آزمون فرض‌های چند متغیره

آزمون میانگین در حالت یک متغیره : X یک متغیر تصادفی

$$\begin{cases} H_0 : \mu = \mu_0 \\ H_1 : \mu \neq \mu_0 \end{cases}$$

برای این کار یک نمونه تصادفی n تایی می‌داریم آماره آزمون :

$$t = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}} \qquad \begin{cases} S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \\ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \end{cases}$$

$$|t| > t_{n-1, \alpha/2} \Rightarrow RH_0$$

رد H_0 وقتی $|t|$ بزرگ است معادل رد H_0 وقتی مربع آن یعنی

$$t^2 = \frac{(\bar{x} - \mu_0)^2}{S^2/n} = n(\bar{X} - \mu_0)(S^2)^{-1}(\bar{X} - \mu_0) \quad (1)$$

بزرگ است یا باشد. یعنی باشد هر دو $\bar{x}$ و S^2 :

$$\text{if} \quad n(\bar{x} - \mu_0)(S^2)^{-1}(\bar{x} - \mu_0) > t^2_{n-1, \alpha/2} \Rightarrow RH_0 \quad (2)$$

و اگر H_0 رد نشود یعنی μ_0 یک مقدار معقول برای میانگین جامعه است. برای بدست آوردن فاصله اطمینان :

$$\left| \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \right| \leq t_{n-1, \alpha/2} \Rightarrow AH_0$$

معادل است با اینکه $\left.\begin{matrix} \mu_0 \text{ در فاصله } 100(1-\alpha)\% \text{ ، } \bar{x} \pm t_{n-1, \alpha/2} \frac{S}{\sqrt{n}} \end{matrix}\right\}$ قرار می‌گیرد)

n کوچک σ^2 نامعلوم

$$\bar{x} - t_{n-1, \alpha/2} \frac{S}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + t_{n-1, \alpha/2} S/\sqrt{n}$$

Subject:

Year. Month. Date. ( )

حال حالت تعمیم این آزمون به حالت p متغیره را فرض کنید هدف آزمون

$$H_0 : \underset{\sim}{\mu}_{p\times 1} = \underset{\sim}{\mu}_{0\, p\times 1} = \begin{pmatrix} \mu_{10} \\ \vdots \\ \mu_{p0} \end{pmatrix}$$

است. یعنی آیا یک بردار $p\times 1$ ، $\underset{\sim}{\mu}_0$ بردار موجه برای میانگین یک توزیع نرمال p متغیره می باشد.

مشابه به حالت یک متغیره عمل می کنیم. یک تعمیم از مربع فاصله ①

$$T^2 = (\underset{\sim}{\bar X} - \underset{\sim}{\mu}_0)' \left(\frac{S}{n}\right)^{-1} (\underset{\sim}{\bar X} - \underset{\sim}{\mu}_0)$$

$$= n(\underset{\sim}{\bar X} - \underset{\sim}{\mu}_0)' S^{-1} (\underset{\sim}{\bar X} - \underset{\sim}{\mu}_0) \sim T^2_{p,n-1}$$

است. که همان آماره T^2 هتلینگ می باشد.

$$\bar X_{p\times 1} = \frac{1}{n}\sum_{j=1}^{n} \underset{\sim}{X}_j \qquad S_{p\times p} = \frac{1}{n-1}\sum_{j=1}^{n} (\underset{\sim}{X}_j - \underset{\sim}{\bar X})(\underset{\sim}{X}_j - \underset{\sim}{\bar X})'$$

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ & & \\ x_{p1} & & x_{pn} \end{bmatrix}_{p\times n} = \frac{1}{n}(\underset{\sim}{X}_1 + \cdots + \underset{\sim}{X}_n) = \begin{pmatrix} \bar x_1 \\ \vdots \\ \bar x_p \end{pmatrix} \longleftarrow \underset{\sim}{\mu}_0 = \begin{pmatrix} \mu_{10} \\ \vdots \\ \mu_{p0} \end{pmatrix}$$

$$\underset{\sim}{x}_1 \quad \cdots \quad \underset{\sim}{x}_n$$

* اگر فاصله تعمیم یافته ی مشاهده شده T^2 خیلی بزرگ شد یعنی خیلی دور از $\underset{\sim}{\mu}$ است که فرض $H_0 : \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ رد می شود.

* چون نشان دادیم توزیع T^2 هتلینگ با فیشر رابطه دارد پس نیازی به جدول نداریم

$$T^2(p,n) = \frac{np}{n-p+1} F_{p,n-p+1} \Longrightarrow$$

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

$p-1 , n$

پس بطور خلاصه :

$\frac{n}{p-1}$ $\cancel{-1}$ $\quad \frac{n}{(n-1)(p-1)} - \frac{p-1}{n-1}$

فرض کنید $\underset{\sim}{X}_1 \; \ldots\ldots \; \underset{\sim}{X}_n$ یک نمونه تصادفی از جامعه $N_p(\underset{\sim}{\mu}, \Sigma)$ است و

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

$$\alpha = P\left(T^2 > \frac{(n-1)p}{(n-p)} F_{p,n-p,\alpha}\right)$$

$$= P\left(n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})' S^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}) > \frac{(n-1)p}{(n-p)} F_{p,n-p,\alpha}\right)$$

که در سطح معنی داری α ، $H_0: \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ در مقابل $H_1: \underset{\sim}{\mu} \neq \underset{\sim}{\mu}_0$

رد می‌شود هرگاه

$$T^2 = n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0)' S^{-1} (\bar{\underset{\sim}{X}} - \underset{\sim}{\mu}_0) > \frac{(n-1)p}{n-p} F_{p,n-p,\alpha} \Rightarrow RH_0$$

مثال: فرض کنید ماتریس داده‌ای مربوط به نمونه تصادفی به حجم $n=3$ از یک جامعه نرمال دو متغیره به صورت زیر باشد.

$$X = \begin{array}{c} \\ 1 \\ 2 \end{array} \begin{array}{c} \text{اول} \quad \text{دوم} \quad \text{سوم} \\ \begin{bmatrix} 6 & 10 & 8 \\ 9 & 6 & 3 \end{bmatrix} \end{array}$$

آزمون مقابل را انجام دهید.

$$\begin{cases} H_0: \underset{\sim}{\mu} = \begin{bmatrix} 9 \\ 5 \end{bmatrix} \\ H_1: \underset{\sim}{\mu} \neq \begin{bmatrix} 9 \\ 5 \end{bmatrix} \end{cases}$$

$$\bar{X} = \begin{bmatrix} \frac{6+10+8}{3} \\ \frac{9+6+3}{3} \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$$

$$S_{11} = \frac{(6-8)^2 + (10-8)^2 + (8-8)^2}{2} = 4 ,$$

$$S_{22} = \frac{(9-6)^2 + (6-6)^2 + (3-6)^2}{2} = 9$$

$$S_{12} = \frac{(6-8)(9-6) + (10-8)(6-6) + (8-8)(3-6)}{2} = -3$$

Subject:
Year. Month. Date. ( )

$$S = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix} \qquad S^{-1} = \frac{1}{36-9}\begin{bmatrix} 9 & 3 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{bmatrix}$$

$$T^2 = 3\,[8-9\,,\,6-5]\begin{bmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{bmatrix}\begin{bmatrix} 8-9 \\ 6-5 \end{bmatrix} = 7/9$$

$$\frac{(n-1)p}{n-p} = \frac{2\times 2}{3-2} = 4$$

$$4F_{2,\underbrace{3-2}_{1}} \qquad 7/9 < 4F_{2,3-2}$$

$$\Rightarrow A H_0$$

نکته: T^2 تحت تبدیلات خطی پایا ست یعنی اگر $\underset{p\times 1}{Y} = \underset{p\times p}{C}\,\underset{p\times 1}{X} + \underset{p\times 1}{d}$ باشد

وارون پذیر

آنگاه: $T^2_Y = T^2_X$ ؛ اثبات C:

$$\bar{Y} = C\bar{X} + d$$

$$S_Y = C S_X C'$$

$$\mu_{0Y} = C\mu_0 + d$$

$$T^2_Y = n(\bar{X}-\mu_0)' S^{-1}(\bar{X}-\mu_0) = T^2_X$$

$$T^2_Y = n(\bar{Y}-\mu_{0Y})' S_Y^{-1}(\bar{Y}-\mu_{0Y})$$

$$n(C\bar{X} + \cancel{d} - C\mu_0 - \cancel{d})'(C S_X C')^{-1}(C\bar{X} + \cancel{d} - C\mu_0 - \cancel{d})$$

$$= n(\bar{X}-\mu_0)' C' C'^{-1} S_X^{-1} C^{-1} C (\bar{X}-\mu_0) = n(\bar{X}-\mu_0)' S^{-1}(\bar{X}-\mu_0) = T^2_X$$

آزمون نسبت درستنمایی و T^2 هتلینگ:

دیدیم:

$$\max_{\mu,\Sigma} L(\mu,\Sigma) = \frac{1}{(2\pi)^{np/2}|\hat{\Sigma}|^{n/2}} e^{-np/2}$$

$$\hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{X})(x_j - \bar{X})' \qquad \hat{\mu} = \bar{X} = \frac{1}{n}\sum_{j=1}^{n} x_j$$

قضیه: T^2 تحت تبدیلات خطی پایاست یعنی اگر $\underset{\sim p\times 1}{Y_j} = \underset{p\times p}{C}\,\underset{\sim p\times 1}{X_j} + \underset{\sim p\times 1}{d}$ که در آن C یک $j=1,\dots,n$ ماتریس معلوم و d یک بردار معلوم باشد آنگاه $T_Y^2 = T_X^2$

اثبات: می دانیم

$$\bar{Y} = C\bar{X} + d \quad , \quad S_Y = C S_X C'$$

$$T_Y^2 = n(\bar{Y} - \mu_Y)' S_Y^{-1} (\bar{Y} - \mu_Y) = n(C\bar{X} + d - C\mu_X - d)'$$

$$(C S_X C')^{-1} (C\bar{X} + d - C\mu_X - d)$$

$$= n(\bar{X} - \mu_X)' C' (C')^{-1} S_X^{-1} C^{-1} C (\bar{X} - \mu_X)$$

$$= n(\bar{X} - \mu_X)' S_X^{-1} (\bar{X} - \mu_X) = T_X^2$$

با توجه به این قضیه آماره T^2 نسبت به یک تبدیل خطی پایا و تغییرناپذیر است و بنابراین اگر واحد اندازه گیری X_j تغییر کند مقدار T^2 تغییر نمی کند.

آزمون در مورد میانگین μ بر اساس آزمون نسبت درستنمایی

فرض کنید $X_1, \dots, X_n$ یک نمونه تصادفی از توزیع $N_p(\mu, \Sigma)$ باشد می خواهیم آزمون فرض $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ انجام دهیم که در آن μ_0 یک بردار معلوم و Σ ماتریس معین مثبت نامعلوم است می خواهیم آزمون فوق را به روش (GLRT) آزمون نسبت درستنمایی انجام دهیم. برای این منظور مراحل زیر را انجام می دهیم

$$\Theta : \{(\mu, \Sigma) \mid -\infty < \mu_i < \infty \;,\; \Sigma \text{ is p.d}\}$$

$$\Theta_0 : \{(\mu_0, \Sigma) \mid \Sigma \text{ is p.d}\} \qquad \Lambda = \frac{\max_{\Theta_0} L(\mu, \Sigma)}{\max_{\Theta} L(\mu, \Sigma)}$$

$$\begin{cases} H_0: \theta \in \Theta_0 \\ H_1: \theta \notin \Theta_0 \end{cases} \quad \Lambda = \frac{\max_{\theta \in \Theta_0} L(\theta)}{\max_{\theta \in \Theta} L(\theta)} < C \Rightarrow RH_0$$

Subject:
Year. Month. Date. ( )

under H_1

$$\max_{\underset{\sim}{\mu},\Sigma} L(\underset{\sim}{\mu}, \Sigma) = \frac{1}{(2\pi)^{np/2}|\hat{\Sigma}|^{n/2}} e^{-\frac{np}{2}}$$ دیدیم

$$\hat{\Sigma} = \frac{1}{n}\sum_{j=1}^{n}(\underset{\sim}{x}_j-\bar{\underset{\sim}{x}})(\underset{\sim}{x}_j-\bar{\underset{\sim}{x}})'$$ که در آن

under H_0. $\hat{\underset{\sim}{\mu}} = \mu_0$

$$-\tfrac{1}{2}\, tr\left(\Sigma^{-1}\sum(x_j-\mu_0)(x_j-\mu_0)'\right)$$

$$L(\underset{\sim}{\mu}_0, \Sigma) = \frac{1}{(2\pi)^{np/2}|\Sigma|^{n/2}} \exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(\underset{\sim}{X}_j-\underset{\sim}{\mu}_0)'\Sigma^{-1}(\underset{\sim}{X}_j-\underset{\sim}{\mu}_0)\right\}$$

چون مقدار $\underset{\sim}{\mu}_0$ ثابت است برآورد درستنمایی Σ مشابه قبل به صورت

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

خواهد شد پس

$$\max_{\Sigma} L(\underset{\sim}{\mu}_0, \Sigma) = \frac{1}{(2\pi)^{np/2}|\hat{\Sigma}_0|^{n/2}} e^{-np/2}$$

در نتیجه

نسبت درستنمایی

$$\Lambda = \frac{\max_{\Theta_0} L(\underset{\sim}{\mu}, \Sigma)}{\max_{\Theta} L(\underset{\sim}{\mu}, \Sigma)} = \frac{L(\underset{\sim}{\mu}_0, \hat{\Sigma}_0)}{L(\hat{\underset{\sim}{\mu}}, \hat{\Sigma})}$$

$$= \frac{(2\pi)^{-np/2}|\hat{\Sigma}_0|^{-n/2}e^{-np/2}}{(2\pi)^{-np/2}|\hat{\Sigma}|^{-n/2}e^{-np/2}} = \frac{|\hat{\Sigma}_0|^{-n/2}}{|\hat{\Sigma}|^{-n/2}}$$

$$\Longrightarrow \Lambda^{2/n} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} = \frac{\left|\sum_{j=1}^{n}(\underset{\sim}{X}_j-\bar{\underset{\sim}{X}})(\underset{\sim}{X}_j-\bar{\underset{\sim}{X}})'\right|}{\left|\sum_{j=1}^{n}(\underset{\sim}{X}_j-\underset{\sim}{\mu}_0)(\underset{\sim}{X}_j-\underset{\sim}{\mu}_0)'\right|}$$

آماره $\Lambda^{2/n}$ را لامبدای ویلکس (wilk's lambda) گویند. حال طبق

آزمون نسبت درستنمایی فرض $H_0: \underset{\sim}{\mu} = \underset{\sim}{\mu}_0$ را رد می‌کنیم اگر

P4PCO

$\Lambda^{2/n} \leq C_\alpha^{2/n}$ یا $\Lambda \leq C_\alpha \Rightarrow RH.$

Subject: (۲۴)۱
Year. Month. Date

```
X1=c(6,10,8)
X2=c(9,6,3)
data=cbind(X1,X2)
n=3
p=2
mu0=c(9,5)
alpha=0.05
xbar=apply(data,2,mean)
S=var(data)
xbar=as.matrix(xbar)
mu0=as.matrix(mu0)
T2=n*t(xbar-mu0)%*%solve(S)%*%(xbar-mu0)
Ta=p*(n-1)/(n-p)*qf(1-alpha,p,n-p)
```

که C_α صدک بالایی (100α)ام توزیع Λ است. $F_{p,n-p,\alpha}$

این ناحیه معادل این است که $T^2 > T^2_\alpha$ باشد. $T^2 = n(\bar{X}-\mu_0)' S^{-1} (\bar{X}-\mu_0)$

package mvnormtest
mshapiro.test(data)

تعریف توزیع لاندای ویلکس:

مستقل
$$\begin{cases} A \sim W_p(I, m) \\ B \sim W_p(I, n) \end{cases} \Rightarrow \Lambda = \frac{|A|}{|A+B|} \sim \Lambda(p, m, n) \qquad m \geq p$$

نشان دهید: if $p=1 \longrightarrow \Lambda(1, m, n) = Beta(m/2, n/2)$

$> A \sim \chi^2_m , \; B \sim \chi^2_n \longrightarrow \Lambda = \frac{A}{A+B} \sim Beta(m/2, n/2)$

قضیه: فرض کنید $X_1 \ldots X_n$ یک نمونه تصادفی از $N_p(\mu, \Sigma)$ باشد در این صورت برای آزمون

$$\begin{cases} H_0: \mu = \mu_0 \\ H_1: \mu \neq \mu_0 \end{cases}$$

داریم

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

$$\longrightarrow T^2 = \frac{(n-1)|\hat{\Sigma}_0|}{|\hat{\Sigma}|} - (n-1)$$

~~$\hat{\Sigma}_0 = \sum (X_j - \mu_0)(X_j - \mu_0)'$~~

اثبات: می دانیم $\Lambda^{2/n} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|}$ از طرفی داریم

$$n\hat{\Sigma}_0 = \sum_{j=1}^{n} (X_j + \bar{X} - \bar{X} - \mu_0)(X_j + \bar{X} - \bar{X} - \mu_0)' = \sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})'$$

