sigma_p \sigma_3 \rho_{p3} & \cdots & \sigma_p^2 \sigma_{pp} \end{vmatrix}$$

$$= \sigma_2 \sigma_3 \cdots \sigma_p \cdot \sigma_2 \sigma_3 \cdots \sigma_p \, P_{1,1}$$

$$= \sigma_2^2 \sigma_3^2 \cdots \sigma_p^2 \, P_{1,1}$$

similarly, $\Sigma_{2,2} = \sigma_1^2 \sigma_3^2 \cdots \sigma_p^2 \, P_{2,2}$

$$\rho_{12.34\cdots p} = \frac{-\Sigma_{21}}{\sqrt{\Sigma_{22} \, \Sigma_{11}}}$$

$$= \frac{-\sigma_1 \sigma_2 \sigma_3^2 \sigma_4^2 \cdots \sigma_p^2 \, P_{2,1}}{\sqrt{\sigma_2^2 \sigma_3^2 \cdots \sigma_p^2 P_{1,1} \cdot \sigma_1^2 \sigma_3^2 \cdots \sigma_p^2 P_{2,2}}}$$

$$= \frac{-P_{2,1}}{\sqrt{P_{1,1} \, P_{2,2}}}$$

## Result

Multiple correlation coefficient is the maximum correlation between $X_1$ and any other linear combination of $X_2, X_3, \ldots, X_p$.

Multiple $\rho_{1.23\ldots p}$ (always larger)

Partial $\rho_{12.34\ldots p}$

Total $\rho_{12}$

$$(1-r^2_{1.23\ldots p}) = (1-r^2_{12.34\ldots p})(1-r^2_{13.45\ldots p})\cdots(1-r^2_{1p})$$

Pf >> Let us observe for any arbitrary $\underset{\sim}{\beta}^* = \begin{bmatrix} \beta_2^* \\ \beta_3^* \\ \vdots \\ \beta_p^* \end{bmatrix}$ from a linear combination, viz, $\underline{\beta}^{*\prime}\underline{X}_{(2)}$

where, $\underline{X}_{(2)} = \begin{pmatrix} X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix}$

$$Cov(X_1, \underline{X}_{(2)}) = Cov\left(X_1, \begin{pmatrix} X_2 \\ X_3 \\ \vdots \\ X_p \end{pmatrix}\right) = \begin{pmatrix} Cov(X_1, X_2) \\ Cov(X_1, X_3) \\ \vdots \\ Cov(X_1, X_p) \end{pmatrix} = \begin{pmatrix} \sigma_{12} \\ \sigma_{13} \\ \vdots \\ \sigma_{1p} \end{pmatrix}$$

Now, $cov.(X_1, \underline{\beta}^{*\prime}\underline{X}_{(2)})$

$= \underline{\beta}^{*\prime}\underline{\sigma}$ where, $\underline{\sigma} = \begin{pmatrix} \sigma_{21} \\ \sigma_{31} \\ \vdots \\ \sigma_{p1} \end{pmatrix}$

$= \underline{\beta}^{*\prime}\Sigma_2\underline{\beta}$

where $\underline{\beta}$ is the vector of least square estimate and from 2nd to pth normal equations in multiple regression setup yield $\underline{\sigma} = \Sigma_2\underline{\beta}$

Also from multiple regression setup we know

$$\mathrm{Cov}(X_1, \underline{\beta}' \underline{X}_{(2)})$$

$$= \underline{\beta}' \Sigma_2 \underline{\beta}$$

$$\boxed{\begin{aligned} &\mathrm{Cov}(X_1, \underline{X}_{(2)}) \\ &= \begin{pmatrix} \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_1, X_3) \\ \mathrm{Cov} \ \vdots \\ \mathrm{Cov}(X_1, X_p) \end{pmatrix} \\ &= \sigma = \Sigma_2 \underline{\beta} \end{aligned}}$$

$$\therefore \rho^2_{X_1, \underline{\beta}^{*\prime} \underline{X}_{(2)}} = \mathrm{Corr}^2(X_1, \underline{\beta}^{*\prime} \underline{X}_{(2)})$$

$$= \frac{\mathrm{Cov}^2(X_1, \underline{\beta}^{*\prime} \underline{X}_{(2)})}{\mathrm{Var}(X_1)\, V(\underline{\beta}^{*\prime} \underline{X}_{(2)})}$$

$$= \frac{(\underline{\beta}^{*\prime} \Sigma_2 \underline{\beta})^2}{\sigma_1^2 \cdot \underline{\beta}^{*\prime} \Sigma_2 \underline{\beta}^*}$$

$$\rho^2_{X_1, \underline{\beta}' \underline{X}(2)} = \mathrm{Corr}^2(X_1, \underline{\beta}' X_{(2)})$$

$$= \frac{\mathrm{Cov}^2(X_1, \underline{\beta}' \underline{X}_{(2)})}{V(X_1)\, V(\underline{\beta}' \underline{X}_{(2)})}$$

$$= \frac{(\underline{\beta}' \Sigma_2 \underline{\beta})^2}{\sigma_1^2 \cdot \underline{\beta}' \Sigma_2 \underline{\beta}}$$

By generalized cauchy schwart inequality,

for $\underset{n\times 1}{\underline{a}}$, $\underset{n\times 1}{\underline{b}}$ and $\underset{n\times n}{A}$ (matrix)

$$(\underset{1\times n}{\underline{a}'} \underset{n\times n}{A} \underset{n\times 1}{\underline{b}})^2 \le (\underline{a}' A \underline{a})(\underline{b}' A \underline{b})$$

$$\boxed{\begin{aligned} &x_1, x_2, \ldots x_n \\ &y_1, y_2, \ldots y_n \\ &\left(\sum x_i y_i\right)^2 \le \sum x_i^2 \sum y_i^2 \\ &(E(XY))^2 \le E(X^2) \cdot E(Y^2) \end{aligned}}$$

Choose $\underline{a} = \underline{\beta}^*$, $\underline{b} = \underline{\beta}$, $A = \Sigma_2$.

$$\left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}\right)^2 \leq \left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}^*\right) \cdot \left(\underline{\beta}' \Sigma_2 \underline{\beta}\right).$$

$$\Rightarrow \frac{\left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}\right)^2}{\left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}^*\right)} \leq \underline{\beta}' \Sigma_2 \underline{\beta}$$

$$\Rightarrow \frac{\left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}\right)^2}{\sigma_1^2 \left(\underline{\beta}^{*'} \Sigma_2 \underline{\beta}^*\right)} \leq \frac{\left(\underline{\beta}' \Sigma_2 \underline{\beta}\right)^2}{\sigma_1^2 \left(\underline{\beta}' \Sigma_2 \cdot \underline{\beta}\right)}$$

$$\Rightarrow \rho^{*2}_{X_1, \underline{\beta}^{*'} \underline{X}_{(2)}} \leq \rho^2_{X_1, \underline{\beta}' \underline{X}_{(2)}}$$

