

TESTING α

Components in α can be tested just like we did for GLMs:

$$t^* = \frac{B'\hat{\alpha}}{\sqrt{B'(X'\hat{\Sigma}X)^{-1}B}}$$
$$F^* = \frac{(B'\hat{\alpha})' [B'(X'\hat{\Sigma}X)^{-1}B]^{-1} (B'\hat{\alpha})}{\text{rank}(B)}$$

B is defined as before: a vector or matrix of contrast coefficients that indicate which components of α are tested.

$$T^* = \left(\begin{array}{c} \text{Reduced model} \\ -2 \log \text{ likelihood} \end{array} \right) - \left(\begin{array}{c} \text{Full model} \\ -2 \log \text{ likelihood} \end{array} \right)$$

For T^* , full and reduced models must be nested and must have the same random effects structure and Σ structure. The degrees of freedom are computed as:

$$r = (\# \text{ parameters in full model's } \alpha) \\ - (\# \text{ parameters in reduced model's } \alpha).$$

As before,

- All these tests are approximate and the test gets better as n gets larger.

↑

subjects
or clusters

- Degrees of freedom for the t^* and F^* must be estimated. Use the SAS defaults. For mixed models (i.e., when a RANDOM statement is used), the default degrees of freedom method is $DDFM = CONTAIN$.
- Use REML for t^* and F^* .
- Use ML for T^* .
- AIC or BIC can be used as before for informal comparisons.

Very Long Aside: F- and t-test degrees of freedom

SAS has five possibilities for the denominator degrees of freedom (DDFM) for the F -test and t -test:

```
MODEL ... / DDFM= CONTAIN ...;  
                BETWITHIN  
                RESIDUAL  
                SATTERTH  
                KENWARDROGER
```

- DDFM=CONTAIN is the default when a RANDOM statement is present.
- DDFM=BETWITHIN is the default when *only* a REPEATED statement is present.
- Thus when both are present, DDFM=CONTAIN is the default.

DDFM = CONTAIN

What does SAS mean by “containment?” A fixed effect (e.g., gender, treatment, etc.) is contained

- in any terms which are interactions with that effect,
- in any terms which are nested within that effect,
- in itself.

SAS checks the RANDOM statment to see if it finds any terms which contain the fixed effect being tested.

EXAMPLE

```
PROC MIXED;  
    CLASS female;  
    MODEL Y = female age female*age;  
    RANDOM int / SUB = child;
```

The fixed intercept is contained within the random intercept.

SAS computes the “rank contribution” of the random intercept to the matrix $[X \ Z]$ and assigns that as the degrees of freedom for the fixed intercept.

The other fixed effects (female, age) are given degrees of freedom $N - \text{rank}[X \ Z]$, the “residual” degrees of freedom, where $N = \text{total } \# \text{ observations}$. female*age gets (df female)*(df age), as in regression.

EXAMPLE

```
PROC MIXED;  
    CLASS female;  
    MODEL Y = female age female * age;  
    RANDOM int age / SUB = child;
```

- The fixed intercept is contained within the random intercept.
- The fixed slope term age is contained within the random slope term age.

SAS then computes the “rank contribution” of each of the random intercept and random age to the matrix $[X \ Z]$. SAS takes the smaller of the two and assigns that as the degrees of freedom for the fixed intercept and fixed age. female*age gets (df female)*(df age), as in regression.

DDFM=BETWITHIN

When only a REPEATED statement is present, the residual degrees of freedom are divided into “between-subject” and “within-subject”. PROC MIXED then checks to see if the fixed effect being tested changes within any subject.

EXAMPLE

```
PROC MIXED;  
  CLASS gender ctrial;  
  MODEL contact = age gender ctrial;  
  REPEATED ctrial/ SUB=subj TYPE = cs;
```

Trial changes within subject, but gender does not.

Denominator degrees of freedom for terms which do change within subject are assigned to be the within-subject residual degrees of freedom. For terms which do not change within subject, SAS assigns the between subject degrees of freedom.

If there are multiple within-subject effects which are class variables, the within-subject residual degrees of freedom are partitioned into subject-by-effect interactions.

EXCEPTION: TYPE=UN

REPEATED trial / SUB = subj TYPE=UN;

In this case, all effects are assigned the between-subject degrees of freedom to provide better small sample approximations.