با اضافه و کم کردن $\bar{X}$ داریم

$$+ n(\bar{X} - \mu_0)(\bar{X} - \mu_0)'$$

یادآوری

$$\hat{\Sigma} = \cancel{\frac{1}{n}} \sum_{j=1}^{n} (X_j - \bar{X})(X_j - \bar{X})' = A$$

$$\longrightarrow A = (n-1)S$$

Subject:
Year. Month. Date. ( )

$$\Lambda^{2/N} = \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_{\circ}|} = \frac{|A|}{|A + n(\bar{X}-\mu_{\circ})(\bar{X}-\mu_{\circ})'|} \quad (1)$$

ماتریس B را به صورت زیر تعریف می‌کنیم

$$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} 1 & \sqrt{n}(\bar{X}-\mu_{\circ})' \\ -\sqrt{n}(\bar{X}-\mu_{\circ}) & \sum_{j=1}^{n}(X_j-\bar{X})(X_j-\bar{X})' \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \sqrt{n}(\bar{X}-\mu_{\circ})' \\ -\sqrt{n}(\bar{X}-\mu_{\circ}) & A \end{bmatrix}$$

با توجه به اینکه

$$|B| = |B_{11}|\,|B_{22} - B_{21}B_{11}^{-1}B_{12}|$$
$$= |B_{22}|\,|B_{11} - B_{12}B_{22}^{-1}B_{21}|$$

پس

$$\underbrace{\begin{vmatrix} 1 & \sqrt{n}(\bar{X}-\mu_{\circ})' \\ -\sqrt{n}(\bar{X}-\mu_{\circ}) & A \end{vmatrix} = \Big|\underbrace{\sum_{j=1}^{n}(X_j-\bar{X})(X_j-\bar{X})'}_{A} + n(\bar{X}-\mu_{\circ})(\bar{X}-\mu_{\circ})'\Big|}^{|B_{11}|\,|B_{22}-B_{21}B_{11}^{-1}B_{12}|}$$

$$= \underbrace{|A|\{1 + n(\bar{X}-\mu_{\circ})'A^{-1}(\bar{X}-\mu_{\circ})\}}_{|B_{22}|\,|B_{11}-B_{12}B_{22}^{-1}B_{21}|} \quad (2)$$

از (1) و (2):

$$\Lambda^{2/N} = \frac{|A|}{|A|\{1 + n(\bar{X}-\mu_{\circ})'A^{-1}(\bar{X}-\mu_{\circ})\}}$$

$$= \frac{1}{1 + \frac{n}{n-1}(\bar{X}-\mu_{\circ})'S^{-1}(\bar{X}-\mu_{\circ})} = \frac{1}{1+\frac{1}{n-1}T^2}$$

$A = (n-1)S$
$A^{-1} = \frac{1}{n-1}S^{-1}$

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

پس با استفاده از این قضیه:

$$\Lambda^{2/n} \leq C_\alpha^{2/n} \iff \left(1+\frac{T^2}{n-1}\right)^{-1} \leq C_\alpha^{2/n} \iff 1+\frac{T^2}{n-1} \geq C_\alpha^{-2/n}$$

$$\iff T^2 \geq T_\alpha^2$$

PAPCO

$$\Lambda^{2/n} = \left(1+\frac{T^2}{n-1}\right)^{-1}$$ نکته ۱: می توان نشان داد که لذا

$$\Rightarrow T^2 = \frac{(n-1)|\Sigma_0|}{|\hat{\Sigma}|} - (n-1)$$

قضیه ۱

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

نکته ۲: اگر حجم نمونه بزرگ باشد

$$-2\ln\Lambda \sim \chi^2$$

$$\begin{pmatrix}\mu_1\\ \vdots \\ \mu_p\end{pmatrix} \quad \begin{pmatrix}\sigma_{11} & \cdots & \\ & \ddots & \\ & & \sigma_{pp}\end{pmatrix}_p \qquad p+\frac{p(p+1)}{2} \leftarrow v - v_0 \rightarrow \frac{p(p+1)}{2} \; (?)$$

$p\times p \longrightarrow$ کل اعضا $\qquad p^2-p \longrightarrow$ اعضا بدون قطر $\qquad p(p-1)/2$

تعداد پارامترهای قطر اصلی بدون در نظر گرفتن تکراری ها:

$$\sum_{i=0}^{p-1} i = (p-1)p/2$$

ناحیه اطمینان برای بردار $\underset{\sim}{\mu}$ (Σ نامعلوم) فرض کنید $X_1, \ldots, X_n$ نمونه از جامعه نرمال چند متغیره باشد

$$P\Big(\underbrace{n(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu})'S^{-1}(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu})}_{T^2} > \underbrace{\frac{(n-1)p}{n-p}F_{p,n-p,\alpha}}_{T^2_{(p,n-1,\alpha)}}\Big) = \alpha$$ دیدیم

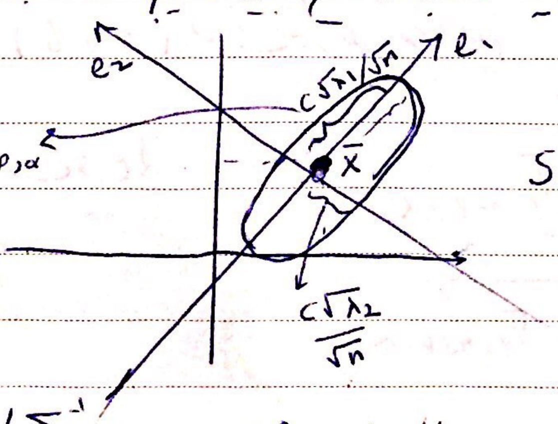
$$\Longrightarrow P\left(n(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu})'S^{-1}(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu}) \leq \frac{(n-1)p}{n-p}F_{p,n-p,\alpha}\right) = 1-\alpha$$

یک ناحیه اطمینان $100(1-\alpha)\%$ برای $\underset{\sim}{\mu}$ از جامعه نرمال p متغیره به صورت

$$n(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu})'S^{-1}(\bar{\underset{\sim}{X}}-\underset{\sim}{\mu}) \leq \underbrace{\frac{(n-1)p}{n-p}F_{p,n-p,\alpha}}_{c^2}$$

است که یک بیضی با مرکز $\bar{\underset{\sim}{X}}$ است و این بیضی ناحیه $100(1-\alpha)\%$ برای $\underset{\sim}{\mu}$ است

$$c^2 = \frac{(n-1)p}{n-p}F_{p,n-p,\alpha}$$



در حالت 2×2 قابل رسم است

e_1, e_2 , λ_1, λ_2

مقادیر ویژه و بردارهای ویژه ماتریس S

$$\begin{cases}|S-\lambda I| = 0\\ Se_i = \lambda_i e_i\end{cases}$$

$e_1' = (x_1, y_1)$

$n(X-\mu)'\Sigma^{-1}(X-\mu) = c^2$ معادله بیضی

با مرکزیت μ است که نیم قطرهای به طول $c\sqrt{\lambda_i}$ می باشد

$\theta = \operatorname{Arctan}(y_1/x_1)$

in R: library(ellipse) $(\bar{x}-mu)' S^{-1} (\bar{x}-mu) = t = c^2/n$

Subject: plot(ellipse(S, centre = xbar, t = sqrt(c^2/n), level = α), type = 'l')

Year. Month. Date. ( )

```
alpha = 0.05
S <- matrix(c(-,-,-,-),2,byrow=T)
xbar <- c(-,-)
c2 = (p*(n-1)/(n-p)) * qf(1-alpha, p, n-p)/n
```

اگر Σ معلوم باشد

$$n(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu) < \underbrace{\chi^2_p(\alpha)}_{c^2}$$

ناحیه اطمینان است.

بیضی اطمینان ~~خواهد~~ در این حالت نیز به مرکزیت $\bar{X}$ خواهد بود و پس از بدست آوردن مقادیر و بردارهای ویژه می‌توانید بیضی اطمینان را رسم کنید نصف طول قطرهای بیضی از رابطه

$$\frac{c\sqrt{\lambda_i}}{\sqrt{n}} = \sqrt{\chi^2(\alpha)\lambda_i/n}$$

بدست می‌آید و نسبت قطر بزرگ به کوچک از رابطه $\frac{\sqrt{\lambda_1}}{\sqrt{\lambda_2}}$ قابل محاسبه است

مثال ۵ ۲۳

مثال: فرض کنید $n = \cancel{4}3$, $p=2$

$$X = \begin{bmatrix} 8 & 6 & 10 \\ 3 & 9 & 6 \end{bmatrix}$$

$$\bar{X} = \begin{bmatrix} 8 \\ 6 \end{bmatrix} \qquad S = \begin{bmatrix} 4 & -3 \\ -3 & 9 \end{bmatrix} \rightarrow S^{-1} = \begin{bmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{bmatrix}$$

بازه اطمینان برای $\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

$$3(8-\mu_1 \quad 6-\mu_2)\begin{pmatrix} 1/3 & 1/9 \\ 1/9 & 4/27 \end{pmatrix}\begin{pmatrix} 8-\mu_1 \\ 6-\mu_2 \end{pmatrix} \le \underbrace{\frac{(n-1)p}{n-p}}_{36} \underbrace{F_{p,n-p,\alpha}}_{199.5}$$

پس مرکز بیضی (8, 6) و نصف قطر بلند و کوتاه بیضی

$$\sqrt{\lambda_1}\,\overbrace{\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p,\alpha}}}^{c/\sqrt{n}}$$

و محورهای بیضوی در جهت e_1 رسم می‌شوند

$$\sqrt{\lambda_2}\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p,\alpha}}$$

PAPCO

پس یعنی اطمینان ۹۵٪ برای $\underline{\mu}$ شامل تمام مقادیر μ_1 و μ_2 است که در نامساوی ★ صدق کنند. ~~الگوی فرض~~ فواصل اطمینان همزمان

حال فرض کنید فرض H_0 رد شده باشد در این صورت کدام یک از μ_1 و μ_2 مخالف μ_{0} بوده است این سوالی است که به کمک نواحی اطمینان همزمان می‌توان پاسخ داد. برای این منظور ابتدا ناحیه اطمینان برای هر ترکیب خطی از مولفه‌های بردار تصادفی $\underline{X}$ را بدست می‌آوریم. فرض کنید

$$\underline{X}_1, \dots, \underline{X}_n \sim N_p(\underline{\mu}, \Sigma)$$

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

آنگاه:

$$l_1 \bar{X}_1 + \dots + l_p \bar{X}_p = \underline{l}' \bar{\underline{X}} \qquad \underline{l} = \begin{pmatrix} l_1 \\ \vdots \\ l_p \end{pmatrix}$$

$$\Rightarrow \begin{cases} \underline{l}' \bar{\underline{X}} \sim N_1\left(\underline{l}'\underline{\mu}, \dfrac{\underline{l}'\Sigma\underline{l}}{n}\right) \\ (n-1)\underline{l}' S \underline{l} \sim W_1(\underline{l}'\Sigma\underline{l}, n-1) \\ \Rightarrow \dfrac{(n-1)\underline{l}' S \underline{l}}{\underline{l}'\Sigma\underline{l}} \sim \chi^2_{n-1} \end{cases}$$

$\bar{\underline{X}}$ و S مستقلند

پس:

$$\frac{\sqrt{n}\,(\underline{l}'\bar{\underline{X}} - \underline{l}'\underline{\mu}) / \sqrt{\underline{l}'\Sigma\underline{l}}}{\sqrt{\dfrac{(n-1)\underline{l}' S \underline{l}}{\underline{l}'\Sigma\underline{l}} \Big/ (n-1)}} \sim t_{n-1}$$

$(n-1)S \sim W(\Sigma, n-1)$

$$\frac{a' M a}{a' \Sigma a} \sim \chi^2_{n-1}$$

و در نتیجه یک ناحیه اطمینان برای $\mu' l$ به صورت

$$l'\bar{X} - t_{n-1,\alpha/2}\sqrt{\frac{l'Sl}{n}} \le l'\mu \le l'\bar{X} + t_{n-1,\alpha/2}\sqrt{\frac{l'Sl}{n}}$$

$$\overset{\text{یا}}{\Longrightarrow} \left|\frac{\sqrt{n}\,(l'\bar{X} - l'\mu)}{\sqrt{l'Sl}}\right| \le t_{n-1,\alpha/2} \qquad (1)$$

مثلاً فرض کنید $l' = [1, 0, 0 \dots 0]$ یعنی در این صورت می‌توانیم ناحیه اطمینان برای μ_1 را تشکیل دهیم

$$\begin{cases} l'\bar{X} = \bar{X}_1 \\ l'Sl = S_{11} \end{cases}$$

و اگر بخواهیم یک فاصله اطمینان برای $\mu_i - \mu_j$ بسازیم باید l را به صورت

$$l' = [0, \dots, 0, \underset{i}{1}, 0, \dots, 0, \underset{j}{-1}, 0 \dots 0]$$

در نظر بگیریم

$$l'\mu = \mu_i - \mu_j \qquad l'\bar{X} = \bar{X}_i - \bar{X}_j$$

$$l'Sl = S_{ii} - 2S_{ij} + S_{jj}$$

نکته: واضح است که با انتخاب بردارهای متفاوت l، فواصل اطمینان متفاوتی برای مؤلفه‌های μ هر یک با ضریب اطمینان $1-\alpha$ ساخته می‌شود و اطمینان مربوط به در نظر گرفتن تمام آنها با هم $1-\alpha$ نیست. برای مثال فرض کنید

مثلاً

$$\Sigma = \begin{pmatrix} \sigma_{11} & & \\ & \ddots & \\ & & \sigma_{pp} \end{pmatrix}$$

$l = [1, 0, 0 \dots]$

$l = [0, 1, \dots]$

$\mu_1 \dots \mu_p$

نکته: با انتخاب بردار متفاوت ℓ، فواصل متفاوت برای مؤلفه‌های μ هر یک با ضریب اطمینان $1-\alpha$ ساخته می‌شود. اطمینان در این حالت به ترتیب کمتر می‌شود تا اگر آن را با هم $1-\alpha$ است مثلاً

Subject:
Year. Month. Date. ( )

$\ell_1' = [1, \cdots]$ $\quad \ell_2' = [0\ 1\ 0 \cdots]$ ...

دراستفاده است

$$P\left(\bar{X}_1 - t_{n-1,\alpha/2}\sqrt{\frac{s_{11}}{n}} \le \mu_1 \le \bar{X}_1 + t_{n-1,\alpha/2}\sqrt{\frac{s_{11}}{n}}\right) = 1-\alpha$$

$$\vdots$$

$$P\left(\bar{X}_p - t_{n-1,\alpha/2}\sqrt{\frac{s_{pp}}{n}} \le \mu_p \le \bar{X}_p + t_{n-1,\alpha/2}\sqrt{\frac{s_{pp}}{n}}\right) = 1-\alpha$$

$P($ تمامی فواصل فوق پارامتر مورد نظر را در بر داشته باشد $) = (1-\alpha)\cdots(1-\alpha) = (1-\alpha)^p \le 1-\alpha$

حال فرض کنید ازای هر ℓ تمامی فواصل اطمینان که ساخته می‌شود شامل پارامتر مورد نظر باشد با تشکیل بردهیم این احتمال خیلی کوچک خواهد شد. مثال فوق تا فقط p ترکیب خطی است.

پس سؤالی که پیش می‌آید این است که چگونه می‌توانیم فواصلی را پیش کنیم که به ازای هر ترکیب خطی توأماً دارای ضریب اطمینان $1-\alpha$ باشد. پس هدف ساخت دادن ضریب اطمینان $1-\alpha$ به تمامی فواصل اطمینان که می‌توان توسط انتخاب‌های مختلف بردار ℓ ساخت می‌باشد. به این منظور می‌توان فواصلی که وسیع‌تر از فواصل معرفی شده در (1) است به وسیله انتخاب یک ℓ مشخص استفاده کرد.

