

BIOS6643 Longitudinal

L7 Random effect

EJC

Department of Biostatistics & Informatics

Random effect

Estimation

The Mt. Kilimanjaro data

Tests for variance components

Summary

Random effect

Estimation

The Mt. Kilimanjaro
data

Tests for variance
components

Summary

Topics for today:

- ▶ Inference for random effects in an LMM
- ▶ Modeling random effects in an LMM, with data
- ▶ Tests for variance components (brief)

Associated reading:

1. Sections 3.3 and 4.1 of 'LMM: inference' course notes)
2. Verbeke (with a focus on Ch. 7)
3. Hedeker (Chapters 4-7)

- ▶ Although we can use ML or REML to estimate variance components, we may be interested in subject-specific random effect estimates.
 - ▶ In particular, they may allow us to determine if there are subjects with unusual trends relative to the rest of the group.
 - ▶ These subject-specific estimates cannot be derived from the marginal model.
 - ▶ A common approach is to use empirical Bayes (EB) estimators. EB estimators have an intuitive appeal since estimates are obtained essentially by taking a weighted average of personal and group-level data.

Example 1: batting averages of Major League Baseball players.

- ▶ At the beginning of the season, averages tend to vary more wildly (between 0.000 and 1.000).
- ▶ As more games are played, the averages tend to settle into range between 0.200 and 0.350.
- ▶ An EB estimate for a particular player near the beginning of the season may use a higher weight for the 'all-player' average and a lower average for that particular player to estimate that player's true average; later in the season the average may be weighted more heavily towards that player's particular average.

Example 2: prevalence of a disease or illness for individual counties in a state.

- ▶ Ideally, the best estimate of prevalence in a county would involve just the county data.
- ▶ However, if collected data is sparse, then it might help to also base the estimate on state data as well.
- ▶ The higher the variability in county data, the more the estimate is based on the state data, while the lower the variability in the county data, the more it is based on county data.

In the Bayesian literature, the marginal distribution of \mathbf{b} is called the prior distribution of the parameters \mathbf{b} since it does not depend on the data \mathbf{Y} . Once observed values of \mathbf{Y} are obtained (\mathbf{y}), the posterior distribution of \mathbf{b} , which is $f(\mathbf{b}|\mathbf{y})$, can be calculated. Considering \mathbf{b}_i and \mathbf{Y}_i as the random effects and outcome data for individual i , the posterior distribution is:

$$f(\mathbf{b}_i|\mathbf{Y}_i = \mathbf{y}_i) = \frac{f(\mathbf{y}_i|\mathbf{b}_i)f(\mathbf{b}_i)}{\int f(\mathbf{y}_i|\mathbf{b}_i)f(\mathbf{b}_i)d\mathbf{b}_i}$$

In the expression above, the dependence of the density function on certain components of $\boldsymbol{\theta}$ is suppressed for notational convenience. The mean of this posterior distribution is a Bayes estimator of \mathbf{b}_i :

$$\hat{\mathbf{b}}_i(\boldsymbol{\theta}) = E(\mathbf{b}_i|\mathbf{Y}_i = \mathbf{y}_i) = \int \mathbf{b}_i f(\mathbf{b}_i|\mathbf{y}_i) d\mathbf{b}_i = \mathbf{G}\mathbf{Z}_i^\top \mathbf{V}_i^{-1}(\boldsymbol{\alpha})(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \quad (6)$$

The EB estimator is then computed by replacing unknown parameters α and β with their ML or REML estimates (and hence the word ‘empirical’).

We'll let $\mathbf{b}_i(\hat{\theta}) = \hat{\mathbf{b}}_i$ denote the Empirical Bayes estimator. For more detail, see **section 7.2 in Verbeke**. In terms of final notation, \mathbf{b} and \mathbf{Y} represent the vector of random effects and data, respectively, for the complete data, where data for subjects are stacked, while \mathbf{b}_i and \mathbf{Y}_i are the data for individual i .

