Chapter 6 Notes

6.2 Paired Comparison and Repeated Measures

Univariate Case – n subjects, 2 treatments, 1 outcome measure

$$D_j = X_{j1} - X_{j2} \quad j = 1, \ldots, n \qquad \bar{D} = \frac{1}{n}\sum_j D_j \qquad s_d^2 = \frac{1}{n-1}\sum_i (D_j - \bar{D})^2$$

$$D_j \sim N(\delta, \sigma_d^2) \Rightarrow t = \frac{\bar{D} - \delta}{s_d/\sqrt{n}} \sim t_{n-1}$$

$$H_0: \delta = 0 \quad H_A: \delta \neq 0 \quad \text{T.S. } t_{obs} = \frac{\bar{D}}{s_d/\sqrt{n}} \quad RR: |t_{obs}| \geq t_{n-1}\left(\frac{\alpha}{2}\right)$$

$$(1-\alpha)100\% \text{ CI for } \delta: \quad \bar{D} \pm t_{n-1}\left(\frac{\alpha}{2}\right)$$

Multivariate Case – n subjects, 2 treatments, p outcome measures

Subject j:

$X_{1j1}, \ldots, X_{1jp}$ Treatment 1

$X_{2j1}, \ldots, X_{2jp}$ Treatment 2

$$\underset{\sim}{D_j} = \begin{bmatrix} D_{j1} = X_{1j1} - X_{2j1} \\ \vdots \\ D_{jp} = X_{1jp} - X_{2jp} \end{bmatrix} \qquad \underset{\sim}{\delta} = E\{\underset{\sim}{D_j}\} = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{bmatrix} \qquad Cov\{\underset{\sim}{D_j}\} = \Sigma_d$$

$$\underset{\sim}{\bar{D}} = \frac{1}{n}\sum_{j=1}^{n} \underset{\sim}{D_j} \qquad S_d = \frac{1}{n-1}\sum_{j=1}^{n} (\underset{\sim}{D_j} - \bar{D})(D_j - \bar{D})'$$

Note: Let $D^* = \begin{bmatrix} \underset{\sim}{D}_1' \\ \underset{\sim}{D}_2' \\ \vdots \\ \underset{\sim}{D}_n' \end{bmatrix} \Rightarrow S_d = \frac{1}{n-1} D^{*\prime}(I_n - \frac{1}{n}J_n)D^*$

If $\underset{\sim}{D}_1, \dots, \underset{\sim}{D}_n \sim$ iid $N_p(\underset{\sim}{\delta}, \Sigma_d) \Rightarrow$ inferences about δ can be obtained from T^2 statistic:

$$T^2 = n(\underset{\sim}{\bar{D}} - \underset{\sim}{\delta})' S_d^{-1} (\underset{\sim}{\bar{D}} - \underset{\sim}{\delta})$$

Result 6.1 under normality $T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$ whatever true values of δ, Σd.

If $n, n-p$ are both large, $T^2 \dot{\sim} \chi^2_p$ regardless of underlying distn of differences.

---

observed differences: $\underset{\sim}{d}_j' = [d_{j1}, \dots, d_{jp}]$

$H_0: \underset{\sim}{\delta} = \underset{\sim}{0}$ $\quad H_A: \underset{\sim}{\delta} \neq \underset{\sim}{0}$ $\quad$ TS: $T^2 = n\,\underset{\sim}{\bar{d}}' S_d^{-1} \bar{d}$

Reject H_0 if $T^2 \geq \frac{p(n-1)}{n-p} F_{p,n-p}(\alpha)$

$(1-\alpha)100\%$ Confidence Region for δ

$\underset{\sim}{\delta}$ s.t. $\quad (\bar{d} - \underset{\sim}{\delta})' S_d (\underset{\sim}{\bar{d}} - \underset{\sim}{\delta}) \leq \frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha)$

$100(1-\alpha)\%$ Simultaneous CI's for Individual Mean Differences δ_i:

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

$\bar{d}_i$, $S_{d_i}^2$ obtained from $\underset{\sim}{\bar{d}}$, S_d

For $n-p$, n large, $\frac{(n-1)p}{n-p} F_{p,n-p}(\alpha) \approx \chi_p^2(\alpha)$ Normality not needed.

Bonferroni Simultaneous $100(1-\alpha)\%$ CI's for p δ's

$$\bar{d}_i \pm t_{n-1}\left(\frac{\alpha}{2p}\right) \sqrt{\frac{S_{d_i}^2}{n}}$$

- Randomization of Treatments to units w/in matched pairs (coin toss, ...)