به عبارتی می‌خواهیم $c > t_{n-1,\alpha/2}$

$$P\left(\left|\frac{\sqrt{n}(\ell'\bar{X} - \ell'\mu)}{\sqrt{\ell' S \ell}}\right| \le c\ ;\ \forall \ell\right) = 1-\alpha$$

آنرا چون انتخاب برای بسیاری از ℓها ... $c^2 > t^2_{n-1,\alpha/2}$ پس از c^2 که از t^2 بزرگتر است ... می‌کنیم

$$\Rightarrow P\left(\forall \ell \quad \frac{n(\ell'(\bar{X}-\mu))^2}{\ell' S \ell} \le c^2\right) = 1-\alpha$$

t^2 به ازای ℓ صفر — بردارهای ناصفر

برای ℓ ماکسیمم کند

$$\Rightarrow P\left(\max_{\ell}\left\{\frac{n(\ell'(\bar{X}-\mu))^2}{\ell' S \ell}\right\} \le c^2\right) = 1-\alpha \qquad \max_{b} \frac{(b'd)^2}{b' B b} = d' B^{-1} d$$

ماکزیمم

نتیجه می‌شود

$$\Rightarrow P\left(n(\bar{X}-\mu)' S^{-1}(\bar{X}-\mu) \le c^2\right) = 1-\alpha$$

T^2

از نامساوی کوشی شوارتز تعمیم یافته می‌توان نشان داد که

$$(b'd)^2 \le (b'Bb)(d'B^{-1}d)$$

$$\begin{cases} B = S \\ d = (\bar{X}-\mu) \\ x = \ell \end{cases}$$

$\frac{p(n-1)}{n-p} F_{p,n-p,\alpha}$

دیگر بردار $p\times 1$

معین مثبت

Subject:
Year. Month. Date. ( )

$$c=\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p,\alpha}}$$

$$\left|\frac{\sqrt{n}\,(\underset{\sim}{l}'\underset{\sim}{\bar{X}}-\underset{\sim}{l}'\underset{\sim}{\mu})}{\sqrt{\underset{\sim}{l}'S\underset{\sim}{l}}}\right|\leq\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p,\alpha}}$$

$$\Longrightarrow \underset{\sim}{l}'\underset{\sim}{\mu}\in\left\{\underbrace{\underset{\sim}{l}'\underset{\sim}{\bar{X}}\pm\sqrt{\frac{p(n-1)}{n(n-p)}F_{p,n-p,\alpha}\,\underset{\sim}{l}'S\underset{\sim}{l}}}\right\}$$

این بازه به طور همزمان برای تمام $\underset{\sim}{l}$ ها با احتمال $1-\alpha$

شامل $\underset{\sim}{\mu}'\underset{\sim}{l}$ است

پس به ازای انتخاب‌های مختلف $\underset{\sim}{l}$ ضریب اطمینان ثابت و $1-\alpha$ می‌ماند

$\underset{\sim}{l}'=[1,0,\dots 0]$ $\quad \underset{\sim}{l}'=[0,1,0,\dots 0]$ $\quad \dots$

$$\bar{X}_1-\sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p,\alpha}}\sqrt{\frac{S_{11}}{n}}\leq\mu_1\leq\bar{X}_1+\sqrt{\frac{p(n-1)}{(n-p)}F_{p,n-p,\alpha}}\sqrt{\frac{S_{11}}{n}}$$

$$\vdots$$

$$\bar{X}_p-\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p,\alpha}}\sqrt{\frac{S_{pp}}{n}}\leq\mu_p\leq\bar{X}_p+\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p,\alpha}}\sqrt{\frac{S_{pp}}{n}}$$

$$\underset{\sim}{l}'=(0,\dots,\underset{i}{1},\dots,\underset{k}{-1},\dots 0)$$

$$\bar{X}_i-\bar{X}_k-\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p,\alpha}}\sqrt{\frac{S_{ii}-2S_{ik}+S_{kk}}{n}}\leq\mu_i-\mu_k$$

$$\leq\bar{X}_i-\bar{X}_k-\sqrt{\frac{p(n-1)}{n-p}F_{p,n-p,\alpha}}\sqrt{\frac{S_{ii}-2S_{ik}+S_{kk}}{n}}$$

که تمام این فواصل همزمان با ضریب $1-\alpha$ برقرارند.

مثال: نمرات بدست آمده از $n=87$ دانشجوی دانشکده ~~در علوم اجتماعی~~ در جدول 2-5

ص ۲۴۲ اطلاعات زیر بدست آمده است: $p=3$

$$\bar{X}=\begin{bmatrix}527.74\\54.69\\25.13\end{bmatrix} \qquad S=\begin{bmatrix}5691.34 & 600.51 & 217.25\\600.51 & 126.05 & 23.37\\217.25 & 23.37 & 23.11\end{bmatrix}$$

فواصل اطمینان همزمان ٪۹۰ برای μ_1 ، μ_2 ، μ_3 :

$$\frac{p(n-1)}{(n-p)}F_{p,n-p,\alpha}=\frac{3(87-1)}{87-3}\underbrace{F_{3,84,0.05}}_{2.7}=8.29$$

$$\bar{x}_1-\sqrt{a}\sqrt{\frac{s_{11}}{n}}\le\mu_1\le\bar{x}_1+\sqrt{a}\sqrt{\frac{s_{11}}{n}}$$

$$527.74-\sqrt{8.29}\sqrt{\frac{5691.34}{87}}\le\mu_1\le \text{"} + \text{"}$$

به همین ترتیب برای μ_2 و μ_3 .

روش بونفرونی:

اغلب تعداد کمی از گزاره های اطمینان مورد توجه است به عبارتی اغلب به ازاء تعداد محدودی از بردارهای l مشخص نسبت ناحیه اطمینان همزمان لازم یعنی مثلا m ترکیب خطی $l_1'\mu, \dots, l_m'\mu$ مورد نظر باشد. در این صورت فواصل اطمینان همزمان را می توان طوری بدست آورد که از فواصل T^2 همزمان کوتاهتر (دقیق تر) باشد.

فرض کنید C_i بیانگر ناحیه اطمینان همزمان باشد

$$P(C_i \text{ درست}) = P(\text{ناحیه اطمینان همزمان ترکیب خطی } l_i'\mu \text{ را در بر داشته باشد}) = 1-\alpha_i$$

در اینصورت

$$P(\overset{\text{درست}}{C_1}\cap\dots\cap\overset{\text{درست}}{C_m})>1-\sum P(C_i \text{ غلط})=1-\sum_{i=1}^{m}\alpha_i$$

پس یک انتخاب مناسب برای α_i ، $\frac{\alpha}{m}$ است در اینصورت $\sum\alpha_i=\alpha$

در این حالت اگر ناحیه اطمینان می‌توانیم گزارش زیر را به دست آوریم
از شرط p ناحیه اطمینان هم‌زمان که توانش ...

$X_1 \dots X_n \sim N_p(\mu, \Sigma)$

$\ell' X \sim N(\ell'\mu, \ell'\Sigma\ell/n)$

$(n-1)\ell' S \ell \sim W_1(\ell'\Sigma\ell, n-1) \quad \frac{\ell' S \ell}{\ell'\Sigma\ell} \sim \chi^2_{n-1}$

$\frac{\sqrt{n}(\ell'\bar{X} - \ell'\mu)}{\sqrt{\ell' S \ell}} \sim t_{n-1}$

$$\bar{X}_1 - t_{n-1, \frac{\alpha}{2p}} \sqrt{\frac{S_{11}}{n}} \le \mu_1 \le \bar{X}_1 + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{11}}{n}}$$

$$\vdots$$

$$\bar{X}_p - t_{n-1, \frac{\alpha}{2p}} \sqrt{\frac{S_{pp}}{n}} \le \mu_p \le \bar{X}_p + t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{pp}}{n}}$$

در مثال ~~عرق~~ تعرق ~~ننتن~~ فواصل اطمینان ۹۵٪ هم‌زمان بونفرونی را برای μ_1، μ_2 و μ_3

بدست آوریم در اسفیورت $\alpha_i = \frac{0.05}{3}$

$$\cancel{\mu_1 \in \bar{X}_1 \pm t_{87-1}(0.0083)\sqrt{\frac{S_{11}}{n}}}$$

$$\cancel{\mu_2 \in \bar{X}_2 \pm t_{86}(0.0083)\sqrt{\frac{S_{22}}{n}}}$$

$$n = 20 \quad \bar{X} = \begin{bmatrix} 4.64 \\ 45.4 \\ 9.965 \end{bmatrix} \quad S = \begin{bmatrix} 2.879, & 10.002, & -1.810 \\ 10.002, & 199.798, & -5.627 \\ -1.810, & -5.627, & 3.628 \end{bmatrix}$$

$$\mu_1 \in \bar{X}_1 \pm t_{19}(0.0083)\sqrt{\frac{S_{11}}{n}} \Rightarrow 3.64 \le \mu_1 \le 5.64$$

$$37.10 \le \mu_2 \le 53.70$$

$$8.85 \le \mu_3 \le 11.08$$

جدول ۵-۴ برای مقایسه

$$\frac{\text{طول فاصله بونفرونی}}{\text{طول فاصله } T^2} = \frac{t_{n-1, \alpha/2p}}{\sqrt{\frac{p(n-1)}{n-p} F_{p, n-p, \alpha}}}$$

۰/۸۸

n	p = 2
۱۵	۰/۸۸
۲۵	۰/۹

$$t_{n-1, \alpha/2} \quad \underbrace{t_{n-1, \alpha/2m} \quad , \quad \sqrt{\frac{p(n-1)}{n-p} F_{p, n-p, \alpha}}}$$

$$\mu_i \in \bar{x}_i \pm c\sqrt{\frac{S_{ii}}{n}}$$

شکل کلی

معمولی، هتلینگ، ونفرونی

$n(\bar{X}-\mu_0)'S^{-1}(\bar{X}-\mu_0)\sim T^2$ ← (جامعه نرمال) Σ نامعلوم و n کوچک

$n(\bar{X}-\mu_0)'\Sigma^{-1}(\bar{X}-\mu_0)\sim \chi^2$ ← Σ معلوم < جامعه نرمال

$n(\bar{X}-\mu_0)'S^{-1}(\bar{X}-\mu_0)\sim \chi^2$ ← Σ نامعلوم و n بزرگ

آزمون برابری ~~نقاطی~~ بردار میانگین: استنباط ها برای یک نمونه ی خیلی بزرگ:

نکته: فرض کنید $X_1, \dots, X_n$ یک نمونه تصادفی از جامعه ای با میانگین $\underset{\sim}{\mu}$ و ماتریس کواریانس معین مثبت Σ است وقتی $n-p$ بزرگ باشد فرض H_0 آزمون

$$\begin{cases} H_0: \underset{\sim}{\mu}=\underset{\sim}{\mu_0} \\ H_1: \underset{\sim}{\mu}\neq \underset{\sim}{\mu_0} \end{cases}$$ در سطح معنی داری α

رد می شود هرگاه

$$n(\underset{\sim}{\bar{X}}-\underset{\sim}{\mu_0})'S^{-1}(\underset{\sim}{\bar{X}}-\underset{\sim}{\mu_0}) > \chi^2_{p,\alpha} \Rightarrow RH_0$$

صدک

$$\overbrace{P\left(n(\underset{\sim}{\bar{X}}-\underset{\sim}{\mu})'S^{-1}(\underset{\sim}{\bar{X}}-\underset{\sim}{\mu}) \le \chi^2_{p,\alpha}\right)}^{\text{ناحیه اطمینان } \mu \text{ بدست می آید}} = 1-\alpha$$

(100α) ام بالای توزیع کای دو

$$n-p \text{ بزرگ} \Rightarrow \frac{(n-1)p}{n-p}F_{p,n-p,\alpha} \approx \chi^2_{p,\alpha}$$

در این صورت ناحیه اطمینان $\underset{\sim}{l}'\underset{\sim}{\mu}$

$$\underset{\sim}{l}'\underset{\sim}{X} \pm \sqrt{\chi^2_{p,\alpha}}\sqrt{\frac{\underset{\sim}{l}'S\underset{\sim}{l}}{n}}$$

خواهد شد، فواصل هم زمان برای p میانگین

$$\mu_1: \bar{x}_1 \pm \sqrt{\chi^2_{p,\alpha}}\sqrt{\frac{s_{11}}{n}}$$

$$\vdots$$

$$\mu_p: \bar{x}_p \pm \sqrt{\chi^2_{p,\alpha}}\sqrt{\frac{s_{pp}}{n}}$$

فصل سوم

مقایسات زوجی

یادآوری: در حالت یک متغیره فرض کنید X_{1j} پاسخ به تیمار ۱ و X_{2j} پاسخ به تیمار ۲

(پاسخ قبل از تیمار) (پاسخ بعد از تیمار)

$$D_j = X_{1j} - X_{2j} \qquad j=1,\dots,n$$

با در نظر گرفتن تفاوت های D_j به صورت مشاهدات مستقل از توزیع $N(\delta, \sigma_d^2)$:

$$t = \frac{\bar{D}-\delta}{\frac{s_d}{\sqrt{n}}} \qquad \bar{D} = \frac{1}{n}\sum_{j=1}^{n} D_j \qquad s_d^2 = \frac{1}{n-1}\sum_{j=1}^{n}(D_j-\bar{D})^2$$

Subject:
Year. Month. Date. ( )

$$\begin{cases} H_0: \delta = 0 \\ H_1: \delta \neq 0 \end{cases} \qquad |t| > t_{n-1,\alpha/2} \Rightarrow RH_0$$

و فاصله اطمینان $100(1-\alpha)\%$ برای δ :

$$\bar{d} - t_{n-1,\alpha/2} \frac{s_d}{\sqrt{n}} \leq \delta \leq \bar{d} + t_{n-1,\alpha/2} \frac{s_d}{\sqrt{n}}$$

$$\delta \to E(X_{1j} - X_{2j})$$

برای تعمیم به حالت چند متغیره

X_{11j} : متغیر ۱ تحت تیمار ۱ $\qquad$ X_{21j} : متغیر ۱ تحت تیمار ۲

X_{12j} : متغیر ۲ تحت تیمار ۱ $\qquad$ X_{22j} : متغیر ۲ تحت تیمار ۲

$\vdots$

X_{1pj} : متغیر p تحت تیمار ۱ $\qquad$ X_{2pj} : متغیر p تحت تیمار ۲

$$\begin{cases} D_{1j} = X_{11j} - X_{21j} \\ D_{2j} = X_{12j} - X_{22j} \\ \vdots \\ D_{pj} = X_{1pj} - X_{2pj} \end{cases}$$

تفاوت زوج ۱ $\qquad$ تفاوت زوج p

$$\underset{\sim}{D}_j' = [D_{1j}, \ldots, D_{pj}] \qquad j = 1, \ldots, n$$

بطوریکه:

$$E(\underset{\sim}{D}_j) = \delta = \begin{bmatrix} \delta_1 \\ \vdots \\ \delta_p \end{bmatrix} \qquad Cov(\underset{\sim}{D}_j) = \Sigma_d$$

$$\underset{\sim}{D}_1, \ldots, \underset{\sim}{D}_n \sim N_p(\underset{\sim}{\delta}, \Sigma_d)$$

اکنون با معلوم بودن تفاضلهای مشاهده شده $\underset{\sim}{d}_j' = (d_{1j}, \ldots, d_{pj})$ $\quad j = 1, \ldots, n$
مربوط به متغیرهای تصادفی بردار $\underset{\sim}{D}_j'$ یک آزمون در سطح α

$$\begin{cases} H_0: \underset{\sim}{\delta} = 0 \\ H_1: \underset{\sim}{\delta} \neq 0 \end{cases}$$

PAPCO

برای یک جامعه $N_p(\underset{\sim}{\delta}, \Sigma_d)$:

$$T^2 \overset{H_0}{=} n\bar{d}' S_d^{-1} \bar{d} > \frac{(n-1)p}{n-p} F_{p, n-p, \alpha} \Rightarrow RH_0$$

و یک ناحیه اطمینان ٪$100(1-\alpha)$ برای $\underset{\sim}{\delta}$

$$(\bar{\underset{\sim}{d}} - \underset{\sim}{\delta})' S_d^{-1} (\bar{\underset{\sim}{d}} - \underset{\sim}{\delta}) \leq \frac{(n-1)p}{n(n-p)} F_{p, n-p, \alpha}$$

و فواصل اطمینان همزمان:

$$\delta_i : \bar{d}_i \pm \sqrt{\frac{(n-1)p}{n-p} F_{p, n-p, \alpha}} \cdot \sqrt{\frac{S_{d_i}^2}{n}}$$

برای $n-p$ (n بزرگ به کوچک) بزرگ $\frac{(n-1)p}{n-p} F_{p, n-p, \alpha} = \chi^2_{p, \alpha}$ و نرمال بودن ضروری نیست.

