Chapter Four
Testing Hypothesis on one or Two Mean vectors

There are two cases in testing hypothesis in univariate case:

ⓐ One population, one variable

$H_0: \mu = \mu_0$, vs. $H_1: \mu \neq \mu_0$

Tested by either Z or t distribution

ⓑ K-population, one variable

$H_0: \mu_1 = \mu_2 = \cdots = \mu_k$, vs.

H_1: at least two are not equal

Tested by the ANOVA F-test.

Generalization

) one population, p-variables

$H_0: \underline{\mu} = \underline{\mu}_0$, vs. $H_1: \underline{\mu} \neq \underline{\mu}_0$

Extendet t-test

) K-population, p-variables

$H_0: \underline{\mu}_1 = \underline{\mu}_2 = \cdots = \underline{\mu}_k$

Multivariate ANOVA $\longrightarrow$ MANOVA

1. Likelihood Ratio Test (LRT)

If the distribution of the random sample observations $\underline{x}_1, \underline{x}_2, ---, \underline{x}_n$ depends upon aparameter vector $\underline{\theta}$ and if $H_0: \underline{\theta} \in \Omega_0$ and $H_1: \underline{\theta} \in \Omega_1$. The likelihood ratio t statistic fortesting H_0 against H_1 is given o follows:

$$\lambda = \frac{L(\hat{\Omega}_0)}{L(\hat{\Omega}_1)}$$

where $L(\hat{\Omega}_0)$ and $L(\hat{\Omega}_1)$ is the larges value which likelihood function taken regior Ω_0 and Ω_1 or may use the test statistic

$$-2\log\lambda = 2[\overset{*}{l}_1 - \overset{*}{l}_0]$$

where $\overset{*}{l}_1 = \log L(\hat{\Omega}_1)$ and $\overset{*}{l}_0 = \log L(\hat{\Omega}_0)$

For which the rejecting region can be expressed as

$$R = \{\underline{x}; -2\log\lambda > k\}$$

The critical value K has to be determined, so that if Ho is true then

$$pr\{-2\log\lambda > K\} = \alpha$$

(2) Tests concerning means $\boxed{\mu}$

Def. 1: if $Y \sim N_p(0, I_p)$ then

$$Y'Y \sim \chi^2_{(p)}$$

since $f(u = Y'X) = \dfrac{u^{p-1} e^{-u/2}}{\Gamma\frac{p}{2}\, 2^{p/2}}$, $u > 0$

$\boxed{\text{central } \chi^2}$

Def. 2: if $Y \sim N_p(\mu, I_p)$ then $Y'Y \sim \chi'^2_{(p,\lambda)}$,

where $\lambda = \frac{1}{2}\mu'\mu$ (non central parameter)

and $$f(u = Y'Y) = e^{-\lambda} \sum \frac{\lambda^k}{k!} \frac{u^{\frac{p+k-1}{2}} e^{-\frac{u}{2}}}{2^{\frac{p+k}{2}} \Gamma\frac{p+k}{2}}$$

$\boxed{\text{non central } \chi^2}$

Def. 3: if $Y \sim N_p(0, \sigma^2 I_p)$ then $Y'Y \sim \chi^2_p$ central

Def 4: if $Y \sim N_p(\underline{\mu}, \sigma^2 I_p)$ then $Y'Y \sim \chi'^2_{(p)}$ non central

Note

$$\chi'^2_{(p,\lambda)} \Longrightarrow \chi^2_{(p)}$$

if $\lambda = 0$
and $\mu = 0$

(A) Test of hypothesis for the mean vectors $\underline{\mu}$.
where Σ is known

or

in the univariate case:

we have
for the multivariate case, suppose that $\underline{x}_1, \dots, \underline{x}_n$ be
arandom sample from $N_p(\underline{\mu}, \Sigma)$ the likelihood
function is:

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

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

$$\ell^*(x; \underline{\mu}, \Sigma) = \log L(\underline{x}; \underline{\mu}, \Sigma)$$

$$\ell^*(\underline{x}; \underline{\mu}, \Sigma) = \frac{-np}{2} \log(2\pi) - \frac{n}{2} \log|\Sigma| - \frac{1}{2} \sum_{j=1}^{n} (\underline{x}_j - \underline{\mu})' \Sigma^{-1} (\underline{x}_j \quad \dots (**1)$$

in the univariate case. $X \sim N(\mu, \sigma^2)$.

for one sample

(1) The sample mean $\bar{X} = \hat{\mu} = \frac{\sum x_i}{n} \sim N(\mu, \frac{\sigma^2}{n})$

(2) The difference between the samples mean $\boxed{\bar{x}}$ and the population mean $\boxed{\mu}$ is

$$\sqrt{n}\,(\bar{x} - \mu) \sim N(0, \sigma^2)$$

(3) To test $H_0: \mu = \mu_0$ we have $\sqrt{n}\,(\bar{x} - \mu_0) \sim N(0, \sigma^2)$