$\Rightarrow$ Correlation between $X_1$ and any other linear combination of $\underline{X}(2)$ $\leq$ multiple correlation (least square principle)

Example

Let $X_1, X_2, X_3, X_4$ be the number of spades, hearts, diamonds and clubs in a five card poker hand dealt from a well-shuffled deck. Find the joint pmf of $(X_1, X_2, X_3, X_4)$

$\rightarrow$ $P(X_1 = x_1 \cap X_2 = x_2 \cap X_3 = x_3 \cap X_4 = x_4)$ $\;\; x_4 = 5 - x_1 - x_2 - x_3$

$$= \frac{\binom{13}{x_1}\binom{13}{x_2}\binom{13}{x_3}\binom{13}{x_4}}{\binom{52}{5}} ; \quad x_i = 0,1,2,3,4,5, \; i = 1(1)4$$

$$x_1 + x_2 + x_3 + x_4 = 5$$

This is a three dimensional multivariate hyper geometric model.

Example

Show that for any $\beta_1, \beta_2, \ldots, \beta_p$.

$$E(X_1 - \beta_1 - \beta_2 X_2 - \beta_3 X_3 - \cdots - \beta_p X_p)^2$$

$$\geq \sigma^2 \frac{(p-1)\rho}{1+(p-2)\rho}$$ where the variance-covariance matrix of $(X_1, X_2, \ldots, X_p)$ is

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix}$$

$$E(\varepsilon_{1.23\cdots p}^2) = V(\varepsilon_{1.23\cdots p})$$

To show,

$$\rightarrow V(\varepsilon_{1.23\cdots p}) \geq \frac{(p-1)\rho}{1+(p-2)\rho}$$

$$\rho^2_{1.23\cdots p} = 1 - \frac{|\rho|}{\rho_{1,1}}$$

$$\boxed{1 - \frac{|R|}{R_{1,1}} = r^2_{1.23\cdots p}}$$

$$\Rightarrow \frac{V(X_{1.23\cdots p})}{V(X_1)} = \frac{V(X_1) - V(X_{1.23\cdots p})}{V(X_1)}$$

$$= \frac{V(\varepsilon_{1.23\cdots p})}{V(X_1)} = \frac{|\rho|}{\rho_{1,1}}$$

$$\Rightarrow V(\varepsilon_{1.23\cdots p}) = \sigma^2 \frac{|\rho|}{\rho_{1,1}} = \frac{\sigma^2 (1-\rho)^{p-1}(1+(p-1)\rho)}{(1-\rho)^{p-2}(1+(p-2)\rho)}$$

$$= \frac{\sigma^2(1-\rho) + (1-\rho)\,\rho(1-\rho)}{1+(p-2)\rho}$$

$$L.H.S \geq \frac{(1-\rho)(p-1)\rho}{1+(p-2)\rho}$$

1. <u>Multivariate Discrete Distribution</u>

Multinomial distribution

Binomial Experiment

$n$ bernoullian trial

↓

Binomial distribution

$X$ : No. of heads

$X \sim Bin(n, p)$

$Y$ : no. of tails

$X + Y = n$

↳ linear constraint

- Each trial has more than 2 possible outcomes, say each trial has $K$ outcomes.
- Probability of having the ith outcome is $p_i$

  obviously, $p_1 + p_2 + \dots + p_K = 1$

  where there are $K$ outcomes/categories.
- Suppose in a trial of $n$ individuals,

$X_1$ being the variable showing the no. of first category outcome, $X_2$ being the variable showing the no. of second category outcome, ..., $X_K$ being the variable showing the no. of $K^{th}$ category outcome. Each outcome is indep.

Joint probability mass function for this experiment can be

$$P[X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_K = x_K]$$

$$= p_1^{x_1} p_2^{x_2} \dots p_K^{x_K} \cdot \frac{n!}{x_1!\, x_2! \dots x_K!}$$

$$= \binom{n}{x_1\ x_2 \dots x_K} p_1^{x_1} p_2^{x_2} \dots p_K^{x_K},$$

$$0 \le x_i \le n, \quad i = 1(1)K$$

$$x_1 + x_2 + \dots + x_K = n$$

$$p_1 + p_2 + \dots + p_K = 1$$

We will say

$$\underset{K\times1}{\underline{X}} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{pmatrix} \sim \text{Multinomial}(n, p_1, p_2, \dots, p_K).$$

<u>Note</u> : For the above format, the dispersion matrix of $X$ will be singular as $X_K = n - X_1 - X_2 - \dots - X_{K-1}$.

thereby leading to create a set of linearly dependent vectors in dispersion matrix

Thus the above $\underline{X}$ would follow a <u>singular</u> multinomial distribution.

<u>Non singular multinomial distribution</u>

- Reduce the variable space from $\mathbb{R}^K$ to $\mathbb{R}^{K-1}$

$\underline{X}_{K\times 1}$

- Define $\underline{X}_{K-1\times 1} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_{K-1} \end{pmatrix}$;

- where $X_1 + X_2 + \dots + X_{K-1} < n$

$P_1 + P_2 + \dots + P_{K-1} < 1$

we'll say $\underline{X}_{K-1\times 1}$ follows a multinomial distribution with parameter $(n, P_1, P_2, \dots, P_{K-1})$ of order $K-1$

$$P[X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_{K-1} = x_{K-1}]$$

$$= \frac{n! \cdot P_1^{x_1} P_2^{x_2} \cdots P_{K-1}^{x_{K-1}} (1 - P_1 - P_2 - \dots - P_{K-1})^{n - x_1 - x_2 - \dots - x_{K-1}}}{x_1! \, x_2! \cdots x_{K-1}! \, (n - x_1 - \dots - x_{K-1})!}$$