~~و فواصل اطمینان همزمان بن‌فرونی:~~

$$\cancel{\delta_i : \bar{d}_i \pm t_{n-p, \frac{\alpha}{p(p-1)}}} \sqrt{\frac{S_{d_i}^2}{n}}$$

مثال ۶-۱ ص ~~۲۷۶~~-۲۷۷

آزمون برابری مؤلفه‌ی بردار میانگین

فرض کنید $\underset{\sim}{X}_1, \dots, \underset{\sim}{X}_n \sim N_p(\underset{\sim}{\mu}, \Sigma)$ و هدف آزمون

$$① \begin{cases} H_0 : \mu_1 = \dots = \mu_p \\ H_1 : \mu_l \neq \mu_k \quad l \neq k \end{cases}$$

در این حالت ماتریس C را که به آن ماتریس مقایسه متقابل گویند به صورت زیر تشکیل می‌دهیم

$$C = (I_{p-1} \quad -\underset{\sim}{J}) = \left(\begin{array}{ccc|c} 1 & \cdots & 0 & -1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -1 \end{array} \right)_{(p-1)\times p}$$

$(p-1)(p-1)$

Subject:

Year. Month. Date. ( )

$$\bar{X} = \begin{bmatrix} ۱ \\ ۳ \\ ۵ \\ ۶ \end{bmatrix} \qquad S = \begin{bmatrix} ۱ & ۲ & ۳ & ۴ \\ ۲ & ۲ & ۱ & ۰ \\ ۳ & ۱ & ۳ & ۵ \\ ۴ & ۰ & ۲ & ۵ \end{bmatrix}$$

$n\bar{x}' S^{-1} \bar{x}$

$n(C\bar{x})'(CSC')^{-1}$

$$C\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 - \mu_p \\ \mu_2 - \mu_p \\ \vdots \\ \mu_{p-1} - \mu_p \end{bmatrix}$$

پس آزمون ① معادل است با آزمون

$$\begin{cases} H_0 : C\underset{\sim}{\mu} = \underset{\sim}{0} \\ H_1 : C\underset{\sim}{\mu} \neq \underset{\sim}{0} \end{cases}$$

$C_{(p-1)\times p}$ $S_{p\times p}$ $C'_{p \times (p-1)}$ $(p-1)\times(p-1)$

پس تعریف می کنیم $\underset{\sim}{Y}_j = \underset{(p-1)\times p}{C} \underset{\sim}{X}_{j\, p\times 1}$ ، $j = 1, \dots, n$

$$\Rightarrow \underset{\sim}{Y}_j \sim N_{p-1}(C\underset{\sim}{\mu}, C\Sigma C')$$

وطبق ~~خاصیت~~ تعریف آماره T^2 ~~تحت فرض~~ و $C\underset{\sim}{\mu} = 0$ (تحت فرض H_0)

$$T^2 = n\bar{Y}' S_Y^{-1} \bar{Y}$$

C را اول بنویسیم

$$= n\bar{X}'C'(CSC')^{-1}C\bar{X} \sim T^2_{p-1, n-1}$$

$n(\bar{x} - \mu)' S^{-1} (\bar{x} - \mu)$

شبیه به این فرم است که در فصل راحت تر است.

$$\Rightarrow \text{if } T^2 > \frac{(n-1)(p-1)}{n-p+1} F_{p-1, n-p+1, \alpha} \Rightarrow RH_0$$

$$T^2(p, n) = \frac{np}{n-p+1} F_{p, n-p+1} \Rightarrow T^2_{(p-1, n-1)} = \frac{(n-1)(p-1)}{(n-1)-(p-1)+1} = \frac{(n-1)(p-1)}{n-p+1}$$

مثال ۶-۲ :

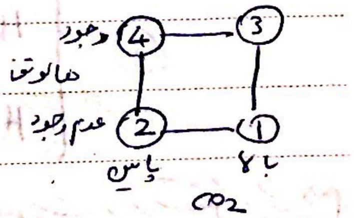
۲۸۱ - ۲۸۴

$n = 19 \qquad p = 4$

$$H_0 : \begin{cases} \underbrace{(\mu_3 + \mu_4)}_{\text{وجود هالوتن}} - \underbrace{(\mu_1 + \mu_2)}_{\text{عدم وجود هالوتن}} = 0 & \leftarrow H \\ (\mu_3 + \mu_1) - (\mu_2 + \mu_4) = 0 & \leftarrow CO_2 \\ (\mu_1 + \mu_4) - (\mu_2 + \mu_3) = 0 & \leftarrow H - CO_2 \end{cases}$$

تفاوت بین فشار بالا و پایین

تأثیر هالوتن روی تفاوت فشار $CO_2 - H$



ماتریس مقایسه :

$$\underset{\sim}{\mu}' = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix} \qquad C = \begin{array}{c} \begin{bmatrix} -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \end{array} \begin{array}{c} C_1 \\ C_2 \\ C_3 \end{array}$$

آزمون برابری مؤلفه‌های بردار میانگین

فرض کنید $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$ که در آن $\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}$ می‌خواهیم آزمون فرض زیر را انجام دهیم

① $H_0: \mu_1 = \mu_2 = \cdots = \mu_p$ vs $H_1: \mu_l \neq \mu_k$ برای حداقل یک $l \neq k$

فرض کنید $\underset{\sim}{X}_1, \ldots, \underset{\sim}{X}_n$ یک نمونه تصادفی از $N_p(\underset{\sim}{\mu}, \Sigma)$ باشند و تبدیل زیر را در نظر بگیرید.

$$C = (I_{p-1} \quad -\underset{\sim}{J}) = \begin{pmatrix} 1 & 0 & \cdots & 0 & -1 \\ 0 & \ddots & & \vdots & \vdots \\ \vdots & & & & \\ 0 & & & 1 & -1 \end{pmatrix}_{(p-1)\times p}$$

به ماتریس C ماتریس مقایسه متقارن گویند در این صورت آزمون ① معادل با آزمون زیر است

$$C\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 - \mu_p \\ \vdots \\ \mu_{p-1} - \mu_p \end{bmatrix} \qquad \begin{cases} H_0: C\underset{\sim}{\mu} = \underset{\sim}{0} \\ H_1: C\underset{\sim}{\mu} \neq \underset{\sim}{0} \end{cases}$$

پس تعریف می‌کنیم

$$\underbrace{\underset{\sim}{Y}_j}_{(p-1)\times 1} = \underbrace{C}_{(p-1)p} \underbrace{\underset{\sim}{X}_j}_{p\times 1} \qquad j = 1, \ldots, n$$

می‌دانیم

$$\Longrightarrow \underset{\sim}{Y}_j \sim N_{p-1}(C\underset{\sim}{\mu}, C\Sigma C')$$

در این صورت

تحت فرض H_0 آماره T^2: $\qquad H_0: C\underset{\sim}{\mu} = 0$

$$T^2 = n\bar{y}'S^{-1}\bar{y} = n\underset{\sim}{\bar{X}}'C'(\underbrace{CSC'}_{\text{C وارون پذیر است}})^{-1}C\underset{\sim}{\bar{X}} \sim T^2_{p-1, n-1}$$

در نتیجه فرض $H_0: C\underset{\sim}{\mu} = \underset{\sim}{0}$ در مقابل $H_1: C\underset{\sim}{\mu} \neq 0$ را در سطح معنی‌داری α رد خواهیم کرد اگر

$$T^2 > \frac{(n-1)(p-1)}{n-p+1} F_{p-1, n-p+1; \alpha} \Longrightarrow RH_0$$

$$T^2(p, n) = \frac{np}{n-p+1} F_{p, n-p+1} \Longrightarrow T^2_{p-1, n-1} = \frac{(n-1)(p-1)}{\underbrace{(n-1)-(p-1)+1}_{n-p+1}} F_{p-1, n-p+1}$$

Subject:
Year. Month. Date. ( )

همچنین یک ناحیه اطمینان $100(1-\alpha)\%$ برای مقایسه $C\mu$ بوسیله کلیه μ های C بدست می آید

که در رابطه زیر قرار گیرند

$$(C\mu - C\bar{X})'(CSC')^{-1}(C\mu - C\bar{X}) \le \frac{(n-1)(p-1)}{n(n-p+1)} F_{\alpha(p-1,n-p+1)}$$

همچنین برای هر بردار مقایسه c' (یک سطر از ماتریس C) یک فاصله اطمینان همزمان برای $c'\mu$ عبارت است از

$$c'\mu \in c'\bar{X} \pm \sqrt{\frac{(n-1)(p-1)}{\cancel{n}-p+1} F_{p-1,n-p+1,\alpha}} \sqrt{\frac{c'Sc}{n}}$$

نکته: اگر ماتریس C هر ماتریس مقایسه دیگری باشد (جمع هر سطر برابر صفر باشد) آزمون و فواصل اطمینان بالا برای هر مقایسه $C\mu$ برقرار می باشد.

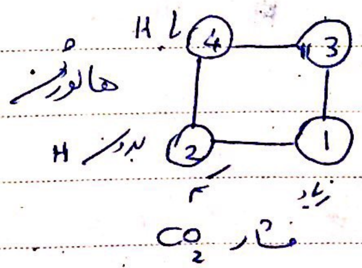
مثال: روشهای بیهوشی بردن جدید معمولاً روی حیوانات مورد آزمایش قرار می گیرد. در یک مطالعه به 19 سگ در ابتدا یک داروی بیهوشی داده شد و سپس تحت فشار گازی اکسید کربن (CO_2) با دو فشار کم و زیاد قرار گرفتند. در آخر گاز هالوتان (H) افزوده شده و دو مرتبه تحت فشار گاز (CO_2) کم و زیاد قرار گرفته است. زمان برحسب میلی ثانیه بین ضربان قلب آنها برای هر یک از چهار تیمار زیر اندازه گیری شده است

تیمار ۱ = H بدون CO_2 فشار زیاد μ_1

تیمار ۲ = H بدون CO_2 فشار کم μ_2

تیمار ۳ = H با CO_2 فشار زیاد μ_3

تیمار ۴ = H با CO_2 فشار کم μ_4



داده ۶ صفحه ۲۸۳

مقایسات زیر مد نظر است:

$$\begin{cases} H \leftarrow \overbrace{(\mu_3+\mu_4)}^{\text{وجود هالوتان}} - \overbrace{(\mu_1+\mu_2)}^{\text{عدم وجود هالوتان}} = 0 \\ H_0: \underbrace{(\mu_3+\mu_1)}_{\text{فشار زیاد}} - \underbrace{(\mu_2+\mu_4)}_{\text{فشار کم}} = 0 \quad CO_2 \\ (\mu_1+\mu_4) - (\mu_2+\mu_3) = 0 \end{cases}$$

$$\mu' = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{bmatrix} \qquad C = \begin{bmatrix} -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

تاثیر هالوتان روی تفاوتهای فشار

اثر متقابل CO_2-H

جمع سطرها صفر است.

Subject:
Year. Month. Date. ( )

$$\bar{X} = \begin{bmatrix} 368.21 \\ 404.63 \\ 479.26 \\ 502.89 \end{bmatrix} \qquad S = \begin{bmatrix} 2819.29 & & & \\ 3568.42 & 7963.14 & & \\ 2943.49 & 5303.98 & 6851.32 & \\ 2295.35 & 4065.44 & 4499.63 & 4878.99 \end{bmatrix}$$

$$C\bar{X} = \begin{bmatrix} 209.31 \\ -60.05 \\ -12.79 \end{bmatrix} \qquad CSC' = \begin{array}{c} c_1' S c_1 \\ \begin{bmatrix} 9432.32 & 1098.92 & 927.62 \\ 1098.92 & 5195.84 & 914.54 \\ 927.62 & 914.54 & 7557.44 \end{bmatrix} \end{array} {}_{2-P}$$

$$T^2 = n(C\bar{X})'(CSC')^{-1}(C\bar{X}) = 19(6.11) = 116$$

$$\alpha = 0.05$$

$$\frac{(n-1)(p-1)}{(n-p+1)} F_{p-1,\, n-p+1,\, \alpha} = \frac{18(3)}{16} \underbrace{F_{3,16,0.05}}_{3.24} = 10.94$$

یعنی فرض است هیچ تیماری تاثیر ندارد رد می‌شود $\Leftarrow$ $116 > 10.94 \Rightarrow RH_0$

* برای اینکه ببینیم کدام تفاوت در رد H_0 موثر بوده

فواصل اطمینان همزمان ۹۵٪ برای این تفاوت‌ها بدست می‌آوریم:

$$c_1' \underset{\sim}{\mu} = (\mu_3 + \mu_4) - (\mu_1 + \mu_2) = \text{ تاثیر مادر}$$

$$: (\bar{x}_3 + \bar{x}_4) - (\bar{x}_1 + \bar{x}_2) \pm \sqrt{10.94} \times \sqrt{\frac{c_1' S c_1}{19}}$$

$$= 209.31 \pm \sqrt{10.94} \times \sqrt{\frac{9432.32}{19}} = (\cancel{114.75, 5.35})$$

$$= (135.61, 283.01)$$

چون فاصله کاملاً مثبت پس با اطمینان ۹۵٪ هالوژن زمان بین ضربانها تشدید را زیادی کند

به همین ترتیب برای سایر گزاره‌ها.

$c_2' \underset{\sim}{\mu}:$ $\qquad c_3' \underset{\sim}{\mu}:$

"مطالعه دو جامعه" مقایسه بردارهای میانگین دو جامعه

متمایز

~~یادآوری~~: فرض کنید یک نمونه تصادفی به حجم n_1 از جامعه ۱ و نمونه تصادفی به حجم n_2 از جامعه ۲ داریم بطوریکه

$p \times 1$ هرکدام بردار

جامعه ۱ $\quad \underset{\sim}{X}_{11}, \underset{\sim}{X}_{12}, \cdots \underset{\sim}{X}_{1n_1}$ $\qquad \bar{\underset{\sim}{x}}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} \underset{\sim}{x}_{1j}$ $\qquad S_1 = \frac{1}{n_1 - 1} \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)(X_{1j} - \bar{X}_1)'$

جامعه ۲ $\quad \underset{\sim}{X}_{21}, \underset{\sim}{X}_{22}, \cdots \underset{\sim}{X}_{2n_2}$ $\qquad \bar{\underset{\sim}{x}}_2 = \frac{1}{n_2} \sum_{j=1}^{n_2} \underset{\sim}{x}_{2j}$ $\qquad S_2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)(X_{2j} - \bar{X}_2)'$

$\underset{\sim}{X}_{1\,p\times 1}$ $\qquad \underset{\sim}{X}_{2\,p\times 1}$

Subject:
Year. Month. Date. ( )

می‌خواهیم آزمون
$$\begin{cases} H_0: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{\delta}_0 \\ H_1: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 \neq \underset{\sim}{\delta}_0 \end{cases}$$
را انجام دهیم.

۱) دقت کنید: $\underset{\sim}{X}_{11}, \dots, \underset{\sim}{X}_{1n_1}$ نمونه تصادفی حجم n_1 از جامعه p متغیره با بردار میانگین $\underset{\sim}{\mu}_1$ و ماتریس کواریانس Σ_1 است.

۲) $\underset{\sim}{X}_{21} \dots \underset{\sim}{X}_{2n_2}$ نمونه تصادفی حجم n_2 از جامعه p متغیره دیگر با بردار میانگین $\underset{\sim}{\mu}_2$ و ماتریس کواریانس Σ_2 است.

۳) همچنین $\underset{\sim}{X}_{11}, \dots, \underset{\sim}{X}_{1n_1}$ مستقل از $\underset{\sim}{X}_{21}, \dots, \underset{\sim}{X}_{2n_2}$ هستند.

۴) اگر n_1 و n_2 کوچک باشد هر دو جامعه نرمال چند متغیره فرض می‌شوند و $\Sigma_1 = \Sigma_2$ در نظر گرفته می‌شود. که برای برآورد کواریانس مشترک یک متغیره ماتریس کواریانس S_p که از ادغام S_1 و S_2 بدست می‌آید داریم:

$$S_p = \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2}$$

ماتریس کوواریانس (ماتریس، ماتریس)

می‌خواهیم آزمون
$$\begin{cases} H_0: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{\delta}_0 \\ H_1: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 \neq \underset{\sim}{\delta}_0 \end{cases}$$
را انجام دهیم.

$$E(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) = \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 \quad , \quad Cov(\bar{\underset{\sim}{X}}_1, \bar{\underset{\sim}{X}}_2) = 0$$

$$Covar(\bar{\underset{\sim}{X}}_1) = \frac{\Sigma}{n_1} \quad \overset{cov}{var}(\bar{\underset{\sim}{X}}_2) = \frac{\Sigma}{n_2}$$

$$\Rightarrow \overset{cov}{var}(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma$$

برآورد ✓ (*) اینجا یکته شود

$$\hat{\overset{cov}{var}}(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)S_p \quad ✓$$

صفحه بعد ←

$$\cancel{S_p \sim W_p(\Sigma, n_1+n_2-2) / n_1+n_2-2}$$

$$(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2) \sim N\left(0, \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma\right) \Rightarrow \quad \left(\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)$$

Subject :

Year . Month . Date . ( )

$\frac{n_1 n_2}{n_1+n_2}$

مشابه آزمون بالا داریم :

$$T^{2\ \text{تحت}\ H_0}_{H_0} = (\bar{X}_1 - \bar{X}_2 - \delta_0)' \left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right) S_p\right]^{-1} (\bar{X}_1 - \bar{X}_2 - \delta_0) > c^2$$

$$\Rightarrow RH_0 \qquad c^2 = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p, n_1+n_2-p-1}$$

$(n_1+n_2-2)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)$ $\frac{n_2+n_1}{n_1 n_2}$

و همین طور به طور کلی اگر $X_{11}, \dots, X_{1n_1}$ یک نمونه تصادفی به حجم n_1 از جامعه $N_p(\mu_1, \Sigma)$ و $X_{21}, \dots, X_{2n_2}$ نمونه n_2 تایی از $N_p(\mu_2, \Sigma)$ به ترتیب باشند آنگاه ناحیه اطمینان $100(1-\alpha)\%$ برای $\mu_1-\mu_2$:

$$P\left([\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2)]' \left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right) S_p\right]^{-1} [\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2)] \leq c^2 \right) = 1-\alpha$$

$$c^2 = \frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p, n_1+n_2-p-1, \alpha}$$

چون (از مطلب قبل)

$$\begin{cases} \bar{X}_1 - \bar{X}_2 \sim N_p\left(\mu_1 - \mu_2, \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\Sigma\right) \\ (n_1-1) S_1 \sim W_p(\Sigma, n_1-1) \\ (n_2-1) S_2 \sim W_p(\Sigma, n_2-1) \end{cases}$$

مستقل

$\left(\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}\right)^{-1} \times$

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

$\sim N(0, \Sigma)$

$$(n_1-1) S_1 + (n_2-1) S_2 \sim W_p(\Sigma, n_1+n_2-2)$$

$$\Rightarrow S_p \sim W_p(\Sigma, n_1+n_2-2)/n_1+n_2-2$$

$$\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2))' S_p^{-1} \times \left(\frac{1}{n_1}+\frac{1}{n_2}\right)^{-1/2} (\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2))$$

(نرمال چندمتغیره)' $\left(\frac{\text{ویشارت}}{\text{درجه آزادی}}\right)^{-1}$ (نرمال چندمتغیره) $= T^2_{p, n_1+n_2-2}$

PAPCO

$$T^2_{p,n} = \frac{np}{n-p+1} F_{p, n-p+1} \qquad T^2_{p, n_1+n_2-2} = \frac{(n_1+n_2-2)p}{(n_1+n_2-p-1)} F_{p, n_1+n_2-p-1}$$

$X \sim N(\mu, \sigma^2)$

$2X \sim N(2\mu, 4\sigma^2)$ $(2X-2\mu) \sim N(0, 4\sigma^2)$

مثال ۶-۳ ص 288

$n_1 = ۵۰$ — روش اول

۵۰ کتاب صحافی شده به روش تقلیدی شوند

$n_2 = ۵۰$ — روش دوم

X_1 : کف

X_2 : نرمی

$p=2$

دو متغیر اندازه گیری می شوند.