There for we can use normal tables carry out the test.

for two sample

(4) Suppose that a sample $\boxed{n_1}$ is drown from $N(\mu_1, \sigma$

and another Sample of size n_2 is drown from $\sim N(\mu_2$

where Common Variance σ^2 is Known

Thus

$$\bar{X}_1 = \mu_1 \sim N(\mu_1, \frac{\sigma^2}{n_1})$$

$$\bar{X}_2 = \mu_2 \sim N(\mu_2, \frac{\sigma^2}{n_2})$$

and $(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2})$

and $(\bar{X}_1 - \bar{X}_2) \sim N(\mu_1 - \mu_2, \frac{n_1 + n_2}{n_1 n_2}\sigma_2)$

$$\Rightarrow \boxed{\sqrt{\frac{n_1 n_2}{n_1 + n_2}}\left\{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)\right\} \sim N(0, \sigma^2)}$$

There fore to test $H_0: \mu_1 = \mu_2$ we base on $\mu_1 - \mu_2 = 0$ نفرض صفر لـ

$$\boxed{\sqrt{\frac{n_1 n_2}{n_1 + n_2}}\,(\bar{X}_1 - \bar{X}_2) \sim N(0, \sigma^2)}$$

and we use normal table to carry out the test.

Know to test $H_0: \underline{\mu} = \underline{\mu}_0$, vs. $H_1: \underline{\mu} \neq \underline{\mu}_0$ with Σ known, we have.

$$H_0: \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} = \begin{pmatrix} \mu_{01} \\ \mu_{02} \\ \vdots \\ \mu_{0p} \end{pmatrix}, \quad H_1: \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} \neq \begin{pmatrix} \mu_{01} \\ \mu_{02} \\ \vdots \\ \mu_{0p} \end{pmatrix}$$

where μ_{0j} specified from previous experience or a target value.

The maximum of log-likelihood under $H_0: \underline{\mu} = \underline{\mu}_0$, with Σ known is:

$$\overset{*}{\ell}_0(\underline{x}; \underline{\mu}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)'\Sigma^{-1}(\underline{x}_j - \underline{\mu}_0)$$

$$= -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)'\Sigma^{-1}(\underline{x}_j - \underline{\mu}_0)\right]$$

$$= -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)(\underline{x}_j - \underline{\mu}_0)'\right]$$

$$= -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j + \bar{\underline{x}} - \bar{\underline{x}} - \underline{\mu}_0)(\underline{x}_j + \bar{\underline{x}} - \bar{\underline{x}} - \underline{\mu}_0)'\right]$$

$$\overset{*}{\ell}_0(\underline{x}; \underline{\mu}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\right] - \frac{n}{2}(\bar{\underline{x}} - \underline{\mu}_0)'\Sigma^{-1}(\bar{\underline{x}} - \underline{\mu}_0)$$

233

Under $H_1: \underline{\mu} \neq \underline{\mu}_0$, we have to maximize with respect to $\underline{\mu}$, the maximum of log-likelihood occus when $\hat{\underline{\mu}} = \bar{\underline{x}}$. Then the maximum of log-likelihood under H_1 is:

$$\overset{*}{\ell}_1(\bar{\underline{x}}, \bar{\underline{x}}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})'\Sigma^{-1}(\underline{x}_j - \bar{\underline{x}})$$

$$= -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}\left[\operatorname{tr}\Sigma^{-1}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\right]$$

$$\therefore \overset{*}{\ell}_1(\bar{\underline{x}}, \bar{\underline{x}}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'\right]$$

The likelihood ratio test statistic can be writ as follows:-

$$-2\log\lambda = 2\left[\overset{*}{\ell}_1 - \overset{*}{\ell}_0\right]$$

$$= \cancel{-np\log(2\pi)} \cancel{-n\log|\Sigma|} \cancel{-\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j} \cancel{+np\log(2\pi)} \cancel{+n\log|\Sigma|} \cancel{+\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j} + n(\bar{\underline{x}} - \underline{\mu}_0)'\Sigma^{-1}(\bar{\underline{x}} - \underline{\mu}_0)$$

$$\therefore -2\log\lambda = n(\bar{\underline{x}} - \underline{\mu}_0)\Sigma^{-1}(\bar{\underline{x}} - \underline{\mu}_0)$$

234

This function has an exact χ^2 distribution under H_0.

i.e, $n(\bar{\underline{X}}-\underline{\mu}_0)'\Sigma^{-1}(\bar{\underline{X}}-\underline{\mu}_0) \sim \chi^2_p$

Therefore, we reject H_0 if:

$$\chi^2_{cal} > \chi^2_{(\alpha,p)}$$

where $\chi^2_{(\alpha,p)}$ is chi-square tabulated with degree of freedom p and level of significance α.