① Marginal distribution of $X_1$

$$P[X_1 = x_1] = \sum_{x_2} \sum_{x_3} \cdots \sum_{x_{k-1}} \frac{n!\, p_1^{x_1} p_2^{x_2} p_3^{x_3} \cdots p_{k-1}^{x_{k-1}} (1-p_1-\cdots-p_{k-1})^{n-x_1-x_2-\cdots-x_{k-1}}}{x_1!\, x_2! \cdots x_{k-1}!\, (n-x_1-x_2-\cdots-x_{k-1})!}$$

$$\sum_{i=2}^{k-1} x_i < n - x_1$$

Remember

$$\sum_{x_1} \sum_{x_2} \cdots \sum_{x_{k-1}} \frac{n!\, p_1^{x_1} p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} (1-p_1-\cdots-p_{k-1})^{n-x_1-x_2-\cdots-x_{k-1}}}{x_1!\, x_2! \cdots x_{k-1}!\, (n-x_1-x_2-\cdots-x_{k-1})!}$$

$$\sum_{i=1}^{k-1} x_i < n, \quad = (p_1 + p_2 + \cdots + p_{k-1} + 1 - p_1 - \cdots - p_{k-1})^n$$

$$\sum_{i=1}^{k-1} p_i < 1 \quad = 1$$

Note

$$P[X_1 = x_1] = \frac{n!\, p_1^{x_1}}{x_1!\,(n-x_1)!} \sum_{x_2} \sum_{x_3} \cdots \sum_{x_{k-1}} \frac{(n-x_1)!\, p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} (1-p_1-\cdots-p_{k-1})^{n-x_1-\cdots-x_{k-1}}}{x_2!\, x_3! \cdots x_{k-1}!\, (n-x_1-\cdots-x_{k-1})!}$$

$$\sum x_i < n - x_1$$

$$= \frac{n!\, p_1^{x_1}}{x_1!\,(n-x_1)!} (p_2 + p_3 + \cdots + p_{k-1} + 1 - p_1 - \cdots - p_{k-1})^{n-x_1}$$

$$P[X_1 = x_1] = \frac{n!\, p_1^{x_1}}{x_1!\,(n-x_1)!} (1-p_1)^{n-x_1} \sim Bin(n, p_1)$$

② Marginal distribution of $(X_1, X_2)$

$$P[X_1 = x_1 \cap X_2 = x_2].$$

$$P[X_1=x_1 \cap X_2=x_2] = \sum_{x_3}\sum_{x_4}\cdots\sum_{x_{k-1}} \frac{p_1^{x_1} p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} (1-p_1-p_2-\cdots-p_{k-1})^{n-x_1-x_2-\cdots-x_{k-1}} \cdot n!}{x_1!\, x_2! \cdots x_{k-1}!\, (n-x_1-x_2-\cdots-x_{k-1})!}$$

$$\left( \sum_{i=3}^{k-1} x_i < n-x_1-x_2 \right)$$

$$= \frac{n!\, p_1^{x_1} p_2^{x_2}}{x_1!\, x_2!\, (n-x_1-x_2)!} \sum_{x_3}\sum_{x_4}\cdots\sum_{x_{k-1}} \frac{(n-x_1-x_2)!\, p_3^{x_3} p_4^{x_4} \cdots p_{k-1}^{x_{k-1}} (1-p_1-p_2-\cdots-p_{k-1})^{n-x_1-\cdots-x_{k-1}}}{x_3!\, x_4! \cdots x_{k-1}!\, (n-x_1-\cdots-x_{k-1})!}$$

$$= \frac{n!\, p_1^{x_1} p_2^{x_2}}{x_1!\, x_2!\, (n-x_1-x_2)!} \left(p_3+p_4+\cdots+p_{k-1}+1-p_1-p_2-p_3-p_4-\cdots-p_{k-1}\right)^{n-x_1-x_2}$$

$$= \frac{n!\, p_1^{x_1} p_2^{x_2}}{x_1!\, x_2!\, (n-x_1-x_2)!} (1-p_1-p_2)^{n-x_1-x_2} \quad \sim \text{Bin}(n, p_1+p_2)$$

$$x_i = 0,1,2,\ldots,n, \quad i=1(1)2$$

Trinomial distribution $(n, p_1, p_2)$

$$E(X_1X_2) = \sum_{x_1=0}^{n}\sum_{x_2=0}^{n} \frac{x_1 x_2\, n!\, p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2}}{x_1!\, x_2!\, (n-x_1-x_2)!}$$

$x_1+x_2 \le n$

$$= \sum_{x_1=1}^{n}\sum_{x_2=1}^{n} \frac{n!\, p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{n-x_1-x_2}}{(x_1-1)!\, (x_2-1)!\, (n-x_1-x_2)!}$$

$$= \sum_{y_1=0}^{n-1}\sum_{y_2=0}^{n-1} \frac{n!\, p_1^{y_1+1} p_2^{y_2+1} (1-p_1-p_2)^{n-y_1-1-y_2-1}}{y_1!\, y_2!\, (n-y_1-1-y_2-1)!}$$

$x_1-1=y_1,$
$x_2-1=y_2$

$y_1+y_2 \le n-2$

$$= n(n-1)p_1p_2 \sum_{\substack{y_1\ y_2 \\ y_1+y_2 \le n-2}}\sum \frac{(n-2)!\ p_1^{y_1}\ p_2^{y_2}\ (1-p_1-p_2)^{n-2-y_1-y_2}}{y_1!\ y_2!\ (n-2-y_1-y_2)!}$$

$$= n(n-1)p_1p_2\left(\cancel{p_1}+\cancel{p_2}+1-\cancel{p_1}-\cancel{p_2}\right)^{n-2}$$

$$= n(n-1)p_1p_2$$

$$Cov(X_1,X_2) = n(n-1)p_1p_2 - np_1 \times np_2 \quad \left[\begin{array}{l} Cov(X_1,X_2) \\ = E(X_1X_2) \\ - E(X_1) \\ \quad E(X_2) \end{array}\right]$$

$$= n^2p_1p_2 - np_1p_2 - n^2p_1p_2$$

$$= -np_1p_2$$

$Cov(X_1, X_2)$ is negative one as $X_1$ and $X_2$ are reversely linearly related.

Thus the dispersion matrix of Multinomial $(n, p_1, p_2, \dots, p_{k-1})$ is

$$\Sigma = \begin{array}{c|c} & \begin{array}{ccccc} x_1 & x_2 & x_3 & \cdots & x_{k-1} \end{array} \\ \begin{array}{c} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k-1} \end{array} & \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & -np_1p_3 & \cdots & -np_1p_{k-1} \\ -np_1p_2 & np_2(1-p_2) & -np_2p_3 & \cdots & -np_2p_{k-1} \\ -np_1p_3 & -np_2p_3 & np_3(1-p_3) & \cdots & -np_3p_{k-1} \\ \vdots & \vdots & & \ddots & \\ -np_1p_{k-1} & -np_2p_{k-1} & \cdots & & np_{k-1}(1-p_{k-1}) \end{pmatrix}_{k-1 \times k-1} \end{array}$$

Result

$\Sigma$ is non singular positive definite matrix.