Predicted values based on EB estimators for \mathbf{b}_i are a weighted average of subject-specific data and group-averaged data, giving it an intuitive appeal:

$$\begin{aligned}\hat{\mathbf{Y}}_i &= \mathbf{X}_i \hat{\boldsymbol{\beta}} + \mathbf{Z}_i \mathbf{b}_i = \mathbf{X}_i \hat{\boldsymbol{\beta}} + \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^{\top} \mathbf{V}_i^{-1} (\mathbf{Y}_i - \mathbf{X}_i \boldsymbol{\beta}) \\ &= (\mathbf{I}_{r_i} - \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^{\top} \mathbf{V}_i^{-1}) \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^{\top} \mathbf{V}_i^{-1} \mathbf{Y}_i = \mathbf{R}_i \mathbf{V}_i^{-1} \mathbf{X}_i \hat{\boldsymbol{\beta}} + (\mathbf{I}_{r_i} - \mathbf{R}_i \mathbf{V}_i^{-1}) \mathbf{Y}_i\end{aligned}$$

$(\mathbf{I}_{r_i} - \mathbf{R}_i \mathbf{V}_i^{-1})$ is a weighted average of the estimated population average profile and the observed data.

This demonstrates that $\hat{\mathbf{Y}}_i$ are shrunk towards the mean (relative to \mathbf{Y}_i).

When residual variability (modeled through R_i) is large in relation to between-subject variability (accounted for in V_i^{-1}), the population-averaged profile ($X_i\beta$) will have more weight, which makes sense since there is less certainty about individual data. (You can think of R_i as the “numerator” and V_i as the “denominator” in the quantity $R_i V_i^{-1}$.)

Alternatively, when residual (within-subject) variability tends to be smaller and between-subject variability greater, then Y_i will have more weight.

The EB estimators themselves exhibit the shrinkage property:

$Var[\mathbf{L}\mathbf{b}_i] \leq Var[\mathbf{L}\hat{\mathbf{b}}_i]$ for any $1 \times q$ real-valued vector \mathbf{L} . Remember also that $E[\mathbf{b}_i] = 0$. Thus the EB estimators are shrunk towards 0. For more detail, see **Verbeke**.

Inference associated with EB estimators

The quantity $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})]$ can be derived easily by substituting the MLE in for $\boldsymbol{\beta}$ and noting that it is a linear form of \mathbf{y}_i . (Laird and Ware, 1982, consider the Bayes estimator as in (6), but with $\boldsymbol{\beta}$ replaced with its MLE; they then derive theoretical results when covariance parameters are known or unknown.) The result is:

$$Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})] = \mathbf{G}_i \mathbf{Z}_i^\top \left(\mathbf{V}_i^{-1} - \mathbf{V}_i^{-1} \mathbf{X}_i (\sum \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i)^{-1} \mathbf{X}_i^\top \mathbf{V}_i^{-1} \right) \mathbf{Z}_i \mathbf{G}_i$$

- ▶ $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})]$ is not the same as $Var[\mathbf{b}_i | \mathbf{Y}_i = \mathbf{y}_i]$; it is $Var[E[\mathbf{b}_i(\boldsymbol{\theta}) | \mathbf{y}_i]]$.
- ▶ For inference, $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta}) - \mathbf{b}_i]$ is used rather than $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})]$ because the former take into account the variability in \mathbf{b}_i . This quantity is: $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta}) - \mathbf{b}_i] = \mathbf{G}_i - Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})] = \mathbf{G}_i - \mathbf{G}_i \mathbf{Z}_i^\top \left(\mathbf{V}_i^{-1} - \mathbf{V}_i^{-1} \mathbf{X}_i (\sum \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i)^{-1} \mathbf{X}_i^\top \mathbf{V}_i^{-1} \right) \mathbf{Z}_i \mathbf{G}_i$

In order to estimate $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta}) - \mathbf{b}_i]$ we typically just 'plug in' numerical values for unknown $\boldsymbol{\theta}$, not accounting for the added variability due to use of estimated values. In light of this, the selection of DF can help control the accuracy of inferential results for random effects, similar to that described previously for inference of fixed effects.

t -tests can be constructed for random effects using relevant approximate t quantities.

