## Assessing fixed effects

In our example so far, we have been concentrating on determining the covariance pattern. Now we'll look at the treatment effects estimates obtained from Model 6. Again, the SAS code for Model 6.

```
proc mixed noclprint data=dbp;
class trt pat visit;
model dbp=trt|visit dbp0/ddfm=satterth;
repeated visit/type=toep subject=pat group=trt r=1,3,4 rcorr=1,3,4;
lsmeans trt/ diff pdiff cl;
run;
```

Type 3 Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| trt | 2 | 184 | 4.05 | 0.0189 |
| visit | 3 | 449 | 12.46 | <.0001 |
| trt*visit | 6 | 339 | 1.75 | 0.1090 |
| dbp0 | 1 | 285 | 29.64 | <.0001 |

Least Squares Means

| Effect | trt | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|---|
| trt | A | 92.7437 | 0.7592 | 96.2 | 122.16 | <.0001 | 0.05 | 91.2367 | 94.2507 |
| trt | B | 91.4931 | 0.6402 | 93.5 | 142.91 | <.0001 | 0.05 | 90.2219 | 92.7644 |
| trt | C | 89.6992 | 0.7630 | 93.3 | 117.56 | <.0001 | 0.05 | 88.1841 | 91.2143 |

Differences of Least Squares Means

| Effect | trt | _trt | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha | Lower | Upper |
|---|---|---|---|---|---|---|---|---|---|---|
| trt | A | B | 1.2506 | 0.9941 | 186 | 1.26 | 0.2100 | 0.05 | -0.7107 | 3.2118 |
| trt | A | C | 3.0445 | 1.0757 | 191 | 2.83 | 0.0051 | 0.05 | 0.9228 | 5.1662 |
| trt | B | C | 1.7939 | 0.9973 | 181 | 1.80 | 0.0737 | 0.05 | -0.1740 | 3.7618 |

There are significant treatment, visit and baseline blood pressure effects. Patients given Treatment C had significantly lower blood pressure than patients given Treatment A.

## Example: Covariance pattern models for Count data

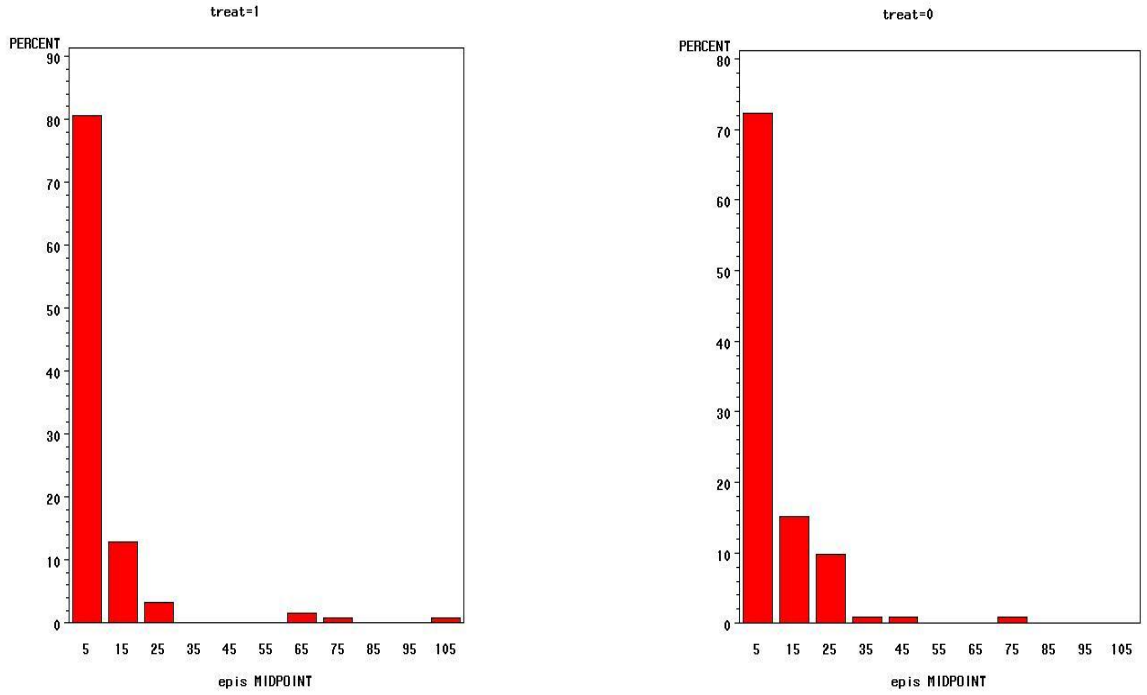
The data are from a study evaluating a new treatment for epilepsy. The trial was a placebo-controlled trial. There were 59 patients. Before treatment, epileptic seizures were counted for 8 weeks. After treatment, the number of seizures were reported every 2 weeks for 8 weeks. The following SAS code reads in the dataset and prints out the first 20 observations. Note that the log of the base count and the log of patient age have been calculated. The number of episodes have also been placed into 1 of 11 categories. The textbook does not include the patient age in their analyses, so the results and conclusions that they present are different. "SAS for Linear Models" by Littell, et al. also analyze these data and include the covariate log(age). Because the covariate has a significant effect on the number of seizures, I've included it in this example.

```
filename ep 'C:\...\epil.dat';
data epil;
infile ep;
input pat time treat epis base lbase age;
lage=log(age);
run;
```