Computation from full sample $\underset{\sim}{\bar{X}}$, S

$$\underset{\sim}{\bar{X}} = \begin{bmatrix} \bar{X}_{11} \\ \vdots \\ \bar{X}_{1p} \\ \bar{X}_{21} \\ \vdots \\ \bar{X}_{2p} \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} p\times 1 \\ \left.\vphantom{\begin{matrix} a \\ b \\ c \end{matrix}}\right\} p\times 1 \end{matrix} \qquad S = \begin{array}{c c c c} & p & p & \\ p & \left[\begin{matrix} S_{11} \\ S_{21} \end{matrix}\right. & \left.\begin{matrix} S_{12} \\ S_{22} \end{matrix}\right] & \begin{matrix} p \\ p \end{matrix} \\ & p & p & \end{array}$$

$$C = \left[\, I_p \;\vdots\; -I_p \,\right] \qquad C\underset{\sim}{\bar{X}} = \begin{bmatrix} \bar{X}_{11} - \bar{X}_{21} \\ \vdots \\ \bar{X}_{1p} - \bar{X}_{2p} \end{bmatrix} = \bar{d}$$

$$S_d = CSC' \Rightarrow T^2 = n\underset{\sim}{\bar{X}}'C'(CSC')^{-1}C\underset{\sim}{\bar{X}}$$

Repeated Measures For Comparing Treatments

q Treatments, single outcome variable (RCBD e.g.)

$$\underset{\sim}{X}_j = \begin{bmatrix} X_{j1} \\ \vdots \\ X_{jq} \end{bmatrix} \qquad j = 1, \dots, n$$

Contrasts Matrices

(1)

$$\begin{bmatrix} \mu_1 - \mu_2 \\ \mu_1 - \mu_3 \\ \vdots \\ \mu_1 - \mu_q \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & -1 \end{bmatrix}}_{q-1 \text{ rows}} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_q \end{bmatrix} = C_1 \underset{\sim}{\mu}$$

(2)

$$\begin{bmatrix} \mu_2 - \mu_1 \\ \mu_3 - \mu_2 \\ \vdots \\ \mu_q - \mu_{q-1} \end{bmatrix} = \underbrace{\begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & & -1 & 1 \end{bmatrix}}_{q-1 \text{ rows}} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \end{bmatrix} = C_2 \underset{\sim}{\mu}$$

$$\mu_1 = \dots = \mu_q \Rightarrow C_1 \underset{\sim}{\mu} = C_2 \underset{\sim}{\mu} = \underset{\sim}{0}$$

$$H_0: C'\underset{\sim}{\mu} = \underset{\sim}{0} \qquad H_A: C'\underset{\sim}{\mu} \neq 0$$

$$\text{T.S. } T^2 = n(C\bar{x})'(CSC')^{-1}C\bar{x}$$

$$RR: T^2 \geq \frac{(n-1)(q-1)}{n-q+1} F_{q-1, n-q+1}(\alpha)$$

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

$$= \frac{1}{n-1} X'\left[I_n - \frac{1}{n} J_n\right] X \qquad X = \begin{bmatrix} X_{11} & \cdots & X_{1q} \\ \vdots & \ddots & \vdots \\ X_{n1} & \cdots & X_{nq} \end{bmatrix}$$

Confidence Region for $C\mu$

$$C\underset{\sim}{\mu} \text{ s.t. } n\left(C\underset{\sim}{\bar{X}}-C\underset{\sim}{\mu}\right)'(CSC')^{-1}\left(C\underset{\sim}{\bar{X}}-C\underset{\sim}{\mu}\right) \le \frac{(n-1)(q-1)}{n-q+1} F_{q-1,n-q+1}(\alpha)$$

Simultaneous $100(1-\alpha)\%$ CI for single contrasts $\underset{\sim}{c}'\underset{\sim}{\mu}$:

$$\underset{\sim}{c}'\underset{\sim}{\bar{X}} \pm \sqrt{\frac{(n-1)(q-1)}{n-q+1} F_{q-1,n-q+1}(\alpha)} \sqrt{\frac{c'Sc}{n}}$$

6.3 Comparing Means from 2 Populations

Sample 1: $\underset{\sim}{X}_{11}, \ldots, \underset{\sim}{X}_{1n_1}$ $\qquad \underset{\sim}{X}_{1j} = \begin{bmatrix} x_{1j1} \\ \vdots \\ x_{1jp} \end{bmatrix}$

