Chapter Three

Estimation of the mean vector and the covariance matrix

1. The maximum likelihood method

Given asample of (vector) observations from p-variate (nondegenerate) normal distribution, we ask for estimators of the mean vector $\underline{\mu}$ and the covariance matrix Σ of the distribution, we shall deduce the maximum likelihood estimators.

Suppose of n observations of p- dimensional vectors $\underline{x}_1, \dots, \underline{x}_n$, $\underline{x}_i \in R^p$, are mutually independent each distribution $N_p(\underline{\mu}, \Sigma)$ where $(n>p)$. The likelihood function is:

$$L(\underline{x}; \underline{\mu}, \Sigma) = \prod_{j=1}^{n} f(\underline{x}_j, \underline{\mu}, \Sigma)$$

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

$$= (2\pi)^{-np/2} |\Sigma|^{-n/2} - \frac{1}{2}\sum_{j=1}^{n} (\underline{x}_j-\underline{\mu})'\Sigma^{-1}(\underline{x}_j$$

The likelihood function after takin the log is:

$$\operatorname{Log} L(\underline{X};\underline{\mu},\Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{X}_j-\underline{\mu})'\Sigma^{-1}(\underline{X}_j-\underline{\mu})$$

In order to maximize the log-likelihood we first have to compute the partial derivative:

$$\frac{\partial \operatorname{Log} L(\underline{X};\underline{\mu}\Sigma)}{\partial \underline{\mu}} = -\frac{1}{2}\sum_{j=1}^{n}(-2)\Sigma^{-1}(\underline{X}_j-\underline{\mu})$$

Setting $\dfrac{\partial \operatorname{Log} L(\underline{X};\underline{\mu},\Sigma)}{\partial \underline{\mu}} = 0$

we obtain the maximum likelihood estimators as follows:

$$\Sigma^{-1}\sum_{j=1}^{n}(\underline{X}_j-\hat{\underline{\mu}})=0$$

$$\sum_{j=1}^{n}(\underline{X}_j-\hat{\underline{\mu}})=0$$

$$\Rightarrow \sum_{j=1}^{n}\underline{X}_j = n\hat{\underline{\mu}}$$

$$\Rightarrow \hat{\underline{\mu}} = \frac{1}{n}\sum_{j=1}\underline{X}_j = \underline{\bar{X}} = \begin{pmatrix}\bar{X}_1\\ \vdots\\ \bar{X}_p\end{pmatrix}$$

where $\bar{X}_i = \dfrac{\sum_{j=1}^{n} x_{ij}}{n}$, $i=1,2,\ldots,p$

we see that the maximum likelihood estimators (MLE) of $\underline{\mu}$ is the sample mean vector $\underline{\bar{X}}$.

In order to obtain the maximum likelihood estimator of Σ, we derivative Log L with respect of Σ as follows:

$$\frac{\partial \operatorname{Log} L(\underline{x}_j;\underline{\mu},\Sigma)}{\partial \Sigma} = -\frac{n}{2}\Sigma^{-1} + \frac{1}{2}\sum_{j=1}^{n}\Sigma^{-1}(\underline{x}_j-\underline{\mu})(\underline{x}_j-\underline{\mu})'\Sigma^{-1}$$

Setting $\frac{\partial \operatorname{Log} L(\underline{x}_j;\underline{\mu},\Sigma)}{\partial \Sigma} = 0$

we obtain,

$$\hat{\Sigma}^{-1}\sum_{j=1}^{n}(\underline{x}_j-\hat{\underline{\mu}})(\underline{x}_j-\hat{\underline{\mu}})' = n$$

$$\Rightarrow \hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j-\underline{\bar{X}})(\underline{x}_j-\underline{\bar{X}})'$$

we see that the maximum likelihood estimator of Σ is the sample covarianc matrix $\hat{\Sigma}$.

where $\underline{x}_j=(x_{1j},\dots,x_{pj})'$ and the matrix of sums of squares and cross products of deviation about the mean b

$$\sum_{j=1}^{n}(\underline{x}_j-\underline{\bar{X}})(\underline{x}_j-\underline{\bar{X}})' = \sum_{j=1}^{n}\underline{x}_j\underline{x}_j' - n\underline{\bar{X}}\underline{\bar{X}}'.$$

"بسم الله الرحمن الرحيم"

Estimation of the mean vector ($\underline{\mu}$) and the variance - covariance matrix (Σ).