| Obs | pat | time | treat | epis | base | lbase | age | lage |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 5 | 11 | 2.39790 | 31 | 3.43399 |
| 2 | 1 | 2 | 0 | 3 | 11 | 2.39790 | 31 | 3.43399 |
| 3 | 1 | 3 | 0 | 3 | 11 | 2.39790 | 31 | 3.43399 |
| 4 | 1 | 4 | 0 | 3 | 11 | 2.39790 | 31 | 3.43399 |
| 5 | 2 | 1 | 0 | 3 | 11 | 2.39790 | 30 | 3.40120 |
| 6 | 2 | 2 | 0 | 5 | 11 | 2.39790 | 30 | 3.40120 |
| 7 | 2 | 3 | 0 | 3 | 11 | 2.39790 | 30 | 3.40120 |
| 8 | 2 | 4 | 0 | 3 | 11 | 2.39790 | 30 | 3.40120 |
| 9 | 3 | 1 | 0 | 2 | 6 | 1.79176 | 25 | 3.21888 |
| 10 | 3 | 2 | 0 | 4 | 6 | 1.79176 | 25 | 3.21888 |
| 11 | 3 | 3 | 0 | 0 | 6 | 1.79176 | 25 | 3.21888 |
| 12 | 3 | 4 | 0 | 5 | 6 | 1.79176 | 25 | 3.21888 |
| 13 | 4 | 1 | 0 | 4 | 8 | 2.07944 | 36 | 3.58352 |
| 14 | 4 | 2 | 0 | 4 | 8 | 2.07944 | 36 | 3.58352 |
| 15 | 4 | 3 | 0 | 1 | 8 | 2.07944 | 36 | 3.58352 |
| 16 | 4 | 4 | 0 | 4 | 8 | 2.07944 | 36 | 3.58352 |
| 17 | 5 | 1 | 0 | 7 | 66 | 4.18965 | 22 | 3.09104 |
| 18 | 5 | 2 | 0 | 18 | 66 | 4.18965 | 22 | 3.09104 |
| 19 | 5 | 3 | 0 | 9 | 66 | 4.18965 | 22 | 3.09104 |
| 20 | 5 | 4 | 0 | 21 | 66 | 4.18965 | 22 | 3.09104 |

The following SAS code produces histograms by treatment of the number of seizures reported by each patient for each 2 week period.

```
proc gchart data=epil;
by treat;
vbar epis/type=percent midpoints=5 to 105 by 10;
run;
```




Notice that the majority of the patients have 10 or fewer seizures during each 2 week period, and that the number of patients in each of the larger categories drops quickly. This L shaped distribution indicates that a Poisson error may be appropriate. Because the periods are strictly 2 weeks, we don't need to use an offset. Also not that the small number of very large frequencies may produce outlying residuals, which could make the Poisson inappropriate.

PROC GENMOD uses "generalized estimating equations" or GEE, a generalized linear model analog of generalized least squares developed by Liang and Zeger (1986). Just like with PROC MIXED for normally distributed data, GEE allows you to fit a variety of correlation models when the data fit one of the distributions from the exponential family, as long as there are no other random-model effects. The first model that will be used to fit the epilepsy data will include the fixed effects of visit, treatment, the covariates of log(baseline) and log(age). The treatment*visit interaction term and the log(baseline)*treatment term to test for heterogeneous slopes are also included.

```
proc genmod;
class pat time treat;
model epis= treat time treat*time lbase treat*lbase lage / dist=p link=log type3;
repeated subject=pat/modelse type=cs corrw;
run;
```

GEE uses a "working correlation matrix" (corrw) to account for correlation among the repeated measures within subjects. The repeated statement is similar to the one used with PROC MIXED, where subject=pat creates a separate correlation matrix for each patient. type=cs defines the correlation pattern as compound symmetry. An equivalent type to CS in SAS is EXCH (exchangeable).