Sample 2: $\underset{\sim}{X}_{21}, \ldots, \underset{\sim}{X}_{2n_2}$ $\qquad \underset{\sim}{X}_{2j} = \begin{bmatrix} x_{2j1} \\ \vdots \\ x_{2jp} \end{bmatrix}$

$$\underset{\sim}{\bar{X}}_1 = \frac{1}{n_1}\sum_j \underset{\sim}{X}_{1j} \qquad S_1 = \frac{1}{n_1-1}\sum_j \left(\underset{\sim}{X}_{1j}-\underset{\sim}{\bar{X}}_1\right)\left(\underset{\sim}{X}_{1j}-\underset{\sim}{\bar{X}}_1\right)'$$

$$\bar{X}_2 = \frac{1}{n_2}\sum_j \underset{\sim}{X}_{2j} \qquad S_2 = \frac{1}{n_2-1}\sum_j \left(\underset{\sim}{X}_{2j}-\underset{\sim}{\bar{X}}_2\right)\left(\underset{\sim}{X}_{2j}-\bar{X}\right)'$$

$\underset{\sim}{X}_{11}, \ldots, \underset{\sim}{X}_{1n_1}$ ≡ random sample from distn with $\underset{\sim}{\mu}_1$, Σ_1

$\underset{\sim}{X}_{21}, \ldots, \underset{\sim}{X}_{2n_2}$ ≡ " " " " " $\underset{\sim}{\mu}_2$, Σ_2

$$\{\underset{\sim}{X}_{1j}\} \perp \{\underset{\sim}{X}_{2j}\}$$

Added Assumptions when n_1, n_2 small

1. Distributions are multivariate normal
2. $\Sigma_1 = \Sigma_2$

Pooled variance covariance matrix:

$$S_{POOLED} = \frac{\sum_{j=1}^{n_1} (\underset{\sim}{X}_{1j} - \bar{\underset{\sim}{X}}_1)(\underset{\sim}{X}_{1j} - \bar{\underset{\sim}{X}}_1)' + \sum_{j=1}^{n_2} (\underset{\sim}{X}_{2j} - \bar{\underset{\sim}{X}}_2)(\underset{\sim}{X}_{2j} - \bar{\underset{\sim}{X}}_2)'}{n_1 + n_2 - 2}$$

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

Testing $H_0: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{\delta}_0$

$$E\{\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2\} = E\{\bar{\underset{\sim}{X}}_1\} - E\{\bar{\underset{\sim}{X}}_2\} = \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$$

independence: $Cov\{\bar{\underset{\sim}{X}}_1, \bar{\underset{\sim}{X}}_2\} = 0 \Rightarrow Cov\{\bar{X}_1 - \bar{X}_2\} = Cov\{\bar{\underset{\sim}{X}}_1\} + Cov\{\bar{\underset{\sim}{X}}_2\}$

$$= \frac{1}{n_1}\Sigma + \frac{1}{n_2}\Sigma = \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma$$

Use $\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{POOLED}$ to estimate $\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\Sigma$

T.S. $$T^2 = \left((\bar{X}_1 - \bar{X}_2) - \delta_0\right)' \left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{pooled}\right]^{-1} \left((\bar{X}_1 - \bar{X}_2) - \delta_0\right)$$

RR: $$T^2 \geq \frac{(n_1 + n_2 - 2)p}{(n_1 + n_2 - p - 1)} F_{p, n_1 + n_2 - p - 1}(\alpha)$$

All $\mu_1 - \mu_2 = \delta$ w/in ellipsoid centred at $\overline{\underline{X}_1 - \underline{X}_2}$

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

are in confidence region for $\underline{\mu}_1 - \underline{\mu}_2$.

Simultaneous CI^s

Define c^2 as above ↑

$(1-\alpha)100\%$ CI for $a'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$

$$a'(\bar{X}_1 - \bar{X}_2) \pm c\sqrt{\underset{\sim}{a}'\left(\frac{1}{n_1} + \frac{1}{n_2}\right) S_{pooled}\, \underset{\sim}{a}}$$

Note: Linear combination w/ largest population difference:

$$\hat{a}'(\underline{\bar{X}}_1 - \underline{\bar{X}}_2) \quad \text{w/} \quad \underline{\hat{a}} \propto S_{pooled}^{-1}(\underline{\bar{X}}_1 - \underline{\bar{X}}_2)$$

