MULTIVARIATE NORMAL DISTRIBUTION

(Wischern & Johnson)

- Multivariate Normal Distribution Functions,
- Conditional Distribution and its relation to regression model,
- Estimation of parameters.

1 → Multivariate normal Dist^n

2 → Multiple Linear Regression

3 → Multivariate Regression

NORMAL DISTRIBUTION (Univariate)

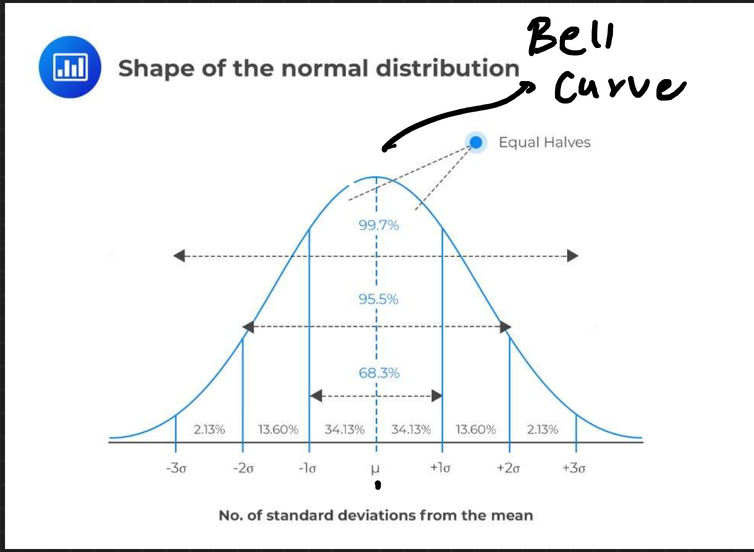
$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{\frac{-1}{2\sigma^2}(x-\mu)^2\right\} \quad : \quad \begin{array}{c} -\infty < x < \infty \\ -\infty < \mu < \infty \\ \sigma^2 > 0 \end{array}$$

where μ = Mean

σ = Standard Deviation

σ^2 = variance

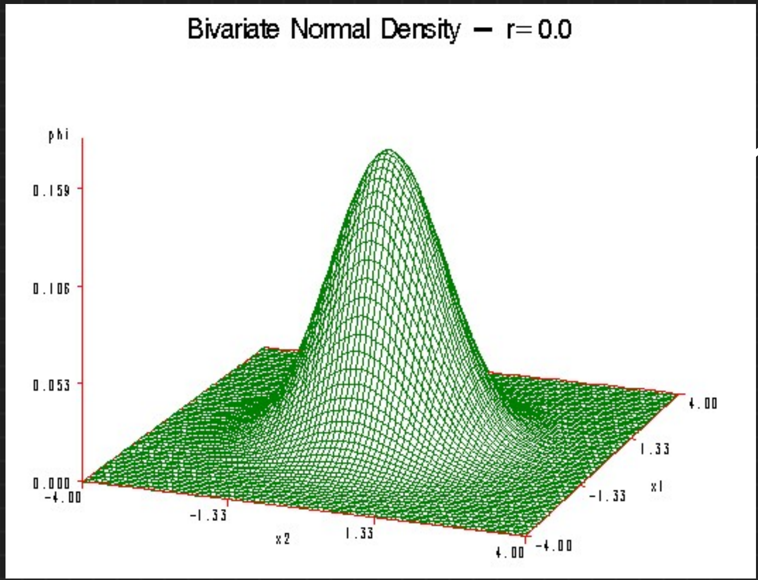
Bell Curve



Mean = Mode = Median

BIVARIATE NORMAL DISTRIBUTION

The “regular” normal distribution has one random variable; A bivariate normal distribution is made up of two independent random variables. The two variables in a bivariate normal are both are normally distributed, and they have a normal distribution when both are added together. Visually, the bivariate normal distribution is a three-dimensional bell curve.



