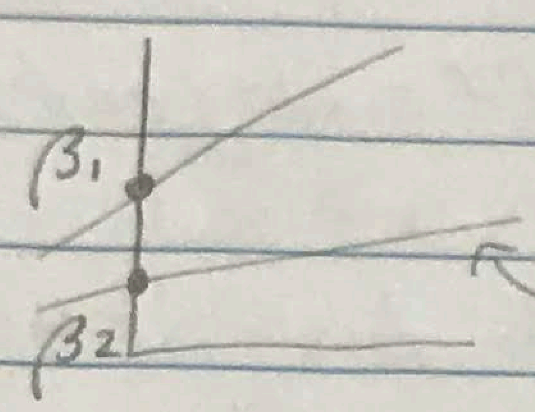
11/25/19

- 1 more hw after this one
- final is the whole class from beginning to end

8.1.8 → int and slopes take into account treatment?
random effects intercept and time

for a given subject



population mean
marginal means for the two groups

every subject has a random deviation from the intercept and the slope.

$\begin{pmatrix} b_{i1} \\ b_{i2} \end{pmatrix}$ – same distribution for both groups

1st group intercept $\quad \beta_1 + \hat{b}_{i1}$
2nd " " $\quad \beta_2 + \hat{b}_{i1}$

There is a difference between predicting
b_{i1}, $\beta_1 + b_{i1}$, and $\beta_2 + b_{i1}$

$\beta_1 I(group=1) + \beta_2 I(group=2)$

linear combination of fixed effects from the population that the subject is drawn

d) 95% CI incorporate intercept

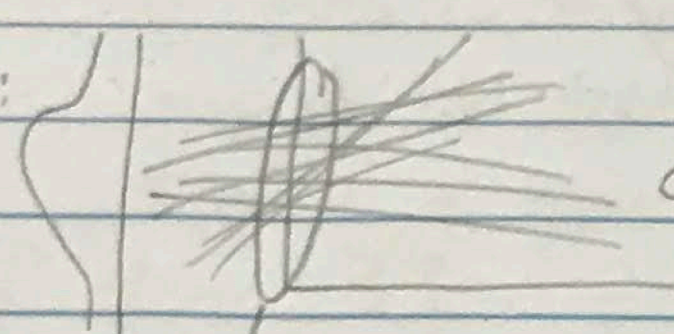
8.1.3 e) variance

- do it for each treatment group → d) and e)

don't incorporate s.e. of fixed effect

$\beta_1 \pm 1.96\sqrt{g_{11}}$

One group:



← diff subjects

Intercepts: a distribution normally distributed

95%

2.5% to 97.5% Quantiles for intercept

$\hat{\beta}_1 \pm 1.96\sqrt{\hat{g}_{11}}$

$\lambda = \beta_1 - 1.96\sqrt{g_{11}}$ ← lower

← this is a parameter → no hats

Final Exam Question → 95% CI ?

Point estimate: $\hat{\lambda} = \hat{\beta}_1 - 1.96\sqrt{\hat{g}_{11}}$

need s.e. of $\hat{\lambda}$

asymptotic → var $(\hat{\lambda})$ → $\hat{\beta}_1$ and $\hat{g}_{11}$ not correlated! great

$a.var(\hat{\lambda}) = \underbrace{a.var(\hat{\beta}_1)}_{\text{get this from output}} + 1.96^2\, a.var(\sqrt{\hat{g}_{11}})$

we know $a.var(\hat{g}_{11})$
use Delta method!

$h(t) = \sqrt{t}$
$h'(t) = 1/2\sqrt{t}$

$a.var(\sqrt{\hat{g}_{11}}) = a.var(\hat{g}_{11}) \cdot \frac{1}{4 \cdot \hat{g}_{11}}$

$a.var(\hat{\lambda}) = a.var(\hat{\beta}_1) + 1.96^2 \left(a.var(\hat{g}_{11}) \cdot \frac{1}{4 \hat{g}_{11}} \right)$

Delta Method
$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \gamma^2)$
$\sqrt{n}\{h(\hat{\theta}) - h(\theta)\} \xrightarrow{d} N(0, \gamma^2 \{h'(\theta)\}^2)$

$h(\theta)$ has to be continuous in neighborhood of θ, $h'(\theta)$ has to be continuous too b/c it needs to exist if you approach from left or right

8.1.6 → asking for formal hyp test?
interpret the estimate, give CI, come up w/summary if difference between 2 groups or not → use CI or hyp test

8.1.7 → $var(Y_{ij})$
model → w/in subject var is constant over time
↳ this is the conditional variance
so incorporate R_i and g
$Var(Y_{ij}) = Var(Y_{ij} | b_i) +$ condiional exp.