Bonferroni $100(1-\alpha)\%$ CI^s for p mean differences:

$$\mu_{1i} - \mu_{2i} : \; (\bar{X}_{1i} - \bar{X}_{2i}) \pm t_{n_1+n_2-2}\left(\frac{\alpha}{2p}\right)\sqrt{s_{ii,pooled}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Large-Sample Case when $\Sigma_1 \neq \Sigma_2$

Result 6.4 Let $n_1 - p$, $n_2 - p$ be large

Approx. $100(1-\alpha)\%$ confidence ellipsoid for $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$:

$\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$ s.t.
$$\left[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)\right]' \left[\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right]^{-1} \left[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)\right] \le \chi_p^2(\alpha)$$

Note: If $n_1 = n_2$, then $\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2 = S_{pooled}\left(\frac{1}{n} + \frac{1}{n}\right)$
(univariate case equivalence).

Approx. to Dist^n of T^2 for Normal populations when n_1, n_2 not large

$$T^2 = \left[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)\right]' \left[\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right]^{-1} \left[(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)\right]$$

Reject $H_0: \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2 = \underset{\sim}{0}$ if

$$T^2 > \frac{\nu p}{\nu - p + 1} F_{p, \nu - p + 1}(\alpha) = c^2$$

$$\nu = \frac{p + p^2}{\sum_{i=1}^{2} \frac{1}{n_i} \left\{ tr\left[\left(\frac{1}{n_i} S_i \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right)^{-1}\right)^2\right] + \left(tr\left[\frac{1}{n_i} S_i \left(\frac{1}{n_1} S_1 + \frac{1}{n_2} S_2\right)^{-1}\right]\right)^2 \right\}}$$

All $\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$ s.t. $T^2 \le c^2$ = confidence region for $\mu_1 - \mu_2$

6.4 1-Way Multivariate ANOVA

Univariate $X_{\ell j} = \mu_\ell + e_{\ell j} = \mu + \tau_\ell + e_{\ell j}$ $\quad \ell = 1, \dots, g$; $j = 1, \dots, n_\ell$

s.t. $\cancel{\sum_{\ell=1}^{g} n_\ell \mu_\ell}$ $\sum_{\ell=1}^{g} n_\ell \tau_\ell = 0$ $\quad e_{\ell j} \sim NID(0, \sigma^2)$, $N = \sum_{\ell=1}^{g} n_\ell$

observed data: $\underbrace{\sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(x_{\ell j} - \bar{x})^2}_{\text{Total SS}} = \underbrace{\sum_\ell \sum_j (\bar{x}_{\ell} - \bar{x})^2}_{\text{TRT SS}} + \underbrace{\sum_\ell \sum_j (x_{\ell j} - \bar{x}_\ell)^2}_{\text{Error SS}}$

ANOVA $\quad H_0: \mu_1 = \dots = \mu_g \quad (\tau_1 = \dots = \tau_g = 0)$

Source	df	SS	MS	F
Trts	$g-1$	SS_{TRTS}	$MS_{TRTS} = \frac{SS_{TRTS}}{g-1}$	$F = \frac{MS_{TRTS}}{MS_{ERR}}$
Error (Residual)	$N-g$	SS_{ERR}	$MS_{err} = \frac{SS_{err}}{N-g}$	
Total (Corr)	$N-1$	SS_{TOTAL}		

Reject H_0 if $F \geq F_{g-1, N-g}(\alpha)$

Multivariate $\underset{\sim}{X}_{\ell j} = \underset{\sim}{\mu} + \underset{\sim}{\tau}_\ell + \underset{\sim}{e}_{\ell j}$ $\quad j = 1, \dots, n_\ell$; $\ell = 1, \dots, g$

$e_{\ell j} \sim iid \ N_p(\underset{\sim}{0}, \Sigma)$

observed data: $\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}} = (\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}}) + (x_{\ell j} - \underset{\sim}{\bar{x}}_\ell)$

$\Rightarrow (\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}})(\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}})' = (\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})(\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})' + (\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})(\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell)'$
$\quad + (\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell)(\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})' + (\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell)(\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell)'$

Summing over all data

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

$$+ \sum_{\ell=1}^{g}(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})\underbrace{\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)}_{0} + \sum_{\ell=1}^{g}\underbrace{\left[\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)\right]}_{0}(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})' + \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)'$$

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

Total Corrected SSCP matrix "B+W"

Between Trt SSCP matrix "B"

within Trt SSCP matrix (Residual/Error) "W"

Within Trt SS: $W = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(x_{\ell j}-\bar{x}_\ell)(x_{\ell j}-\bar{x})' = \sum_{\ell=1}^{g}(n_\ell-1)\, S_\ell$

$S_\ell \equiv$ Sample Covariance matrix for ℓ_{th} Trt/Sample.

