## Repeated Measures Data – Random Coefficients Models

The most common applications of random coefficients models are those where a linear relationship is assumed between the outcome variable of interest and time, for example test-day data in dairy cows. Usually, how the outcome variables differ between "treatment" groups are of interest. For example, the effect of BST on milk production in dairy cows.

In the textbook they go over examples of a linear random coefficients model and a polynomial random coefficients model. However, the data they use for the linear random coefficient model is not available from their website. So we will cover the Polynomial random coefficients model for now and find another dataset to go over a linear random coefficient model later.

## Polynomial random coefficients model

This study involved 45 children suffering from either solid lump tumor cancer (18) or leukemia (27). Antibody levels to a herpes virus were measured during hospital visits during chemotherapy treatment. Treatment ranged from one month to three years and the time between treatments varied among the children. The objective of the study was to determine whether virus antibody levels were affected by chemotherapy treatment and whether any change was related to type of cancer. The herpes virus antibody was present in all children in the study and it's known that the average level of herpes virus antibody decreases with age. The book presents plots of individual patients for herpes virus antibodies over time. The plots indicate that the levels of antibody fluctuate widely in some children. The non-linear relationship of virus level with time indicates that a model using polynomials of times should be considered because we want to assess whether there is a time trend, or whether the trend varies across treatment groups.

A polynomial model is built up by adding polynomials of increasing order to the model one at a time. Time is included both as fixed effects and random coefficients. The initial linear model fitted is:

Fixed effects: type of cancer, age, time
Random coefficients: patient(intercepts), patient*time (slopes):

```
OPTIONS LS=80;
filename ep 'i:\kathy mixed model\lyall.dat';
DATA a; INFILE ep;
INPUT pat age agecat month month2 month3 month4 virus type $;

PROC MIXED; CLASS type pat;
MODEL virus=age type month/ S DDFM=SATTERTH;
RANDOM int month/ SUB=pat TYPE=UN;
run;
```

Model Information

| | |
|---|---|
| Data Set | WORK.A |
| Dependent Variable | virus |
| Covariance Structure | Unstructured |
| Subject Effect | pat |
| Estimation Method | REML |
| Residual Variance Method | Profile |
| Fixed Effects SE Method | Model-Based |
| Degrees of Freedom Method | Satterthwaite |

| Class | Levels | Values |
|---|---|---|
| type | 2 | AL ST |
| pat | 44 | 3 5 13 61 63 65 67 69 71 73 77 79 81 85 87 91 93 95 97 101 103 105 107 109 113 115 117 119 123 127 129 141 142 143 144 145 146 147 148 149 150 151 152 155 |

Dimensions

| | |
|---|---|
| Covariance Parameters | 4 |
| Columns in X | 5 |
| Columns in Z Per Subject | 2 |
| Subjects | 44 |
| Max Obs Per Subject | 33 |

Number of Observations

| | |
|---|---|
| Number of Observations Read | 625 |
| Number of Observations Used | 625 |

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.4433 |
| UN(2,1) | pat | 0.01326 |
| UN(2,2) | pat | 0.004238 |
| Residual | | 0.5640 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1588.4 |
| AIC (smaller is better) | 1596.4 |
| AICC (smaller is better) | 1596.4 |
| BIC (smaller is better) | 1603.5 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 3 | 465.66 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.6544 | 0.2331 | 47.3 | 15.68 | <.0001 |
| age | | -0.04570 | 0.03653 | 43.9 | -1.25 | 0.2176 |
| type | AL | -0.2341 | 0.2465 | 43.8 | -0.95 | 0.3475 |
| type | ST | 0 | . | . | . | . |
| month | | -0.03161 | 0.01324 | 16.3 | -2.39 | 0.0293 |