Final Question: - computed empirical BLUP $\rightarrow \hat{b}_{i1}$

compute sample variance of $\hat{b}_{i1}$

in treatment group: sample variance of $\hat{b}_{i1}$.
Should it be close to $\hat{g}_{11}$?

$g_{11} = var(b_{i1})$ ← we know this
subjects are independent of each other

True BLUP definition:
$E[b_i \mid Y_i] = \hat{b}_i$
Y_i ← observed / observable data

sample variance → approximation of population var

expect this to be smaller than g_{11} → $var(\hat{b}_{i1}) = var(E[b_{i1} \mid Y_i])$

$g_{11} = var(E[b_{i1} \mid Y_i]) + E[var(b_{i1} \mid Y_i)]$

Lecture 11/25

fat002.sas

- mixed models → to estimate variance components
  - to do subject-specific predictions

The 3 models depend on the random parts

last model, G matrix is 3×3 : $G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ & g_{22} & g_{23} \\ & & g_{33} \end{bmatrix}$

degenerate R.V. -> a constant

$H_0: g_{33} = 0$

-> means there is no quadratic variable in random effects

LRT -> use REML or full ML

# Random Effects	LRT	
$q+1=3$	6112.3	model (w/3 random effects w/ quadratic)
$q=2$	6124.4	<- H_0 model w/o quadratic
	12.1	

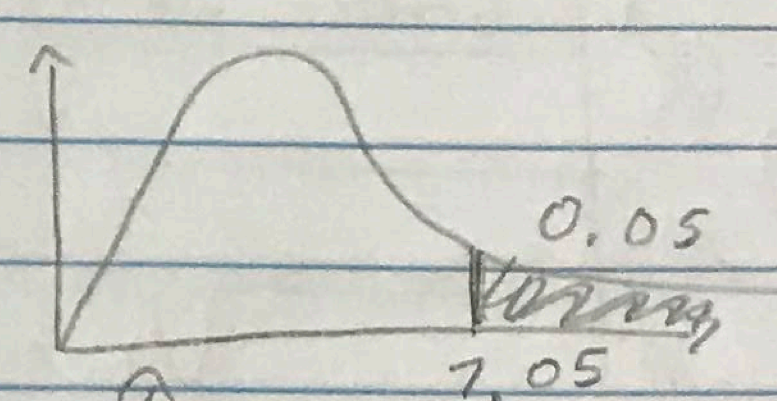
-> test statistic

What is the reference distribution under H_0?

50-50 mixture of χ^2_q and χ^2_{q+1}

cdf is an average of the dist. functions of χ^2_q and χ^2_{q+1}. Get pvalues from both + get avg.

- Critical point to reject at 0.05 -> table in back of book.



the mixture distribution

	Critical point
0.05	7.05
0.005	11.97
0.0025	13.43

Don't write $\frac{\chi^2_q + \chi^2_{q+1}}{2}$ <- very different!! It means you avg the variables themselves

$Y_{ij} = \beta_1 + \beta_2 \cdot time_{ij} + b_{1i} + b_{2i} \cdot time_{ij}$

Subject specific slope: $\beta_2 + b_{2i}$

$P(\beta_2 + b_{2i} < 0)$ ← reasonable question to ask since b_{2i} is random

$$P\Big(\underbrace{\frac{b_{2i}}{\sqrt{g_{22}}}}_{\sim N(0,1)} < \frac{-\beta_2}{\sqrt{g_{22}}}\Big)$$

$var(b_{2i}) = g_{22}$

$$\Phi\left(\frac{-\beta_2}{\sqrt{g_{22}}}\right)$$

← probnorm in sas

← pnorm in R

w/ data, we plug in estimates

But what if we want CI? Need s.e. Rely on 2 things

asymptotic
1) a variance: $\hat{\beta}_2$, $\hat{g}_{22}$ is 0
2) Delta method

$$a.\hat{var}\left(\frac{-\hat{\beta}_2}{\sqrt{\hat{g}_{22}}}\right) = \frac{V_1}{\hat{g}_{22}} + \frac{\hat{\beta}_2}{4\hat{g}_{22}^3} \cdot V_2$$

on avg ppl are gaining weight, but some are losing weight

$V_1 = a.\hat{var}(\hat{\beta}_2)$

$V_2 = a.\hat{var}(\hat{g}_{22})$

95% CI for $\Phi\left(\frac{-\beta_2}{\sqrt{g_{22}}}\right)$:

$$\Phi\left(\text{est.} \pm 1.96 \sqrt{a.\hat{var}(\,)}\right)$$

$\uparrow$

$\frac{-\hat{\beta}_2}{\sqrt{\hat{g}_{22}}}$

tlc011.sas

proc mixed

can output:

- fitted marginal mean $X\hat{\beta}$ $\boxed{\text{OUTPM}}$ ← in model

($E[\underset{\sim}{Y}] = X\underset{\sim}{\beta}$)

- subject specific mean: $X_i\hat{\underset{\sim}{\beta}} + Z_i\hat{\underset{\sim}{b}}_i$ $\boxed{\text{OUTP}}$

↳ prediction since random

- $\hat{\underset{\sim}{b}}_i$ → "random" statement → "solution"

1.) fixed effects are the same as before

"model" → "s" for $\hat{\beta}$ (?)

tlc011.pdf

"Solution for Random Effects" → b_i

Std Err Pred: est. of root mean sq error

25.9325 → est. of g_{11}

var g_{11} → have to say "covtest" in "random" statement after the slash

↳ all the variance components

$\hat{g}_{11} = 25.9$

$\boxed{\text{Sample Final Exam Q}}$ have subject w/ 4 observations missing → no data, they will get a prediction for that β subject and Std Error for prediction

?

prediction: the marginal mean = 0

Std. error of prediction: $\text{var}(\hat{b}_i - b_i) = \hat{g}_{11}$ ← ?

MSE since unbiased predictor, so Std error: $\sqrt{25.9}$

$var(b_i \mid \underset{\sim}{Y_i}) \approx 4$

EBLUP dataset → just b_i (not $Z_i b_i$)

MEANS procedure

sum of predicted values: 0
(orthogonal to intercept)

sample S.D. of $\hat{b_i}$ = 4.6989 = $\hat{S.D}(\hat{b_i})$

$\hat{var}(\hat{b_i}) = (4.7)^2 = 24$

$\hat{g}_{11} = 25.9$ ↳ marginal variance of $\hat{b_i}$

conditional variance

Std Error Pred = 2

they are different

$\hat{var}(b_i \mid \underset{\sim}{Y_i}) = 2^2 = 4 < 4.7 = \hat{S.D.}(\hat{b_i})$

$var(E(b_i \mid Y_i))$

$\hat{var}(\hat{E}(b_i \mid Y_i)) = (4.7)^2$

?

var of mean < marginal var

OUTPM $X*Beta$

marginal means (subj. 2 and 3 in same group → same mean)

Std Error Pred: est. std error of fitted value

OUTP $X*Beta + Z*b_i$

subject 1: 28.17 close to 26.4, the fitted mean

4 obs per subject: prediction for each time point per subject.

(?) b is random there is no standard error
↳ have to do conditional expectation (?)

$var(E[b_i | Y_i])$ → can't compute this don't have parameters

$\hat{var}(\hat{E}[b_i | Y_i])$ → estimate since don't

Std Error prediction : $\hat{var}(b_i | Y_i)$

We assume G is the same for the 2 groups in this model.

We could have diff G for 2 groups: G_A, G_P

$\underset{\sim}{b_i} \sim N(\underset{\sim}{0}, G)$ → so var doesn't depend on i, the group

could be: $\underset{\sim}{b_i} \sim N(\underset{\sim}{0}, G_i)$
← Proc mixed allows this

Distribution of random effects

Marginal models: $E[Y_i] = X_i \underset{\sim}{\beta}$ ← one cluster

$E[Y_{ij}] = \underset{\sim}{x_{ij}}^T \underset{\sim}{\beta}$
↖ fixed, not random

↳ $= E[Y_{ij} | x_i]$

An example where this assumption is suspicious: FEV1 and used height as covariate. $E[Y_{ij} | \text{all height data}] \neq E[Y_{ij} | \text{height at time } 1]$

including height in future   current time

$\text{cov}(\text{FEV1}, \text{height})$

FEV1 $\longrightarrow$ $\begin{pmatrix} \underset{\sim}{Y_i} \\ \underset{\sim}{T_i} \end{pmatrix}$
height $\longrightarrow$

$E\begin{pmatrix} \underset{\sim}{Y_i} \\ \underset{\sim}{T_i} \end{pmatrix} =$

model their covariance structure

This is a problem w/ all time-dependent covariates and exactly what you are conditioning on.

We won't go into this.

exogenous covariates: assumption
$E[Y_{i1} \mid \text{all height data}] = E[Y_{i1} \mid \text{Height time } 1]$

Generalized Mixed Models
$g(\mu_{ij}) \rightarrow$ link function

One cluster: $g(\mu_i) \rightarrow$ apply $g()$ to each component

Main Difference
- link function
- we assume that condition

$Y_{ij} \mid b_i$ are independent
↳ conditional independence assumption
(b/c otherwise it is too difficult)