Def: Suppose we have a r.s of size n from normal dist. $N_p(\mu, \Sigma)$ then Liklehood fun. can be written as follows:

$$L(x; \underline{\mu}, \Sigma) = \prod_{j=1}^{n} f(x_j; \underline{\mu}, \Sigma)$$

$$= \prod_{j=1}^{n} \left[\frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left\{ -\frac{1}{2} (\underline{x}_j - \underline{\mu})' \Sigma^{-1} (\underline{x}_j - \right.\right.$$

$$= (2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\left\{ -\frac{1}{2} \Sigma (\underline{x}_j - \underline{\mu})' \Sigma^{-1} (\underline{x}_j - \right.$$

وبأخذ اللوغاريتم ln [illegible]

$$\ln L = -\frac{np}{2} \ln(2\pi) - \frac{n}{2} \ln|\Sigma| - \frac{1}{2} \{ \Sigma (\underline{x}_j - \underline{\mu})' \Sigma^{-1} (\underline{x}_j - $$

.... (1)

بإيجاد المشتقة بالنسبة الى μ [illegible] للمعادلة (1) ينتج.

$$\frac{\partial \ln L}{\partial \underline{\mu}} = 0 - 0 - \frac{1}{2} \sum_{j=1}^{n} (-2) \Sigma^{-1} (\underline{x}_j - \underline{\mu})$$

$$\therefore \frac{\partial \ln L}{\partial \underline{\mu}} = \sum_{j=1}^{n} (\Sigma^{-1}) (\underline{x}_j - \underline{\mu})$$

وبجعل $\frac{\partial \ln L}{\partial \underline{\mu}} = 0$ ينتج أن.

$$0 = \sum_{j=1}^{n} \Sigma^{-1} (\underline{x}_j - \hat{\underline{\mu}})$$

$$\sum_{j=1}^{n} (\underline{x}_j - \hat{\underline{\mu}}) = 0$$

[illegible]

$$\sum_{j=1}^{n} x_j - n\hat{\mu} = 0$$

$$\sum_{j=1}^{n} x_j = n\hat{\mu}$$

$$\hat{\mu} = \frac{\sum_{j=1}^{n} x_j}{n}$$

$$\boxed{\hat{\mu} = \bar{x}}$$

$\therefore \bar{x}$ is the m.l.e. for μ.

[illegible] 1-W.M

ولإيجاد مقدر الإمكان الأعظم [illegible] التباين المشترك Σ.
وذلك بالتعويض [illegible] عن μ بـ $\bar{x}$ في المعادلة (1)
[illegible] بالنسبة الى Σ
[illegible]

now replace $\hat{\mu} = \bar{x}$ is equ.(1) and then,

$$\ln L^* = \frac{-np}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})'\Sigma^{-1}(\underline{x}_j - \bar{\underline{x}})$$

$$\frac{\partial \ln L^*}{\partial \Sigma} = 0 - \frac{n}{2}\Sigma^{-1} - \frac{1}{2}\sum_{j=1}^{n} -\Sigma^{-1}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\Sigma^{-1}$$

$$\frac{\partial \ln L^*}{\partial \Sigma} = -\frac{n}{2}\Sigma^{-1} + \frac{1}{2}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\Sigma^{-1}$$

بوضع $\frac{\partial \ln L^*}{\partial \Sigma} = 0$ ينتج أن.

$$0 = \frac{-n}{2}\hat{\Sigma}^{-1} + \frac{1}{2}\hat{\Sigma}^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\hat{\Sigma}^{-1}$$

let $\boxed{\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})' = A}$

$$0 = \frac{-n}{2}\hat{\Sigma}^{-1} + \frac{1}{2}\hat{\Sigma}^{-1} A\, \hat{\Sigma}^{-1}$$

بضرب الطرفين بـ $\hat{\Sigma}$ ينتج أن

$$0 = \frac{-n}{2}\hat{\Sigma}^{-1}\hat{\Sigma} + \frac{1}{2}\hat{\Sigma}^{-1} A\, \hat{\Sigma}^{-1}\hat{\Sigma}$$

$$0 = \frac{-n}{2} I + \frac{1}{2}\hat{\Sigma}^{-1} A$$

$$n = \hat{\Sigma}^{-1} A$$

$$\boxed{A = n\hat{\Sigma}}$$

$\Rightarrow \therefore \boxed{\hat{\Sigma} = \frac{A}{n}}$ is the m.l.e of Σ, where A is the matrix of sums of squares and cross products.