روش اول $\underset{\sim}{\bar{X}_1} = \begin{bmatrix} 8.3 \\ 4.1 \end{bmatrix}$ $S_1 = \begin{bmatrix} 2 & 1 \\ 1 & 6 \end{bmatrix}$

روش دوم $\underset{\sim}{\bar{X}_2} = \begin{bmatrix} 10.2 \\ 3.9 \end{bmatrix}$ $S_2 = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$

یک ناحیه اطمینان ۹۵٪

برای $\mu_1-\mu_2$ بدست آورید. استدلال کنید S_1 و S_2 تقریباً با هم برابرند، فرض

را می‌پذیریم: $\Sigma_1=\Sigma_2=\Sigma$

$$S_p = \frac{(50-1)S_1+(50-1)S_2}{100-2} = \begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}$$

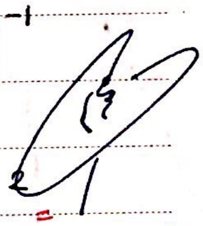
$\begin{bmatrix} \mu_{11}-\mu_{21} \\ \mu_{12}-\mu_{22} \end{bmatrix}$

$\underset{\sim}{\mu_1}-\underset{\sim}{\mu_2}$

$$\bar{X}_1-\bar{X}_2 = \begin{bmatrix} -1.9 \\ 0.2 \end{bmatrix}$$

$$\left(\begin{bmatrix} -1.9 \\ 0.2 \end{bmatrix}-\begin{bmatrix} \mu_{11}-\mu_{21} \\ \mu_{12}-\mu_{22} \end{bmatrix}\right)'\left[\left(\frac{1}{۵۰}+\frac{1}{۵۰}\right)\begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix}\right]^{-1}$$

$$\left(\begin{bmatrix} -1.9 \\ 0.2 \end{bmatrix}-\begin{bmatrix} \mu_{11}-\mu_{21} \\ \mu_{12}-\mu_{22} \end{bmatrix}\right) \le c^2$$



برای رسم بیضی:

$$X'AX = c^2 \qquad \left(\frac{1}{n_1}+\frac{1}{n_2}\right)c^2 = \left(\frac{1}{50}+\frac{1}{50}\right)\underbrace{\frac{\overbrace{(n_1+n_2-2)}^{98}\overbrace{2}^{p}}{\underbrace{n_1+n_2-p-1}_{97}}\underbrace{F_{2,97,\alpha}}_{3.1}}_{c^2}$$

$$= 0.25$$

$$0=|S_p-\lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 5-\lambda \end{vmatrix} = \lambda^2-7\lambda+9=0 \Rightarrow \begin{matrix} \lambda_1=5.303 \\ \lambda_2=1.697 \end{matrix}$$

$$S_p e_i = \lambda_i e_i \Rightarrow \quad e_1=\begin{bmatrix} 0.290 \\ 0.957 \end{bmatrix} \quad e_2=\begin{bmatrix} 0.957 \\ -0.290 \end{bmatrix}$$ ۱۰

PAPCO

$\mathrm{Arctan}\,\frac{0.957}{0.29} \simeq 73$; $\frac{0.957}{0.29} = 3.3$

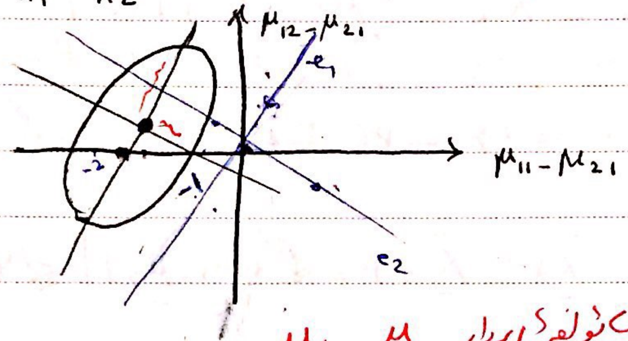
$c\sqrt{\lambda_i}/\sqrt{n}$

$\max \frac{(\underset{\sim}{x}'d)^2}{\underset{\sim}{x}'B\underset{\sim}{x}} = d'B^{-1}d$ $(b'd)^2 \le (b'Bb)(d'B^{-1}d)$

$\bar{X}_1 - \bar{X}_2 = \begin{bmatrix} -1.9 \\ 0.2 \end{bmatrix}$ پس در این حالت ناحیه اطمینان یک بیضی به مرکزیت

$\sqrt{\lambda_i}\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})c^2} = \sqrt{\lambda_i}\sqrt{0.25}$ $\begin{cases} 1.15 & \text{نصف قطرای} \\ 0.65 & \text{بیضی} \end{cases}$

در جهت محور e_1 و e_2



فواصل اطمینان هم زمان برای مولفه‌ی بردار $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$

فرض نرمال بودن دو جامعه و کواریانس مشترک Σ است.

$\underset{\sim}{X}_{11} \dots \underset{\sim}{X}_{1n_1} \sim N_p(\underset{\sim}{\mu}_1, \Sigma)$ $\Sigma_1 = \Sigma_2 = \Sigma$

$\underset{\sim}{X}_{21}, \underset{\sim}{X}_{22}, \dots \underset{\sim}{X}_{2n_2} \sim N_p(\underset{\sim}{\mu}_2, \Sigma)$

$\ell' = (\ell_1 \dots \ell_p)$ $\cancel{\ell_1(\bar{X}_{11}-\bar{X}_{21}) + \dots +} \ell_p(\bar{X}_{1p}-\bar{X}_{2p})$

$\ell'(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) \rightarrow \begin{cases} \underset{\sim}{\ell}'\bar{\underset{\sim}{X}}_1 \sim N_p(\underset{\sim}{\ell}'\underset{\sim}{\mu}_1, \underset{\sim}{\ell}'\Sigma\underset{\sim}{\ell}) \\ \underset{\sim}{\ell}'\bar{\underset{\sim}{X}}_2 \sim N_p(\underset{\sim}{\ell}'\underset{\sim}{\mu}_2, \underset{\sim}{\ell}'\Sigma\underset{\sim}{\ell}) \end{cases}$

$$P\left(\frac{[\ell'(\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2) - \ell'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)]}{\sqrt{(\frac{1}{n_1}+\frac{1}{n_2})\ell' S_p \ell}} \le c \cancel{t_{n_1+n_2-2}} ; \forall \ell\right) = 1-\alpha$$

$c > t_{n_1+n_2-2}$

$$\Rightarrow \frac{\{\ell'[(\bar{X}_1 - \bar{X}_2) - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)]\}^2}{\ell'(\frac{1}{n_1}+\frac{1}{n_2})S_p \ell} = t_\ell^2 \le c^2$$

و مشابه حالت تک جامعه‌ای مطابق لم ماکزیمم سازی داریم

$B = (\frac{1}{n_1}+\frac{1}{n_2})S_p$ و $d = [\bar{X}_1 - \bar{X}_2 - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)]$

$x' = \ell'$

Subject:
Year. Month. Date. ( )

$t_l^2 \le \max t_l^2 \Rightarrow$ α'

برای ~~...~~ داریم: $\underset{\sim}{l}$ B

$$t_l^2 \le (\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2))' \left[\left(\frac{1}{n_1}+\frac{1}{n_2}\right) S_p\right]^{-1}$$

$$(\bar{X}_1 - \bar{X}_2 - (\mu_1-\mu_2)) = T^2_{p, n_1+n_2-2} \Rightarrow c = \sqrt{\ }$$

$$\Rightarrow \cancel{p(T^2 \le c^2) = p(t_l^2 \le c^2 \ \forall l) = 1-\alpha}$$

$$\Rightarrow p\left[\left|l'(\bar{X}_1-\bar{X}_2) - l'(\mu_1-\mu_2)\right| \le c\sqrt{l'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_p l}\right.$$

$$\left.\forall l\right] = 1-\alpha \quad \text{با} \quad c = \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p, n_1+n_2-p-1}}$$

$$l'(\mu_1-\mu_2): \ l'(\bar{X}_1-\bar{X}_2) \pm c\sqrt{l'\left(\frac{1}{n_1}+\frac{1}{n_2}\right)S_p l}$$

مثال ۴۰۶ ص ۲۹۲

نمونه‌ای به حجم $n_1=45$ (دارند) و $n_2=55$ (ندارند) از صاحبان منازل شهری (ویسکانسین) که هوا دارند یا ندارند انتخاب کردیم در دو اندازه مربوط به استفاده برق

X_1: اندازه کل مصرف در حالت پیک در ماه جولای

X_2: " " " غیر پیک "

را در نظر گرفتیم. آماره‌ها بصورت زیر نتیجه شد:

$$\bar{X}_1 = \begin{bmatrix} 204.4 \\ 556.6 \end{bmatrix} \quad S_1 = \begin{bmatrix} 13825.3 & 23823.4 \\ 23823.4 & 73107.4 \end{bmatrix} \quad n_1 = 45$$

$$\bar{X}_2 = \begin{bmatrix} 130.0 \\ 355.0 \end{bmatrix} \quad S_2 = \begin{bmatrix} 8632.0 & 19616.7 \\ 19616.7 & 55964.5 \end{bmatrix} \quad n_2 = 55$$

فواصل اطمینان همزمان ۹۵٪ را برای تفاوت بردار میانگین‌ها پیدا کنیم

$$S_p = \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2} = \begin{bmatrix} 10963.7 & 21505.5 \\ 21505.5 & 63661.3 \end{bmatrix}$$

P4PCO

$$C^2 = \frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p,n_1+n_2-p-1,\alpha} = \frac{98(2)}{97} \underbrace{F_{2,97,0.05}}_{3.1}$$

$$= 6.26$$

$$\underset{\sim}{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix} \qquad \underset{\sim}{\mu}_2 = \begin{bmatrix} \mu_{21} \\ \mu_{12} \end{bmatrix}$$

$$(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)' = \begin{pmatrix} \mu_{11}-\mu_{21} & \mu_{12}-\mu_{22} \end{pmatrix}$$

فواصل اطمینان همزمان عبارتند از :

$$\mu_{11} - \mu_{21} : (204.4-130) \pm \sqrt{6.26}\sqrt{\left(\frac{1}{45}+\frac{1}{55}\right) 10963.7}$$

$$\Rightarrow 21.7 \leq \mu_{11} - \mu_{21} \leq 127.1$$

$$\mu_{12} - \mu_{22} : (556.6-355) \pm \sqrt{6.26}\sqrt{\left(\frac{1}{45}+\frac{1}{55}\right) 63661.3}$$

$$\Rightarrow 74.7 \leq \mu_{12} - \mu_{22} \leq 328.5$$

که نشان می دهند بین منازلی که تهویه دارند و ندارند در مصرف برق متفاوت هستند.

وضعیت دو نمونه وقتی $\Sigma_1 \neq \Sigma_2$

وقتی $\Sigma_1 \neq \Sigma_2$ نمی توان ~~تشکیل ... پیدا کنیم که~~ یک اندازه فاصله‌ای مانند T^2 که توزیعش به Σ_1 و Σ_2 نامعلوم بستگی نداشته باشد پیدا کنیم (آندربرگ Anderberg 1973)

از آزمون بارتلت برای آزمون کردن تساوی Σ_1 و Σ_2 برحسب واریانس تعمیم یافته استفاده کرد اما می دانیم اگر جامعه نرمال نباشند استفاده از این آزمون گمراه کننده است.

~~اگر بخواهیم با چنین مواردی مواجه شویم که S_1 و S_2 متفاوتند به خصوص ...~~

فقط ~~...~~ اگر n_1 و n_2 بزرگ باشند و تعداد متغیرها کم یعنی $n_1 - p$ و $n_2 - p$ بزرگ باشند می توان ~~بیضوی~~ ناحیه اطمینان $100(1-\alpha)\%$ تقریبی برای $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$ به صورت زیر نوشت :

$$\left[\bar{\underset{\sim}{X}}_1 - \bar{X}_2 - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)\right]' \left[\frac{1}{n_1}S_1 + \frac{1}{n_2}S_2\right]^{-1} \left[\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2 - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)\right] \leq \chi^2_{p,\alpha} \quad (*)$$

Subject:
Year. Month. Date. ( )

و فواصل اطمینان همزمان $100(1-\alpha)\%$ برای تمام ترکیبات خطی $\underset{\sim}{l}'(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2)$

$$\underset{\sim}{l}'(\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2)\pm\sqrt{\chi^2_p(\alpha)}\sqrt{\underset{\sim}{l}'(\frac{1}{n_1}S_1+\frac{1}{n_2}S_2)\underset{\sim}{l}}$$

می‌دانیم $E(\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2)=\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2$ ، $Cov(\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2)=\frac{1}{n_1}\Sigma_1+\frac{1}{n_2}\Sigma_2$

وقتی n_1-p و n_2-p بزرگ شوند بنا به قضیه حد مرکزی

$$\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2\sim N_p(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2\,,\ \frac{1}{n_1}\Sigma_1+\frac{1}{n_2}\Sigma_2)$$

است که اگر Σ_1 و Σ_2 معلوم باشند می‌دانیم

$$(\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2-(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2))'(\frac{1}{n_1}\Sigma_1+\frac{1}{n_2}\Sigma_2)^{-1}(\bar{\underset{\sim}{X}}_1-\bar{\underset{\sim}{X}}_2-(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2))$$

تقریباً یک توزیع χ^2_p است. وقتی n_1 و n_2 بزرگ هستند احتمال زیاد S_1 به Σ_1

و S_2 به Σ_2 نزدیک است و S_1 و S_2 جایگزین Σ_1 و Σ_2 گذاشته می‌شوند و رابطه ★

برقرار خواهد شد.

تبصره نکته: اگر $n_1=n_2=n$ پس $\frac{n-1}{n+n-2}=\frac{1}{2}$ و بنابراین

$$\frac{1}{n_1}S_1+\frac{1}{n_2}S_2=\frac{1}{n}(S_1+S_2)=\frac{(n-1)S_1+(n-1)S_2}{n+n-2}(\frac{1}{n}+\frac{1}{n})$$

$$=S_p(\frac{1}{n}+\frac{1}{n})$$

پس اگر حجم نمونه‌ها برابر باشد روش نمونه‌ای بزرگ مشابه برابری کواریانس‌ها انجام شده است.

مثال ۶-۵ : مثال مصرف برق را فرض کنید $\Sigma_1\neq\Sigma_2$ و $n_1-p=43$ ، $n_2-p=53$

داریم پس:

$$\frac{1}{n_1}S_1+\frac{1}{n_2}S_2=\begin{bmatrix}464.17 & 886.08\\ 886.08 & 2642.15\end{bmatrix}$$

$$\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2=\begin{array}{c} \\ 1\\ 2\end{array}\begin{array}{c}\text{جامعه ۱ \quad جامعه ۲}\\ \begin{bmatrix}\mu_{11}-\mu_{21}\\ \mu_{12}-\mu_{22}\end{bmatrix}\end{array}$$

$\underset{\sim}{l}'(\underset{\sim}{\mu}_1-\underset{\sim}{\mu}_2)$ ، $\underset{\sim}{l}'=[1,0]\,,\ [0,1]$

Subject:
Year. Month. Date. ( )

مقایسه میانگین‌های چند جامعه چندمتغیره multivariate Anova (MANOVA)

اغلب بیش از دو جامعه برای مقایسه لازم است فرض کنید g جامعه مد نظر است: (البته طوری که تمام جامعه‌ها ماتریس کواریانس مشترک دارند ۲- هر جامعه نرمال چندمتغیره است ۳- نمونه‌های تصادفی از جامعه‌های مختلف مستقلند)

$$
\begin{array}{llll}
\text{جامعه ۱} : & \underset{\sim}{X}_{11} \; \cdots\cdots \; \underset{\sim}{X}_{1n_1} & \underset{\sim}{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \vdots \\ \mu_{1p} \end{bmatrix} & \bar{\underset{\sim}{X}}_1 \\
\text{جامعه ۲} : & \underset{\sim}{X}_{21} \; \cdots\cdots \; \underset{\sim}{X}_{2n_2} & \vdots & \vdots \\
\vdots & & & \\
\text{جامعه } g : & \underset{\sim}{X}_{g1} \; \cdots\cdots \; \underset{\sim}{X}_{gn_g} & \underset{\sim}{\mu}_g = \begin{bmatrix} \mu_{g1} \\ \vdots \\ \mu_{gp} \end{bmatrix} & \bar{\underset{\sim}{X}}_g
\end{array}
$$

که برای بررسی فرضیه ① $\begin{cases} H_0 : \underset{\sim}{\mu}_1 = \cdots = \underset{\sim}{\mu}_g \\ H_1 : Not\ H_0 \end{cases}$ معمولاً به یک متغیره از ANOVA استفاده می‌شود در اینجا از MANOVA استفاده می‌کنیم.

