2nd Midterm: Nov 6

TLC study

tlc 32 ← sas program on Sakai

- Analysis of 2 variables: Y_1 = week 0, $Y_2 = \frac{\text{week } 1 + 4 + 6}{3}$ $\quad d = Y_2 - Y_2$ ← estimate of mean change over time

- if you regress Y_2 on Y_1, you get a slope (since both are normal → linear)

$$\underbrace{Y_2 | Y_2}_{\text{This is actually } Y_1 | d} \sim N\left(\mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(Y_1 - \mu_1)\right)$$

↖ slope! $= \frac{17.86}{25.2} = \beta$ ← In Active Group

$$E[Y_2 | Y_1] = \underbrace{(\mu_2 - \beta\mu_1)}_{\text{intercept}} + \beta Y_i$$

- Regression of d on baseline Y_1

Group A: $\frac{-7.35}{25.21} = -0.2914$ se: (0.1675), Placebo Group $\frac{-2.43}{25.24} = -0.0964$ (s.e. 0.0765)

Compare Group A and Placebo: $\frac{-0.2914 + 0.0964}{\sqrt{(0.0765)^2 + (0.1675)^2}}$ ← large sample, compare to normal

↖ no covariance b/c Active ⊥ Placebo

- Regression of d on Group

Estimated mean of Group A: $-2.1467 + -7.7940$

" " " " P: -2.1467

- Can do an exact t test if assume normality, etc.
- Or large sample: Wald

- standard error is valid → independent, only 1 obs (?)
- this is a valid comparison whether they were randomized or not

Estimate of sigma: "scale" on SAS output → $\hat{\sigma} = 4.7017$

- Regression of d on group, adj for baseline
  - Implicitly assuming that interaction term is zero
  - So, interpretation of coefficient for group A: expected difference holding baseline constant

$\hat{\sigma} = 4.6019$ ← this is what we expect: conditional var < marginal var, it's good that the estimates are in keeping with this.

$$E[\underbrace{Y_{i2}-Y_{i1}}_{D_i} \mid Y_{i1}] = \beta_1 + \beta_2 \cdot \overset{\overset{P/A}{0/1}}{x_i} + \beta_3 Y_{i1}$$

$\hat{\beta}_3 = -0.1938$

$$= E[Y_{i2} \mid Y_{i1}] - \underbrace{Y_{i1}}$$

b/c $E[Y_{i1} \mid Y_{i1}] = Y_{i1}$

looking at estimates from SAS, we can see that these are the same model!

$$= E[Y_{i2} \mid Y_{i1}] = \beta_1 + \beta_2 x_i + \underbrace{(1+\beta_3)} Y_{i1}$$

$\widehat{(1+\beta_3)} = 0.8062$

$= 1 + -0.1938$

Interpretation: Group A has a lower expected value between 6-10 microLiter

....

- Model #4

model d = group $\underbrace{\text{y1 (group)}}$

= y1 nested in group → slope of y1 varies by group

→ equivalent to model d = $\underbrace{\text{group | y1}}$

key word

identical to group y1 group*y1

→ include interactions

y1 (group)
y2 (group) → same slopes calculated earlier

y1 → slope of placebo → same -0.0964

y1*group → " " active → -0.0964 + -0.1951 ≈ -0.29

same

You can test the assumption of equal slope!

This is for 2 var. Independent. Could have done in 663

When you go to 3 var?

- Everything applies to each column separately
- Assume dist is multivariate normal (In Ch 5, all outcomes assumed normal)
- We are mostly modeling the means. The mean structure, how you parametrize, model is exactly the same as ANOVA.
- The difference is: for estimation we can't just multiply

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{pmatrix} \sim N_4\left(\begin{bmatrix} \mu_{i1} \\ \mu_{i2} \\ \mu_{i3} \\ \mu_{i4} \end{bmatrix}, \Sigma_{4\times4} \right)$$

$i = 1, \dots, 100$

parameters $\xrightarrow{\text{determine}} \Sigma_{4\times4}$

parameters $\xrightarrow{\text{determine}} \underset{\sim}{\mu}_i \rightarrow \underset{\sim}{\beta}$

$\underset{\sim}{\mu}_{400\times1}$

$$\underset{\sim}{\mu} = [X]_{400\times p} (\underset{\sim}{\beta})_{p\times1}$$

Big Covariance Matrix:

$$\begin{bmatrix} [\Sigma_{4\times4}] & 0_{4\times4} & 0_{4\times4} \\ & [\Sigma] & \\ & & \ddots \\ & & & [\Sigma] \end{bmatrix}_{400\times400}$$