Ex: Suppose that a sample of size (20) originated from bivariate normal $N_2(\underline{\mu},\Sigma)$, where

$$\Sigma = \begin{pmatrix} 20 & 100 \\ 100 & 1000 \end{pmatrix} \text{ and } \bar{\underline{X}} = \begin{pmatrix} 71.45 \\ 164.7 \end{pmatrix}$$

(a) Test $H_0: \underline{\mu} = \begin{pmatrix} 70 \\ 170 \end{pmatrix}$, v.s $H_1: \underline{\mu} \neq \begin{pmatrix} 70 \\ 170 \end{pmatrix}$ at $\alpha = 0.05$

(b) If H_0 is rejected, test each variable separately.

235

Sol

(a) $\chi^2 = n(\bar{\underline{x}} - \underline{\mu}_0)' \Sigma^{-1} (\bar{\underline{x}} - \underline{\mu}_0)$

$$= 20 \begin{pmatrix} 71.45 - 70 \\ 164.7 - 170 \end{pmatrix} \begin{pmatrix} 20 & 100 \\ 100 & 1000 \end{pmatrix}^{-1} \begin{pmatrix} 71.45 - 70 \\ 164.7 - 170 \end{pmatrix}$$

$$= 8.402$$

using $\alpha = 0.05 \rightarrow \chi^2_{(0.05, 2)} = 5.99$

Therefore, we reject $H_0: \underline{\mu} = \begin{pmatrix} 70 \\ 170 \end{pmatrix}$

because $\chi^2 = 8.402 > 5.99$

(b) $H_0: \mu_{01} = 70$, vs. $H_1: \mu_{01} \neq 70$

$H_0: \mu_{02} = 170$, vs $H_1: \mu_{02} \neq 170$

$$Z_1 = \frac{\bar{x}_1 - \mu_{01}}{\sigma_1 / \sqrt{n}} = \frac{71.45 - 70}{\sqrt{20/20}} = 1.45$$

$$Z_2 = \frac{\bar{x}_2 - \mu_{02}}{\sigma_2 / \sqrt{n}} = \frac{164.7 - 170}{\sqrt{1000/20}} = -0.7495$$

using $Z_{(\frac{\alpha}{2})} = 1.96$ for $\alpha = 0.05$.

we have $|Z_1| < 1.96$ and $|Z_2| < 1.96$

236

Thus both test accept the hypothesis.

in this case neither of the $\bar{X}_1$ and $\bar{X}_2$ is far enough from the hypothesis value to cause rejection. But when the positive correlation between X_1 and X_2 is taken into account in the multivariate test. the evidences against $\underline{\mu}_0$ combine to cause rejection.

(ii) <u>similarly for the multivariate case.</u>

we have:

(1) $\underline{X} \sim N_p(\underline{\mu}, \Sigma)$ then the sample mean

$$\bar{\underline{X}} = \underline{\hat{\mu}} = \frac{1}{n}\sum_{\alpha=1}^{n} x_{i\alpha} \sim N_p(\underline{\mu}, \frac{\Sigma}{n}) \quad , \forall\, i = 1, 2, \dots, p$$

(2) $\sqrt{n}(\bar{\underline{X}} - \underline{\mu}) \sim N_p(\underline{0}, \Sigma)$

(3) To test $H_0: \underline{\mu} = \underline{\mu}_0$ we have (one sample)

$$\sqrt{n}(\bar{X} - \mu_0) \sim N_p(0, \Sigma)$$

but we can not use the normal table because only a bivariate normal table are available.

(4) There for we have to use another tables where

$$\sqrt{n}(\bar{\underline{X}} - \underline{\mu}) = \underline{Y}$$

where $\underline{Y} \sim N_p(0, \Sigma)$

since Σ is p.d. matrix $\exists$ a non singular mat

let C is a non singular matrix

$\ni C \Sigma C' = I$

Define.

$$\underline{Z} = CY$$

$\therefore Z \sim N_p(0, C \Sigma C')$

$\rightarrow \therefore Z \sim N_p(0, I)$

$\rightarrow \therefore Z_i \sim N(0, 1) \quad , \forall\, i = 1, 2, \dots, p$

$$z'z = \sum_{i=1}^{p} z_i^2 \sim \chi^2_{(p)}$$

$$z'z = (cy)(cy)$$

$$z'z = y'c'cy$$

but $c\Sigma c' = I$

$$\boxed{\Sigma^{-1} = c'c}$$

$$\underline{z}'\underline{z} = \underline{y}'c'c\underline{y} \sim \chi^2_{(p)}$$

$$= \underline{y}'\Sigma^{-1}\underline{y} \sim \chi^2_{(p)}$$

$\longrightarrow$ but $\underline{y} = \sqrt{n}(\bar{\underline{x}} - \underline{\mu})$ then

$$\sqrt{n}(\bar{\underline{x}} - \underline{\mu})'\Sigma^{-1}\sqrt{n}(\bar{\underline{x}} - \underline{\mu}) \sim \chi^2_{(p)}$$

$$\Rightarrow \boxed{n(\bar{\underline{x}} - \underline{\mu})'\Sigma^{-1}(\bar{\underline{x}} - \underline{\mu}) \sim \chi^2_{(p)}}$$

Know to test

$H_0: \underline{\mu} = \underline{\mu}_0$ with size α

$n(\bar{x} - \underline{\mu}_0)'\Sigma^{-1}(\bar{x} - \underline{\mu}_0) \sim \chi^2_{(p)}$, α under H_0

and we reject H_0 with size α iff

$$n(\bar{\underline{x}} - \underline{\mu}_0)'\Sigma^{-1}(\bar{\underline{x}} - \underline{\mu}_0) \geq \chi^2_{(p)}(\alpha)$$ 1-sample

for one sample Σ known.