$$\Sigma = n \begin{pmatrix} p_1(1-p_1) & -p_1 p_2 & \cdots & -p_1 p_{k-1} \\ -p_1 p_2 & p_2(1-p_2) & \cdots & -p_2 p_{k-1} \\ \vdots & \vdots & & \vdots \\ -p_1 p_{k-1} & -p_2 p_{k-1} & \cdots & p_{k-1}(1-p_{k-1}) \end{pmatrix}$$

$$= n \left[ \begin{pmatrix} p_1 & 0 & 0 & \cdots & 0 \\ 0 & p_2 & & & \vdots \\ 0 & & p_3 & & \vdots \\ \vdots & & & \ddots & \\ 0 & \cdots & & & p_{k-1} \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{k-1} \end{pmatrix} (p_1 \; p_2 \cdots p_{k-1}) \right]$$

$$= n \left[ \text{diag}(p_1 \; p_2 \ldots p_{k-1}) - \underline{p}\,\underline{p}' \right]$$

$$|\Sigma| = n^{k-1} \left| \text{diag}(p_1 \; p_2 \ldots p_{k-1}) - \underline{p}\,\underline{p}' \right|$$

$$= n^{k-1} \left| \text{diag}(p_1 \; p_2 \ldots p_{k-1}) \right| \left| I - \text{diag}^{-1}(p_1 p_2 \ldots p_{k-1}) \underline{p}\,\underline{p}' \right|$$

$$\therefore = n^{k-1}\, p_1 \times p_2 \times \cdots p_{k-1} \left| I - \text{diag}^{-1}(p_1, p_2, \ldots p_{k-1}) \underline{p}\,\underline{p}' \right|$$

$\longleftarrow$ (1)

Consider a block matrix.

$$\begin{bmatrix} 1 & \underline{P}' \\ \underline{P} & D_{K-1} \end{bmatrix} \quad \text{where } \underline{P} = \begin{pmatrix} P_1 \\ P_2 \\ \vdots \\ P_{K-1} \end{pmatrix}.$$

$$D_{K-1} = \text{diag}\,(P_1, P_2 \cdots P_{K-1})$$

Therefore $\begin{vmatrix} 1 & \underline{P}' \\ \underline{P} & D_{K-1} \end{vmatrix}$

$$= |D_{K-1}| \, (1 - \underline{P}' D_{K-1}^{-1} \underline{P})$$ (By Block determinant result)

Now from $\begin{vmatrix} 1 & \underline{P}' \\ \underline{P} & D_{K-1} \end{vmatrix}$ we get

$$|D_{K-1} - \underline{P}\,\underline{P}'|$$

$$\therefore |D_{K-1} - \underline{P}\,\underline{P}'| = |D_{K-1}| \, (1 - \underline{P}' D_{K-1}^{-1} \underline{P})$$

$$\Rightarrow |D_{K-1}| \, |I - D_{K-1}^{-1} \underline{P}\,\underline{P}'| = |D_{K-1}| \, (1 - \underline{P}' D_{K-1}^{-1} \underline{P})$$

$$\Rightarrow I - D_{K-1}^{-1} \underline{P}\,\underline{P}' = 1 - \underline{P}' D_{K-1}^{-1} \underline{P}$$

Now using the above in (1)

$$|\Sigma| = n^{K-1} P_1 P_2 \cdots P_{K-1} \, (1 - \underline{P}' D_{K-1}^{-1} \underline{P})$$

$$1 - \underline{p}' D_{k-1}^{-1} \underline{p}$$

$$= 1 - (p_1 \; p_2 \; \cdots \; p_{k-1}) \begin{pmatrix} 1/p_1 & 0 & \cdots & 0 \\ 0 & 1/p_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & & 1/p_{k-1} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{k-1} \end{pmatrix}$$

$$= \cancel{1 \quad 1 \quad 0} \quad 0$$

$$= 1 - (1 \; 1 \; \cdots \; 1) \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_{k-1} \end{pmatrix}$$

$$= 1 - p_1 - p_2 - \cdots - p_{k-1}$$

$$\therefore |\Sigma| = n^{k-1} p_1 p_2 \cdots p_{k-1} (1 - p_1 - p_2 - \cdots p_{k-1})$$

$$\neq 0$$

$$|\Sigma| > 0$$

Nonsingularity deflects when $\sum_{i=1}^{k-1} p_i = 1$

## MGF of multinomial vector

Suppose $\underline{X}_{p\times 1}$ be a random vector having some probability distribution.

Then the m.g.f of $\underline{X}_{p\times 1}$ is

$$M_{\underline{X}}(\underline{t}) = E\left(e^{\underline{t}'\underline{x}}\right) = E\left(e^{t_1 x_1 + t_2 x_2 \cdots + t_p x_p}\right)$$

where, $\underline{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_p \end{pmatrix} \in \mathbb{R}^p$, an arbitrary vector of constants.

$$\cancel{M_{\underline{X}}(\underline{t}) = \sum_{x_1}\sum_{x_2}\cdots\sum_{x_{k-1}} e^{t_1x_1+t_2x_2+\cdots+x_{k-1}} \cdot p_1^{x_1}p_2^{x_2}\cdots p_{k-1}^{x_{k-1}}\cdot n!}$$

$$M_{\underline{X}}(\underline{t}) = \sum_{x_1}\sum_{x_2}\cdots\sum_{x_{k-1}} e^{t_1x_1+t_2x_2+\cdots+t_{k-1}x_{k-1}} \frac{n!\, p_1^{x_1} p_2^{x_2}\cdots p_k^{x_{k-1}} (1-p_1-p_2-\cdots-p_{k-1})^{n-x_1-x_2-\cdots-x_{k-1}}}{x_1!\, x_2! \cdots x_{k-1}!\, (n-x_1-x_2-\cdots-x_{k-1})!}$$

$\sum x_i < n$

$\sum p_i < 1$

$$= \sum_{x_1}\sum_{x_2}\cdots\sum_{x_{k-1}} n! \frac{(e^{t_1}p_1)^{x_1}(e^{t_2}p_2)^{x_2}\cdots(e^{t_{k-1}}p_{k-1})^{x_{k-1}} \cdot (1-p_2-p_3-\cdots-p_{k-1})^{n-x_1-x_2-\cdots-x_{k-1}}}{x_1!\, x_2! \cdots x_{k-1}!\, (n-x_1-x_2-\cdots-x_{k-1})!}$$

$$= \left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{k-1}}p_{k-1} + 1 - p_1 - \cdots - p_{k-1}\right)^n$$

Find out $E(X_1)$ from m.g.f.