212

Not $\hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j - \hat{\underline{\mu}})(\underline{x}_j - \hat{\underline{\mu}})'$
$= \frac{1}{n}\sum x_j x_j' - \bar{x}_j \bar{x}_j' = \sum x_j x_j' - n\bar{\underline{x}}\,\bar{\underline{x}}'$

Since $E(\underline{\bar{X}}) = \frac{1}{n}\sum_{j=1}^{n} E(\underline{X}_j) = \frac{1}{\cancel{n}}\cancel{n}\cdot\underline{\mu}$

$$\Longrightarrow E(\underline{\bar{X}}) = \underline{\mu}$$

$\therefore$ The sample mean vector $\underline{\bar{X}}$ is an unbiased estimator of $\underline{\mu}$.

However,

$$E(\hat{\Sigma}) = \frac{1}{n} E\left[\sum_{j=1}^{n}(\underline{x}_j - \underline{\bar{x}})(\underline{x}_j - \underline{\bar{x}})'\right]$$

$$= \frac{1}{n} E\left[\sum_{j=1}^{n}\underline{x}_j\underline{x}_j' - n\underline{\bar{x}}\,\underline{\bar{x}}'\right]$$

$$= \frac{1}{n}\left[\sum_{j=1}^{n} E(\underline{X}_j\underline{X}_j') - nE(\underline{\bar{X}}\,\underline{\bar{X}}')\right]$$

Since $E(\underline{X}_j\underline{X}_j') = \Sigma + \underline{\mu}\,\underline{\mu}'$

Therefore, we have

$$E(\hat{\Sigma}) = \frac{1}{n}\left[(n\Sigma + n\underline{\mu}\,\underline{\mu}') - (n\frac{\Sigma}{n} + n\underline{\mu}\,\underline{\mu}')\right]$$

$$= \frac{1}{n}\left[n\Sigma + \cancel{n\underline{\mu}\,\underline{\mu}'} - \Sigma - \cancel{n\underline{\mu}\,\hat{\underline{\mu}}}\right]$$

$$\therefore E(\hat{\Sigma}) = \frac{(n-1)}{n}\Sigma$$

Thus $\hat{\Sigma}$ is biased estimator of Σ.

We shall therefore define:

$$S = \frac{1}{n-1}\sum_{j=1}^{n}(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_j - \bar{\underset{\sim}{x}})'$$

as the sample covariance matrix. It is an unbiased estimator of Σ, where

$$S = \begin{pmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}$$

Also the (MLE) of the partial correlation is,

$$\hat{\rho}_{12.3\cdots p} = \frac{\hat{\sigma}_{12.3\cdots p}}{\sqrt{\hat{\sigma}_{11.3\cdots p}}\sqrt{\hat{\sigma}_{22.3\cdots p}}}$$

And the (MLE) of the multiple correlation is

$$\hat{R}_{1.2,\cdots p} = \sqrt{\frac{S_{(1)} S_{22}^{-1} S_{(1)}'}{S_{11}}}$$

Also

$$\hat{\rho}_{11.2} = (\hat{D}^{-\frac{1}{2}})\, S_{11.2}\, (\hat{D}^{-\frac{1}{2}})$$

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

The properties of m.l.e. of $\underline{\mu}$ & Σ

① Unbiased :

(i) for $\hat{\underline{\mu}}$. we have $x \sim N_p(\underline{\mu}, \Sigma)$

$$\hat{\mu} = \bar{x}$$

$$E\hat{\underline{\mu}} = E\bar{\underline{x}}$$

$$= E\ \frac{1}{n}\sum_{j=1}^{n} \underline{x}_j$$

$$= \frac{1}{n}\sum_{j=1}^{n} E x_j$$

$$= \frac{1}{n}(n)(\underline{\mu})$$

$$= \frac{\cancel{n}}{\cancel{n}}\ \mu$$

$$E\hat{\mu} = \mu$$

$\therefore$ $\hat{\mu}$ is unbiased estimator for μ.