الاختبار.

يستخدم هذا الاختبار في حالة 1-sample وتكون Σ معلومة known عند اختبار المتوسطات mean.

Suppose we have a sample of size $\boxed{n_1}$ from $\boxed{N_p(\underline{\mu}_1, \Sigma)}$ and another sample $\boxed{n_2}$ from $\boxed{N_p(\underline{\mu}_2, \Sigma)}$, where

The common covariance matrix is known.

How to test H_0:

$H_0: \mu_1 = \mu_2$ (2-sample)

الان في حالة 2-sample

we have

$$\bar{X}_1 \sim N_p\left(\underline{\mu}_1, \frac{\Sigma}{n_1}\right)$$

$$\bar{X}_2 \sim N_p\left(\underline{\mu}_2, \frac{\Sigma}{n_2}\right)$$

بما ان $\therefore (\bar{X}_1 - \bar{X}_2) \sim N_p\left[(\underline{\mu}_1 - \underline{\mu}_2), \frac{n_1 n_2}{n_1+n_2}\Sigma\right]$

Therefore $\sqrt{\frac{n_1 n_2}{n_1+n_2}}\left[(\bar{X}_1 - \bar{X}_2) - (\underline{\mu}_1 - \underline{\mu}_2)\right] \sim N_p(0, \Sigma)$

then $\boxed{\frac{n_1 n_2}{n_1+n_2}\left[(\bar{\underline{X}}_1 - \bar{\underline{X}}_2) - (\underline{\mu}_1 - \underline{\mu}_2)\right]' \Sigma^{-1} \left[(\bar{\underline{X}}_1 - \bar{\underline{X}}_2) - (\underline{\mu}_1 - \underline{\mu}_2)\right] \sim \chi^2_{(p)}}$

we want to test.

$H_0: \mu_1 = \mu_2$

$\therefore H_1: \mu_1 \neq \mu_2$

Know $H_0: \mu_1 = \mu_2 \Rightarrow \mu_1 - \mu_2 = 0$

we can base our test on.

نعوض في العلاقة بـ صفر لان.

$$\frac{n_1 n_2}{n_1+n_2}(\bar{\underline{X}}_1 - \bar{\underline{X}}_2)' \Sigma^{-1} (\bar{\underline{X}}_1 - \bar{\underline{X}}_2) \sim \chi^2_{(p)} \quad \text{under } H_0$$

e reject Ho with size α iff:

$$\boxed{\frac{n_1 n_2}{n_1+n_2} (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' \Sigma^{-1} (\bar{\underline{x}}_1 - \bar{\underline{x}}_2) \geq \chi^2_{(p)\alpha}}$$

under Ho for (2-Sample)
when (Σ known)

اذا Σ موجودة في السؤال
فهذا يعني بانها معلومة known
اما اذا اعطانا S فهذا يعني
بان ال Σ مجهولة غير معرفة
unknown.
ويجب تقديرها.

i.e $\frac{n_1 n_2}{n_1+n_2} D^2 \geq \chi^2_{(p)\alpha}$

Remark

$D^2 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\mu_1 - \mu_2)$
is called Mahalanobis Distance $= D^2$
The distance between two multivariate
population.

ex: let $x \sim N_2(\underline{\mu}, \Sigma)$

where $\Sigma = \begin{pmatrix} 7.92 & 5.68 \\ & 6.992 \end{pmatrix}$

Test the hypothesis

$H_0: \mu = \begin{pmatrix} 39 \\ 60.43 \end{pmatrix}$,

since $\bar{x} = \begin{pmatrix} 36.4 \\ 62.6 \end{pmatrix}$, $n = 5$.

sol

we reject $H_0: \mu = \begin{pmatrix} 39 \\ 60.43 \end{pmatrix}$ if.

$$n(\underline{\bar{x}} - \underline{\mu_0})' \Sigma^{-1} (\underline{\bar{x}} - \underline{\mu_0}) \geq \chi^2_{(p)^\alpha}$$

$$(\underline{\bar{x}} - \underline{\mu_0}) = \begin{pmatrix} 36.4 \\ 62.6 \end{pmatrix} - \begin{pmatrix} 39 \\ 60.43 \end{pmatrix} = \begin{pmatrix} -2.6 \\ 2.17 \end{pmatrix}$$

$$\Rightarrow (5) \begin{pmatrix} -2.6 & 2.17 \end{pmatrix} \begin{pmatrix} 7.92 & 5.68 \\ & 6.992 \end{pmatrix}^{-1} \begin{pmatrix} -2.6 \\ 2.17 \end{pmatrix} = 40.96$$