- ▶ For example, if \mathbf{b}_i contains just a random intercept (i.e., $\mathbf{b}_i = b_{0i}$) then we can use $t = \frac{(\hat{b}_{0i} - b_{0i}) - E[\hat{b}_{0i} - b_{0i}]}{\widehat{SE}(\hat{b}_{0i} - b_{0i})}$, which reduces to $t = \frac{\hat{b}_{0i}}{\widehat{SE}(\hat{b}_{0i} - b_{0i})}$ under the null, for the test of $H_0 : b_{0i} = 0$.

- ▶ For models with multiple random effect terms, we can carry out t -tests separately for each component of \mathbf{b}_i (and subject). As before, the DF (ν) is ideally chosen to get the correct distribution of the test statistic under H_0 ; available methods to do this are as previously described.
- ▶ Theory also exists for tests $H_0 : \mathbf{Lb} = 0$ versus $H_1 : \mathbf{Lb} \neq 0$. However, in practice, I have not yet found the need to use this.
- ▶ A $100 \times (1-\alpha)\%$ confidence interval for an element b_{hi} of \mathbf{b}_i , is $\hat{b}_{hi} \pm t_{\nu, \frac{\alpha}{2}} \times \widehat{SE}(\hat{b}_{hi} - b_{hi})$.

In SAS, when you request a solution for the random effects, the 'Estimate' will be numerical versions of (6), while 'Std Err Pred' is the square root of (diagonal elements of) $Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta}) - \mathbf{b}_i]$. The calculated variance of the random effect estimates (using the 'population' version) will be the same as $\widehat{Var}[\hat{\mathbf{b}}_i(\boldsymbol{\theta})]$ (here, the hat on 'Var' indicates that estimated values of $\boldsymbol{\theta}$ are 'plugged into' the calculation) and will be somewhat less than σ_b^2 , reflecting the shrinking of the estimates back to the estimated population mean.

Computation of estimates and associated variances for random effects

BIOS6643
Longitudinal

EJC

Random effect

Estimation

The Mt. Kilimanjaro
data

Tests for variance
components

Summary

see course notes

Empirical Bayes estimators for LMMs with random intercepts

We have discussed Empirical Bayes estimators of random effects in mixed models. They have an intuitive appeal because they can be expressed as weighted averages of subject-specific information and population-average information.

- ▶ The greater the variability of the subject data, the higher the weight is placed on the population average;
- ▶ The more consistent the subject data is, the higher the weight is placed on the subject portion.
- ▶ In previous notes, the weighted average was expressed for predicted values (\hat{Y}_i) from an LMM.
- ▶ It was briefly mentioned that the random effects estimates themselves ($\hat{\mathbf{b}}_i$) are shrunk towards the population mean (relative to \mathbf{b}_i), such that $Var[\hat{\mathbf{b}}_i] \leq Var[\mathbf{b}_i]$.
- ▶ The amount of shrinkage depends on residual variance relative to subject variance. To study this further, we'll consider LMMs with random intercept terms.

It was mentioned that the random effects estimates ($\hat{\mathbf{b}}_i$) are shrunk towards the population mean (relative to \mathbf{b}_i), such that

$Var[\mathbf{L}\hat{\mathbf{b}}_i] \leq Var[\mathbf{L}\mathbf{b}_i]$ for a $1 \times q$ real-valued vector \mathbf{L} .