بدیهی است $\underset{\sim}{X}_{lj} = \underset{\sim}{\mu}_l + \underset{\sim}{e}_{lj}$ و $\underset{\sim}{\mu}_l = \underset{\sim}{\mu} + \underbrace{(\underset{\sim}{\mu}_l - \underset{\sim}{\mu})}_{\underset{\sim}{\tau}_l}$ (*)

$$
\Rightarrow \underset{\sim}{X}_{lj} = \underset{\text{میانگین کل}}{\underset{\sim}{\mu}} + \underset{\text{اثر تیمار ام } l}{\underset{\sim}{\tau}_l} + \underset{\text{باقیمانده}}{\underset{\sim}{e}_{lj}} \quad (**) \quad \begin{array}{l} j = 1, \ldots, n_l \\ l = 1, \ldots, g \end{array}
$$

و از * نتیجه می‌شود که آزمون ① معادل با آزمون $\begin{cases} H_0 : \underset{\sim}{\tau}_1 = \cdots = \underset{\sim}{\tau}_g = 0 \\ H_1 : Not\ H_0 \end{cases}$ می‌باشد

به طوری که $\underset{\sim}{e}_{lj}$ ها متغیرهای مستقل $N_p(\underset{\sim}{0}, \Sigma)$ هستند و تحت H_0: $\sum_{l=1}^{g} n_l \underset{\sim}{\tau}_l = 0$

و برآورد مدل (**):

$$
\underset{\sim}{x}_{lj} = \underbrace{\bar{\underset{\sim}{x}}}_{\hat{\underset{\sim}{\mu}}} + \underbrace{(\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})}_{\hat{\underset{\sim}{\tau}}_l} + \underbrace{(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l)}_{\hat{\underset{\sim}{e}}_{lj}}
$$

و از حاصلضرب متقاطع $(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}})'$:

$$
(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}})' = [(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l) + (\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})][(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l) + (\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})]'
$$

$$
= (\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l)(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l)' + \underbrace{(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l)(\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})'}_{\text{جمع روی } j}
$$

$$
+ \underbrace{(\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l)'}_{\sum_j \text{ جمع روی } j} + (\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_l - \bar{\underset{\sim}{x}})'
$$

روی j جمع ببندیم

$$
\sum_{j=1}^{n_l} (\underset{\sim}{x}_{lj} - \bar{\underset{\sim}{x}}_l) = 0
$$

Subject:

Year. Month. Date. ( )

و بنابراین با جمع روی ℓ داریم :

$$\underbrace{\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})'}_{B+W}=\underbrace{\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)'}_{W}$$

$$+\underbrace{\sum_{\ell=1}^{g}\underbrace{\sum_{j=1}^{n_\ell}(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})'}_{n_\ell(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})'}}_{B}$$

و درنهایت :

مجموع مربعات کل $\quad B+W=\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})'$

مجموع مربعات باقیمانده $\quad W=\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)'=(n_1-1)S_1+\dots+(n_g-1)S_g$

مجموع مربعات تیمار $\quad B=\sum_{\ell=1}^{g}n_\ell(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})'$

و درنهایت جدول MANOVA :

منبع تغییرات	ماتریس مجموع مربعات	df
تیمار	B	$g-1$
خطا	W	$\sum_{\ell=1}^{g}n_\ell-g$
کل	$B+W$	$\sum_{\ell=1}^{g}n_\ell-1$

و فرآیند آزمون :

$$\Lambda^*=\frac{|W|}{|B+W|}$$

چند حالت خاص :

$$\text{if } \Lambda^*<c \Rightarrow RH.$$

تعداد متغیرها	تعداد گروه‌ها	آماره
$P=1$	$g\geq 2$	$F_0=\left(\frac{\sum n_\ell-g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right)\sim F_{g-1,\sum n_\ell-g}$
$P=2$	$g\geq 2$	$F_0=\left(\frac{\sum n_\ell-g-1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)\sim F_{2(g-1),2(\sum n_\ell-g-1)}$
$P\geq 1$	$g=2$	$F_0=\left(\frac{\sum n_\ell-P-1}{P}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right)\sim F_{P,\sum n_\ell-P-1}$
$P\geq 1$	$g=3$	$\cancel{F_0=\left(\frac{\sum n_\ell-P-2}{P}\right)}\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)\sim F_{2P,2(\sum n_\ell-P-2)}$

Subject:

Year. Month. Date. ( )

نکته: اگر $\sum n_\ell = n$ بزرگ باشد می‌توان نشان داد تحت فرض H_0 :

$$\chi_0^2 = -\left(n-1-\frac{p+g}{2}\right)\ln \Lambda^* = -\left(n-1-\frac{p+g}{2}\right)\ln\frac{|W|}{|B+W|}$$

دارای توزیع $\chi^2_{p(g-1)}$ است و اگر $\chi_0^2 > \chi^2_{p(g-1),\alpha} \Rightarrow RH_0$

مثال:

$$\begin{array}{c|ccc|l} & \underset{\sim}{x}_{11} & \underset{\sim}{x}_{12} & \underset{\sim}{x}_{1n_1} & \\ g=1 & \begin{pmatrix}9\\3\end{pmatrix} & \begin{pmatrix}6\\2\end{pmatrix} & \begin{pmatrix}9\\7\end{pmatrix} & \to x_1,\ x_2 \\ g=2 & \begin{pmatrix}0\\4\end{pmatrix} & \begin{pmatrix}2\\0\end{pmatrix} & & \\ g=3 & \begin{pmatrix}3\\8\end{pmatrix} & \begin{pmatrix}1\\9\end{pmatrix} & \begin{pmatrix}2\\7\end{pmatrix} & \\ & \underset{\sim}{x}_{31} & \underset{\sim}{x}_{32} & \underset{\sim}{x}_{33} & \end{array}$$

$\bar{\underset{\sim}{X}}_1 = \begin{pmatrix}8\\4\end{pmatrix}$

$\bar{\underset{\sim}{X}}_2 = \begin{pmatrix}1\\2\end{pmatrix}$ $\qquad \bar{\underset{\sim}{X}} = \begin{pmatrix}4\\5\end{pmatrix}$

$\bar{\underset{\sim}{X}}_3 = \begin{pmatrix}2\\8\end{pmatrix}$

$n_1=۳$

$n_2=۲$

$n_3=۳$

$$\underbrace{B+W}_{\sum_\ell\sum_j (\underset{\sim}{x}_{\ell j}-\underset{\sim}{\bar x})(\underset{\sim}{x}_{\ell j}-\underset{\sim}{\bar x})'} = \left(\begin{pmatrix}9\\3\end{pmatrix}-\begin{pmatrix}4\\5\end{pmatrix}\right)\left(\begin{pmatrix}9\\3\end{pmatrix}-\begin{pmatrix}4\\5\end{pmatrix}\right)' + \left(\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}4\\5\end{pmatrix}\right)\left(\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}4\\5\end{pmatrix}\right)' + \cdots = \begin{pmatrix}88 & -11\\ -11 & 72\end{pmatrix}$$

$$B+W = \begin{pmatrix}5\\-2\end{pmatrix}(5\ -2) + \begin{pmatrix}2\\-3\end{pmatrix}(2\ -3) + \begin{pmatrix}5\\2\end{pmatrix}(5\ 2) + \begin{pmatrix}-4\\-1\end{pmatrix}(-4\ -1)$$

$$+ \begin{pmatrix}-2\\-5\end{pmatrix}(-2\ -5) + \begin{pmatrix}-1\\3\end{pmatrix}(-1\ 3) + \begin{pmatrix}-3\\4\end{pmatrix}(-3\ 4) + \begin{pmatrix}-2\\2\end{pmatrix}(-2\ 2)$$

$$= \begin{pmatrix}88 & -11\\ -11 & 72\end{pmatrix}$$

$$B = \sum n_\ell (\underset{\sim}{\bar x}_\ell - \underset{\sim}{\bar x})(\underset{\sim}{\bar x}_\ell - \underset{\sim}{\bar x})'$$

$$\cancel{W = (n_1-1)S_1 + (n_2-1)S_2 + (n_3-1)S_3}$$

$$B = ۳\begin{pmatrix}\overset{4}{8-4}\\ \underset{-1}{4-5}\end{pmatrix}(\underset{4}{8-4}\quad \overset{-1}{4-5}) + 2\begin{pmatrix}\overset{-3}{1-4}\\ \underset{-3}{2-5}\end{pmatrix}(\overset{-3}{1-4}\quad \overset{-3}{2-5})$$

$$+ 3\begin{pmatrix}2-4\\8-5\end{pmatrix}(\overset{-2}{2-4}\quad \overset{3}{8-5}) = 3\begin{pmatrix}16 & -4\\ -4 & 1\end{pmatrix} + 2\begin{pmatrix}9 & 9\\ 9 & 9\end{pmatrix}$$

$$+ 3\begin{pmatrix}4 & -6\\ -6 & 9\end{pmatrix} = \begin{pmatrix}78 & -12\\ -12 & 48\end{pmatrix}$$

Subject:
Year. Month. Date. ( )

$$W = (B+W) - B = \begin{pmatrix} 10 & 1 \\ 1 & 24 \end{pmatrix}$$

جدول MANOVA

منبع تغییرات	ماتریس مجموع مربع	درجه آزادی
تیمار	$\begin{bmatrix} 78 & -12 \\ -12 & 48 \end{bmatrix}$	$g-1=2$
خطا	$\begin{bmatrix} 10 & 1 \\ 1 & 24 \end{bmatrix}$	$\Sigma n_e - g = 5$
کل	$\begin{bmatrix} 88 & -11 \\ -11 & 72 \end{bmatrix}$	7

$$\Lambda^* = \frac{|W|}{|W+B|} = 0.0385$$

$P=2$, $\underbrace{g=3}_{g \geq 2}$

$$F_0 = \frac{\Sigma n_e - g - 1}{g-1} \cdot \frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}} = 8.19$$

$$F_{\underbrace{2(g-1)}_{4},\ \underbrace{2(\Sigma n_e - g - 1)}_{8},\ 0.05} = 7.01$$

$F_0 > F_\alpha \Rightarrow RH_0$

مثال ۹-۹ صفحه ۳۰۸ بررسی شود.

دستورات:

```
Y1 <-c (9, 6, 9, 0, 2, 3, 1, 2)
Y2 <-c (3, 2, 7, 4, 0, 8, 9, 7)
Factor <-c (1, 1, 1, 2, 2, 3, 3, 3)
response =c bind (Y1, Y2)
Factor = factor (rep(c(1,2,3), c(3,2,3)))
fit = manovA(response ~ Factor)
summary(fit, test="wilks")
Summary (fit, test = "Pillia") -> V= tr(B(B+W)^-1)
Summary (fit, test = "Roy") -> Max(λi)   λi -> BW^-1
summary (fit, test = "Hotelling-lawely") -> T0^2 = tr(BW^-1)
a <- summary(fit)
a$SS -> B -> treatment
        W -> Residuals
```

Subject :

Year . Month . Date . ( )

مثال ۶-۹

براساس خلاصه آماره‌های زیر جدول MANOVA را تشکیل دهید

گروه

$l=1$ خصوصی $n_1=271$ $\bar{x}_1=[2.066\quad 0.480\quad 0.082\quad 0.360]'$

$l=2$ غیرانتفاعی $n_2=138$ $\bar{x}_2=[2.167\quad 0.596\quad 0.124\quad 0.418]'$

$l=3$ دولت $n_3=107$ $\bar{x}_3=[2.273\quad 0.521\quad 0.125\quad 0.383]'$

$\sum n_l = 516$

ماتریس کوواریانس :

$$S_1=\begin{bmatrix}0.291 & & & \\ -0.001 & 0.011 & & \\ 0.002 & 0.000 & 0.001 & \\ 0.010 & 0.003 & 0.000 & 0.010\end{bmatrix}$$

$$S_2=\begin{bmatrix}0.561 & & & \\ 0.011 & 0.025 & & \\ 0.001 & 0.004 & 0.005 & \\ 0.037 & 0.007 & 0.002 & 0.019\end{bmatrix}\qquad S_3=\begin{bmatrix}0.261 & & & \\ 0.030 & 0.017 & & \\ 0.003 & -0.000 & 0.004 & \\ 0.018 & 0.006 & 0.001 & 0.013\end{bmatrix}$$

$$W=(n_1-1)S_1+(n_2-1)S_2+(n_3-1)S_3=\begin{bmatrix}182.962 & & & \\ 4.408 & 8.2 & & \\ 1.695 & 0.633 & 1.484 & \\ 9.581 & 2.428 & 0.394 & 0.54\end{bmatrix}$$

$$\bar{x}=\frac{n_1\bar{x}_1+n_2\bar{x}_2+n_3\bar{x}_3}{n_1+n_2+n_3}=\begin{bmatrix}2.136\\0.516\\0.102\\0.380\end{bmatrix}$$

$$B=\sum n_l(\bar{x}_l-\bar{x})(\bar{x}_l-\bar{x})'=\begin{bmatrix}3.475 & & & \\ 1.111 & 1.225 & & \\ 0.821 & 0.453 & 0.235 & \\ 0.584 & 0.610 & 0.230 & 0.304\end{bmatrix}$$

$$\Lambda^*=\frac{|W|}{|B+W|}=0.7714$$

$$\left(\frac{\sum n_l-p-2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)=17.67>F_{8,1020}\Rightarrow RH.$$

$$-(n-1-(p+g)/2)\ln\frac{|W|}{|B+W|}=-511.5\times\ln(0.7714)=132.76>\chi^2_{8,0.01}=20.09$$

$\alpha=0.01$

Subject:
Year. Month. Date. ( )

فواصل اطمینان همزمان:

فرض کنید هدف بررسی مقایسه‌ای دو به دو باشد یعنی فواصل اطمینان همزمان برای مؤلفه‌ای اختلافهای $\tau_k - \tau_e$ یا $\mu_k - \mu_e$ فرض می‌کنند.

برآورد τ_k، $\hat{\tau}_k$ دیدیم $\bar{X}_k - \bar{X}$ است فرض کنید τ_{ki} مؤلفه iام τ_k باشد

$$\hat{\tau}_{ki} = \bar{X}_{ki} - \bar{X}_i$$

و به همین ترتیب $\hat{\tau}_e$ برآورد τ_e به صورت $\bar{X}_e - \bar{X}$ است و τ_{ei}

$$\hat{\tau}_{ei} = \bar{X}_{ei} - \bar{X}_i$$

(فرض کردیم نمونه‌ها مستقلند و کواریانس مشترک)

$$var(\hat{\tau}_{ki} - \hat{\tau}_{ei}) = var(\bar{X}_{ki} - \bar{X}_{ei}) = \left(\frac{1}{n_k} + \frac{1}{n_e}\right)\sigma_{ii}$$

عضو قطری Σ ام i

$W = (n_1-1)S_1 + \cdots + (n_g-1)S_g$

$\frac{W}{n-g} \rightarrow \hat{\Sigma}$ برآورد

$\hat{\sigma}^2 = \frac{SSE}{n-a}$

$$\Rightarrow \hat{var}(\bar{X}_{ki} - \bar{X}_{ei}) = \left(\frac{1}{n_k} + \frac{1}{n_e}\right)\frac{w_{ii}}{n-g}$$

و اگر هدف به دست آوردن فواصل اطمینان همزمان از روش بونفرونی برای تمام اختلافهای دو به دو باشد بین $\binom{g}{2}$ تفاوت دو به دو وجود دارد

$$\binom{g}{2} = \frac{g(g-1)}{2} \qquad \frac{\alpha}{2m} = \frac{\alpha}{pg(g-1)}$$

و p متغیر

و فواصل اطمینان همزمان بونفرونی به صورت زیر به دست می‌آید

$$\tau_{ki} - \tau_{ei}: \bar{X}_{ki} - \bar{X}_{ei} \pm t_{n-g,\left(\frac{\alpha}{pg(g-1)}\right)} \sqrt{\frac{w_{ii}}{n-g}\left(\frac{1}{n_k} + \frac{1}{n_e}\right)}$$

مثال ۶-۱۰: در مثال قبل:

$$\hat{\tau}_1 = (\bar{X}_1 - \bar{X}) = \begin{pmatrix} -0.070 \\ -0.039 \\ -0.020 \\ -0.020 \end{pmatrix}$$

$$\hat{\tau}_3 = (\bar{X}_3 - \bar{X}) = \begin{pmatrix} 0.137 \\ 0.002 \\ 0.023 \\ 0.003 \end{pmatrix}$$

$$\hat{\tau}_{13} - \hat{\tau}_{33} = -0.020 - 0.023 = -0.043$$

تیمار اول مؤلفه سوم متغیر سوم

تیمار سوم مؤلفه سوم متغیر سوم

PAPCO

۲۸-۱

Subject:

Year. Month. Date. ( )

$$\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_3}\right)\frac{W_{33}}{n-g}} = \sqrt{\left(\frac{1}{۲۷۱}+\frac{1}{۱۰۷}\right)\cdot\frac{۱,۴۸۸}{516-3}} = 0.00614$$

$$\tau_{13}-\tau_{33} : -0.043 \pm \overbrace{t_{513,\,0.00208}}^{2.87} \cdot 0.00614$$

$$= (-0.061\ ,\ -0.025)$$

هزینه نگهداری و کارمزد برای بیمارستان‌هایی که مالکیت آنها دولتی است ۲۵٪ تا ۶۱٪ ساعت به ازای هر مریض در روز بالاتر از بیمارستان‌های خصوصی است. به همین ترتیب می‌توانید برای بقیه تفاوت‌های $\tau_{k3}-\tau_{l3}$ فاصله اطمینان ۹۵٪ بدست آورید.