$\rightarrow$ from table we have $\chi^2_{(p)^\alpha} = \chi^2_{(p,\alpha)} = \chi^2_{(2,0.05)} = 5.9$

$$\therefore \chi^2_{cal} = 40.96 > \chi^2_{table} = 5.99$$

So we reject H_0

وهذا يعني i.e $\underline{\mu} \neq \begin{pmatrix} 39 \\ 60.43 \end{pmatrix}$.

ex) let $\underline{X}_1 \sim N_2(\underline{\mu}_1, \Sigma)$ and $\underline{X}_2 \sim N_2(\underline{\mu}_2, \Sigma)$

The Covariance matrix Σ is given by:

$$\Sigma = \begin{pmatrix} 7.92 & 5.68 \\ & 6.292 \end{pmatrix}, \quad (\Sigma \text{ known})$$

and $n_1 = 5$, $n_2 = 7$, $\bar{X}_1 = \begin{pmatrix} 36.4 \\ 62.6 \end{pmatrix}$ and $\bar{X}_2 = \begin{pmatrix} 39 \\ 60.43 \end{pmatrix}$

Test the hypothesis

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

<u>sol</u> By using the statistics

$$\frac{n_1 n_2}{n_1 + n_2} (\bar{\underline{X}}_1 - \bar{\underline{X}}_2)' \Sigma^{-1} (\bar{\underline{X}}_1 - \bar{\underline{X}}_2) \geq \chi^2_{(p),\alpha}$$

$$(\bar{\underline{X}}_1 - \bar{\underline{X}}_2) = \begin{pmatrix} -2.6 \\ 2.17 \end{pmatrix}$$

$$\Rightarrow \frac{5(7)}{(5+7)} (-2.6 \quad 2.17) \begin{pmatrix} 7.92 & 5.68 \\ & 6.292 \end{pmatrix}^{-1} \begin{pmatrix} -2.6 \\ 2.17 \end{pmatrix} = 23.891$$

$$\chi^2_{(-1)} = 23.891 \geq \chi^2_{table} = 5.99 \; (2, 0.01)$$

so we reject H_0

$\Rightarrow \underline{\mu}_1 \neq \underline{\mu}_2$

Case B

Test the hypothesis $\underline{\mu} = \underline{\mu}_0$ for the mean vector μ and Σ is unknown

in order to calculate the maximum of likelihood under $H_0: \underline{\mu} = \underline{\mu}_0$ we have to ~~find~~ maximum (w.r.to Σ) the expression:

$$l_0(\underline{x}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)'\Sigma^{-1}(\underline{x}_j - \underline{\mu}_0)$$

$$= -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\left[\operatorname{tr}\Sigma^{-1}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)(\underline{x}_j - \underline{\mu}_0)'\right]$$

$$\frac{\partial l_0(\underline{x}, \Sigma)}{\partial(\Sigma^{-1})} = \frac{n}{2}\Sigma - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)(\underline{x}_j - \underline{\mu}_0)'$$

setting $\frac{\partial l_0(\underline{x}, \Sigma)}{\partial(\Sigma^{-1})} = 0$

we have,

$$\hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu}_0)(\underline{x}_j - \underline{\mu}_0)'$$

$$= \frac{1}{n}\sum_{j=1}^{n}\{(\underline{x}_j - \bar{\underline{x}} + \bar{\underline{x}} - \underline{\mu}_0)(\underline{x}_j - \bar{\underline{x}} + \bar{\underline{x}} - \underline{\mu}_0)'$$

$$= \frac{1}{n}\sum_{j=1}^{n}\{[(\underline{x}_j - \bar{\underline{x}}) + (\bar{\underline{x}} - \underline{\mu}_0)][(\underline{x}_j - \bar{\underline{x}}) + (\bar{\underline{x}} - \underline{\mu}_0)]'$$

$$= \frac{1}{n}\sum_{j=1}^{n}\{(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})' + (\bar{\underline{x}} - \underline{\mu}_0)(\bar{\underline{x}} - \underline{\mu}_0)' + 2(\bar{\underline{x}} - \underline{\mu}_0)(\underline{x}_j - \bar{\underline{x}})'\}$$

$$= \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})' + \frac{1}{n}\sum_{j=1}^{n}(\bar{\underline{x}} - \underline{\mu}_0)(\bar{\underline{x}} - \underline{\mu}_0)'$$

$$= \frac{1}{n}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})' + (\bar{\underline{x}} - \underline{\mu}_0)(\bar{\underline{x}} - \underline{\mu}_0)'$$

$$\therefore \hat{\Sigma} = S + \underline{d}\,\underline{d}'$$

where $\underline{d} = (\bar{\underline{x}} - \underline{\mu}_0)$.