3d

BIVARIATE NORMAL DISTRIBUTION

$$f(x_1) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left\{ \frac{-1}{2\sigma_1^2} (x_1 - \mu_1)^2 \right\}$$

$$f(x_2) = \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left\{ \frac{-1}{2\sigma_2^2} (x_2 - \mu_2)^2 \right\}$$

pdf of Bivariate Normal dist^n is

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{z}{2(1-\rho^2)} \right\}$$

where $z = \frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2}$

ρ = Correlation coefficients

$= \frac{\text{Cov}(x_1, x_2)}{\sigma_1 \sigma_2}$

BIVARIATE NORMAL DISTRIBUTION

Matrix approach to Bivariate Dist^n

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \mu = E(x) = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} V(x_1) & cov(x_1 x_2) \\ cov(x_1 x_2) & V(x_2) \end{bmatrix} = \text{var-covariance matrix} = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12} \\ \sigma_{21} & \sigma_{22}^2 \end{bmatrix}$$

$$\rho = \frac{\sigma_{12}}{\sigma_{11}\sigma_{22}} \qquad |\rho| \leq 1$$

BIVARIATE NORMAL DISTRIBUTION

p = No. of independent variables

$$\frac{(x-\mu)^2}{\sigma^2} = (x-\mu)(\sigma^2)^{-1}(x-\mu)$$

$$\Downarrow$$

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

pdf of Bivariate Normal distribution (Matrix form)

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

or

$$f(x) = \frac{1}{2\pi \sqrt{\det V}} \exp\left\{ -\frac{1}{2} (x-\mu)' V^{-1} (x-\mu) \right\} \text{ ——— } (1)$$

BIVARIATE NORMAL DISTRIBUTION

$$\det V = \det \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = \sigma_1^2 \sigma_2^2 (1-\rho^2)$$

Hence, the inverse of V is

$$V^{-1} = \frac{1}{\det(V)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)} \begin{pmatrix} \sigma_2^2 & -\rho\sigma_1\sigma_2 \\ -\rho\sigma_1\sigma_2 & \sigma_1^2 \end{pmatrix}$$

$$= \frac{1}{(1-\rho^2)} \begin{pmatrix} \sigma_1^{-2} & -\rho\sigma_1^{-1}\sigma_2^{-1} \\ -\rho\sigma_1^{-1}\sigma_2^{-1} & \sigma_2^{-2} \end{pmatrix}$$

$$f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \exp\left(\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right)$$

BIVARIATE NORMAL DISTRIBUTION

If X and Y are bivariate normal, what is the necessary and sufficient condition for X and Y to be independent?

$$\rightarrow f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}\right]$$

when $\rho = 0$,

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} \exp -\frac{1}{2}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}$$

$$f(x,y) = \frac{1}{\sqrt{2\pi}\,\sigma_1} \exp\left\{\frac{-1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right\} \cdot \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\left\{\frac{-1}{2}\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}$$

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

Hence X & Y are independent.

BIVARIATE NORMAL DISTRIBUTION

Result:

If X is distributed as Np(μ, Σ) , then any linear combination of variables a'X = $a_1x_1 + a_2X_2 + \ldots + a_pX_p$ is distributed as N( ). Also, if a'X is distributed as N( ) for every a, then X must be Np(μ, Σ)

$a'\mu,\ a'\Sigma a$ $a'\mu, a'\Sigma a$

$$f(x; \mu, \Sigma) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} (x-\mu)' \Sigma^{-1} (x-\mu) \right\}$$

BIVARIATE NORMAL DISTRIBUTION

Consider the linear combination a’X of a multivariate normal random vector determined by the choice a’ = [1, 0, . . . ,0].

$$a'X = [1\ 0\ 0 \cdots 0]\begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix} = x_1$$

$$a'\mu = [1\ 0 \cdots 0]\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} = \mu_1$$

$$a'\Sigma a = [1\ 0 \cdots 0]\begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & & & \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \sigma_{11}$$

x_1 is distributed as $N(\mu_1, \sigma_{11})$

x_i is distributed as $N(\mu_i, \sigma_{ii})$

BIVARIATE NORMAL DISTRIBUTION

Example: (The distribution of two linear combinations of the components of a normal random vector) For X distributed as N3(u, Σ) . find the distn of