MANOVA Table $H_0: \underset{\sim}{\tau}_1 = \dots = \underset{\sim}{\tau}_g = \underset{\sim}{0}$

Source	df	SSCP
Trt	$g-1$	$B = \sum_{\ell=1}^{g} n_\ell(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})(\bar{\underset{\sim}{x}}_\ell-\bar{\underset{\sim}{x}})'$
Residual (Error)	$N-g$	$W = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}}_\ell)'$
Total (Corr)	$N-1$	$B+W = \sum_{\ell=1}^{g}\sum_{j=1}^{n_\ell}(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})(\underset{\sim}{x}_{\ell j}-\bar{\underset{\sim}{x}})'$

Test based on Wilk's Lambda

$$\Lambda^* = \frac{|W|}{|B+W|} \qquad \text{Reject } H_0: \underset{\sim}{\tau}_1 = \cdots = \underset{\sim}{\tau}_g = \underset{\sim}{0} \text{ if } \Lambda^* \le c^*$$

Distribution of Wilk's Lambda: $\Lambda^* = \dfrac{|W|}{|W+B|}$

# of Variables (p)	# of groups (g)	Sampling Dist^n for MVN Data
$p=1$	$g \ge 2$	$\left(\frac{N-g}{g-1}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{g-1, N-g}$
$p=2$	$g \ge 2$	$\left(\frac{N-g-1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2(g-1), 2(N-g-1)}$
$p \ge 1$	$g=2$	$\left(\frac{N-p-1}{p}\right)\left(\frac{1-\Lambda^*}{\Lambda^*}\right) \sim F_{p, N-p-1}$
$p \ge 1$	$g=3$	$\left(\frac{N-p-2}{p}\right)\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right) \sim F_{2p, 2(N-p-2)}$

Large-Sample results w/ Ho True and N large:

$$-\left(N-1-\frac{(p+g)}{2}\right)\ln\Lambda^* = -\left(N-1-\frac{(p+g)}{2}\right)\ln\left(\frac{|W|}{|B+W|}\right) \;\dot\sim\; \chi^2_{p(g-1)}$$

$$\Rightarrow \text{Reject } H_0 \text{ if } -\left(N-1-\left(\frac{p+g}{2}\right)\right)\ln\left(\frac{|W|}{|B+W|}\right) \ge \chi^2_{p(g-1)}(\alpha)$$

(Bartlett's Correction)

6.5 Simultaneous CI's for Trt Effects

Univariate CI's: $(\bar{X}_{k} - \bar{X}_{l}) \pm t_{N-g}(\alpha)\sqrt{MS_{Err}\left(\frac{1}{n_k}+\frac{1}{n_l}\right)}$

Multivariate CI's: $(\bar{X}_{ki} - \bar{X}_{li}) \pm t_{N-g}\left(\frac{\alpha}{2m}\right)\sqrt{\frac{w_{ii}}{N-g}\left(\frac{1}{n_k}+\frac{1}{n_l}\right)}$

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

6.6 Testing for Equality of Covariance Matrices

$H_0: \Sigma_1 = \Sigma_2 = \cdots = \Sigma_g = \Sigma$

LR Test: $\Lambda = \prod_{l=1}^{g}\left(\frac{|S_l|}{|S_{pooled}|}\right)^{\frac{n_l-1}{2}}$

$$S_{pooled} = \frac{1}{N-g}\left[(n_1-1)S_1 + \cdots + (n_g-1)S_g\right]$$

$$M = -2\ln\Lambda = -2\left[\sum_{l=1}^{g}\left(\frac{n_l-1}{2}\right)\left[\ln|S_l| - \ln|S_{pooled}|\right]\right]$$

$$= (N-g)\ln|S_{pooled}| - \sum_{l=1}^{g}(n_l-1)\ln|S_l|$$

Box's Test (works well when $n_\ell > 20 \; \forall \ell$, $p, g \le 5$)

$$U = \left[\sum_\ell \frac{1}{n_\ell - 1} - \frac{1}{N-g} \right] \left[\frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \right]$$

$$C = (1-U)M = (1-U)\left[(N-g)\ln|S_{pooled}| - \sum_{\ell=1}^{g} (n_\ell - 1)\ln|S_\ell| \right]$$

$$C \sim \chi^2_\nu \qquad \nu = \tfrac{1}{2} g\, p(p+1) - \tfrac{1}{2} p(p+1) = \tfrac{1}{2} p(p+1)(g-1)$$