GEE Model Information

| | |
|---|---|
| Correlation Structure | Exchangeable |
| Subject Effect | pat (59 levels) |
| Number of Clusters | 59 |
| Correlation Matrix Dimension | 4 |
| Maximum Cluster Size | 4 |
| Minimum Cluster Size | 4 |

Algorithm converged.

Working Correlation Matrix

| | Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|---|
| Row1 | 1.0000 | 0.3579 | 0.3579 | 0.3579 |
| Row2 | 0.3579 | 1.0000 | 0.3579 | 0.3579 |
| Row3 | 0.3579 | 0.3579 | 1.0000 | 0.3579 |
| Row4 | 0.3579 | 0.3579 | 0.3579 | 1.0000 |

Exchangeable Working Correlation

| | |
|---|---|
| Correlation | 0.3579450915 |

The "GEE Model Information" lets us know the number of patients and the dimension of each block. The working correlation matrix and exchangeable working correlation are next. The observations at any two visits for the same patient have a correlation of 0.3579.

Score Statistics For Type 3 GEE Analysis

| Source | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|
| treat | 1 | 4.92 | 0.0265 |
| time | 3 | 5.04 | 0.1692 |
| time*treat | 3 | 1.54 | 0.6724 |
| lbase | 1 | 6.34 | 0.0118 |
| lbase*treat | 1 | 3.55 | 0.0595 |
| lage | 1 | 6.72 | 0.0095 |

The time*treatment interaction is not significant, so the analysis will be rerun without that term. Additional statements are added to test for equal slopes. Note that the alternative form of the regressions over log(base) for each treatment is used lbase(treat). The **e** and **diff** options have been added to the lsmeans statement. The **e** option requests that the coefficients used to compute the lsmeans be printed, while the **diff** requests the test for the treatment differences in their lsmeans.

```
proc genmod data=epil; class pat time treat;
model epis= treat time lbase(treat) lage / dist=p link=log type3;
repeated subject=pat/modelse type=cs corrw;
 lsmeans treat/e diff;
contrast 'lbase slopes=' lbase(treat) 1 -1;
```

Working Correlation Matrix

| | Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|---|
| Row1 | 1.0000 | 0.3552 | 0.3552 | 0.3552 |
| Row2 | 0.3552 | 1.0000 | 0.3552 | 0.3552 |
| Row3 | 0.3552 | 0.3552 | 1.0000 | 0.3552 |
| Row4 | 0.3552 | 0.3552 | 0.3552 | 1.0000 |

Exchangeable Working Correlation

| | |
|---|---|
| Correlation | 0.3551679728 |

Notice that dropping the treatment*visit interaction out of the model did not impact the correlation impact (.3579 vs. .3552).

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| Intercept | | -6.4597 | 1.2031 | -8.8178 | -4.1016 | -5.37 | <.0001 |
| **treat** | **0** | **2.1457** | **0.6601** | **0.8518** | **3.4395** | **3.25** | **0.0012** |
| treat | 1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| time | 1 | 0.2030 | 0.0987 | 0.0096 | 0.3964 | 2.06 | 0.0397 |
| time | 2 | 0.1344 | 0.0762 | -0.0149 | 0.2837 | 1.76 | 0.0776 |
| time | 3 | 0.1445 | 0.1228 | -0.0963 | 0.3852 | 1.18 | 0.2395 |
| time | 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| lbase(treat) | 0 | 0.9500 | 0.0986 | 0.7567 | 1.1432 | 9.64 | <.0001 |
| lbase(treat) | 1 | 1.5202 | 0.1423 | 1.2413 | 1.7992 | 10.68 | <.0001 |
| lage | | 0.9194 | 0.2773 | 0.3759 | 1.4630 | 3.32 | 0.0009 |