$$\left.\frac{dM_{\underline{X}}(\underline{t})}{dt_1}\right|_{t_1=0} = E(X_1)$$

$$\Rightarrow \frac{dM_{\underline{X}}(\underline{t})}{dt_1} = n\left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{k-1}}p_{k-1} + 1 - p_1 - p_2 - \cdots - p_{k-1}\right)^{n-1} \cdot p_1 e^{t_1}$$

at $t_1 = 0$

$$\left.\frac{d M_{\underline{X}}(\underline{t})}{dt_1}\right|_{t_1=0} = n p_1 \left(p_1 + e^{t_2} p_2 + \cdots + e^{t_{k-1}} p_{k-1} + 1 - \cancel{p_1} - p_2 - \cdots - p_{k-1}\right)^{n-1}$$

$$= n p_1 \left(1 + e^{t_2} p_2 + \cdots + e^{t_{k-1}} p_{k-1} - p_2 - \cdots - p_{k-1}\right)^{n-1}$$

Therefore,

$$E(X_1) = n p_1 (1 + 0) \quad [\because \forall\, t_i = 0,\ i = 2(1)k-1]$$

$$= n p_1$$

$$\frac{d^2 M_{\underline{X}}(\underline{t})}{dt_1^2} = \cancel{\frac{n p_1^2 e^{2t_1}}{2}}$$

$$= n(n-1)\left(e^{t_1} p_1 + e^{t_2} p_2 + \cdots + e^{t_{k-1}} p_{k-1} + 1 - p_1 - p_2 - \cdots - p_{k-1}\right)^{n-2} \left(p_1 e^{t_1}\right)^2$$

$$+ n\left(e^{t_1} p_1 + e^{t_2} p_2 + \cdots + e^{t_{k-1}} p_{k-1} + 1 - p_1 - p_2 - \cdots - p_{k-1}\right)^{n-1} \cdot p_1 e^{t_1}$$

$[\forall\, t_i = 0\ ;\ i = 2(1)k-1]$

we have,

$$\cancel{\frac{d^2 M_{\underline{X}}(\underline{t})}{dt_1^2} = n(n-1)\left(e^{t_1} p_1 + 1 - p_1\right)^{n-2} \left(p_1 e^{t_1}\right)^2 + n\left(e^{t_1} p_1 + 1 - p_1\right)^{n-1} p_1 e^{t_1}}$$

Now; $Var(X_1) = \frac{d^2 M_{\underline{x}}(t)}{dt_1^2}\Big|_{t_1=0}$

$= \cancel{n(n-1)\, p_1^{n-2} \cdot p_1^2 + n\, p_1^{n-1} p_1}$

$= n(n-1)\, p_1^n + n p_1^n$

$\cancel{= n p_1}$ $= n(n-1)\, p_1^2 + n p_1$

---

$M_{\underline{x}}(\underline{t}) = E(e^{\underline{t}'\underline{x}})$

$= \left(e^{t_1}p_1 + e^{t_2}p_2 + \dots + e^{t_{k-1}}p_{k-1} + 1 - p_1 - \dots - p_{k-1}\right)^n$

$E(X_1)$

$= \left[\frac{\partial}{\partial t_1}\left(e^{t_1}p_1 + \dots - p_{k-1}\right)^n\right]_{\underline{t}=0}$

$\underline{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

$\cancel{= n(n-1)}$ $= n p_1$

$E(x_1^2) = \frac{\partial^2}{\partial x_2}$

H.W $cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$

$$\frac{\partial^2}{\partial t_1 \partial t_2} (e^{t_1} p_1 + \cdots - p_{k-1})^n \Big|_{\underline{t} = \underline{0}}$$

$$= E(X_1 X_2)$$

Conditional distribution :

Conditional distribution of $X_1$ given $X_2, X_3, \ldots, X_{k-1}$

is $P(X_1 = x_1 \mid X_2 = x_2 \cap X_3 = x_3 \cap \cdots \cap X_{k-1} = x_{k-1})$

$$= \frac{P(X_1 = x_1 \cap X_2 = x_2 \cap X_3 = x_3 \cap \cdots \cap X_{k-1} = x_{k-1})}{P(X_2 = x_2 \cap X_3 = x_3 \cap \cdots \cap X_{k-1} = x_{k-1})}$$

We know,

$$P_{X_1, X_2, \ldots, X_{k-1}}(x_1, x_2, \ldots, x_{k-1})$$

$$= \frac{n!\, p_1^{x_1} p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} (1 - p_1 - \cdots - p_{k-1})^{n - x_1 - \cdots - x_{k-1}}}{x_1!\, x_2! \cdots x_{k-1}!\, (n - x_1 - x_2 - \cdots - x_{k-1})!}$$

$\sum_{i=1}^{k-1} p_i < 1$

$\sum_{i=1}^{k-1} x_i < n$

Then, $P_{X_2, X_3, \ldots, X_{k-1}}(x_2, x_3, \ldots, x_{k-1})$

$$= \frac{n!\, p_2^{x_2} p_3^{x_3} \cdots p_{k-1}^{x_{k-1}} (1 - p_2 - \cdots - p_{k-1})^{n - x_2 - \cdots - x_{k-1}}}{x_2!\, x_3! \cdots x_{k-1}!\, (n - x_2 - x_3 - \cdots - x_{k-1})!}$$

Check it on $P(x_2, x_3, \ldots, x_{k-1})$

$$= \sum_{x_1=0}^{n} P(x_1, x_2, \ldots, x_{k-1})$$

$$= \frac{n!\, p_1^{x_1} p_2^{x_2} \cdots p_{k-1}^{x_{k-1}} \left(1 - \sum_{i=1}^{k-1} p_i\right)^{n - \sum_{i=1}^{k-1} x_i}}{x_1!\, x_2! \cdots x_{k-1}!\, \left(n - \sum_{i=1}^{k-1} x_i\right)!} \times \frac{x_2!\, x_3! \cdots x_{k-1}!\, \left(n - \sum_{i=2}^{k-1} x_i\right)!}{n!\, p_2^{x_2} p_3^{x_3} \cdots p_{k-1}^{x_{k-1}} \cdot \left(1 - \sum_{i=2}^{k-1} p_i\right)^{n - \sum_{i=2}^{k-1} x_i}}$$

$$= \frac{\cancel{p_1^{x_1}} \cdot \left(n - \sum_{i=2}^{k-1} x_i\right)!}{x_1!\, \left(n - \sum_{i=1}^{k-1} x_i\right)!} \cdot p_1^{x_1} \cdot \frac{\left(1 - \sum_{i=1}^{k-1} p_i\right)^{n - \sum_{i=1}^{k-1} x_i}}{\left(1 - \sum_{i=2}^{k-1} p_i\right)^{n - \sum_{i=2}^{k-1} x_i - x_1}}$$