It is clear that the maximum likelihood under H_0 is $l_0^* = \max l_0(\underline{x}; \Sigma) = l_0(\underline{x}; \underline{\mu}_0, S + \underline{d}\,\underline{d}')$.

In order to calculate the maximum likelihood under $H_1: \underline{\mu} \neq \underline{\mu}_0$, we have to maximize (w.r.t $(\underline{\mu}, \Sigma)$).

$$f_1(\underline{x}; \underline{\mu}, \Sigma) = -\frac{np}{2}\log(2\pi) - \frac{n}{2}\log|\Sigma| - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \underline{\mu})'\Sigma^{-1}(\underline{x}_j - \underline{\mu})$$

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

$$\frac{\partial f_1(\underline{x}; \Sigma)}{\partial \underline{\mu}} = +\frac{1}{2}\Sigma^{-1}\sum_{j=1}^{n}(+2)(\underline{x}_j - \underline{\mu})$$

we see that the MLE $\hat{\underline{\mu}} = \bar{\underline{x}}$.

Let us now maximize the function $f_1(\underline{x}; \underline{\mu}, \Sigma)$ interms of the parameter Σ.

$$\frac{\partial f_1(\underline{x}; \Sigma)}{\partial(\Sigma^{-1})} = \frac{n}{2}\Sigma - \frac{1}{2}\sum_{j=1}^{n}(\underline{x}_j - \bar{\underline{x}})(\underline{x}_j - \bar{\underline{x}})'$$

$$\Longrightarrow \frac{\partial f_1(\underline{x}; \Sigma)}{\partial(\Sigma^{-1})} = 0$$

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

$$f_1^* = f_1(\underline{x}; \bar{\underline{x}}, S)$$

$$-2\log\lambda = 2\left[\ell_1(\underline{x}; \bar{\underline{x}}, S) - \ell_0(\underline{x}; \underline{\mu}, S + \underline{d}\,\underline{d}')\right]$$

$$= n\log|S| + n\log|S + \underline{d}\,\underline{d}'| - tr\left[S^{-1}\sum_{j=1}^{n}(x_j - \bar{x})(x_j - \bar{x})'\right]$$

$$+ tr\left[(S + \underline{d}\,\underline{d}')^{-1}\sum_{j=1}^{n}(x_j - \underline{\mu}_0)(x_j - \underline{\mu}_0)'\right]$$

$$= n\log\left[|S|(1 + \underline{d}'S^{-1}\underline{d})\right] - n\log|S|$$

$$\therefore -2\log\lambda = n\log(1 + \underline{d}'S^{-1}\underline{d})$$

This statistic depend upon $n\underline{d}'S^{-1}\underline{d}$.
This lead to test statistic with Hotelling T^2-distribution, we have

$$n\underline{d}'S^{-1}\underline{d} \sim T^2(p, n-1)$$

we reject H_0 if $T^2 \geq T^2(\alpha, p, n-1)$ and accept otherwise.
If $H_0: \underline{\mu} = \underline{\mu}_0$ is true, then:

$$\frac{(n-p)}{p(n-1)}T^2 \sim F(\alpha, p, n-p)$$

The rejection region of the test is given as

$$\frac{(n-p)}{p(n-1)}T^2 \geq F(\alpha, p, n-p)$$

Ex If $\underline{X} \sim N_2(\underline{\mu}, \Sigma)$ where Σ is unknown and the sample mean vector is

$\bar{\underline{x}} = \begin{pmatrix} 55.24 \\ 39.97 \end{pmatrix}$ and the sample covariance matrix with $S = \begin{pmatrix} 210.54 & 126.99 \\ & 119.68 \end{pmatrix}$

and $n = 101$.

Test the hypothesis $H_0: \underline{\mu} = \begin{pmatrix} 60 \\ 50 \end{pmatrix}$ vs $H_1: \underline{\mu} \neq \begin{pmatrix} 60 \\ 50 \end{pmatrix}$

Sol

$$T^2 = n(\bar{\underline{x}} - \underline{\mu}_0)' S^{-1} (\bar{\underline{x}} - \underline{\mu}_0)$$

$$= (101) \begin{pmatrix} 55.24 - 60 \\ 39.97 - 50 \end{pmatrix}' \begin{pmatrix} 210.54 & 126.99 \\ & 119.68 \end{pmatrix}^{-1} \begin{pmatrix} 55.24 - 60 \\ 39.97 - 50 \end{pmatrix}$$

$$= 357.43$$

$$F_{cal} = \frac{(n-p)}{p(n-1)} T^2 = \frac{(101-2)}{2(101-1)} (357.43)$$

$$= 176.92$$

$F(\alpha, p, n-p) = F(0.01, 2, 99) = 4.83$

Since $F_{cal} > F_{(0.01, 2, 99)} = 4.83$.