$$\begin{bmatrix} X_1 - X_2 \\ X_2 - X_3 \end{bmatrix}_{2\times 1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}_{2\times 3} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}_{3\times 1} = a'X$$

$$a'\mu = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \mu_1 - \mu_2 \\ \mu_2 - \mu_3 \end{bmatrix}$$

$$a'\Sigma a = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{bmatrix} =$$

$$\begin{bmatrix} \sigma_{11} - 2\sigma_{12} + \sigma_{22} & \sigma_{12} + \sigma_{23} - \sigma_{22} - \sigma_{13} \\ \sigma_{12} + \sigma_{23} - \sigma_{22} - \sigma_{13} & \sigma_{22} - 2\sigma_{23} + \sigma_{33} \end{bmatrix}$$

BIVARIATE NORMAL DISTRIBUTION

Result:

a. If X_1 and X_2 are independent, then $Cov(x_1, x_2) = 0$, a $q_1 \times q_2$ matrix of zeros.

b. If $\begin{bmatrix} x_1 \\ \vdots \\ x_2 \end{bmatrix}$ is $N_{q_1+q_2}\left(\begin{bmatrix} \mu_1 \\ \vdots \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)$ then x_1 & x_2 are ind. iff

$\Sigma_{12} = 0$

$$a' \Sigma a = \begin{bmatrix} \sigma_{11} - 2\sigma_{12} + \sigma_{22} & \sigma_{12} + \sigma_{23} - \sigma_{22} - \sigma_{13} \\ \sigma_{12} + \sigma_{23} - \sigma_{22} - \sigma_{13} & \sigma_{22} - 2\sigma_{23} + \sigma_{33} \end{bmatrix}$$

BIVARIATE NORMAL DISTRIBUTION

Example:
(The equivalence of zero covariance and independence for normal variables)
Let X be $N_3(\mu, \Sigma)$ with

$$\Sigma = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Are X1 and X2 independent? What about (X1, X2) and X3?

→ x_1 & x_2 are not ind. since $\sigma_{12} = \sigma_{21} \neq 0$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \hline x_3 \end{bmatrix}, \quad \Sigma = \left[\begin{array}{cc|c} 4 & 1 & 0 \\ 1 & 3 & 0 \\ \hline 0 & 0 & 2 \end{array}\right] = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ & X_3 have cov. matrix $\Sigma_{12} = 0$

(X_1, X_2) & X_3 are independent.

BIVARIATE NORMAL DISTRIBUTION

The conditional density of X_1, given that $X_2 = x_2$ for any bivariate dist^n is defined by

$$f(x_1 | x_2) = \begin{pmatrix}\text{conditional probability of } X_1 \\ \text{given that } X_2 = x_2\end{pmatrix} = \frac{f(x_1, x_2)}{f(x_2)}$$

$$f(x_1 | x_2) \sim N\left(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2),\ \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}\right)$$

BIVARIATE NORMAL DISTRIBUTION

Consider a bivariate normal popl with $\mu_1 = 0$, $\mu_2 = 2$

$\sigma_{11} = 2$, $\sigma_{22} = 1$ & $\rho_{12} = 0.5$

1) write down the bivariate normal density

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

$\rightarrow$ $p = 2$, $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\mu = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$, $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} 2 & \sqrt{2}/2 \\ \sqrt{2}/2 & 1 \end{pmatrix}$

$$\sigma_{12} = \sigma_{21} = \rho_{12} \sqrt{\sigma_{11}} \sqrt{\sigma_{22}} = \frac{\sqrt{2}}{2}$$

$|\Sigma| = 2 - \frac{2}{4} = \frac{3}{2}$ $\quad |\Sigma|^{1/2} = \sqrt{\frac{3}{2}}$ $\quad \Sigma^{-1} = \frac{2}{3} \begin{pmatrix} 1 & -\sqrt{2}/2 \\ -\frac{\sqrt{2}}{2} & 2 \end{pmatrix}$