DDFM=RESIDUAL

Residual degrees of freedom are assigned to all effects, $N - \text{rank}[X Z]$. $N = \#$ observations total.

DDFM=SATTERTH

The Satterthwaite approximation is used in two-sample t -tests with unknown variances which are not pooled:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1-1} \left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2-1} \left(\frac{s_2^2}{n_2}\right)^2}$$

There is a generalization of this for fixed effects F -tests. See the SAS documentation of PROC MIXED for details.

DDFM=KENWARDROGER

This method inflates the variance/covariance matrix of the vector α , and then uses this in a Satterthwaite approximation.

This is the most important thing to remember about degrees of freedom:

IF ALL THE DENOMINATOR DEGREES OF FREEDOM IN YOUR F-TESTS ARE BIGGER THAN ~ 25 (no matter what your covariance structure is or which DF option is used) THEN IT DOES NOT REALLY MATTER WHICH DF OPTION IS USED.

End of Very Long Aside

TESTING β_i

Do we really want to test β_i ? Think about the Orthodontic Study:

$$Y_i = X_i\alpha + \beta_{i1} + \beta_{i2} * age + \delta_i$$

random random
(subject-specific) (subject-specific)
intercept slope

$$\begin{bmatrix} \beta_{i1} \\ \beta_{i2} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, D = \begin{bmatrix} D_{11} & \\ & D_{22} \end{bmatrix} \right)$$

What do we want to test to decide whether, e.g., the random slopes are needed?

TESTING Σ

Remember that which random effects we include dictates what structure Σ will take. When comparing two structures to each other:

you MUST understand whether going from your full (more complex) to your reduced (simpler) model involves testing whether a variance (on the diagonal of Σ or D or R_i) is equal to 0.

IF SO, LIKELIHOOD RATIO TESTS CANNOT BE CARRIED OUT AS BEFORE.

IF NOT, we use the same likelihood ratio test as for GLMs:

$$T^* = \left(\begin{array}{c} \text{Reduced model} \\ -2 \log \text{ likelihood} \end{array} \right) - \left(\begin{array}{c} \text{Full model} \\ -2 \log \text{ likelihood} \end{array} \right)$$

The degrees of freedom are computed as

$$r = (\# \text{ parameters in full model's } \Sigma) \\ - (\# \text{ parameters in reduced model's } \Sigma).$$

and T^* is compared to a χ_r^2 .

As for GLMs this must be done with REML and your full and reduced models must be nested. Lastly, you must have the same $X\alpha$ (MODEL statement) in both the full and reduced models.

IF YOU DO NEED TO TEST WHETHER A VARIANCE IS 0, determine whether the test you are trying to do is one of the following special cases:

SPECIAL CASE #1 Testing one random effect (full model) vs. no random effect (reduced model)

$$\text{Full Model: } Y_{ij} = X_{ij}\alpha + \beta_{i1} + \delta_{ij}$$

$$\beta_{i1} \sim N(0, D)$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

$$\text{Reduced Model: } Y_{ij} = X_{ij}\alpha + \delta_{ij}$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

This is really a test of H_0 : vs. H_a :

For this case, $T^* \overset{\text{approx}}{\sim} \frac{1}{2} \chi_0^2 + \frac{1}{2} \chi_1^2$

NOT $T^* \sim \chi_1^2$ as you might expect.

How do we carry out this test?

- Compute T^* as before.
- Get a p-value from comparing T^* to a χ_1^2 .
- Cut the p-value in half. This is the appropriate p-value for the test.

SPECIAL CASE #2 Testing two random effects (full model) vs. one random effect (reduced model)

EXAMPLE

$$\text{Full Model: } Y_{ij} = X_{ij}\alpha + \beta_{i1} + \beta_{i2} * \text{time} + \delta_{ij}$$

$$\begin{bmatrix} \beta_{i1} \\ \beta_{i2} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, D = \begin{bmatrix} D_{11} & \\ & D_{22} \end{bmatrix} \right)$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

$$\text{Reduced Model: } Y_{ij} = X_{ij}\alpha + \beta_{i1} + \delta_{ij}$$

$$\beta_{i1} \sim N(0, D)$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

This is a test of H_0 :
vs. H_a :

For this case, $T^* \overset{\text{approx}}{\sim} \frac{1}{2} \chi_1^2 + \frac{1}{2} \chi_2^2$.