- ▶ A special case of this is $Var[\hat{b}_{hi}] \leq Var[b_{hi}]$, for $h = 1, \dots, q$. This is easy to prove, since $Var[\hat{\mathbf{b}}_i - \mathbf{b}_i] + Var[\hat{\mathbf{b}}_i] = \mathbf{G}_i$, and the diagonal elements must be nonnegative.
- ▶ The only time equality holds, such that $Var[\hat{\mathbf{b}}_{hi}] = Var[b_{hi}]$, is when $Var[\hat{\mathbf{b}}_{hi} - \mathbf{b}_{hi}] = 0$. The amount of shrinkage in estimators depends on residual variance relative to subject variance. To study this further, we'll consider LMMs with random intercept terms.

If the only random term in the model is an intercept term (for subjects) and $\mathbf{R}_i = \sigma^2 \mathbf{I}$, (6) will reduce, since \mathbf{G} only has one element (the variance of the random intercepts, call it σ_b^2), and \mathbf{Z}_i^\top is a row vector of 1's, call it $\mathbf{J}_{(1 \times r)}$. For this case:

$$\mathbf{V}_i^{-1} = (\sigma_b^2 \mathbf{J}_{(r_i \times r_i)} - \sigma_\epsilon^2 \mathbf{I}_{(r_i \times r_i)})^{-1} = \frac{1}{\sigma_\epsilon^2} \left(\mathbf{I}_{(r_i \times r_i)} - \frac{(\sigma_b^2)}{(\sigma_\epsilon^2 + r_i \sigma_b^2)} \mathbf{J}_{(r_i \times r_i)} \right)$$

Ultimately, the Bayes estimator reduces to:

$$\hat{b}_i(\boldsymbol{\theta}) = \lambda \left(\bar{Y}_i - \frac{\sum_j \mathbf{X}_{ij}^r \boldsymbol{\beta}}{r_i} \right) \quad (7)$$

► where \bar{Y}_i is the mean response for subject i , \mathbf{X}_{ij}^r is the j th row of \mathbf{X}_i , and $\lambda = \frac{r_i \sigma_b^2}{\sigma_\epsilon^2 + r_i \sigma_b^2}$.

Note that λ is between 0 and 1; it is the weight used in the averaging of subject-specific and population average statistics. (Note also that **there is no u, and I have no idea** is unbolded since it involves just one estimator.) Greater between-subject variability relative to within-subject variability will yield larger values of λ (just like the ICC), but so will increasing the number of repeated measures.

For practice, show that (7) holds, starting with (6). (You can use the given result for V_i^{-1} .) When there is only a random intercept term and fixed intercept in the model $[Y_{ij} = 0 + b_i + \epsilon_{ij}]$;

$b_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_b^2)$, independently of $\epsilon_{ij} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$; call it the 'simple random intercept model', (7) becomes $\hat{b}_i(\boldsymbol{\theta}) = \lambda[\bar{Y}_i - \beta_0]$ (8)

We can consider λ as the shrinkage factor. What is being shrunk is the difference between the estimate of the random intercept for subject i and the population mean. If we add the population mean, β_0 , we get the estimate for subject i in context of the population:

$$\lambda(\bar{Y}_i - \beta_0) + \beta_0 = \lambda\bar{Y}_i + (1 - \lambda)\beta_0 \quad (9)$$

which is a weighted average of \bar{Y}_i and β_0 .

In practice we typically replace unknown parameters λ (which involves σ_b^2 and σ_ϵ^2) and β_0 in (8) and (9) with their estimators, yielding EB estimators.

For the simple random intercept model, the variance of the Bayes estimator is

$$Var[\hat{b}_i(\boldsymbol{\theta})] = \frac{n-1}{n} \lambda \sigma_b^2 \quad (10)$$

As noted earlier, the variance quantity normally used in inference to account for randomness in b_i is:

$$Var[\hat{b}_i(\boldsymbol{\theta}) - b_i] = \sigma_b^2 - Var[\hat{b}_i(\boldsymbol{\theta})] = \sigma_b^2 - \frac{n-1}{n} \lambda \sigma_b^2 = \frac{(1 - \lambda(n-1))}{n} \sigma_b^2 \quad (11)$$

Since the variance of EB estimators is more difficult to tackle, we usually work with the variance quantities of the Bayes estimators, in (10) and (11). But in practice we do then typically plug in values of unknown variances in the quantity, which I will denote as $\widehat{Var}[\hat{b}_i(\boldsymbol{\theta})]$ and $\widehat{Var}[\hat{b}_i(\boldsymbol{\theta}) - b_i]$.