(ii) for $\hat{\Sigma}$ we have $x_j \sim N_p(\underline{\mu}, \Sigma)$ and $\bar{\underline{x}} \sim N_p(\underline{\mu}, \frac{1}{n}\Sigma)$

$$\hat{\Sigma} = \frac{1}{n}\Sigma x_j x_j' - \bar{\underline{x}}\,\bar{\underline{x}}'$$

$$E(\hat{\Sigma}) = E\left\{\frac{1}{n}\sum_{j=1}^{n} \underline{x}_j \underline{x}_j' - \bar{\underline{x}}\,\bar{\underline{x}}'\right]$$

$$E(\hat{\Sigma}) = \frac{1}{n}\sum_{j=1}^{n} E\underline{x}_j\underline{x}_j' - E\bar{\underline{x}}\,\bar{\underline{x}}'$$

$$= \frac{1}{n}\sum_{j=1}^{n}\Sigma - \frac{1}{n}\Sigma$$

$$E(\hat{\Sigma}) = \frac{1}{\cancel{n}}(n)\Sigma - \frac{1}{n}\Sigma$$

$$E(\hat{\Sigma}) = \frac{\cancel{n}}{\cancel{n}}\Sigma - \frac{1}{n}\Sigma$$

$$E(\hat{\Sigma}) = \Sigma\left(1 - \frac{1}{n}\right)$$

$$\boxed{E(\hat{\Sigma}) = \left(\frac{n-1}{n}\right)\Sigma}$$

215

$\therefore$ $\hat{\Sigma}$ is biased est. for Σ.

notes

① if $x \sim N_p(\underset{\sim}{\mu}, \Sigma)$ then $\bar{\underset{\sim}{x}} \sim N_p(\underset{\sim}{\mu}, \frac{1}{n}\Sigma)$

proof:

$$M_{\bar{\underset{\sim}{x}}}(t) = E\, e^{\underset{\sim}{t}'\bar{\underset{\sim}{x}}}$$

$$= E\, e^{\underset{\sim}{t}' \frac{1}{n}\sum_{j=1}^{n} x_j}$$

$$= E\, e^{\frac{t'}{n}(x_1+x_2+\cdots+x_n)}$$

$$= E\left[e^{\frac{t'}{n}x_1} \cdot e^{\frac{t'}{n}x_2} \cdots e^{\frac{t'}{n}x_n}\right]$$

$$= E\left(e^{x_1\left(\frac{t'}{n}\right)}\right) \cdot \left(E\, e^{x_2\left(\frac{t'}{n}\right)}\right) \cdots \left(E\, e^{x_n\left(\frac{t'}{n}\right)}\right)$$

$$M_{\bar{\underset{\sim}{x}}}(t) = M_{x_1}\left(\frac{t'}{n}\right) \cdot M_{x_2}\left(\frac{t'}{n}\right) \cdots M_{x_n}\left(\frac{t'}{n}\right)$$

The m.g.f. of $\bar{x}$ is. عبارة الدالة المولدة للعزوم

$$\boxed{M_{\bar{\underset{\sim}{x}}}(t) = \left\{e^{\left(\frac{t'}{n}\right)\underset{\sim}{\mu} + \frac{1}{2}\left(\frac{t'}{n}\right)\Sigma\left(\frac{t}{n}\right)}\right\}^n}$$

$$M_{\bar{\underset{\sim}{x}}}(t) = e^{t'\underset{\sim}{\mu} + \frac{1}{2}\underset{\sim}{t}'\Sigma\underset{\sim}{t}}$$

$$\therefore \boxed{\bar{\underset{\sim}{x}} \sim N_p\left(\mu, \frac{\Sigma}{n}\right)}$$

② $S = \frac{A}{n-1}$ unbiased est. of Σ.

$$\hat{\Sigma} = \frac{A}{n} = \frac{1}{n}\sum_{\alpha=1}^{n}(x_\alpha - \bar{x})(x_\alpha - \bar{x})'$$

$$\Rightarrow \boxed{E(S) = \hat{\Sigma}}$$

proof it. H.W

② Consistency

ⓘ for $\hat{\underline{\mu}}$

we have $\hat{\underline{\mu}} = \bar{\underline{x}}$ and $\bar{\underline{x}} \sim N_p(\underline{\mu}, \frac{\underline{\Sigma}}{n})$

① $$\lim_{n\to\infty} E(\hat{\underline{\mu}} - \underline{\mu}) = \lim_{n\to\infty} (E\hat{\mu} - \mu) = \lim_{n\to\infty} [\mu - \mu] = \lim_{n\to\infty} (0) = \boxed{0}$$

② $$\lim_{n\to\infty} var(\hat{\mu}) = \lim_{n\to\infty} \left(\frac{\Sigma}{n}\right) = \frac{1}{\infty}(\Sigma) = (0)(\Sigma) = 0$$

$\therefore$ $\hat{\mu} = \bar{x}$ is a consistency estimator for μ.