- We noticed var was increasing over time
- " " var was different between A and P

$L(\underset{\sim}{\beta}, \Sigma; \underset{\sim}{Y}) \rightarrow$ at the end get $\hat{\beta}, \hat{\Sigma}, \widehat{cov}$

- Cov matrix is the inverse of the expected information matrix

$\Sigma \rightarrow \underset{\sim}{\hat{\theta}}_{10\times1}$ b/c $\begin{bmatrix} \sigma_{11} & \sigma_{12} & \\ & \sigma_{22} & \\ & & \sigma_{33} \end{bmatrix} \leftarrow 10\ \sigma$'s

- Property of multivariate Normals: Expected (Fisher) Information matrix

$\hat{\beta}_{p\times1}$, $\hat{\theta}_{q\times1}$

$$\begin{array}{c|c} & \underset{\sim}{\beta} \quad \underset{\sim}{\theta} \\ \hline \begin{array}{c} \underset{\sim}{\beta} \\ \underset{\sim}{\theta} \end{array} & \begin{bmatrix} & 0 \\ 0 & \end{bmatrix} \end{array}$$

The asymptotic covariance between $\hat{\beta}$ parameters and $\hat{\theta}$ parameters is zero (So they are independent)

$\hat{\beta} \perp \hat{\sigma}^2$

$\hat{\beta} \perp (\underset{\sim}{Y} - \underset{\sim}{\hat{\mu}})$

Normally: deviance $\perp \hat{\beta}$ (except for Bernoulli as we talked about)

In the normal: $\bar{Y} \perp S^2 \rightarrow$ a special case of $\hat{\beta} \perp \hat{\sigma}^2$

Standard Maximum Likelihood

Next time: $\boxed{\text{REML}}$ $\rightarrow$ not explained well in book.

- Remember, our main focus is the mean structure.

$\boxed{\text{profile analysis}}$ $\leftarrow$ jargon, especially in social sciences

We have 2×4 table as far as mean structure is concerned

	1	2	3	4
P	μ_{11}	μ_{12}	μ_{13}	μ_{14}
A	μ_{21}	μ_{22}		

- What are the treatment effects? The 3 interaction contrasts:

$$\left.\begin{array}{l}\mu_{12} - \mu_{11} \\ -(\mu_{22} - \mu_{21})\end{array}\right] \rightarrow \gamma_2$$

(between 1 and 2

Repeat: make these contrasts between $\underbrace{\text{1 and 3}}_{\gamma_3}$, and $\underbrace{\text{1 and 4}}_{\gamma_4}$

There are other ways to make these contrasts, this is just one way.

$$\underset{\sim}{\gamma} = \begin{pmatrix} \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}$$

Sakai $\rightarrow$ tlc34.sas

5.3 - 5.5 in book recreated here

General 2 Way ANOVA model

"model" statement $\rightarrow$ defines the mean structure

"repeated" " $\rightarrow$ correlation. Important here: type=un

$\hookrightarrow \Sigma$ is unstructured

time $\rightarrow$ variable used to line up observations time: have to line things up properly

$$\begin{array}{c} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{array} \begin{bmatrix} G_{11} & & & \\ & & & \\ & & & \\ & & & \end{bmatrix}$$

(for example, if 1 subject is missing the 3rd visit, SAS will move the 4th visit observation up, and it will not line up will the right covar)

	1	2	3	4
P	26.27	26.27	26.27 −2.2	26.27 −2.63
A	26.27 +.268	26.27 +.268	26.27 +.268 −2.2 −8.8	26.27 +.268 − 2.63 −3.15

26.27 = int

.268 = Group A

−3.15 = group A * 4

Time 4 (6) A − P = 0.268 − 3.15

Time 0 A − P = 0.268

Meaning of interaction: difference between the groups varies between the weeks

Type 3 Tests of Fixed Effects ∈ SAS Table
- ↳ Num d.f., denominator d.f.

actual distribution is not actually an F distribution
- ↳ use the Chi-Sq quantities (column) b/c this is based on large sample theory [F value depends on normality + other things, so not valid]

Look at group * time to assess treatment effect.

Compare with Hotelling's

Next time will continue w/ this program.