Therefore we reject $H_0: \underline{\mu} = \begin{pmatrix} 60 \\ 50 \end{pmatrix}$ and accept $H_1: \underline{\mu} \neq \begin{pmatrix} 60 \\ 50 \end{pmatrix}$

248

(b) If H_0 is rejected, test each variable separately.

$H_0: \mu_{01} = 60$ vs. $H_1: \mu_{01} \neq 60$

$H_0: \mu_{02} = 50$ vs. $H_1: \mu_{02} \neq 50$

$$t_1 = \frac{\bar{x}_1 - \mu_{01}}{s_1/\sqrt{n}} = \frac{55.24 - 60}{\sqrt{\frac{210.54}{101}}} = -3.299$$

$$t_2 = \frac{\bar{x}_2 - \mu_{02}}{s_2/\sqrt{n}} = \frac{39.97 - 50}{\sqrt{\frac{119.68}{101}}} = -9.210$$

using $t(\frac{\alpha}{2}, n-1) = t(0.005, 100) = 2.617$

we have $|t_1| > 2.617$ and $|t_2| > 2.617$

Thus both test reject the hypothesis because the T^2-test reject H_0.

H.w: Suppose a random sample from $N_2(\underline{\mu}, \Sigma)$. using the data

$$\underline{X}' = \begin{pmatrix} 3 & 6 & 5 & 10 \\ 10 & 12 & 14 & 9 \end{pmatrix}$$ to test

$H_0: \underline{\mu}' = (6 \quad 11)$ at $\alpha = 0.05$

and determine the (S.C.I) for μ_1 and μ_2

ii) Confidence region for $\underline{\mu}$

In univariate case, a confidence interval (C.I) for μ is the set of all points m satisfying

$$\frac{n(\bar{\underline{x}}-\mu_0)^2}{s^2} \leq t^2\left(1-\frac{\alpha}{2}, n-1\right)$$

In multivariate case, a confidence region (C.R) for $\underline{\mu}$ is the set of all vector $\underline{\mu}$ satisfying

$$n(\bar{\underline{x}}-\underline{\mu}_0)' \bar{S}^{-1}(\bar{\underline{x}}-\underline{\mu}_0) \leq T^2_{(\alpha,p,n-p)}$$

Simultaneous confidence interval (S.C.I) for the component $\underline{\mu}$ at level $(1-\alpha)=95\%$ calculate by the interval

$$\bar{x}_j \pm \sqrt{\frac{p(n-1)}{n(n-p)} F_{(1-\alpha,p,n-p)}\, S_{jj}}$$

For the last example determine S.C.I for μ_1 and μ_2.

$$50.73 < \mu_1 < 59.75$$

$$51.58 < \mu_2 < 58.35$$

250

case B. Test of the hypothesis for the mean vector μ and Σ is unknown

GLR. $\frac{L(\hat{\omega})}{L(\hat{\Omega})}$

(i) in the Univariate case

$X \sim N_p(\mu, \sigma^2)$, where $\sigma^2 > 0$, is unknown

To test $H_0: \underset{\sim}{\mu} = \underset{\sim}{\mu_0}$

$H_1: \underset{\sim}{\mu} \neq \underset{\sim}{\mu_0}$

if H_0 is true then

$$\hat{\sigma}^2 = \frac{\Sigma(x_i - \mu_0)^2}{n}$$

where $\hat{\sigma}^2$ is an unbiased estimator

But if H_1 is true then

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

$$\hat{\sigma}^2 = \frac{\Sigma(x_i - \bar{x})^2}{n}, \text{ where}$$

where $\hat{\sigma}^2$ is an biased estimator

(1) عند الفرضية H_0

$$\omega = \{\mu = \mu_0 \ , \ \sigma^2 > 0\} \quad \text{for } H_0$$

$$\because \hat{\omega} = \left\{\mu = \mu_0 \ , \ \hat{\sigma}^2 = \frac{\sum (x_i - \mu_0)^2}{n}\right\} \text{ under } H_0 .$$

$$L(\hat{\omega}) = (2\pi)^{-\frac{n}{2}} \left(\hat{\sigma}^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2}\frac{\sum (x_i - \mu_0)^2}{\hat{\sigma}^2}}$$

نعوض بقيمة $\hat{\sigma}^2$ عند فرضية H_0 فيبقى

$$= (2\pi)^{-\frac{n}{2}} \left(\frac{\sum (x_i - \mu_0)^2}{n}\right)^{-\frac{n}{2}} e^{-\frac{1}{2}\frac{\sum (x_i - \mu_0)^2}{\frac{\sum (x_i - \mu_0)^2}{n}}}$$

$$= (2\pi)^{-\frac{n}{2}} \left(\frac{1}{n} \sum (x_i - \mu_0)^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}}$$

$$\boxed{L(\hat{\omega}) = \left(\frac{2\pi}{n} \sum (x_i - \mu_0)^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}} \quad \text{----} (1)}$$