BIVARIATE NORMAL DISTRIBUTION

$$f(x) = \frac{1}{(2\pi)^{2/2}\sqrt{3/2}} \exp\left\{ \frac{-1}{2} \begin{pmatrix} x_1 & x_2 - 2 \end{pmatrix} \frac{2}{3} \begin{pmatrix} 1 & -\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 - 2 \end{pmatrix} \right\}$$

$$= \frac{1}{(2\pi)\sqrt{3/2}} \exp\left\{ \frac{-1}{3} \left(x_1^2 - \sqrt{2}\, x_1 (x_2 - 2) + 2(x_2 - 2)^2 \right] \right.$$

MULTIVARIATE NORMAL DISTRIBUTION

Result: Let $x_1, x_2 \ldots x_n$ be mutually independent with X_j distributed as $N_p(\mu, \Sigma)$ then

$\Rightarrow V_1 = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$

is distributed as $N_p\left(\sum_{j=1}^{n} c_j \mu_j, \left(\sum_{j=1}^{n} c_j^2\right)\Sigma\right)$

&

$\Rightarrow V_2 = b_1 X_1 + b_2 X_2 + \cdots + b_n X_n \sim N_p\left(\sum_{j=1}^{n} b_j \mu_j, \left(\sum_{j=1}^{n} b_j^2\right)\Sigma\right)$

are jointly multivariate normal with covariance matrix

$$\begin{bmatrix} \left(\sum_{j=1}^{n} c_j^2\right)\Sigma & (b'c)\Sigma \\ (b'c)\Sigma & \left(\sum_{j=1}^{n} b_j^2\right)\Sigma \end{bmatrix}$$

V_1 & V_2 are independent if $(b'c)\Sigma = 0$

MULTIVARIATE NORMAL DISTRIBUTION

Let X1, X2, X3 and X4 be independent and identically distributed (3x1) random vectors with

$$\mu = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \qquad \Sigma = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

Consider the linear combinations of random vectors

$$\frac{1}{2}X_1 + \frac{1}{2}X_2 + \frac{1}{2}X_3 + \frac{1}{2}X_4$$

$$\& \qquad X_1 + X_2 + X_3 - 3X_4 .$$

Find the mean vector and covariance matrix for each linear combination of vectors and also the covariance between them.

$\longrightarrow$ $c_1X_1 + c_2X_2 + c_3X_3 + c_4X_4$ .. Comparing with a linear combⁿ

$$c_1 = c_2 = c_3 = c_4 = \frac{1}{2}$$

MULTIVARIATE NORMAL DISTRIBUTION

To find mean vector & covariance matrix $\sim N_p\left(\sum_{j=1}^{n} c_j \mu_j \ , \left(\sum_{j=1}^{n} c_j^2\right)\Sigma\right)$

$$\sum_{j=1}^{n} c_j \mu_j = (c_1 + c_2 + c_3 + c_4)\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \left(\frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}\right)\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}$$

$$= 2\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ -2 \\ 2 \end{pmatrix}$$

$$\left(\sum_{j=1}^{n} c_j^2\right)\Sigma = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^2 \Sigma = 1 . \Sigma = \begin{bmatrix} 3 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix}$$

MULTIVARIATE NORMAL DISTRIBUTION

$x_1 + x_2 + x_3 - 3x_4$ comparing it with $b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4$

$$b_1 = b_2 = b_3 = 1 \qquad b_4 = -3$$

$$Y_2 \sim N_p\left(\sum_{j=1}^{n} b_j \mu_j ,\ \left(\sum_{j=1}^{n} b_j^2\right)\Sigma\right)$$

$$(b_1 + b_2 + b_3 + b_4)\begin{bmatrix}3\\-1\\1\end{bmatrix}$$

$$= 0\begin{bmatrix}3\\-1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

$$b_1^2 + b_2^2 + b_3^2 + b_4^2\begin{bmatrix}3 & -1 & 1\\-1 & 1 & 0\\1 & 0 & 2\end{bmatrix}$$

$$= 12\begin{bmatrix}3 & -1 & 1\\-1 & 1 & 0\\1 & 0 & 2\end{bmatrix} = \begin{bmatrix}36 & -12 & 12\\-12 & 12 & 0\\12 & 0 & 24\end{bmatrix}$$

MULTIVARIATE NORMAL DISTRIBUTION

$$c_1 = c_2 = c_3 = c_4 = \frac{1}{2}$$

$$b_1 = b_2 = b_3 = 1 \quad , \; b_4 = -3$$

$$(b_1c_1 + b_2c_2 + b_3c_3 + b_4c_4)\Sigma = 0$$

v_1 & v_2 are independent.