(ii) for $\hat{\Sigma} = \frac{A}{n}$

① $$\lim_{n\to\infty} E(\hat{\Sigma} - \Sigma) = \lim_{n\to\infty} \left(E(\hat{\Sigma}) - \Sigma\right)$$
$$= \lim_{n\to\infty} \left(\frac{n-1}{n}\Sigma - \Sigma\right)$$
$$= \lim_{n\to\infty} \{\Sigma \left(\frac{n-1}{n} - 1\right)$$
$$= \lim_{n\to\infty} \{\Sigma \left(\frac{\cancel{n}-1-\cancel{n}}{n}\right)$$
$$= \lim_{n\to\infty} \{\left(\frac{-1}{n}\right)\Sigma\} = \Sigma \lim_{n\to\infty}\left(\frac{-1}{n}\right)$$
$$= \left(\frac{1}{\infty}\right)(\Sigma)$$
$$= (0)(\Sigma)$$
$$= 0$$

② $$\lim_{n\to\infty} var(\hat{\Sigma}) = \lim_{n\to 0}\left(\frac{1}{n^2}\Sigma_A\right) = 0.$$

now we want to find

$$Var(\hat{\Sigma}) = Var(S) = E\left(\hat{\Sigma} - E\hat{\Sigma}\right)\left(\hat{\Sigma} - E\hat{\Sigma}\right)'$$
$$= E\left[\frac{A}{n} - \frac{n-1}{n}\Sigma\right]\left[\frac{A}{n} - \frac{n-1}{n}\Sigma\right]'$$
$$= \frac{1}{n^2} E\left[A - (n-1)\Sigma\right]\left[A - (n-1)\Sigma\right]'$$

$$Var(\hat{\Sigma}) = \frac{1}{n^2}\Sigma$$

$\therefore$ $\hat{\Sigma} = \frac{A}{n}$ is aconsistent est. for Σ.

③ Sufficient

(i) for $\hat{\mu}$

if Σ is known

then $\bar{x}$ is sufficient statistics for μ.

$$\ln L(\underline{\mu}, \Sigma) = \underbrace{K - \frac{1}{2} \operatorname{tr} \Sigma^{-1} A}_{h(x)} - \underbrace{\frac{n}{2} (\underline{\bar{x}}' - \underline{\mu})' \Sigma^{-1} (\underline{\bar{x}} - \underline{\mu})}_{g(\bar{x}, \mu)}$$

where the function of observation $h(x_1, \dots, x_n)$ only does not depend on μ.

By using factrization thm.

$$\prod_{\alpha=1}^{n} f(x; \mu) = g(\bar{x}, \underline{\mu}) . h(x_1, x_2, \dots, x_n)$$

$\therefore$ $\hat{x} = \hat{\mu}$ is asufficient statistic for μ.

(ii) for $\hat{\Sigma}$

if μ is known, $\hat{\Sigma} = \frac{A}{n}$ is not sufficient statistic

for Σ but. $\sum_{\alpha=1}^{n} (X_\alpha - \underline{\mu})' (X_\alpha - \underline{\mu})'$ is asufficie

statistics for Σ.

④ Efficient.

for $\underline{\hat{\mu}}, \bar{x}$ is an efficient statistic for μ,

$\bar{x}$ K s have efficientcy

eff. $\frac{V(\bar{x})}{V(S)} \leq 1$

~~[illegible]~~

Ex: Suppose that a random sample of size (10) originated from bivariate normal $N_2(\underline{\mu}, \Sigma)$ is given as hollows:

X_1: 1.9 0.8 1.1 0.1 −0.1 4.4 5.5 1.6 4.6 3.4
X_2: 0.7 −1.6 −0.2 −1.2 −0.1 3.4 3.7 0.8 0 2

Compute the sample mean vector and Sample covariance matrix

Sol

$$\underline{\hat{\mu}} = \underline{\bar{X}} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}$$

$$\bar{x}_1 = \frac{\sum_{j=1}^{10} x_{1j}}{10} = \frac{23.3}{10} = 2.32$$

$$\bar{x}_2 = \frac{\sum_{j=1}^{10} x_{2j}}{10} = \frac{7.5}{10} = 0.75$$

$$\Rightarrow \underline{\bar{X}} = \begin{pmatrix} 2.32 \\ 0.75 \end{pmatrix}$$

$$S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \begin{pmatrix} 4.011 & 2.849 \\ 2.849 & 3.211 \end{pmatrix}$$

and $$\hat{\Sigma} = \begin{pmatrix} 3.61 & 2.564 \\ 2.564 & 2.89 \end{pmatrix}$$

H.w: If you have asample of siZ

(10) as follows:

X_1: 35 35 40 10 6 20 35 35 35 30

X_2: 35 49 30 27 28 25 15 22 30 32

X_3: 22 18 25 30 32 23 18 15 20 1

Compute $\underline{\bar{X}}$, S, $\hat{\Sigma}$ and $\hat{R}$

2. Distribution of $\bar{\underline{X}}$ and S

For the distribution of $\bar{\underline{x}} = \sum_{j=1}^{n} \underline{x}_j / n$, we can distinguish two cases:

① when $\bar{\underline{X}}$ is based on a random sample $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ from a multivariate normal distribution $N_p(\underline{\mu}, \Sigma)$, then $\bar{\underline{X}} \sim N_p(\underline{\mu}, \Sigma/n)$.

② when $\bar{\underline{X}}$ is based on a random sample $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ from a nonnormal multivariate population with mean vector $\underline{\mu}$ and covariance matrix Σ, then for large n, $\bar{\underline{X}}$ is approximately $N_p(\underline{\mu}, \Sigma/n)$.

More formally, this result is known as the multivariate central limit theorem.

If $\bar{\underline{X}}$ is the mean vector of a random sample $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ from a population with mean vector $\underline{\mu}$ and covariance matrix Σ, then as $n \longrightarrow \infty$, the distribution of $\sqrt{n}(\bar{\underline{X}} - \underline{\mu})$ approaches $N_p(\underline{0}, \Sigma)$.

③ There are $\textcircled{p}$ variances in S and $\binom{p}{2}$ covariances, for a total of

$$p + \binom{p}{2} = p + \frac{p(p-1)}{2} = \frac{p(p+1)}{2}$$

distinct entries. The joint distribution of these distinct variables in $A = (n-1)S = \sum_{j=1}^{n} (\underline{x}_j - \bar{\underline{x}})$ $(\underline{x}_j -$

3- The wishart distribution

any symmtric positive definite matrix A of Quadratic ~~form~~ and bilinear form which can be transformed to the sum $\sum_{i=1}^{n} y_i y_i'$, whose p- component vectors y_i are independent distributed according to the distribution $N(0, \Sigma)$ is said to have wishart dist. (1928)

توزيع ويشرت
هو توزيع للمصفوفة A .

The wishart dist. denoted by $\boxed{W_p(n-1, \Sigma)}$,
where $(n-1)$ is the degree of freedom.

The wishart dist. is the multivariate analogue of χ^2- dist. and it has similar uses.

as we known a χ^2 random variable is the sum of squares of independent standard normal (univariate) r.v.s

$$\sum_{i=1}^{n} z_i^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu_i}{\sigma}\right)^2 \sim \chi^2_{(n)}$$

if $\bar{x}$ substituted for μ, then

$$\sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{\sigma}\right)^2 \sim \chi^2_{(n-1)}$$

ويساوي $\checkmark = \dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)}$

223

Similarly the formal dist. of a wishart r.v.