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| Intercept | | -6.4597 | 1.4685 | -9.3380 | -3.5814 | -4.40 | <.0001 |
| treat | 0 | 2.1457 | 0.7356 | 0.7039 | 3.5874 | 2.92 | 0.0035 |
| treat | 1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| time | 1 | 0.2030 | 0.1105 | -0.0136 | 0.4196 | 1.84 | 0.0663 |
| time | 2 | 0.1344 | 0.1122 | -0.0855 | 0.3543 | 1.20 | 0.2309 |
| time | 3 | 0.1445 | 0.1119 | -0.0749 | 0.3639 | 1.29 | 0.1967 |
| time | 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| lbase(treat) | 0 | 0.9500 | 0.1325 | 0.6903 | 1.2096 | 7.17 | <.0001 |
| lbase(treat) | 1 | 1.5202 | 0.1397 | 1.2465 | 1.7940 | 10.89 | <.0001 |

```
      lage            0.9194   0.3540   0.2256   1.6132    2.60   0.0094
      Scale           2.1172    .        .        .         .      .
NOTE: The scale parameter for GEE estimation was computed as the square root of
      the normalized Pearson's chi-square.
```

Both the empirical and model based estimates are presented. Although the difference between the empirical and model-based standard errors are not huge, the small difference may indicate that a more complex covariance pattern may be required or that the Poisson may not be the correct distribution. The treatment 0 parameter GEE parameter estimates presented above are the estimate of the treatment effect at 0 baseline epileptic episodes. None of the patients enrolled in the study had 0 baseline episodes, so this value would be outside of the inference space. The Type 3 analysis presented next is based on the Score statistics, while the difference of least squares means presented on the next page is based on the Wald statistics (see the empirical results above). The treatment difference of least squares means is calculated using the coefficients presented below. Prm1 is the intercept coefficient, Prm2 and Prm 3 are Treatment 0 and 1 coefficients, Prm4-Prm7 are the coefficients for the 4 times, Prm 8 is the log(baseline) for Treatment 0 coefficient, Prm 9 is the log(baseline) for Treatment 1 coefficient, and Prm 10 is the log(age) coefficient. The Prm8 coefficient for Treatment 0 and the Prm9 coefficient for Treatment 0 are the average log(baseline) values. So the treatment difference of least squares means is calculated using the average log(baseline) (about 23.4 baseline epileptic seizures), rather than the 0 baseline epileptic seizures used to calculate the empirical treatment effects.

Score Statistics For Type 3 GEE Analysis

| Source | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|
| treat | 1 | 5.00 | 0.0253 |
| time | 3 | 4.71 | 0.1941 |
| lbase(treat) | 2 | 9.94 | 0.0070 |
| lage | 1 | 6.47 | 0.0110 |

Coefficients for treat Least Squares Means

| Label | Row | Prm1 | Prm2 | Prm3 | Prm4 | Prm5 | Prm6 | Prm7 | Prm8 | Prm9 | Prm10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| treat | 1 | 1 | 1 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 3.1542 | 0 | 3.3198 |
| treat | 2 | 1 | 0 | 1 | 0.25 | 0.25 | 0.25 | 0.25 | 0 | 3.1542 | 3.3198 |

Least Squares Means

| Effect | treat | Estimate | Standard Error | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|
| treat | 0 | 1.8552 | 0.1047 | 1 | 313.92 | <.0001 |
| treat | 1 | 1.5084 | 0.1480 | 1 | 103.94 | <.0001 |

Differences of Least Squares Means

| Effect | treat | _treat | Estimate | Standard Error | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|---|
| treat | 0 | 1 | 0.3469 | 0.1798 | 1 | 3.72 | 0.0536 |

Contrast Results for GEE Analysis

| | | Chi- |
|---|---|---|

| Contrast | DF | Square | Pr > ChiSq | Type |
|---|---|---|---|---|
| lbase slopes= | 1 | 3.64 | 0.0565 | Score |

The contrast result shows some evidence of unequal slopes for the regression over log(base) for each treatment. This indicates that the size and statistical significance of the treatment effect will vary with log(base).

We can investigate further the differences between the treatment effects at 0 base and at the mean base by adding estimate and contrast statements to our SAS code.

```
proc genmod data=epil;
class pat time treat;
model epis= treat time lbase(treat) lage / dist=p link=log type3;
repeated subject=pat/modelse type=cs corrw;
lsmeans treat/e diff;
contrast 'lbase slopes=' lbase(treat) 1 -1;
estimate 'lsm trt diff at 0 base' treat 1 -1;
estimate 'lsm trt diff at mean base' treat 1 -1 lbase(treat) 3.1542
-3.1542;
contrast 'lsm trt diff at 0 base' treat 1 -1;
run;
```

The 3.1542 and -3.1542 values in the estimate statement are the Prm8 and Prm9 coefficient values used to calculate the treatment least squares means at the mean value of the baseline. Following are selected output from the analysis.