There are no tables for the distribution of this mixture $\frac{1}{2} \chi_1^2 + \frac{1}{2} \chi_2^2$.

Option 1: Compare T^* to a χ_2^2 . This will always be conservative, i.e. the p-value will be too big (less significant).

Option 2: Simulate your own table value for $\frac{1}{2} \chi_1^2 + \frac{1}{2} \chi_2^2$. Compare it to your T^* . See the *R* program on the class R-plotting web page.

Option 3: Use AIC or BIC to compare.

SPECIAL CASE #3 Testing $q+l$ random effects (full model) vs. q random effects (reduced model)

EXAMPLE $q = 2$ and $l = 1$

Full Model:

$$Y_{ij} = X_{ij}\alpha + \beta_{i1} + \beta_{i2} * time + b_{i2} * (time)^2 + \delta_{ij}$$

$$\begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ b_{i2} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, D = \begin{bmatrix} D_{11} & & \\ D_{21} & D_{22} & \\ D_{31} & D_{32} & D_{33} \end{bmatrix} \right)$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

Reduced Model:

$$Y_{ij} = X_{ij}\alpha + \beta_{i1} + \beta_{i2} * time + \delta_{ij}$$

$$\begin{bmatrix} \beta_{i1} \\ \beta_{i2} \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, D = \begin{bmatrix} D_{11} & \\ D_{21} & D_{22} \end{bmatrix} \right)$$

$$\delta_{ij} \sim N(0, \sigma^2)$$

This is a test of

H_0 :

H_a :

For this case, $T^* \overset{approx}{\sim} \frac{1}{2} \chi_q^2 + \frac{1}{2} \chi_{q+1}^2$.

There are no tables for this distribution.

Option 1: Compare T^* to a χ_{q+1}^2 . This will always be conservative (p-value will be too big).

Option 2: Simulate your own table value. Replace the $df = 1$ and $df = 2$ with $df = q$ and $df = q + 1$ for your q .

As always, use REML when computing T^* for all of these cases.

- If the test you are trying to do is not a standard LRT and is not one of the special cases, then there is no formal statistical test. You can compare AIC and BIC values instead.
- You can always use AIC and BIC to compare non-nested models.

EXAMPLE

Model #1 is GLM with Toeplitz.

Model #2 is GLMM with random intercept and slope.

Take the model with the smaller AIC or BIC.

Words of warning:

- These tests for the special cases are based on many layers of approximations. In practice this means that I will use a more complex covariance structure only if it is *strongly* significantly different than a less complex covariance structure.
- These tests also were derived assuming $R_i = \sigma^2 I$. Later we will learn about models where R_i has e.g. an AR(1) structure. In such cases, these tests for the special cases may not be appropriate, and you should rely on standard LRTs or AIC or BIC.

Words of encouragement:

- The modeling process is the same as it was for GLMs.
- Don't worry too much if your n is reasonably large ($n \geq 100$).

OVERALL MODELING STRATEGY

- Create summary statistics and EDA plots to explore the mean structure.
- Decide upon a reasonable starting (full) model for X_α . It is best to start with a large model, so that covariance structures can be compared with all reasonable covariates already taken into account (included in X_α).
- Create summary statistics and EDA plots to explore the variance/covariance structure.
- Decide upon several reasonable covariance structures to consider. They could be from either GLMs or GLMMs or both.
- Fit one model for each covariance structure using REML and the same full model for X_α in each. Compare the models using LRTs (if possible), AIC, or BIC to determine the best structure.

- Refit the model with the chosen covariance structure and the full model for X_α using ML. For model reduction, use ML if you will be using LRTs or REML if you will be using t -tests or F -tests.
- Reduce the full model for X_α to the most parsimonious model appropriate for the scientific questions. Compare nested model with LRTs (using ML), F -tests (using REML), or t -tests (using REML). It will likely take several stepwise reductions to get to a final model (just like in regression).
- Compare your full model for X_α to your final model for X_α using ML and a LRT.
 - If they are not significantly different, re-fit your final model with REML.
 - If they are, then likely there is collinearity or confounding among the covariates. Reconsider your stepwise reductions.
- Carry out model diagnostics.