adjust

$$= \binom{n - \sum_{i=2}^{k-1} x_i}{x_1} \left(\frac{p_1}{1 - p_2 - \cdots - p_{k-1}}\right)^{x_1} \left(\frac{1 - p_1 - p_2 - \cdots - p_{k-1}}{1 - p_2 - \cdots - p_{k-1}}\right)^{n - x_2 - \cdots - x_{k-1} - x_1}$$

$$= \binom{n - \sum_{i=2}^{k-1} x_i}{x_1} \left(\frac{p_1}{1 - p_2 - \cdots - p_{k-1}}\right)^{x_1} \left(1 - \frac{p_1}{1 - p_2 - \cdots - p_{k-1}}\right)^{n - x_1 - x_2 - \cdots - x_{k-1}}$$

$$\sim \text{Bin}\left(n - x_2 - x_3 - \cdots - x_{k-1},\ \frac{p_1}{1 - p_2 - \cdots - p_{k-1}}\right)$$

$E(X_1 | X_2, X_3, \ldots, X_{k-1})$

$$= (n - x_2 - x_3 - \ldots - x_{k-1}) \frac{P_1}{1 - P_2 - \ldots - P_{k-1}}$$

Remark

From regression theory,

$$X_{1.23\ldots k-1} = \frac{nP_1}{1 - P_1 - \ldots - P_{k-1}} - \frac{P_1 x_2}{1 - P_2 - \ldots - P_{k-1}}$$

$$- \ldots - \frac{P_1 x_{k-1}}{1 - P_2 - \ldots - P_{k-1}}$$

Find out the intercept of the equation.

$$\rightarrow \frac{nP_1}{1 - P_1 - \ldots - P_{k-1}}$$

Remark

$$V(X_1 | X_2, X_3, \ldots, X_{k-1}) = (n - x_2 - \ldots - x_{k-1}) \frac{P_1}{1 - P_2 - \ldots - P_{k-1}} \left(1 - \frac{P_1}{1 - P_2 - \ldots - P_{k-1}}\right)$$

Is this a homoscedastic variance?

→ NO, because conditional variance depends on $x_2, x_3, \ldots, x_{k-1}$ i.e. conditioning variables.

So its heteroscedastic variance.

Inverse of variance covariance matrix and multiple correlation

Dispersion matrix for Multinomial distribution :

$$\Sigma = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & \cdots & -np_1p_{k-1} \\ -np_1p_2 & np_2(1-p_2) & \cdots & -np_2p_{k-1} \\ \vdots & \vdots & & \vdots \\ -np_1p_k & -np_2p_k & \cdots & np_k(1-p_k) \end{pmatrix}_{k-1 \times k-1}$$

To deduce $\Sigma^{-1}$, we use a result from linear algebra.

$$(\underset{p\times p}{A} + \underset{p\times 1}{\underline{u}}\,\underset{1\times p}{\underline{v}'})^{-1} = A^{-1} - \frac{(A^{-1}\underline{u})(\underline{v}'A^{-1})}{1+\underline{v}'A^{-1}\underline{u}}$$

In $\Sigma$,

$$\Sigma = \underbrace{n\begin{pmatrix} p_1 & & & \\ & p_2 & & \\ & & \ddots & \\ & & & p_{k-1} \end{pmatrix}}_{A} + n\underbrace{\begin{pmatrix} -p_1 \\ -p_2 \\ \vdots \\ -p_{k-1} \end{pmatrix}}_{\underline{u}}\underbrace{(p_1 \; p_2 \cdots p_{k-1})}_{\underline{v}'}$$

$$A^{-1} = \begin{bmatrix} \frac{1}{P_1} & & & \\ & \frac{1}{P_2} & & \\ & & \ddots & \\ & & & \frac{1}{P_{K-1}} \end{bmatrix}$$

$$A^{-1}\underline{u} = \begin{bmatrix} \frac{1}{P_1} & 0 & \cdots & 0 \\ \vdots & \frac{1}{P_2} & & \\ \vdots & & \ddots & \\ 0 & \cdots & & \frac{1}{P_{K-1}} \end{bmatrix} \begin{pmatrix} -P_1 \\ -P_2 \\ \vdots \\ -P_{K-1} \end{pmatrix}$$

$$= \begin{pmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}$$

$$\underline{v}' A^{-1} = (P_1 \; P_2 \cdots P_{K-1}) \begin{bmatrix} 1/P_1 & & & \\ & 1/P_2 & & \\ & & \ddots & \\ & & & \frac{1}{P_{K-1}} \end{bmatrix}$$

$$= (1 \; 1 \; \cdots \; 1)$$

$$\underline{v}' A^{-1} \underline{u} = (1 \; 1 \cdots 1) \begin{pmatrix} -P_1 \\ -P_2 \\ \vdots \\ -P_{K-1} \end{pmatrix}$$

$$= -P_1 - P_2 - \cdots - P_{K-1}$$

$$1 + \underline{v}' A^{-1} \underline{u} = 1 - P_1 - P_2 - \cdots - P_{K-1}$$

$$\Sigma^{-1} = \frac{1}{n}\left( \begin{bmatrix} \frac{1}{p_1} & & & \\ & \frac{1}{p_2} & & \\ & & \ddots & \\ & & & \frac{1}{p_{k-1}} \end{bmatrix} - \frac{1 \begin{pmatrix} -1 \\ -1 \\ \vdots \\ -1 \end{pmatrix} (1, 1, \cdots 1)}{1 - p_1 - p_2 - \cdots - p_{k-1}} \right)$$

$$= \frac{1}{n}\left[ \begin{pmatrix} \frac{1}{p_1} & & & \\ & \frac{1}{p_2} & & \\ & & \ddots & \\ & & & \frac{1}{p_{k-1}} \end{pmatrix} + \frac{1}{1 - p_1 - p_2 - \cdots - p_{k-1}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & \cdots & & 1 \\ \vdots & & & \vdots \\ 1 & \cdots & & 1 \end{bmatrix} \right]$$

$$= \frac{1}{n} \begin{bmatrix} \frac{1}{p_1} + \frac{1}{1 - \sum_{i=1}^{k-1} p_i} & \frac{1}{1 - \sum_{i=1}^{k-1} p_i} & \cdots & \frac{1}{1 - \sum_{i=1}^{k-1} p_i} \\ \frac{1}{1 - \sum_{i=1}^{k-1} p_i} & \frac{1}{p_2} + \frac{1}{1 - \sum_{i=1}^{k-1} p_i} & \cdots & \frac{1}{1 - \sum_{i=1}^{k-1} p_i} \\ \vdots & \vdots & & \vdots \\ & & & \frac{1}{p_{k-1}} + \frac{1}{1 - \sum_{i=1}^{k-1} p_i} \end{bmatrix}$$