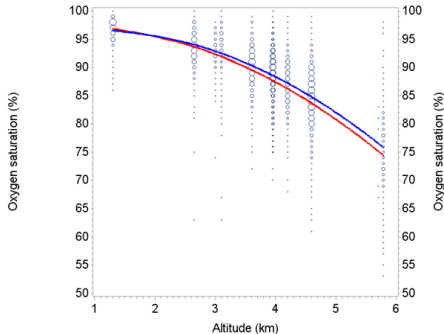
For the random intercept model we know that

$Var[\hat{b}_i(\boldsymbol{\theta}) - b_i] + Var[\hat{b}_i(\boldsymbol{\theta})] = \sigma_b^2$ (more generally, that

$Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta}) - b_i] + Var[\hat{\mathbf{b}}_i(\boldsymbol{\theta})] = \mathbf{G}_i$). For fixed variances, we know that $\lambda \rightarrow 1$ as the number of repeated measures, r_i is increased (and also n), in which case $Var[\hat{b}_i(\boldsymbol{\theta})] \rightarrow \sigma_b^2$, and hence $Var[\hat{b}_i(\boldsymbol{\theta}) - b_i] \rightarrow 0$.

- ▶ Oxygen saturation, or SAO₂, can be measured as the percentage of hemoglobin molecules which are oxygenated (oxyhemoglobin) in arterial blood.
- ▶ The normal range is $> 95\%$, however at higher altitudes this percentage tends to go down.
- ▶ This measure was taken on hundreds of subjects that climbed Mt. Kilimanjaro (the tallest mountain on the continent of Africa).
- ▶ The following graph shows SAO₂ versus altitude, along with a quadratic fit using a linear mixed model.
- ▶ The bubble plot was used because many values occurred on the same (x, y) location; bubbles indicate how many subjects occurred at each point, the bigger the bubble, the more subjects at that location.

- ▶ Although these are repeated measures data, lines connect points are suppressed due to the large amount of data.
- ▶ Superimposed on the bubble plot are two curves, one showing subjects that were taking a medication to help prevent symptoms of high-altitude sickness (blue), and those that were not (not).
- ▶ Those taking the medication are able to maintain slightly higher oxygen levels (which may also help reduce symptoms of high altitude sickness), with greater differences at higher altitudes.
- ▶ These differences are statistically significant after about 3km, but it may be somewhat subjective as to whether the small differences are worth taking the medication.



The bubble plot was generated by using the 'bubble' statement in PROC GPLOT. I then overlaid the fitted curves from the fitted linear mixed model in a subsequent 'plot2' statement.

Below is the SAS code used to fit the mixed model:

```
proc mixed data=alldata; class id recnum;
model oxygen_sat= x x*x diamox_ever x*diamox_ever x*x*diamox_ever
/ outpm=outypm outp=outyp solution;
random intercept x x*x / subject=id v solution g type=un;
estimate 'diamox, alt=5km'
intercept 1 x 5 x*x 25 diamox_ever 1 x*diamox_ever 5 x*x*diamox_ever 25;
estimate 'no diamox, alt=5km' intercept 1 x 5 x*x 25;
estimate 'diff at alt=1km' diamox_ever 1 x*diamox_ever 1 x*x*diamox_ever 1;
estimate 'diff at alt=2km' diamox_ever 1 x*diamox_ever 2 x*x*diamox_ever 4;
estimate 'diff at alt=3km' diamox_ever 1 x*diamox_ever 3 x*x*diamox_ever 9;
estimate 'diff at alt=4km' diamox_ever 1 x*diamox_ever 4 x*x*diamox_ever 16;
estimate 'diff at alt=5km' diamox_ever 1 x*diamox_ever 5 x*x*diamox_ever 25;
estimate 'diam intercept' intercept 1 diamox_ever 1;
estimate 'diam x term' x 1 x*diamox_ever 1;
estimate 'diam x*x term' x*x 1 x*x*diamox_ever 1;
contrast 'interaction' x*diamox_ever 1, x*x*diamox_ever 1;
contrast 'curve comparison'
diamox_ever 1, x*diamox_ever 1, x*x*diamox_ever 1; run;
```