The variance components corresponding to the patient and patient*time random coefficients are both positive, indicating that there is more variation between the regression lines for each patient than expected by chance. In other words, patients vary in their rates of decay. So we now go ahead and add the quadratic time effects to both the fixed and random coefficients.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2/ S DDFM=SATTERTH;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.5864 |
| UN(2,1) | pat | -0.04333 |
| UN(2,2) | pat | 0.02451 |
| UN(3,1) | pat | 0.001586 |
| UN(3,2) | pat | -0.00072 |
| UN(3,3) | pat | 0.000022 |
| Residual | | 0.5275 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1583.4 |
| AIC (smaller is better) | 1597.4 |
| AICC (smaller is better) | 1597.6 |
| BIC (smaller is better) | 1609.9 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 482.42 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.7011 | 0.2417 | 48.5 | 15.32 | <.0001 |
| age | | -0.05071 | 0.03695 | 43.5 | -1.37 | 0.1770 |
| type | AL | -0.08073 | 0.2485 | 42.6 | -0.32 | 0.7469 |
| type | ST | 0 | . | . | . | . |
| month | | -0.08140 | 0.02987 | 18 | -2.72 | 0.0139 |
| month2 | | 0.002468 | 0.001004 | 14.2 | 2.46 | 0.0274 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| age | 1 | 43.5 | 1.88 | 0.1770 |
| type | 1 | 42.6 | 0.11 | 0.7469 |
| month | 1 | 18 | 7.42 | 0.0139 |
| month2 | 1 | 14.2 | 6.04 | 0.0274 |