Reject H_0 if $C \ge \chi^2_{\frac{p(p+1)(g-1)}{2}}(\alpha)$

6.7 2-Way ANOVA (Balanced w/ n reps/trt)

Univariate $\quad X_{\ell k r} = \underset{\sim}{\mu} + \tau_\ell + \beta_k + \gamma_{\ell k} + e_{\ell k r}$ $\qquad \ell = 1, \ldots, g$; $k = 1, \ldots, b$; $r = 1, \ldots, n$

$$\sum_\ell \tau_\ell = \sum_k \beta_k = \sum_\ell \gamma_{\ell k} + \sum_k \gamma_{\ell k} = 0$$

observed data

$$(x_{\ell k r} - \bar{x}) = (\bar{x}_{\ell \cdot} - \bar{x}) + (\bar{x}_{\cdot k} - \bar{x}) + (\bar{x}_{\ell k} - \bar{x}_{\ell \cdot} - \bar{x}_{\cdot k} + \bar{x}) + (x_{\ell k r} - \bar{x}_{\ell k})$$

$$\Rightarrow \sum_\ell \sum_k \sum_r (x_{\ell k r} - \bar{x})^2 = nb \sum_\ell (\bar{x}_{\ell \cdot} - \bar{x})^2 + gn \sum_k (\bar{x}_{\cdot k} - \bar{x})^2$$

$$+ n \sum_\ell \sum_k (\bar{x}_{\ell k} - \bar{x}_{\ell \cdot} - \bar{x}_{\cdot k} + \bar{x})^2 + \sum_\ell \sum_k \sum_r (x_{\ell k r} - \bar{x}_{\ell k})^2$$

$\Rightarrow SS_{TOTAL} = SS_{TRT1} + SS_{TRT2} + SS_{INT} + SS_{Err}$

Multivariate: p variables

SS:

$$\sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{lkr}-\bar{x})(x_{lkr}-\bar{x})' = bn\sum_{l=1}^{g}(\bar{x}_{l.}-\bar{x})(\bar{x}_{l.}-\bar{x})'$$

$$+ gn\sum_{k=1}^{b}(\bar{x}_{.k}-\bar{x})(\bar{x}_{.k}-\bar{x})' + n\sum_{l}\sum_{k}(\bar{x}_{lk}-\bar{x}_{l.}-\bar{x}_{.k}+\bar{x})(\bar{x}_{lk}-\bar{x}_{l.}-\bar{x}_{.k}+\bar{x})'$$

$$+ \sum_{l=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}(x_{lkr}-\bar{x}_{lk})(x_{lkr}-\bar{x}_{lk})'$$

df: $gbn-1 = (g-1)+(b-1)+(g-1)(b-1)+gb(n-1)$

Manova

Source	df	SSCP
Factor 1	$g-1$	$SSCP_{Fac1} = bn\sum_{l}(\bar{x}_{l.}-\bar{x})(\bar{x}_{l.}-\bar{x})'$
Factor 2	$b-1$	$SSCP_{Fac2} = gn\sum_{k}(\bar{x}_{.k}-\bar{x})(\bar{x}_{.k}-\bar{x})'$
Interaction	$(g-1)(b-1)$	$SSCP_{int} = n\sum_{l}\sum_{k}(\bar{x}_{lk}-\bar{x}_{l.}-\bar{x}_{.k}+\bar{x})(\bar{x}_{lk}-\bar{x}_{l.}-\bar{x}_{.k}+\bar{x})'$
Residual (Error)	$gb(n-1)$	$SSCP_{res} = \sum_{l}\sum_{k}\sum_{r}(x_{lkr}-\bar{x}_{lk})(x_{lkr}-\bar{x}_{lk})'$
Total (Corrected)	$gbn-1$	$SSCP_{tot} = \sum_{l}\sum_{k}\sum_{r}(x_{lkr}-\bar{x})(x_{lkr}-\bar{x})'$