Let, $p_k = 1 - \sum_{i=1}^{k-1} p_i$

$$\Sigma^{-1} = \frac{1}{n} \begin{bmatrix} \frac{1}{p_1} + \frac{1}{p_k} & \frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \frac{1}{p_k} & \frac{1}{p_2} + \frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \vdots & \vdots & & \vdots \\ & & & \frac{1}{p_{k-1}} + \frac{1}{p_k} \end{bmatrix}$$

multiple correlation,

$$\rho_{1.23\cdots p} = \left(1 - \frac{1}{\sigma^{11}\sigma_{11}}\right)^{1/2}$$

$\sigma^{11} =$ (1,1) element of $\Sigma^{-1}$

$\sigma_{11} =$ (1,1) element " $\Sigma$

$$\rho_{12.34\ldots p} = \frac{-\sigma^{12}}{\sqrt{\sigma^{11}\sigma^{22}}}$$

$$\rho_{1.23\ldots p} = \left(1 - \frac{1}{\frac{1}{n}\left(\frac{1}{p_1}+\frac{1}{p_k}\right) n p_1(1-p_1)}\right)^{1/2}$$

$$= \left(\frac{\left(\frac{1}{p_1}+\frac{1}{p_k}\right) n p_1(1-p_1) - 1}{\frac{1}{n}\left(\frac{1}{p_1}+\frac{1}{p_k}\right) n p_1(1-p_1)}\right)^{1/2}$$

$$= \left(\frac{n(1-p_1) + \frac{1}{p_k} n p_1(1-p_1) - 1}{\left(\frac{1}{p_1}+\frac{1}{p_k}\right) n p_1(1-p_1)}\right)^{1/2}$$

$$= \left(\frac{\cancel{1}-p_1 + \frac{1}{p_k}p_1 - \frac{p_1^2}{p_k} - \cancel{1}}{\left(\frac{1}{p_1}+\frac{1}{p_k}\right) p_1(1-p_1)}\right)^{1/2}$$

$$= \left(\frac{-1 + \frac{1}{p_k} - \frac{p_1}{p_k}}{\left(\frac{1}{p_1}+\frac{1}{p_k}\right)(1-p_1)}\right)^{1/2}$$

$\rho_{12.34\ldots p} = \dfrac{-1/p_K}{\sqrt{\phantom{x}}}$

Mean: $E(X_1)$ & Variance: $V(X_1)$

$$M_{\underline{X}}(t) = \left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{K-1}}p_{K-1} + 1 - p_1 - \cdots - p_{K-1}\right)^n$$

$$\frac{\partial M_{\underline{X}}(\underline{t})}{\partial t_1} = np_1 e^{t_1}\left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{K-1}}p_{K-1} + 1 - p_1 - \cdots - p_{K-1}\right)^{n-1}$$

$$\frac{\partial^2 M_{\underline{X}}(\underline{t})}{\partial t_1^2} = np_1 e^{t_1}\left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{K-1}}p_{K-1} + 1 - p_1 - \cdots - p_{K-1}\right)^{n-1} + n(n-1)p_1^2 e^{2t_1} \cdot \left(e^{t_1}p_1 + e^{t_2}p_2 + \cdots + e^{t_{K-1}}p_{K-1} + 1 - p_1 - \cdots - p_{K-1}\right)^{n-2}$$

$$E(X_1) = \left[\frac{\partial M_{\underline{X}}(\underline{t})}{\partial t_1}\right]_{\underline{t}=0}, \quad \underline{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\boxed{E(X_1) = np_1}$$

$$E(X_1^2) = \left.\frac{\partial^2 M_{\underline{X}}(\underline{t})}{\partial t_1^2}\right|_{\underline{t}=0} = np_1 + n(n-1)p_1^2$$

$$V(X_1) = E(X_1^2) - [E(X_1)]^2 = np_1 + \cancel{n^2p_1^2} - np_1^2 - \cancel{n^2p_1^2}$$

$$\boxed{V(X_1) = np_1(1-p_1)}$$

$\underline{Cov(X_1, X_2)}$

$$Cov(X_1, X_2) = E(X_1 X_2) - E(X_1) E(X_2)$$

$$E(X_1 X_2) = \left. \frac{\partial^2 M_{\underline{X}}(\underline{t})}{\partial t_1 \partial t_2} \right|_{\underline{t}=0}$$

$$E(X_1) = nP_1 \quad , \quad E(X_2) = nP_2$$

$$M_{\underline{X}}(\underline{t}) = \left(e^{t_1} P_1 + e^{t_2} P_2 + \cdots + e^{t_{k-1}} P_{k-1} + 1 - P_1 - \cdots - P_{k-1}\right)^n$$

$$\frac{\partial M_{\underline{X}}(\underline{t})}{\partial t_1} = nP_1 e^{t_1} \left(e^{t_1} P_1 + e^{t_2} P_2 + \cdots + e^{t_{k-1}} P_{k-1} + 1 - P_1 - \cdots - P_{k-1}\right)^{n-1}$$

$$\frac{\partial^2 M_{\underline{X}}(\underline{t})}{\partial t_1 \partial t_2} = n(n-1) P_1 P_2 e^{t_1} e^{t_2} \left(e^{t_1} P_1 + e^{t_2} P_2 + \cdots + e^{t_{k-1}} P_{k-1} + 1 - P_1 - \cdots - P_{k-1}\right)^{n-2}$$

$$E(X_1 X_2) = \left. \frac{\partial^2 M_{\underline{X}}(\underline{t})}{\partial t_1 \partial t_2} \right|_{\underline{t}=0}$$

$$= n(n-1) P_1 P_2$$

$$\therefore Cov(X_1, X_2) = n(n-1) P_1 P_2 - n^2 P_1 P_2$$

$$\boxed{Cov(X_1, X_2) = -n P_1 P_2}$$

Multiple correlation

$$\rho_{1.23\cdots p} = \left(1 - \frac{1}{\sigma^{11}\,\sigma_{11}}\right)^{1/2}$$

where, $\sigma^{11} = (1,1)$ element of $\Sigma^{-1}$

$\sigma_{11} = (1,1)$ element of $\Sigma$.