Again, the three variance components (patient, patient*time, and patient*time$^2$) are all positive, indicating more variation between the quadratic curves for each patient than expected by chance. Also, the month2 fixed effect was significant. So we go ahead and fit the cubic term for both fixed and random.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3/ S DDFM=SATTERTH;
RANDOM int month month2 month3/ SUB=pat TYPE=UN;
run;
```

WARNING: Stopped because of infinite likelihood.

Covariance Parameter Values
At Last Iteration

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.4271 |
| UN(2,1) | pat | 0.02673 |
| UN(2,2) | pat | 0.02403 |
| UN(3,1) | pat | -0.00196 |
| UN(3,2) | pat | -0.00130 |
| UN(3,3) | pat | 0.000087 |
| UN(4,1) | pat | 0.000051 |
| UN(4,2) | pat | 0.000015 |
| UN(4,3) | pat | -1.26E-6 |
| UN(4,4) | pat | 2.121E-8 |
| Residual | | 0.5238 |

In other words, the model did not converge trying to include the cubic variance component. When we look at the estimates of the covariance parameter values from the last interaction performed, it appears that the time-cubed variance component was heading toward zero. So we will run the model dropping the month3 term from the random statement, but still including the month3 fixed effect.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3/ S DDFM=SATTERTH;
```

```
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.5960 |
| UN(2,1) | pat | -0.04470 |
| UN(2,2) | pat | 0.02356 |
| UN(3,1) | pat | 0.001705 |
| UN(3,2) | pat | -0.00071 |
| UN(3,3) | pat | 0.000022 |
| Residual | | 0.5258 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1597.5 |
| AIC (smaller is better) | 1611.5 |
| AICC (smaller is better) | 1611.7 |
| BIC (smaller is better) | 1624.0 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 464.77 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.7385 | 0.2427 | 50 | 15.41 | <.0001 |
| age | | -0.04852 | 0.03683 | 43.3 | -1.32 | 0.1946 |
| type | AL | -0.06019 | 0.2485 | 43 | -0.24 | 0.8098 |
| type | ST | 0 | . | . | . | . |
| month | | -0.1182 | 0.03533 | 39.6 | -3.35 | 0.0018 |
| month2 | | 0.006543 | 0.002345 | 123 | 2.79 | 0.0061 |
| month3 | | -0.00011 | 0.000055 | 79.5 | -1.93 | 0.0569 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| age | 1 | 43.3 | 1.74 | 0.1946 |
| type | 1 | 43 | 0.06 | 0.8098 |
| month | 1 | 39.6 | 11.20 | 0.0018 |
| month2 | 1 | 123 | 7.78 | 0.0061 |
| month3 | 1 | 79.5 | 3.73 | 0.0569 |

The time3 fixed effect was almost significant (p=.0569), so we will run a model with a time4 fixed effect term added.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3 month4/ S DDFM=SATTERTH;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.5729 |
| UN(2,1) | pat | -0.03886 |
| UN(2,2) | pat | 0.02148 |
| UN(3,1) | pat | 0.001537 |
| UN(3,2) | pat | -0.00066 |
| UN(3,3) | pat | 0.000022 |
| Residual | | 0.5235 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1616.1 |
| AIC (smaller is better) | 1630.1 |
| AICC (smaller is better) | 1630.3 |
| BIC (smaller is better) | 1642.6 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 464.67 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.8012 | 0.2423 | 51.4 | 15.69 | <.0001 |
| age | | -0.04334 | 0.03673 | 44 | -1.18 | 0.2444 |
| type | AL | -0.05877 | 0.2468 | 43.3 | -0.24 | 0.8129 |
| type | ST | 0 | . | . | . | . |
| month | | -0.1908 | 0.04753 | 144 | -4.01 | <.0001 |
| month2 | | 0.01777 | 0.005493 | 497 | 3.24 | 0.0013 |
| month3 | | -0.00064 | 0.000241 | 490 | -2.64 | 0.0086 |
| month4 | | 7.455E-6 | 3.33E-6 | 565 | 2.24 | 0.0256 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| age | 1 | 44 | 1.39 | 0.2444 |
| type | 1 | 43.3 | 0.06 | 0.8129 |
| month | 1 | 144 | 16.12 | <.0001 |
| month2 | 1 | 497 | 10.47 | 0.0013 |
| month3 | 1 | 490 | 6.97 | 0.0086 |
| month4 | 1 | 565 | 5.01 | 0.0256 |

From our analysis, the fixed quartic term (month4) was significant. In the book, they said they did not have a significant quartic term, so did not continue. I went and added a month5 term to the dataset and ran the model with the quintic term (month5)

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3 month4 month5/ S
DDFM=SATTERTH;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.5930 |
| UN(2,1) | pat | -0.04490 |
| UN(2,2) | pat | 0.02364 |
| UN(3,1) | pat | 0.001770 |
| UN(3,2) | pat | -0.00075 |
| UN(3,3) | pat | 0.000025 |
| Residual | | 0.5200 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1643.2 |
| AIC (smaller is better) | 1657.2 |
| AICC (smaller is better) | 1657.3 |
| BIC (smaller is better) | 1669.7 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 465.04 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.7651 | 0.2462 | 52.9 | 15.29 | <.0001 |
| age | | -0.04517 | 0.03693 | 44 | -1.22 | 0.2277 |
| type | AL | -0.05772 | 0.2476 | 43.2 | -0.23 | 0.8167 |
| type | ST | 0 | . | . | . | . |
| month | | -0.1373 | 0.06920 | 332 | -1.98 | 0.0480 |
| month2 | | 0.005511 | 0.01239 | 558 | 0.44 | 0.6566 |
| month3 | | 0.000340 | 0.000902 | 497 | 0.38 | 0.7068 |
| month4 | | -0.00002 | 0.000028 | 438 | -0.85 | 0.3957 |
| month5 | | 3.303E-7 | 0 | 0 | . | . |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| age | 1 | 44 | 1.50 | 0.2277 |
| type | 1 | 43.2 | 0.05 | 0.8167 |
| month | 1 | 332 | 3.94 | 0.0480 |
| month2 | 1 | 558 | 0.20 | 0.6566 |
| month3 | 1 | 497 | 0.14 | 0.7068 |
| month4 | 1 | 438 | 0.72 | 0.3957 |
| month5 | 1 | 425 | 1.28 | 0.2585 |

We see that a fixed quintic effect was not significant, so we'll stop with the previous model and go ahead and test for interactions between time and type of cancer.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3 month4 type*month type*month2
type*month3 type*month4/
S DDFM=SATTERTH;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

WARNING: Stopped because of infinite likelihood.

Covariance Parameter Values
At Last Iteration

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.4143 |
| UN(2,1) | pat | 0.02576 |
| UN(2,2) | pat | 0.008990 |
| UN(3,1) | pat | -0.00103 |
| UN(3,2) | pat | -0.00020 |
| UN(3,3) | pat | 6.402E-6 |
| Residual | | 0.5096 |

Again, with this more complex model, we run into a problem with the month3 variance component heading toward zero. So we will back up a step and run a model without month4 included as a fixed effect.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3 type*month type*month2
type*month3/
```

```
S DDFM=SATTERTH;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.5972 |
| UN(2,1) | pat | -0.04689 |
| UN(2,2) | pat | 0.02376 |
| UN(3,1) | pat | 0.001740 |
| UN(3,2) | pat | -0.00071 |
| UN(3,3) | pat | 0.000022 |
| Residual | | 0.5248 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1618.4 |
| AIC (smaller is better) | 1632.4 |
| AICC (smaller is better) | 1632.5 |
| BIC (smaller is better) | 1644.9 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 463.50 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.9512 | 0.2794 | 56.3 | 14.14 | <.0001 |
| age | | -0.04903 | 0.03665 | 43.5 | -1.34 | 0.1879 |
| type | AL | -0.3603 | 0.3058 | 45.8 | -1.18 | 0.2448 |
| type | ST | 0 | . | . | . | . |
| month | | -0.2079 | 0.1296 | 410 | -1.60 | 0.1094 |
| month2 | | 0.006063 | 0.02236 | 579 | 0.27 | 0.7864 |
| month3 | | 0.000253 | 0.001083 | 558 | 0.23 | 0.8157 |
| month*type | AL | 0.1303 | 0.1360 | 299 | 0.96 | 0.3388 |
| month*type | ST | 0 | . | . | . | . |
| month2*type | AL | -0.00134 | 0.02250 | 582 | -0.06 | 0.9524 |
| month2*type | ST | 0 | . | . | . | . |
| month3*type | AL | -0.00034 | 0.001085 | 560 | -0.31 | 0.7544 |
| month3*type | ST | 0 | . | . | . | . |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| age | 1 | 43.5 | 1.79 | 0.1879 |
| type | 1 | 45.8 | 1.39 | 0.2448 |
| month | 1 | 299 | 4.40 | 0.0367 |
| month2 | 1 | 581 | 0.23 | 0.6320 |
| month3 | 1 | 560 | 0.02 | 0.8786 |
| month*type | 1 | 299 | 0.92 | 0.3388 |
| month2*type | 1 | 582 | 0.00 | 0.9524 |
| month3*type | 1 | 560 | 0.10 | 0.7544 |

There are no significant interactions between time effects and cancer type. Therefore, we can reasonably assume that changes in antibody level are similar for the two cancer types. We can also test for interactions between time effects and age.