(2) عند الفرضية H_1 البديلة .

now

$$\Omega = \{\mu \neq \mu_0 < \infty \ , \ \sigma^2 > 0\} . \quad \text{for } H_1$$

$$\hat{\Omega} = \left\{\hat{\mu} = \bar{x} \ , \ \hat{\sigma}^2 = \frac{\sum (x_i - \bar{x})^2}{n}\right\} \text{ under } H_1$$

$$L(\hat{\Omega}) = (2\pi)^{-\frac{n}{2}} \left(\hat{\sigma}^2\right)^{-\frac{n}{2}} e^{-\frac{1}{2}\frac{\sum (x_i - \bar{x})^2}{\hat{\sigma}^2}}$$

$$= (2\pi)^{-\frac{n}{2}} \left(\frac{\sum (x_i - \bar{x})^2}{n}\right)^{-\frac{n}{2}} e^{-\frac{1}{2}\frac{\sum (x_i - \bar{x})^2}{\frac{\sum (x_i - \bar{x})^2}{n}}}$$

$$\boxed{L(\hat{\Omega}) = \left(\frac{2\pi}{n} \sum (x_i - \bar{x})^2\right)^{-\frac{n}{2}} e^{-\frac{n}{2}} \quad \text{-----} (2)}$$

By using generalized likelihood Ratio test. (GLR)

We have

$$\lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

عند الفرضية H_0

عند الفرضية H_1

$$\Rightarrow \lambda = \frac{\left[\frac{2\pi}{n}\sum(x_i-\mu_0)^2\right]^{-\frac{n}{2}} e^{-\frac{n}{2}}}{\left[\frac{2\pi}{n}\sum(x_i-\bar{x})^2\right]^{-\frac{n}{2}} e^{-\frac{n}{2}}}$$

$$\lambda = \left[\frac{\sum(x_i-\mu_0)^2}{\sum(x_i-\bar{x})^2}\right]^{-\frac{n}{2}}$$

$$\lambda = \left[\frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\mu_0)^2}\right]^{\frac{n}{2}}$$

نرفع للقوة $\frac{-2}{n}$ ضخ

نضيف ونطرح $\bar{x}$ ضخ

$$\lambda^{2/n} = \frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\mu_0+\bar{x}-\bar{x})^2}$$

$$\lambda^{2/n} = \frac{\sum(x_i-\bar{x})^2}{\sum\left((x_i-\bar{x})+(\bar{x}-\mu_0)\right)^2}$$

$$= \frac{\sum(x_i-\bar{x})^2}{\sum\left[(x_i-\bar{x})^2+(\bar{x}-\mu_0)^2\right]}$$

$$\lambda^{2/n} = \frac{\sum(x_i-\bar{x})^2}{\sum(x_i-\bar{x})^2+n(\bar{x}-\mu_0)^2}$$

$$\lambda^{2/n} = \frac{1}{1+\frac{n(\bar{x}-\mu_0)}{\sum(x_i-\bar{x})^2}}$$

$$\lambda^{\frac{2}{n}} = \frac{1}{1 + \frac{n(\bar{x} - \mu_0)^2}{\Sigma(x_i - \bar{x})^2}}$$

But we have $s = \frac{\Sigma(x_i - \bar{x})^2}{(n-1)}$

then $\Sigma(x_i - \bar{x})^2 = (n-1)s$

$\Rightarrow$

$$\lambda^{\frac{2}{n}} = \frac{1}{1 + \frac{n(\bar{x} - \mu_0)^2}{(n-1)s^2}}$$

but we know that $t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{\sqrt{n}(\bar{x} - \mu)}{s}$

then.

$$\lambda^{\frac{2}{n}} = \frac{1}{1 + \frac{n(\bar{x} - \mu_0)^2}{(n-1)s^2}}$$

$$= \frac{1}{1 + \frac{\sqrt{n}(\bar{x} - \mu_0)^2}{(n-1)s^2}}$$

$$= \frac{1}{1 + \frac{1}{n-1}\left[\frac{\sqrt{n}(\bar{x} - \mu_0)}{s}\right]^2}$$

$$= \frac{1}{1 + \frac{1}{n-1}(t)^2}$$

$$\lambda^{2/n} = \frac{1}{1 + \frac{t^2}{n-1}}$$

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

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

$$1+\frac{t^2}{n-1}=\frac{1}{\lambda^{2/n}}$$

$$1+\frac{t^2}{n-1}=\lambda^{-2/n}$$

$$\frac{t^2}{n-1}=\lambda^{-2/n}-1$$

$$\therefore \boxed{t^2=(n-1)\left[\lambda^{-2/n}-1\right]}$$

$$\boxed{\therefore |t|<K_0}$$

The critical region to test $H_0: \mu=\mu_0$ is

$$\boxed{C=\{x_1,\dots,x_n;\ |t|\leq K_0\}}$$