Abbreviated output:

BIOS6643
Longitudinal

EJC

Random effect

Estimation

The Mt. Kilimanjaro
data

Tests for variance
components

Summary

The Mixed Procedure		Solution for Fixed Effects	
Dependent Variable	Oxygen_Sat	Effect	Est. Std Err DF t Value Pr> t
Covariance Structure	Unstructured	Intercept	97.062 0.737 914 131.74 <.0001
Subject Effect	id	x	0.954 0.555 914 1.72 0.0857
Estimation Method	REML	x*x	-0.840 0.093 914 -9.01 <.0001
Residual Variance Method	Profile	diamox_ever	-1.096 0.775 11E3 -1.41 0.1575
Fixed Effects SE Method	Model-Based	x*diamox_ever	0.663 0.584 11E3 1.14 0.2561
Degrees of Freedom Method	Containment	x*x*diamox_ever	-0.040 0.098 11E3 -0.41 0.6840
Dimensions		Solution for Random Effects	
Covariance Parameters	7	Effect	Id Estimate Std Err DF t Value Pr> t
Columns in X	6		Fixed
Columns in Z Per Subject	3	Intercept 1	-6.2686 0 11E3 -Infy <.0001
Subjects	916	x	1 3.5771 0 11E3 -Infy <.0001
Max Obs Per Subject	20	x*x	1 -0.8219 0.320 11E3 -2.57 0.0102
Number of Observations Used	13369	Intercept 2	-2.8734 0 11E3 -Infy <.0001
Estimated G Matrix		x	2 1.5678 0 11E3 -Infy <.0001
Row Effect Id Col1 Col2 Col3		x*x	2 -0.1304 0.119 11E3 -1.09 0.2744
1 Int. 1	-6.434 1.572	Intercept 3	1.9668 0 11E3 -Infy <.0001
2 x	1 -6.434 9.116 -1.790	x	3 -1.2063 0 11E3 -Infy <.0001
3 x*x	1 1.572 -1.790 0.353	x*x	3 0.1812 0.119 11E3 1.52 0.1290
Residual Variance Estimate:	8.8320	...	
Fit Statistics		Intercept 921	-8.8468 0 11E3 -Infy <.0001
-2 Res Log Likelihood	69592.3	x	921 5.2530 0 11E3 -Infy <.0001
AIC (smaller is better)	69604.3	x*x	921 -1.1837 0.135 11E3 -8.75 <.0001
AICC (smaller is better)	69604.3	Estimates	
BIC (smaller is better)	69633.2	Label	Est. Std Err DF t Value Pr> t
Type 3 Tests of Fixed Effects		Diamox, alt=5km	82.060 0.152 11E3 541.43 <.0001
		no diamox, alt=5km	80.839 0.466 914 175.49 <.0001
		diff at alt=1km	-0.472 0.335 11E3 -1.41 0.1580
		diff at alt=2km	0.071 0.304 11E3 0.35 0.7281
		diff at alt=3km	0.534 0.252 11E3 2.12 0.0341
		diff at alt=4km	0.918 0.290 11E3 3.16 0.0016
		diff at alt=5km	1.221 0.490 11E3 2.49 0.0127
		diam intercept	95.967 0.240 11E3 399.32 <.0001
		diam x term	1.617 0.182 11E3 8.87 <.0001
		diam x*x term	-0.880 0.031 11E3 -28.66 <.0001
		Contrasts	
		Label	Num DF Den DF F Value Pr > F
		Interaction	2 11E3 6.21 0.0020
		curve comparison	3 11E3 4.37 0.0044