Factor 1

$$H_0: \underset{\sim}{\tau}_1 = \cdots = \underset{\sim}{\tau}_g = \underset{\sim}{0}$$

$$\Lambda^*_{Fac1} = \frac{|SSCP_{Res}|}{|SSCP_{Fac1} + SSCP_{Res}|}$$

Reject H_0 if $-\left[gb(n-1) - \frac{p+1-(g-1)}{2}\right] \ln \Lambda^*_{Fac1}$

$$> \chi^2_{(g-1)p}(\alpha)$$

Factor 2 $\quad H_0: \underset{\sim}{\beta}_1 = \cdots = \underset{\sim}{\beta}_b = \underset{\sim}{0}$

$$\Lambda^*_{Fac2} = \frac{|SSCP_{Res}|}{|SSCP_{Fac2} + SSCP_{Res}|}$$

Reject H_0 if $-\left[gb(n-1) - \frac{p+1-(b-1)}{2}\right] \ln \Lambda^*_{Fac2} \geq \chi^2_{(b-1)p}(\alpha)$

Interaction $\quad H_0: \underset{\sim}{\gamma}_1 = \cdots \underset{\sim}{\gamma}_{gb} = \underset{\sim}{0}$

$$\Lambda^*_{INT} = \frac{|SSCP_{Res}|}{|SSCP_{INT} + SSCP_{Res}|}$$

Reject H_0 if $-\left[gb(n-1) - \frac{p+1-(g-1)(b-1)}{2}\right] \ln \Lambda^*_{INT} \geq \chi^2_{(g-1)(b-1)p}(\alpha)$

$(1-\alpha)100\%$ CI's (Simultaneous)

$i = 1, \dots, p$

Factor 1: $\tau_{li} - \tau_{mi}$: $(\bar{x}_{l.i} - \bar{x}_{m.i}) \pm t_\nu\left(\frac{\alpha}{2p\frac{g(g-1)}{2}}\right)\sqrt{\frac{E_{ii}}{\nu}\frac{2}{bn}}$

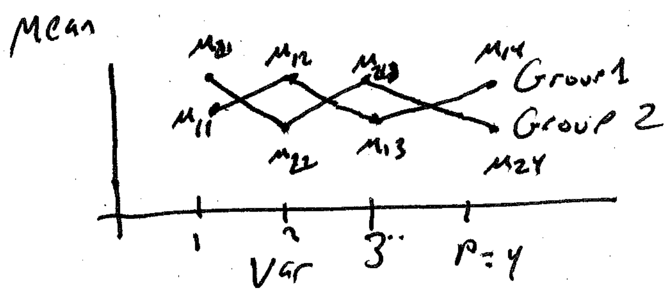
$\nu = gb(n-1)$ E_{ii} from $E = SSCP_{res}$:

$i = 1, \dots, p$

Factor 2: $\beta_{ki} - \beta_{qi}$: $(\bar{x}_{.ki} - \bar{x}_{.qi}) \pm t_\nu\left(\frac{\alpha}{2p\frac{b(b-1)}{2}}\right)\sqrt{\frac{E_{ii}}{\nu}\frac{2}{gn}}$

ν and E_{ii} from above.

6.8 Profile Analysis



1. Test for Parallel profiles - 2 populations

$H_{01}: \mu_{1i} - \mu_{1,i-1} = \mu_{2i} - \mu_{2,i-1}$ $i = 2, \dots, p$

$$C_1 \underset{\sim}{\mu}_1 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \underset{\sim}{\mu}_1 = C_1 \underset{\sim}{\mu}_2$$

T.S. $T_1^2 = \left[C_1(\bar{x}_1 - \bar{x}_2)\right]'\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right) C_1 S_{pooled} C_1'\right]^{-1} C_1(\bar{x}_1 - \bar{x}_2)$

RR: $T_1^2 \geq \frac{(n_1+n_2-1)(p-1)}{n_1+n_2-p} F_{p-1, n_1+n_2-p}(\alpha)$

2. Assuming parallel profiles, are profiles coincident

$H_{02}: \mu_{1i} = \mu_{2i} \quad i=1,\dots,p \qquad H_{02}: \mathbf{1}'\underset{\sim}{\mu}_1 = \underset{\sim}{\mathbf{1}}'\underset{\sim}{\mu}_2 \Rightarrow \underset{\sim}{\mathbf{1}}'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2) = 0$

$$\text{T.S. } T_2^2 = \left(\mathbf{1}'(\bar{X}_1 - \bar{X}_2)\right)\left[\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\underset{\sim}{\mathbf{1}}' S_{pooled}\underset{\sim}{\mathbf{1}}\right]^{-1}\left(\mathbf{1}'(\underset{\sim}{\bar{X}}_1 - \underset{\sim}{\bar{X}}_2)\right)$$

$$= \left(\frac{\mathbf{1}'(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)\underset{\sim}{\mathbf{1}}' S_{pooled}\underset{\sim}{\mathbf{1}}}}\right)^2 > t^2_{n_1+n_2-2}(\alpha) = F_{1,n_1+n_2-2}(\alpha)$$