$$\underbrace{\sum_{j=1}^{n} (\underline{x}_j - \underline{\mu})(\underline{x}_j - \underline{\mu})'}_{\cdot A \text{ نضع}} \sim W_p(n, \Sigma)$$

where $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_n$ are independent distributed as $N_p(\underline{\mu}, \Sigma)$, when

when $\bar{x}$ substituted for $\underline{\mu}$, the dist. remains wishart dist. with one less degree of freedom

$$(n-1) S = \sum_{j=1}^{n} (\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})' \sim W_p(n-1, \Sigma)$$

where $(n-1)$ is the d.o.f. of the distribution

The density fun. of Wishart dist. is

$$W(A; n-1, \Sigma) = \begin{cases} \dfrac{|A|^{\frac{n-p-1}{2}} \, e^{-\frac{1}{2} \operatorname{tr} \Sigma^{-1} A}}{2^{\frac{np}{2}} \, \pi^{\frac{p(p-1)}{4}} \, |\Sigma|^{\frac{n}{2}} \prod_{j=1}^{p} \Gamma\left(\frac{1}{2}(n+1-j)\right)} , & \text{if } A \text{ is p.d.} \\ 0 , & \text{o.w.} \end{cases}$$

where A is a positive definite matrix
n is d.f.

هذه الدالة هي الدالة
العامة لتوزيع χ^2
وعندما $p=1$, $\Sigma=1$ تكون توزيع χ^2 .

Notes

① if the number of variables $\boxed{p}$ equal one $\boxed{p=1}$ and $\Sigma = \sigma^2$ then

$$W_p(n-1, \Sigma) \longrightarrow \chi^2_{(n-1)}$$

and

$$A = \sum_{i=1}^{n} (x_i - \bar{x})^2 \sim \sigma^2 \chi^2_{(n-1)}$$

② if $\boxed{p=1}$ and $\boxed{\Sigma=1}$, the wishart density becomes that of the chi-squared dist. with $\boxed{n}$ d.f.

③ if $A_1, A_2, \dots, A_k$ are indep. distributed as wishart. matrices Σ with common parameter matrix Σ and respective d.f. $n_1, n_2, \dots, n_k$ their sum has the wishart dist. with parameters Σ and $n = n_1 + n_2 + \dots + n_k$

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

إذا لدي عدد $A_k, \dots, A_1$ من المتقلات
وهي مستقلة فيما بينها و مجموعها يتوزع
ايضاً ويشرت.

④ The wishart dist. is used in the multivariat generalization of the <u>t-test</u> called Hotelling's T^2 — test, for comparing the mean vectors of two multivariate normal dist. it is also used in comparing disperion matrices of two multivariate normal dist. and in comparing the mean vectors of several multivariate normal populations in (MANOVA) problem.

وأيضاً يستخدم في حالات الاختبار Hotelling T^2 test
هوتلنك T^2 . عندما نريد أن نقارن
بين المتوسطات
او عند المقارنة بين S ومصفوفة التباين والتباينات المشتركة
او عند المقارنة بين المتوسطات في حالة الـ (MANOVA)
وهذه هي الاستعمالات المختلفة لتوزيع ويشرت .

استعمالات توزيع ويشرت

⑤ For the problem of T^2 and other multivariate normal tests, The non-central wishart dist. must be substituted for the central wishart dist. when the null hypothesis is not true.

T.W. Anderson (Ann, Math, stat, vol (117), 1946, pp(409-431)

has obtaind some properties of the non central wishart dist.

H_0: true → Central wishart
H_1: true → non Central wishart

اذا تحققت صحة الفرضية H_0 هو توزيع ويشرت المركزي
تكون فرضية H_1 .
اما خصائص الـ non central فكان اول من وضعها العالم الفلاني
وان H_1 اذا تحققت هذه الفرضية فالتوزيع هو توزيع ويشرت غير المركزي

⑥ The wishart dist. is amultivariate generalization of the χ^2 dist.

توزيع ويشرت هو تعميم لتوزيع الـ χ^2
والـ χ^2 هو حالة خاصة من الويشرت .

3. Some of basic properties

(i) If $A_1 \sim W_p(n_1-1,\Sigma)$ and $A_2 \sim W_p(n_2-1,\Sigma)$ A_1 and A_2 are independent, then

$$A_1 + A_2 \sim W_p(n_1+n_2-2,\Sigma)$$

(ii) If $A \sim W_p(n-1,\Sigma)$ and B $(p\times q)$ matrix then $B'AB \sim W_q(n-1, B'\Sigma B)$

(iii) If $A \sim W_p(n-1, I)$ and $B_{(p\times q)}$ orthogonal matrix, $B'B = I_q$, then

$$B'AB \sim W_q(n-1, I).$$

Finally, we note that when sampling from a multivarite normal distribution $\underline{\bar{X}}$ and S are independent.