```
PROC MIXED; CLASS type pat;
MODEL virus=age type month month2 month3 age*month age*month2
age*month3/ S;
RANDOM int month month2/ SUB=pat TYPE=UN;
run;
```

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate |
|---|---|---|
| UN(1,1) | pat | 0.6206 |
| UN(2,1) | pat | -0.05316 |
| UN(2,2) | pat | 0.02614 |
| UN(3,1) | pat | 0.002006 |
| UN(3,2) | pat | -0.00080 |
| UN(3,3) | pat | 0.000026 |
| Residual | | 0.5248 |

Fit Statistics

| | |
|---|---|
| -2 Res Log Likelihood | 1639.9 |
| AIC (smaller is better) | 1653.9 |
| AICC (smaller is better) | 1654.1 |
| BIC (smaller is better) | 1666.4 |

Null Model Likelihood Ratio Test

| DF | Chi-Square | Pr > ChiSq |
|---|---|---|
| 6 | 458.47 | <.0001 |

Solution for Fixed Effects

| Effect | type | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | 3.7860 | 0.2688 | 42 | 14.09 | <.0001 |
| age | | -0.06084 | 0.04618 | 490 | -1.32 | 0.1883 |
| type | AL | -0.05328 | 0.2488 | 490 | -0.21 | 0.8306 |
| type | ST | 0 | . | . | . | . |
| month | | -0.1463 | 0.06602 | 42 | -2.22 | 0.0321 |
| month2 | | 0.008624 | 0.004999 | 42 | 1.73 | 0.0919 |
| month3 | | -0.00015 | 0.000129 | 490 | -1.19 | 0.2332 |
| age*month | | 0.006353 | 0.01331 | 490 | 0.48 | 0.6332 |
| age*month2 | | -0.00049 | 0.001087 | 490 | -0.45 | 0.6552 |
| age*month3 | | 0.000011 | 0.000029 | 490 | 0.38 | 0.7016 |

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| age | 1 | 490 | 1.74 | 0.1883 |
| type | 1 | 490 | 0.05 | 0.8306 |
| month | 1 | 42 | 4.91 | 0.0321 |
| month2 | 1 | 42 | 2.98 | 0.0919 |
| month3 | 1 | 490 | 1.42 | 0.2332 |
| age*month | 1 | 490 | 0.23 | 0.6332 |
| age*month2 | 1 | 490 | 0.20 | 0.6552 |
| age*month3 | 1 | 490 | 0.15 | 0.7016 |

There also does not appear to be a significant age*time interaction and the overall age effect was not significant in any of the models that we tested. However, because it is known that average antibody levels decrease with age, age should be retained in the model. However, we can drop the age*time interactions.