Subject:
Year. Month. Date. ( )

MANOVA دوطرفه :

$$\underset{\sim}{x}_{lkr} = \underset{\sim}{\mu} + \underset{\sim}{\tau}_l + \underset{\sim}{\beta}_k + \underset{\sim}{\gamma}_{lk} + \underset{\sim}{e}_{lkr}$$

$$l = 1, \dots, g \qquad k = 1, \dots, b \qquad r = 1, \dots, n$$

$e_{lkr} \sim N_p(0, \Sigma)$, $\sum_{l=1}^{g} \tau_l = \sum_{k=1}^{b} \beta_k = \sum_{l=1}^{g} \gamma_{lk} = \sum_{k=1}^{b} \gamma_{lk} = 0$ به طوری که

بردار تصادفی

جدول MANOVA :

منبع تغییرات	مجموع مربعات SSP	درجات آزادی
عامل ۱	$\sum_{l=1}^{g} bn(\bar{\underset{\sim}{x}}_{l.} - \bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_{l.} - \bar{\underset{\sim}{x}})'$	$g-1$
عامل ۲	$\sum_{k=1}^{b} gn(\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}})'$	$b-1$
متقابل	$\sum_{l=1}^{g}\sum_{k=1}^{b} n(\bar{\underset{\sim}{x}}_{lk} - \bar{\underset{\sim}{x}}_{l.} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_{lk} - \bar{\underset{\sim}{x}}_{l.} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}})'$	$(g-1)(b-1)$
خطا	$\sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (\underset{\sim}{x}_{lkr} - \bar{\underset{\sim}{x}}_{lk})(\underset{\sim}{x}_{lkr} - \bar{\underset{\sim}{x}}_{lk})'$	$(gb)(n-1)$
کل	$\sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n} (\underset{\sim}{x}_{lkr} - \bar{\underset{\sim}{x}})(\underset{\sim}{x}_{lkr} - \bar{\underset{\sim}{x}})'$	$gbn-1$

$$H_0: \underset{\sim}{\gamma}_{11} = \underset{\sim}{\gamma}_{12} = \dots = \underset{\sim}{\gamma}_{gb} = 0$$

$$H_1: \underset{\sim}{\gamma}_{lk} \neq 0$$

$$\Lambda^* = \frac{|SSP_{res}|}{|SSP_{int} + SSP_{res}|}$$

برای نمونه‌ی بزرگ

$$-\left[gb(n-1) - \frac{p+1-(g-1)(b-1)}{2}\right] \ln \Lambda^* > \chi^2_{(g-1)(b-1)\times p, \alpha}$$

$$\Rightarrow RH_0$$

Subject:

Year. Month. Date. ( )

$$H_0: \underset{\sim}{\tau_1} = \dots = \underset{\sim}{\tau_g} = 0 \qquad \Lambda^* = \frac{|SSP_{res}|}{|SSP_{fac1} + SSP_{res}|}$$

$$-\left[gb(n-1) - \frac{p+1-(g-1)}{2}\right]\ln \Lambda^* > \chi^2_{(g-1)p,\alpha}$$

$$\Longrightarrow RH_0$$

$$H_0: \underset{\sim}{\beta_1} = \dots = \underset{\sim}{\beta_b} = 0 \qquad \Lambda^* = \frac{|SSP_{res}|}{|SSP_{fac2} + SSP_{res}|}$$

$$-\left[gb(n-1) - \frac{p+1-(b-1)}{2}\right]\ln \Lambda^* > \chi^2_{(b-1)p,\alpha} \Longrightarrow RH_0$$

~~فواصل اطمینان همزمان~~ :

$$\cancel{\tau_{li} - \tau_{mi} : \bar{x}_{l \cdot i} - \bar{x}_{m \cdot i} \pm t\left(\frac{\alpha}{pg(g-1)}\right)\sqrt{\frac{E_{ii}}{\nu} \cdot \frac{2}{bn}}}$$

مثال ۱۲.۶

Fit = manova (response ~ Factor1 + Factor2)

Factor1 * Factor2

مؤلفه‌های اصلی :

گاهی در عمل در نظر گرفتن p مؤلفه جهت بررسی مدنظر است اما می‌توان با تحلیل (برای مطالعه کل سیستم) مؤلفه اصلی با تعداد مؤلفه کمتر یعنی k مؤلفه اصلی کل سیستم را مورد مطالعه قرار داد.

پس تحلیل مؤلفه اصلی وسیله‌ای برای رسیدن به اهداف در تعاریک مطالعه است که هدف از استفاده از مؤلفه‌های اصلی ۱) کاهش حجم داده‌ها ۲) تعبیر و تفسیر صحیح آنها است و بر اساس ساختار واریانس ـ کواریانس به کمک چند ترکیب خطی از متغیرها (p متغیر اولیه) سر و کار دارد.

تعریف : فرض کنید Σ ماتریس کواریانس بردار تصادفی $\underset{\sim}{X}' = [X_1 \dots X_p]$ باشد و Σ دارای زوج مقدار ویژه - بردار ویژه $(\lambda_1, \underset{\sim}{e}_1) \dots (\lambda_p, e_p)$ باشد بطوریکه داریم $\lambda_1 \geq \lambda_2 \geq \dots \lambda_p \geq 0$ ، مؤلفه اصلی i ام با :

$$\underset{\sim}{Y}_i = \underset{\sim}{e}_i' \underset{\sim}{X} = e_{1i} X_1 + e_{2i} X_2 + \dots + e_{pi} X_p$$

$i = 1, \dots, p$

نشان داده می‌شود و داریم :

$$\begin{cases} V(\underset{\sim}{Y}_i) = \underset{\sim}{e}_i' \Sigma \underset{\sim}{e}_i = \lambda_i & i = 1, \dots p \\ cov(Y_i, Y_k) = e_i' \Sigma e_k = 0 & i \neq k \end{cases}$$

اثبات: چون $\underset{\sim}{e}_i' \underset{\sim}{e}_i = 1$ ، $e_i' e_k = 0$ $\quad i \neq k$

$\Sigma \underset{\sim}{e}_k = \lambda_k \underset{\sim}{e}_k \Rightarrow e_i' \Sigma e_k = e_i' \lambda_k e_k$

$\Sigma \underset{\sim}{e}_i = \lambda_i \underset{\sim}{e}_i \Rightarrow \underset{\sim}{e}_i' \Sigma \underset{\sim}{e}_i = \underset{\sim}{e}_i' \lambda_i \underset{\sim}{e}_i = \lambda_i$

نکته : از تعریف بالا نتیجه می‌شود که مؤلفه‌های اصلی ناهمبسته‌اند و واریانس آنها برابر مقادیر ویژه Σ است.

Subject:
Year. Month. Date. ( )

نکته ۲ : فرض کنید $\underset{\sim}{X}' = (X_1, \ldots, X_p)$ دارای ماتریس کوواریانس Σ است. فرض کنید

با زوج مقادیر ویژه ـ بردار ویژه $(\lambda_1, e_1), \ldots, (\lambda_p, e_p)$ است. فرض کنید

$$\underset{\sim}{Y_1} = \underset{\sim}{e_1}' \underset{\sim}{X}, \ldots, \underset{\sim}{Y_p} = \underset{\sim}{e_p}' \underset{\sim}{X}$$

p مؤلفه‌ی اصلی با شرط آن‌ها می‌دانیم

$$\sum_{i=1}^{p} var(X_i) = \sigma_{11} + \cdots + \sigma_{pp} = tr(\Sigma) = \sum_{i=1}^{p} \lambda_i = \sum_{i=1}^{p} var(Y_i)$$

قبلاً تعریف کردیم : واریانس کل : $\sigma_{11} + \cdots + \sigma_{pp} = \sum_{i=1}^{p} \lambda_i$

پس

سهم کل واریانس جامعه مربوط به مؤلفه اصلی k ام $= \dfrac{\lambda_k}{\lambda_1 + \cdots + \lambda_p} \quad k = 1, 2, \ldots, p$

اگر برای p بزرگ بیشتر واریانس کل جامعه مثلاً ۸۰ تا ۹۰ درصد آن را بتوان

به سه مؤلفه اول نسبت داد در آن صورت این مؤلفه‌ها را بدون این که اطلاعات زیادی

از دست بدهیم می‌توان جانشین p متغیر اول کرد.

نکته ۳ اگر $\underset{\sim}{Y_1} = \underset{\sim}{e_1}' \underset{\sim}{X}, \ldots, Y_p = e_p' \underset{\sim}{X}$ ، $\underset{\sim}{X} = (X_1, \ldots, X_p)$ و $\underset{\sim}{e_i}$ ها هم بردارهای

ویژه ماتریس Σ) مؤلفه‌های اصلی به دست آمده از ماتریس کوواریانس که به شرط آن‌ها

$$\rho_{Y_i, X_k} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} \qquad i, k = 1, \ldots, p$$

ضریب همبستگی بین مؤلفه‌ی $\underset{\sim}{Y_i}$ و متغیر X_k

$l_k = (0, \ldots, \overset{k}{1}, 0, \ldots, 0)$

$$cov(\underset{\sim}{e_i}' \underset{\sim}{X}, \underset{\sim}{l_k}' \underset{\sim}{X}) = l_k' \underbrace{\Sigma \underset{\sim}{e_i}}_{\lambda_i e_i} = \lambda_i e_{ki}$$

اثبات: $$\rho_{Y_i, X_k} = \frac{Cov(\underset{\sim}{Y_i}, X_k)}{\sqrt{var(\underset{\sim}{Y_i})}\sqrt{var(X_k)}} = \frac{\lambda_i e_{ki}}{\sqrt{\lambda_i}\sqrt{\sigma_{kk}}} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}}$$

مثال ۸ ـ ۱

$Cov(A\underset{\sim}{X}) = A'\Sigma A$ $\quad Cov(\underset{1\times p}{A}\underset{\sim}{X}, \underset{1\times p}{B}\underset{\sim}{X}) = \underset{1\times p}{B}\Sigma \underset{p\times 1}{A'}$

Subject:

Year. Month. Date. ( )

مثال ۸-۱ فرض کنید متغیرهای تصادفی X_1، X_2 و X_3 دارای ماتریس کواریانس زیر باشند

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$|\overset{\Sigma}{A} - \lambda I| = 0 \Rightarrow \begin{vmatrix} 1-\lambda & -2 & 0 \\ -2 & 5-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 5-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} + 2\begin{vmatrix} -2 & 0 \\ 0 & 2-\lambda \end{vmatrix} + 0$$

$$5-\lambda-5\lambda+\lambda^2$$

$$= (1-\lambda)(5-\lambda)(2-\lambda) + 2(-2)(2-\lambda)$$

$$= -\lambda^3 + 6\lambda^2 - 5\lambda + 10 - 12\lambda + 2\lambda^2 - 8 + 4\lambda$$

$$= \underbrace{-\lambda^3}_{-۱۹۸۲} + \underbrace{8\lambda^2}_{۲۷۱٫۹} - ۱۳\lambda + 2 = 0 \quad \begin{cases} \lambda_1 = 5.83 \\ \lambda_2 = 2 \\ \lambda_3 = 0.17 \end{cases}$$

$$\Sigma e_i = \lambda_i e_i \Rightarrow \begin{cases} e_1' = [0.383, -0.924, 0] \\ e_2' = [0, 0, 1] \\ e_3' = [0.924, 0.383, 0] \end{cases}$$

پس مؤلفه‌های اصلی عبارتند از:

$$Y_1 = e_1' X = 0.383 X_1 - 0.942 X_2$$

$$Y_2 = e_2' X = X_3$$

$$Y_3 = e_3' X = 0.924 X_1 + 0.383 X_2$$

پس متغیر X_3 یکی از مؤلفه‌های اصلی است (دست نخورده X_3 از X_1، X_2 ناهمبسته است)

سهم کل واریانس مربوط به مؤلفه اصلی اول

$$= \frac{\lambda_1}{\lambda_1+\lambda_2+\lambda_3} = \frac{5.83}{8} = 0.73$$

$$\frac{\lambda_2}{\lambda_1+\lambda_2+\lambda_3} = \frac{2}{8} = 0.25$$

$$\Big\} \; 0.98$$

PAPCO

$$var(Y_1) = var(0.383 X_1 - 0.924 X_2) = 5.83 = \lambda_1$$

$$Cov(Y_1, Y_2) = Cov(0.383 X_1 - 0.924 X_2 , X_3) = 0$$

$$\rho_{Y_1, X_1} = \frac{e_{11}\sqrt{\lambda_1}}{\sqrt{\sigma_{11}}} = \frac{0.383\sqrt{5.83}}{\sqrt{1}} = 0.925$$

$$\rho_{Y_1, X_2} = -0.998$$

اهمیت هریک از متغیرهای X_1 و X_2 در اولین مؤلفه اصلی تقریباً یکسان است

$$\rho_{Y_2, X_1} = \rho_{Y_2, X_2} = 0 \quad , \quad \rho_{Y_2, X_3} = 1$$

پس با بکارگیری Y_1 و Y_2 به جای سه متغیر اولیه اطلاعات ناچیزی از دست می‌دهیم.

بدست آوردن مؤلفه‌های اصلی از متغیرهای استاندارد:

مؤلفه‌های اصلی را می‌توان برای متغیرهای استاندارد شده بصورت زیر بدست آورد

$$Z_1 = \frac{X_1 - \mu_1}{\sqrt{\sigma_{11}}} \quad \dots \quad Z_p = \frac{X_p - \mu_p}{\sqrt{\sigma_{pp}}}$$

$$\underset{\sim}{Z} = (V^{1/2})^{-1}(\underset{\sim}{X} - \underset{\sim}{\mu})$$

$$V = diag(\sigma_{11} \dots \sigma_{pp}) = \begin{bmatrix} \sqrt{\sigma_{11}} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\sigma_{pp}} \end{bmatrix}$$

$$Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix}$$

که قبلاً دیدیم

$$V^{1/2} \rho V^{1/2} = \Sigma \quad , \quad \begin{cases} \rho = (V^{1/2})^{-1} \Sigma (V^{1/2})^{-1} \\ E(\underset{\sim}{Z}) = 0 \quad , \quad cov(Z) = \rho \end{cases}$$

که ماتریس کوواریانس

پس مؤلفه‌های اصلی Z را از بردارهای ویژه ماتریس همبستگی ρ به دست آورد.

نتیجه: مؤلفه اصلی i ام متغیرهای استاندارد شده $\underset{\sim}{Z}' = [Z_1, \dots, Z_p]$ با $Cov(Z) = \rho$ بصورت

$$\underset{\sim}{Y_i} = \underset{\sim}{e}_i' \underset{\sim}{Z} = \underset{\sim}{e}_i' (V^{1/2})^{-1} (\underset{\sim}{X} - \underset{\sim}{\mu}) \qquad i = 1, \dots, p$$

می باشد و $\sum_{i=1}^{p} var(Y_i) = \underbrace{\sum_{i=1}^{p} var(Z_i)}_{I\ [1, \dots, 1]} = p$ و $\rho_{Y_i, Z_k} = e_{ki} \sqrt{\lambda_i}$ $\quad i = 1, \dots, p$ $\quad e_i' \Sigma_Z e_i$

در این حالت $(\lambda_1, e_1), \dots, (\lambda_p, e_p)$ زوجهای مقدار ویژه - بردار ویژه $\underline{\rho}$ با شرط $\lambda_1 \geq \dots \geq \lambda_p \geq 0$ هستند.

می دانیم واریانس کل جامعه متغیرهای استاندارد شده برابر p یعنی مجموع اعضای قطری ماتریس ρ است.

پس

$$\text{سهم کل واریانس جامعه مربوط به مؤلفه اصلی } k \text{ ام} = \frac{\lambda_k \ (\text{مقدار ویژه } \rho)}{\underbrace{p}_{\sum \lambda_i = p}} \qquad k = 1, 2, \dots, p$$

مثال: ماتریس کوواریانس $\Sigma = \begin{bmatrix} 1 & 4 \\ 4 & 100 \end{bmatrix}$ و ماتریس همبستگی $\rho = \begin{bmatrix} 1 & 0.4 \\ 0.4 & 1 \end{bmatrix}$ را در نظر بگیرید:

$$\begin{vmatrix} 1-\lambda & 4 \\ 4 & 100-\lambda \end{vmatrix} = 0 \Rightarrow (1-\lambda)(100-\lambda) - 16 = 0$$

$$\Rightarrow \lambda^2 - \underbrace{101\lambda}_{10116.16} - 84 = 0 \quad \begin{cases} \lambda_1 = 100.16 \\ \lambda_2 = 0.84 \end{cases}$$

$$\begin{cases} \underset{\sim}{e}_1' = [0.040 \quad 0.999] \\ \underset{\sim}{e}_2' = [0.999 \quad -0.040] \end{cases}$$

و برای ماتریس ρ:

$$\begin{vmatrix} 1-\lambda & 0.4 \\ 0.4 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 0.16 = 0 \Rightarrow \lambda^2 + 1 - 2\lambda - 0.16 = 0 \Rightarrow$$

$$\begin{cases} \lambda_1 = 1.4 & \underset{\sim}{e}_1' = [0.707 \quad 0.707] \\ \lambda_2 = 0.6 & \underset{\sim}{e}_2' = [0.707 \quad -0.707] \end{cases}$$

$\rightarrow 2 = p$

براساس Σ:

$$Y_1 = 0.04 X_1 + 0.999 X_2$$

$$Y_2 = 0.999 X_1 - 0.040 X_2$$

Subject.
Year. Month. Date. ( )

$$P:\begin{cases} Y_1 = 0.707\,Z_1 + 0.707\,Z_2 = 0.707\overset{\frac{X_1-\mu_1}{\sqrt{\sigma_{11}}}}{(X_1-\mu_1)} + 0.0707(X_2-\mu_2) \\ Y_2 = 0.707\,Z_1 - 0.707\,Z_2 = 0.707(X_1-\mu_1) - 0.0707(X_2-\mu_2) \end{cases}$$

$$\Sigma:\ \frac{\lambda_1}{\lambda_1+\lambda_2} = \frac{100.16}{101} = 0.992 \longrightarrow$$ مؤلفه اصلی اول ۰٫۹۹ از واریانس کل جامعه را بیان می‌کند

پس مؤلفه اصلی اول به جای دو متغیر اول به کار می‌رود.

$$P:\ \frac{\lambda_1}{P} = \frac{1.4}{2} = 0.7 \Longrightarrow$$ مؤلفه اصلی اول ۰٫۷۰ از واریانس کل جامعه است نوار استاندارد شده را بیان می‌کند.