MULTIVARIATE NORMAL DISTRIBUTION

Ex. 2) Let X_1, X_2, X_3, X_4 be independent $N_p(\mu, \Sigma)$ random vectors

$$V_1 = \frac{1}{4}X_1 - \frac{1}{4}X_2 + \frac{1}{4}X_3 - \frac{1}{4}X_4$$

&

$$V_2 = \frac{1}{4}X_1 + \frac{1}{4}X_2 - \frac{1}{4}X_3 - \frac{1}{4}X_4$$

whether V_1 & V_2 are independent.

$\rightarrow$ $c_1 = c_3 = \frac{1}{4}$ $\quad c_2 = c_4 = -\frac{1}{4}$

$$\Sigma c_j \mu_j = (c_1 + c_2 + c_3 + c_4)(\mu)$$
$$= 0$$

MULTIVARIATE NORMAL DISTRIBUTION

$v_1 \sim N(0, \frac{1}{4}\Sigma) \qquad v_2 \sim N(0, \frac{1}{4}\Sigma)$

$$b'c\,\Sigma = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}$$

$$= 0$$

v_1 & v_2 are independent.

MULTIVARIATE NORMAL DISTRIBUTION

Ex. Let X be $N_3(\mu, \Sigma)$ with $\mu' = (2 \quad -3 \quad 1)$ and $\Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$

a) Find the distribution of $3X_1 - 2X_2 + X_3$

b) find a 2×1 vector a such that X_2 and $X_2 - a'\begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$ are independent.

a) Let $a = (3 \quad -2 \quad 1)'$, then $a'X = 3X_1 - 2X_2 + X_3$

$$a'X \sim N_p(a'\mu, a'\Sigma a)$$

$$a'\mu = (3 \quad -2 \quad 1)\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} = 13$$

$$a'\Sigma a = (3 \quad -2 \quad 1)\begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

$$= 9$$

MULTIVARIATE NORMAL DISTRIBUTION

The dist^n of $3x_1 - 2x_2 + x_3$ is $N_3(13, 9)$

b) Let $a' = (a_1 \quad a_2)$ then $y = x_2 - a'\begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = -a_1 x_1 + x_2 - a_3 x_3$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -a_1 & 1 & -a_3 \end{bmatrix} \quad \text{then} \quad AX = \begin{bmatrix} x_2 \\ y \end{bmatrix} \sim N(A\mu, A'\Sigma A)$$

$$A'\Sigma A = \begin{pmatrix} 0 & 1 & 0 \\ -a_1 & 1 & -a_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & -a_1 \\ 1 & 1 \\ 0 & -a_3 \end{pmatrix}$$

$$\checkmark \quad A'\Sigma A = \begin{array}{cc} \sigma_{11} & \\ \left[\begin{array}{cc} 3 & -a_1 - 2a_2 + 3 \quad (\sigma_{12}) \\ -a_1 - 2a_2 + 3 & a_1^2 - 2a_1 - 4a_2 + 2a_1 a_2 + 2a_2^2 + 3 \end{array}\right] & \\ (\sigma_{21}) & \sigma_{22} \end{array}$$

MULTIVARIATE NORMAL DISTRIBUTION

since, we want to have X_2 & Y ind., this implies that $-a_1 - 2a_2 + 3 = 0$

$$a = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c \begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad \text{for } c \in \mathbb{R}.$$

MULTIVARIATE NORMAL DISTRIBUTION

Conditional Distribution of Bivariate Normal Distribution

$$f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f(x_2)}$$

Result:-

If $f(x_1, x_2)$ is the bivariate normal density, then

$$f(x_1 \mid x_2) \sim N\left(\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2),\ \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}}\right)$$