and
$$\Sigma = \begin{pmatrix} np_1(1-p_1) & -np_1p_2 & \cdots & -np_1p_{k-1} \\ -np_1p_2 & np_2(1-p_2) & \cdots & -np_2p_{k-1} \\ \vdots & & & \\ -np_1p_k & -np_2p_k & \cdots & np_k(1-p_k) \end{pmatrix}$$

$$\Sigma^{-1} = \frac{1}{n}\begin{pmatrix} \frac{1}{p_1}+\frac{1}{p_k} & \frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \frac{1}{p_k} & \frac{1}{p_2}+\frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \vdots & & & \\ \frac{1}{p_k} & \frac{1}{p_k} & \cdots & \frac{1}{p_{k-1}}+\frac{1}{p_k} \end{pmatrix}$$

$$\rho_{1.23\cdots p} = \left(1 - \frac{1}{\frac{1}{\cancel{n}}\left(\frac{1}{p_1}+\frac{1}{p_k}\right)\cdot \cancel{n}\, p_1(1-p_1)}\right)^{1/2}$$

$$= \left(1 - \frac{1}{\left(\frac{1}{p_1}+\frac{1}{p_k}\right)\cdot p_1(1-p_1)}\right)^{1/2}$$

$$= \left(\frac{\left(\frac{1}{p_1}+\frac{1}{p_k}\right)p_1(1-p_1) - 1}{\left(\frac{1}{p_1}+\frac{1}{p_k}\right)p_1(1-p_1)}\right)^{1/2}$$

$$= \left( \frac{\left(1+\frac{P_1}{P_K}\right)(1-P_1) - 1}{\left(\frac{1}{P_1}+\frac{1}{P_K}\right) P_1(1-P_1)} \right)^{1/2}$$

$$= \left( \frac{\cancel{1} + \frac{P_1}{P_K} - P_1 - \frac{P_1^2}{P_K} - \cancel{1}}{\left(\frac{1}{P_1}+\frac{1}{P_K}\right)(1-P_1)P_1} \right)^{1/2}$$

$$\boxed{\rho_{1\cdot 23\cdots p} = \left( \frac{\frac{1}{P_K} - \frac{P_1}{P_K} - 1}{\left(\frac{1}{P_1}+\frac{1}{P_K}\right)(1-P_1)} \right)^{1/2}}$$

<u>Partial Correlation</u>

$$\rho_{12\cdot 34\cdots p} = \frac{-\sigma^{12}}{\sqrt{\sigma^{11}\sigma^{22}}}$$

where $\sigma^{ij} = (i,j)$ element of $\Sigma^{-1}$

$$\rho_{12\cdot 34\cdots p} = -\frac{\frac{1}{nP_K}}{\sqrt{\frac{1}{n}\left(\frac{1}{P_1}+\frac{1}{P_K}\right)\cdot\frac{1}{n}\left(\frac{1}{P_2}+\frac{1}{P_K}\right)}}$$

$$= -\frac{1/P_K}{\sqrt{\left(\frac{1}{P_1}+1/P_K\right)\cdot\left(\frac{1}{P_2}+\frac{1}{P_K}\right)}}$$

$$\boxed{\rho_{12\cdot 34\cdots p} = \frac{-P_1 P_2}{\sqrt{(P_1+P_K)(P_2+P_K)}}}$$

check it on

$$P(x_2, x_3, \ldots, x_{K-1})$$

$$= \sum_{x_1=0}^{n} P(x_1, x_2, \ldots, x_{K-1})$$

$$P(x_1, x_2, \ldots x_{K-1})$$

$$= \frac{n!\, P_1^{x_1} P_2^{x_2} \cdots P_{K-1}^{x_{K-1}} (1-P_1-P_2-\cdots-P_{K-1})^{n-x_1-x_2-\cdots-x_{K-1}}}{x_1!\, x_2! \cdots x_{K-1}!\, (n-x_1-x_2-\cdots-x_{K-1})!},$$

$$\sum_{i=1}^{K-1} P_i < 1, \quad \sum_{i=1}^{K-1} x_i < n$$

$$\sum_{x_1=0}^{n} P(x_1, x_2, \ldots, x_{K-1})$$

$$= \frac{n!\, P_2^{x_2} P_3^{x_3} \cdots P_{K-1}^{x_{K-1}} (1-P_1-P_2-\cdots-P_{K-1})^{n-x_2-x_3-\cdots-x_{K-1}}}{x_2!\, x_3! \cdots x_{K-1}!\, (n-x_2-\cdots-x_{K-1})!},$$

## Multivariate Normal Distribution

Def >> A random vector $\underline{X} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}$ is called a p-variate multivariate normal if the p.d.f of $\underline{X}_{p\times 1}$ can be written in the following exponential expression. quadratic function of an

$$f_{\underline{X}}(\underline{x}) = k e^{-\frac{1}{2}(\underline{x}-\underline{b})' A (\underline{x}-\underline{b})}$$

$$x^2 + 4yx + 4y^2 = (x \;\; y)\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

where $k$ being the normalizing constant

$A$ is a symmetric positive definite matrix

$\underline{b}$ is a vector of constants,

such that $(\underline{x}-\underline{b})' A (\underline{x}-\underline{b}) > 0 \;\; \forall \;\; \underline{x} \neq \underline{b}$

and $(\underline{x}-\underline{b})' A (\underline{x}-\underline{b}) = 0$ when $\underline{x} = \underline{b}$

### Determination of the constants $k$, $\underline{b}$, $A$

Determination of $k$ in terms of $A_{p\times p}$

$$k \mapsto \int_{\underline{x} \in \mathbb{R}^p} f_{\underline{x}}(\underline{x})\, d\underline{x} = 1$$

Since $A$ is a p.d. matrix from linear algebra result there exists ($\exists$) a non singular (n.s) matrix.

$$C \rightarrow C' A C = I$$

$$\Rightarrow A^{-1} = CC'$$

$$\int_{\underline{x} \in \mathbb{R}^p} f_X(\underline{x}) = 1$$

$$\Rightarrow K \int_{\underline{x} \in \mathbb{R}^p} e^{-\frac{1}{2}(\underline{x}-\underline{b})' A (\underline{x}-\underline{b})} \, d\underline{x} = 1$$

$$\Rightarrow K \int_{\underline{x}} e^{-\frac{1}{2}(\underline{x}-\underline{b})' (CC')^{-1} (\underline{x}-\underline{b})} \, d\underline{x} = 1$$

Let $\underline{x} - \underline{b} = C\underline{y}$

$d\underline{x} = |C| d\underline{y}$

$$\cancel{\Rightarrow K \int_{\underline{y}} e^{-\frac{1}{2}\underline{y}'(CC')^{-1}\underline{y}} \, d\underline{y}} = 1$$

$$\Rightarrow K \int_{\underline{y}} e^{-\frac{1}{2}\underline{y}' C' (CC')^{-1} C \underline{y}} \, |C| \, d\underline{y} = 1$$