3. Assuming Coincident Profiles, Test for Level Profiles

$H_{03}: C_1 \underset{\sim}{\mu} = 0 \qquad \underset{\sim}{\mu} = (\underset{\sim}{\mu}_1 + \underset{\sim}{\mu}_2)\frac{1}{2} \qquad C_1 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$

$$\underset{\sim}{\bar{X}} = \frac{1}{n_1+n_2}\left(\sum_{j=1}^{n_1} \underset{\sim}{X}_{1j} + \sum_{j=1}^{n_2} \underset{\sim}{X}_{2j}\right) = \frac{n_1 \underset{\sim}{\bar{X}}_1 + n_2 \underset{\sim}{\bar{X}}_2}{n_1+n_2}$$

$$\text{TS. } T^2 = (n_1+n_2)\left(C\bar{X}\right)'\left[CSC'\right]^{-1} C\bar{X}$$

$$RR: T^2 \geq \frac{(n_1+n_2-1)(p-1)}{n_1+n_2-p+1} F_{p-1,n_1+n_2-p+1}(\alpha)$$

6.9 Repeated Measures Designs and Growth Curves

1) Measurements on subjects are under different trt conditions (e.g. RBD) w/ measurements w/in subjects being correlated
2) Measurements on each subject are under same condition @ multiple time points. Can be one or more trts (Between Subj trts). Growth Curve

Potthoff and Roy Model for growth Curves

1) Single Treatment - each subject observed @ time points $t_1, \dots, t_p$ (Quadratic)

$$E[X] = E\begin{Bmatrix} X_1 \\ \vdots \\ X_p \end{Bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 t_1 + \beta_2 t_1^2 \\ \vdots \\ \beta_0 + \beta_1 t_p + \beta_2 t_p^2 \end{bmatrix}$$

2) g Trt Groups

$\underset{\sim}{X}_{lj}$ $\quad l = 1, \dots, g$; $\quad j = 1, \dots, n_l$

$$E\{\underset{\sim}{X}_{lj}\} = \begin{bmatrix} \beta_{l0} + \beta_{l1} t_1 + \beta_{l2} t_1^2 \\ \vdots \\ \beta_{l0} + \beta_{l1} t_p + \beta_{l2} t_p^2 \end{bmatrix} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_p & t_p^2 \end{bmatrix} \begin{bmatrix} \beta_{l0} \\ \beta_{l1} \\ \beta_{l2} \end{bmatrix}$$

$$= B \underset{\sim}{\beta_l}$$

q-order Polynomial

$$B = \begin{bmatrix} 1 & t_1 & \cdots & t_1^q \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_p & \cdots & t_p^q \end{bmatrix} \qquad \beta_l = \begin{bmatrix} \beta_{l0} \\ \beta_{l1} \\ \vdots \\ \beta_{lq} \end{bmatrix}$$

ML / EGLS Esitimator for β_ℓ:

$$\hat{\beta}_\ell = \left(B' S_{pooled}^{-1} B \right)^{-1} B' S_{pooled}^{-1} \bar{X}_\ell \quad \ell = 1, \dots, g$$

$$S_{pooled} = \frac{1}{N-g} \left[(n_1 - 1) S_1 + \dots + (n_g - 1) S_g \right] = \frac{1}{N-g} W$$

$S_i \equiv$ Variance - covariance matrix of the p observations w/in each trt group.

$$\hat{Cov}\{\hat{\beta}_\ell\} = \frac{k}{n_\ell} \left(B' S_{pooled}^{-1} B \right)^{-1} \quad \ell = 1, \dots, g$$

$$K = \frac{(N-g)(N-g-1)}{(N-g-p+q)(N-g-p+q+1)}$$

$$Cov\{\hat{\beta}_\ell, \hat{\beta}_h\} = 0 \quad \ell \neq h$$

H_0: q-order polynomial is adequate

$W \equiv$ w/in groups $SSCP_{Err}$ w/ $N-g$ df

$$W_q = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} \left(X_{\ell j} - B\hat{\beta}_\ell \right)\left(X_{\ell j} - B\hat{\beta}_\ell \right)' \quad \text{w/ } n_g - g + p - q - 1 \text{ df}$$

$$\Lambda^* = \frac{|W|}{|W_q|}$$

TS: $-\left(N - \frac{1}{2}(p-q+g)\right) \ln \Lambda^*$ RR: $TS \geq \chi^2_{(p-q-1)g}(\alpha)$