می‌دانیم در بعضی مسائل که مقیاس‌های متغیرها باهم خیلی تفاوت دارند و به درجه‌ها دستاندارد شوند.

پس بعضی مواقع لازم است از متغیرهای استاندارد شده استفاده شود و در نتیجه نیاز به مؤلفه‌های اصلی داریم که از این متغیرها به دست می‌آیند.

نکته:

در عمل Σ مجهول است و اگر $S=\{s_{ik}\}$ ماتریس واریانس نمونه $p\times p$ باشد مقادیر و بردارهای ویژه S را به دست می‌آوریم و مؤلفه‌های اصلی نمونه را با

$$\hat{Y}_i = \hat{e}_i'\,\underset{\sim}{X} \qquad i=1,\dots,p$$

داده می‌شود که $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_p \geq 0$.

$$V(\hat{Y}_k) = \hat{\lambda}_k \qquad \operatorname{cov}(\hat{Y}_i,\hat{Y}_k)=0 \quad i\neq k$$

$$\sum_{i=1}^{p} s_{ii} = \cancel{\sum_{k=1}^{p}\hat{\lambda}_k} = \hat{\lambda}_1+\dots+\hat{\lambda}_p$$

$$r_{\hat{Y}_i,X_k} = \frac{\hat{e}_{ki}\sqrt{\hat{\lambda}_i}}{\sqrt{s_{kk}}} \qquad i,k=1,\dots,p$$

به همین ترتیب برای داده‌های استاندارد شده ماتریس کواریانس R استفاده می‌شود

P4PCO

چند مثال کتاب

①

Subject:

Year. Month. Date. ( )

۲۲۵ ص

۹

۱) فرض کنید ماتریس داده‌ی مربوط به نمونه تصادفی به حجم ۲۰ از یک جامعه نرمال دو متغیره به صورت زیر باشد

$$X = \begin{bmatrix} 8 & 12 & 5 & 3 & 6 & 7 & 8 & 4 & 9 & 11 & 10 & 7 & 8 & 9 & 5 & 15 & 12 & 8 & 7 & 10 \\ 4 & 8 & 7 & 3 & 4 & 5 & 6 & 7 & 2 & 3 & 4 & 5 & 2 & 7 & 6 & 3 & 4 & 2 & 1 & 4 \end{bmatrix}$$

الف) در سطح معنی‌داری ۰/۰۵ فرض $H_0: \mu' = (8\ 6)$ را با استفاده از T^2 و لاندای ویلکس آزمون کنید

ب) در سطح معنی‌داری ۰/۰۵ فرض $\mu_1 - \mu_2 = 0$ را آزمون کنید

ج) ناحیه اطمینان ۹۵٪ را برای میانگین جامعه به دست آورید

$$\begin{cases} H_0: \mu = \begin{bmatrix} 8 \\ 6 \end{bmatrix} \\ H_1: \underset{\sim}{\mu} \neq \begin{bmatrix} 8 \\ 6 \end{bmatrix} \end{cases} \qquad \bar{X} = \begin{bmatrix} 8.3 \\ 4.25 \end{bmatrix} \qquad S = \begin{bmatrix} 8.642 & -1.553 \\ -1.553 & 3.355 \end{bmatrix}$$

$$S^{-1} = \frac{1}{26.282} \begin{bmatrix} 3.355 & 1.553 \\ 1.553 & 8.642 \end{bmatrix} = \begin{bmatrix} 0.128 & 0.059 \\ 0.059 & 0.329 \end{bmatrix}$$

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

$$= 20 \underbrace{\left(\begin{pmatrix} 8.3 \\ 4.25 \end{pmatrix} - \begin{pmatrix} 8 \\ 6 \end{pmatrix} \right)'}_{(0.3\ \ -1.75)_{1\times 2}} \begin{pmatrix} 0.128 & 0.059 \\ 0.059 & 0.329 \end{pmatrix}_{2\times 2} \begin{pmatrix} 0.3 \\ -1.75 \end{pmatrix}_{2\times 1}$$

$$= ۱۹/۱۴۱ \qquad \frac{(n-1)p}{n-p} F_{\underbrace{p}_{۲},\ \underbrace{n-p}_{۱۸},\ \alpha} = \frac{۱۹ \times ۲}{۱۸} F_{2, 18, 0.05} = ۷/۵۵$$

$$\Rightarrow T^2 > ۷/۵۰۵ \Rightarrow RH_0$$

$$\Lambda^{*\,2/n} = \left(1 + \frac{T^2}{n-1}\right)^{-1} = \left(1 + \frac{۱۹/۱۴۱}{۱۹}\right)^{-1} = ۰/۴۹۸$$

$$\Lambda^{*\,2/n} = \left(\frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right) = \qquad \hat{\Sigma}_0 = \frac{1}{n} \sum (X_i - \mu)(X_j - \mu)'$$

$$\rightarrow \hat{\Sigma}_0 = \begin{bmatrix} 8.3 & -2 \\ -2 & 6.25 \end{bmatrix}$$

د) فواصل اطمینان همزمان ۹۵٪ برای μ_2, μ_1

$\mu_1 - \mu_2$ ~~...~~ با فرض ... به دست آورید

Subject:

Year. Month. Date. ( )

$C = [1 \quad -2]$ $\qquad C\underset{\sim}{\mu} = \mu_1 - 2\mu_2$ $\qquad \begin{cases} H_0: C\underset{\sim}{\mu} = 0 \\ H_1: C\underset{\sim}{\mu} \neq 0 \end{cases}$ ب)

$$\Rightarrow T^2 = n(C\bar{X})'(\underset{1\times2\,2\times2\,2\times1}{CSC'})^{-1}(C\bar{X}) \sim T^2_{p-1,n-1}$$

$$= 0.102$$

$$\frac{(n-1)p}{n-p+1} F_{p-1,n-p+1} = ۲ F_{2,19,0.05} = ۷/۰۴۴$$

$$T^2 < \frac{(n-1)p}{n-p+1} F_{p-1,n-p+1} \Rightarrow A H_0$$

ج)

$$T^2 = 20\left(\begin{bmatrix} 8.3 \\ 4.25 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}\right)' \begin{bmatrix} 0.128 & 0.059 \\ 0.059 & 3.555 \end{bmatrix} \left(\begin{bmatrix} 8.3-\mu_1 \\ 4.25-\mu_2 \end{bmatrix}\right)$$

$$= 20(8.3-\mu_1 \quad 4.25-\mu_2)\left(\quad '' \quad \right)\left[\quad '' \quad \right] < \underbrace{\frac{(n-1)p}{n-p} F_{p,n-p,\alpha}}_{۷/۵۰۵}$$

د)

$$\mu_i : \bar{X}_i \pm \sqrt{\frac{(n-1)p}{n-p} F_{p,n-p,\alpha}} \sqrt{\frac{S_{ii}}{n}}$$

$$\mu_1 : 8.3 \pm \sqrt{\underbrace{7.505}_{2.739}} \sqrt{\frac{8.642}{20}} = (6.5, 10.1)$$

$$\mu_2 : 4.25 \pm 2.7505 \sqrt{\frac{3.355}{20}} = (3.1, 5.4)$$

$$\mu_1 - \mu_2 : \bar{X}_1 - \bar{X}_2 \pm \sqrt{7.505} \sqrt{\frac{S_{11}+S_{22}-۲S_{12}}{n}} = (1/67, 6.43)$$

P4PCO

۲

Subject:

Year. Month. Date. ( )

روش بونفرونی

$$l'\bar{X} \pm t_{n-1, \alpha/2m} \sqrt{\frac{l'Sl}{n}}$$

$$\mu_1 : \bar{X}_1 \pm t_{19, \frac{0/05}{4}} \sqrt{\frac{S_{11}}{20}} = 8/3 \pm 2/925 \sqrt{\frac{8/448}{20}}$$

= (۶٫۵۷۹ , ۱۰٫۰۲۶)

$$\mu_2 : \bar{X}_2 \pm t_{19, \frac{0/05}{4}} \sqrt{\frac{S_{22}}{20}}$$ = (۳٫۱۷۶ , ۵٫۳۲۴)

$$\mu_1 - \mu_2 : (\bar{X}_1 - \bar{X}_2) \pm t_{n-1, \alpha/2m} \sqrt{\frac{S_{11} + S_{22} - 2S_{12}}{n}}$$

= (۱٫۷۶۹ , ۶٫۳۳۱)

فواصل اطمینان به روش بونفرونی از T^2 همگی کوتاهتر است .

۲

$$\bar{X} = \begin{bmatrix} 19 \\ 2 \end{bmatrix} \qquad S = \begin{bmatrix} 70 & 6 \\ 6 & 1 \end{bmatrix}$$

$$|S - \lambda I| = 0 \rightarrow \begin{vmatrix} 70-\lambda & 6 \\ 6 & 1-\lambda \end{vmatrix}$$

$$\Rightarrow \cancel{7/\dots +} \lambda^2 - 71\lambda + 34 = 0 \rightarrow \lambda_1 = 70.518, \quad \lambda_2 = 0.482$$

$$S\underset{\sim}{x} = \lambda_1 \underset{\sim}{x} \Rightarrow \begin{bmatrix} 70 & 6 \\ 6 & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 70/518 \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$\Rightarrow \begin{cases} 70X_1 + 6X_2 = 70.518X_1 \\ 6X_1 + X_2 = 70.518X_2 \end{cases} \rightarrow X_1 = 1, \quad X_2 = 0.084$$

$$\underset{\sim}{x} = [1 \quad 0.084]$$

P4PCO

$$\rightarrow e_1' = \left[\frac{X_1}{\sqrt{\underset{\sim}{x}'\underset{\sim}{x}}} , \frac{X_2}{\sqrt{\underset{\sim}{x}'\underset{\sim}{x}}}\right] = [0.996 \quad 0.084]$$

Subject:
Year. Month. Date. ( )

$$e_2' = [0.084 \quad -0.996]$$

$$y_1 = \underset{\sim}{e_1} \underset{\sim}{X} = [0.996 \quad 0.084]\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 0.996 X_1 + 0.084 X_2$$

$$y_2 = \underset{\sim}{e_2} \underset{\sim}{X} = 0.084 X_1 - 0.996 X_2$$

$$var(y_1) = \lambda_1 = 70.518$$

$$var(y_2) = \lambda_2 = 0.4825$$

نسبت واریانس مؤلفه اصلی اول $\dfrac{\lambda_1}{\lambda_1 + \lambda_2} = 0.993$ $\qquad \dfrac{\lambda_2}{\lambda_1 + \lambda_2} = 0.007$

مؤلفه اصلی اول ۹۹٪ واریانس کل را بیان می کند پس می توان بدون آنکه اطلاعاتی از دست بدهیم از مؤلفه اصلی اول استفاده کنیم.

$$\rho_{y_1, x_1} = \frac{e_{11}\sqrt{\lambda_1}}{\sqrt{\sigma_{11}}} = \frac{0.996\sqrt{70.518}}{\sqrt{70}} = 0.999$$

$$\rho_{y_2, x_2} = 0.7 \qquad \rho_{y_2, x_1} = 0.0069 \qquad \rho_{y_2, x_2} = -0.083$$

۳

Subject:

Year. Month. Date. ( )

(۳

$n_1 = ۵$

(الف

$$\begin{cases} H_0: \underset{\sim}{\mu_1} - \underset{\sim}{\mu_2} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \\ H_1: \underset{\sim}{\mu_1} - \underset{\sim}{\mu_2} \neq \begin{bmatrix} 5 \\ 2 \end{bmatrix} \end{cases}$$

$$\bar{X}_1 = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

$$\bar{X}_2 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

$$S_1 = \begin{bmatrix} 2.5 & -1.5 \\ -1.5 & 2 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 1 & -1.5 \\ -1.5 & 3 \end{bmatrix}$$

$$S_p = \frac{(n_1-1)S_1 + n_2-1)S_2}{n_1+n_2-2} = \begin{bmatrix} 1.33 & -0.5 \\ -0.5 & 0.33 \end{bmatrix}$$

$$S_p^{-1} = \begin{bmatrix} 1.75 & 2.64 \\ 2.64 & 7.0357 \end{bmatrix}$$

$$\left(\frac{1}{n_1} + \frac{1}{n_2}\right) = 0.53 \rightarrow \frac{1}{0.53} = 1/887$$

$$T^2 = (\underset{\sim}{\bar{X}_1} - \underset{\sim}{\bar{X}_2} - (\underset{\sim}{\mu_1} - \underset{\sim}{\mu_2}))' \left(\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_p\right)^{-1} (\underset{\sim}{\bar{X}_1} - \underset{\sim}{\bar{X}_2} - (\hat{\mu}_1 - \hat{\mu}_2))$$

$= ۳۹/۸۱۵۴$

$$\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{p, n_1+n_2-p-1} = ۱۳/۸۸۶$$

$$T^2 > ۱۳/۸۸۶ \rightarrow RH_0$$

P4PCO

Subject: $\begin{pmatrix}\mu_{11}\\ \mu_{12}\end{pmatrix}$ $\begin{pmatrix}\mu_{21}\\ \mu_{22}\end{pmatrix}$ $\underset{1\times 2}{(1\ \ 0)}\underset{2\times 1}{\begin{pmatrix}\bar{x}_{11}-\bar{x}_{22}\\ \bar{x}_{12}-\bar{x}_{22}\end{pmatrix}}$

Year. Month. Date. ( )

$$\underset{\sim}{\ell}'(\underset{\sim}{\bar{X}}_1-\bar{X}_2) \pm \sqrt{\frac{p(n_1+n_2-2)}{(n_1+n_2-p-1)} F_{p, n_1+n_2-p-1}}$$ (C.

$(\bar{x}_{11} -$ $$\times \sqrt{\underset{\sim}{\ell}'\left(\frac{1}{n_1}+\frac{1}{n_2}\right) S_p \underset{\sim}{\ell}}$$ $\ell' = (1\ \ 0)$

$$\mu_{11}-\mu_{21} : \bar{X}_{11}-\bar{X}_{21} \pm \sqrt{13.886}\sqrt{0.709}$$

$$\mu_{12}-\mu_{22} : \bar{X}_{12}-\bar{X}_{22} \pm \sqrt{13.886}\sqrt{0.175}$$

$$\begin{cases} H_0 : \underset{\sim}{\mu}_1 = \underset{\sim}{\mu}_2 = \underset{\sim}{\mu}_3 \\ H_1 : \underset{\sim}{\mu}_i \neq \mu_j \end{cases} \quad \text{يا} \quad \begin{cases} \underset{\sim}{\tau}_1 = \underset{\sim}{\tau}_2 = \underset{\sim}{\tau}_3 = \underset{\sim}{0} \end{cases}$$

$$\underset{\sim}{\bar{X}}_1 = \begin{bmatrix}8\\4\end{bmatrix} \quad \underset{\sim}{\bar{X}}_2 = \begin{bmatrix}1\\2\end{bmatrix} \quad \underset{\sim}{\bar{X}}_3 = \begin{bmatrix}2\\8\end{bmatrix} \quad \underset{\sim}{\bar{X}} = \begin{bmatrix}4\\5\end{bmatrix}$$

$$S_1 = \begin{bmatrix}3 & 3\\ 3 & 7\end{bmatrix} \quad S_2 = \begin{bmatrix}2 & -4\\ -4 & 8\end{bmatrix} \quad S_3 = \begin{bmatrix}1 & -0.5\\ -0.5 & 1\end{bmatrix}$$

$n_1 = 3 \quad n_2 = 2 \quad n_3 = 3$

$$W = (n_1-1)S_1 + (n_2-1)S_2 + (n_3-1)S_3 = \begin{bmatrix}10 & 1\\ 1 & 24\end{bmatrix}$$

$$B = \sum n_\ell (\bar{X}_\ell - \underset{\sim}{\bar{X}})(\underset{\sim}{\bar{X}}_\ell - \underset{\sim}{\bar{X}})' = 3\begin{bmatrix}4\\-1\end{bmatrix}[4\ \ -1]$$

$$+\,2\begin{bmatrix}-3\\-3\end{bmatrix}[-3\ \ -3] + 3\times\begin{bmatrix}-2\\3\end{bmatrix}[-2\ \ 3]$$

$$= \begin{bmatrix}78 & -12\\ -12 & 48\end{bmatrix}$$

df

$g-1$	$=2$	B	?
$\sum n_\ell - g$	$=5$	W	?
	7	B+W	?

PAPCO

f

$$\Lambda^* = \frac{|W|}{|W+B|} = 0.038\cancel{46}$$

$g = 3 \qquad p \geq 1$

$$\frac{\sum n_\ell - p - 2}{p} \cdot \frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}} \sim \underbrace{F_{\underbrace{2p}_{4},\ \underbrace{2(\sum n_\ell - p - 2)}_{8}}}_{3.833}$$

$$\frac{8-2-2}{2} \cdot \frac{1-\sqrt{0.038}}{\sqrt{0.038}} \cong 8.2$$

$8.2 > F_{4,8}$

$\Rightarrow$ RH.