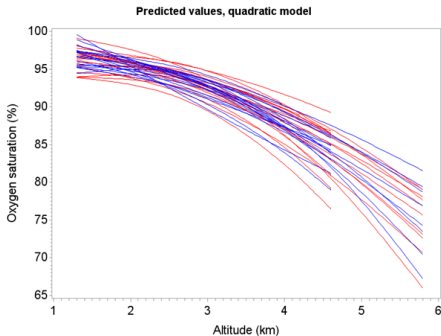
Some interesting things to point out from the output:

The variance of the intercept was estimated to be 0. However, no penalty was added for this in the AIC; essentially, that parameter is removed from the model. You can tell this is the case because the difference between $-2\log(\text{Restricted Likelihood})$ and the AIC is 12, so 6 parameters are accounted for (5 in \mathbf{G} matrix, plus residual variance).

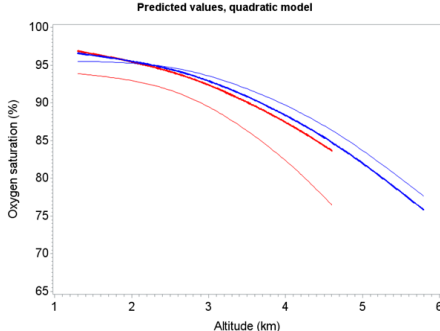
Notice also that subject estimates of intercept and some linear terms have predicted standard errors of 0; our interpretation should be that these standard errors could not be estimated, rather than that they were true 0's.

Estimates included demonstrate that although differences between medication users and non-users appeared to be minor, visually, they were statistically significant, with greater significance at higher elevations.

The contrasts indicate that there were differences between curves that could not be accounted for by intercept differences alone (see 'interaction' test, $p = 0.0020$), and that the curves were not the same (including both 'interaction' and y-intercept differences, $p = 0.0044$).



The graph above shows predicted values for subjects, from the mixed model, using only 20 per group (no medication – red / medication – blue) are plotted. Differences for subjects are due to the use of the intercept, linear and quadratic random effect terms. Notice that the predicted curves tend to fan out at higher altitudes, just like the raw data.



This graph shows population-averaged estimates (thick red and blue) and two subject curves (thin red and blue: see subject ID's 1 (red) and 2 (blue) on previous SAS output).

Recall that random effect estimates are deviations from fixed effect estimates. Notice how the red subject curve has much greater curvature than the thick red curve, compared with the thin and thick blue lines. This is echoed in the tests on the previous output, where the t -tests indicate that subject 1 has significant difference in quadratic effect compared with its population counterpart ($p = 0.0102$), while the blue does not (0.2744).

Both subject curves have lower intercepts and higher coefficients for the first order term, although we are unable to conduct tests to see if they

Tests for variance components

We can use 'COVTEST' as an option in the PROC MIXED statement for tests involving covariance parameters, using Wald Z tests.

We can also test for a 'significant additions' of random terms to a model (e.g., when including the random slope to an LMM with a random intercept) using likelihood ratio test methods. Here we compare changes in $-2\ln(\mathcal{L})$ between models, which has an asymptotic χ^2 distribution with $DF =$ difference in the number of covariance parameters between the 2 models.

For both approaches, tests are more valid when certain regularity conditions hold. See **Verbeke, pages 64-66** for more detail.

Summary

BIOS6643
Longitudinal

EJC

Random effect

Estimation

The Mt. Kilimanjaro
data

Tests for variance
components

Summary