$\Downarrow$ Mean

$E(X_1 \mid X_2)$

$\Downarrow$ Variance

$V(X_1 \mid X_2)$

MULTIVARIATE NORMAL DISTRIBUTION

Relation to Conditional Distribution to Regression Model

If the joint distⁿ of x & y is a normal distⁿ, then

$$E(y|x) = \alpha + \beta x$$

$$x^2 - 2\rho xy + y^2$$

$$x^2 - 2\rho xy + y^2 - \rho^2 x^2 + \rho^2 x^2$$

$$(y^2 - 2\rho xy + \rho^2 x^2) - \rho^2 x^2 - x^2$$

$$(y - \rho x)^2$$

consider the bivariate distⁿ with PDF as

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}$$

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_y}{\sigma_y} - \rho\,\frac{x-\mu_x}{\sigma_x}\right)^2 - (1-\rho^2)\left(\frac{x-\mu_x}{\sigma_x}\right)^2\right]\right\}$$

MULTIVARIATE NORMAL DISTRIBUTION

$$f(x,y) = f(y|x) \cdot f(x)$$

$$\mu_{y|x} = \mu_y + \frac{\rho \sigma_y}{\sigma_x}(x - \mu_x) \Rightarrow \quad a + bx$$

MULTIVARIATE NORMAL DISTRIBUTION

suppose that weight (lbs) & height (inches) of undergraduate college men have MVN with

$$\mu = \begin{pmatrix} 175 \\ 71 \end{pmatrix} \quad \& \quad \Sigma = \begin{pmatrix} 550 & 40 \\ 40 & 8 \end{pmatrix}$$

Find mean & variance of conditional distn

$$\longrightarrow \quad \text{Mean} = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(x_2 - \mu_2) = 175 + \frac{40}{8}(x_2 - 71) = 5x_2 - 180$$

$$\text{variance} = \sigma_{11} - \frac{\sigma_{12}^2}{\sigma_{22}} = 550 - \frac{40^2}{8} = 350$$

MULTIVARIATE NORMAL DISTRIBUTION

$x_2 = 70$, $E(y|x) = -180 + 5x_2 \Rightarrow \beta_0 + \beta_1 x$

$= 170$

(height)

(weight)

MULTIVARIATE NORMAL DISTRIBUTION

Candy company makes 3 size candy bars.

X_1 : Regular $\quad\quad$ X_3 : Big size

X_2 : Fun size

$$\mu = \begin{bmatrix} 5 \\ 3 \\ 7 \end{bmatrix} \quad \& \quad \Sigma = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{bmatrix}$$

What is the prob. that

i) $X_1 > 8$

ii) $(X_1 \mid X_2 = 1, X_3 = 10)$

iii) $P(4X_1 - 3X_2 + 5X_3 < 63)$

MULTIVARIATE NORMAL DISTRIBUTION

1) $X_1 \sim N(5, 4)$

The prob. that the regular bar is more than 8oz is

$$P(X_1 > 8) = P\left(Z > \frac{8-5}{2}\right)$$

$$= P(Z > 1.5)$$ Normal table / using calculator

$$P(X_1 > 8) = 0.0668$$

MULTIVARIATE NORMAL DISTRIBUTION

III) $4x_1 - 3x_2 + 5x_3$ is normally distributed

$$a' = (4 \quad -3 \quad 5) \qquad a'x = 4x_1 - 3x_2 + 5x_3$$

$$a'\mu = (4 \quad -3 \quad 5)\begin{pmatrix} 5 \\ 3 \\ 7 \end{pmatrix} = 46$$

$$a'\Sigma a = (4 \quad -3 \quad 5)\begin{pmatrix} 4 & -1 & 0 \\ -1 & 4 & 2 \\ 0 & 2 & 9 \end{pmatrix}\begin{pmatrix} 4 \\ -3 \\ 5 \end{pmatrix}$$

$$= (4 \quad -3 \quad 5)\begin{pmatrix} 19 \\ -6 \\ 39 \end{pmatrix} = 289$$

MULTIVARIATE NORMAL DISTRIBUTION

$$P(4x_1 - 3x_2 + 5x_3 < 63) = P\left(Z < \frac{63-46}{\sqrt{289}}\right)$$

$$= P(Z < 1)$$

$$= \phi(1)$$

$$= 0.841347$$