Differences of Least Squares Means

| Effect | treat | _treat | Estimate | Standard Error | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|---|
| treat | 0 | 1 | 0.3469 | 0.1798 | 1 | 3.72 | 0.0536 |

Contrast Estimate Results

| Label | Estimate | Standard Error | Alpha | Confidence Limits | |
|---|---|---|---|---|---|
| lsm trt diff at 0 base | 2.1457 | 0.6601 | 0.05 | 0.8518 | 3.4395 |
| lsm trt diff at mean base | 0.3469 | 0.1798 | 0.05 | -0.0054 | 0.6992 |

Contrast Estimate Results

| Label | Chi-Square | Pr > ChiSq |
|---|---|---|
| lsm trt diff at 0 base | 10.56 | 0.0012 |
| lsm trt diff at mean base | 3.72 | 0.0536 |

Contrast Results for GEE Analysis

| Contrast | DF | Chi-Square | Pr > ChiSq | Type |
|---|---|---|---|---|
| lbase slopes= | 1 | 3.64 | 0.0565 | Score |
| lsm trt diff at 0 base | 1 | 5.00 | 0.0253 | Score |

Note that the Type III Score treatment test is the treatment difference at a baseline number of episodes at 0. This is outside the parameter space, because none of the patients had a baseline number of episodes of 0. We can subtract the log(average number of baseline episodes from the log(base) to center the zero base.

```
data epil;
infile ep;
input pat time treat epis base lbase age;
lage=log(age);
lbas2=lbase-3.1542;

title 'Compound Symmetry -lbase-mean check for hetero. slope';
proc genmod data=epil;
class pat time treat;
model epis= treat time lbas2(treat) lage / dist=p link=log type3;
repeated subject=pat/modelse type=cs corrw;
lsmeans treat/e diff;
contrast 'lbase slopes=' lbas2(treat) 1 -1;
estimate 'lsm trt diff at mean base' treat 1 -1;
estimate 'lsm trt diff at zero base' treat 1 -1 lbas2(treat) -3.1542 3.1542;
contrast 'lsm trt diff at mean base' treat 1 -1;
run;
```

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| Intercept | | -1.6645 | 0.9566 | -3.5395 | 0.2104 | -1.74 | 0.0819 |
| treat | 0 | 0.3469 | **0.1798** | **-0.0054** | **0.6992** | **1.93** | **0.0536** |
| treat | 1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| time | 1 | 0.2030 | 0.0987 | 0.0096 | 0.3964 | 2.06 | 0.0397 |
| time | 2 | 0.1344 | 0.0762 | -0.0149 | 0.2837 | 1.76 | 0.0776 |
| time | 3 | 0.1445 | 0.1228 | -0.0963 | 0.3852 | 1.18 | 0.2395 |
| time | 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| lbas2(treat) | 0 | 0.9500 | 0.0986 | 0.7567 | 1.1432 | 9.64 | <.0001 |
| lbas2(treat) | 1 | 1.5202 | 0.1423 | 1.2413 | 1.7992 | 10.68 | <.0001 |
| lage | | 0.9194 | 0.2773 | 0.3759 | 1.4630 | 3.32 | 0.0009 |

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| Intercept | | -1.6645 | 1.1980 | -4.0125 | 0.6834 | -1.39 | 0.1647 |
| treat | 0 | 0.3469 | **0.1855** | **-0.0166** | **0.7104** | **1.87** | **0.0614** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| treat | 1 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| time | 1 | 0.2030 | 0.1105 | -0.0136 | 0.4196 | 1.84 | 0.0663 |
| time | 2 | 0.1344 | 0.1122 | -0.0855 | 0.3543 | 1.20 | 0.2309 |
| time | 3 | 0.1445 | 0.1119 | -0.0749 | 0.3639 | 1.29 | 0.1967 |
| time | 4 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | . | . |
| lbas2(treat) | 0 | 0.9500 | 0.1325 | 0.6903 | 1.2096 | 7.17 | <.0001 |
| lbas2(treat) | 1 | 1.5202 | 0.1397 | 1.2465 | 1.7940 | 10.89 | <.0001 |
| lage | | 0.9194 | 0.3540 | 0.2256 | 1.6132 | 2.60 | 0.0094 |
| Scale | | 2.1172 | . | . | . | . | . |

Score Statistics For Type 3 GEE Analysis

| Source | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|
| treat | 1 | 3.74 | 0.0531 |
| time | 3 | 4.71 | 0.1941 |
| lbas2(treat) | 2 | 9.94 | 0.0070 |
| lage | 1 | 6.47 | 0.0110 |

Coefficients for treat Least Squares Means

| Label | Row | Prm1 | Prm2 | Prm3 | Prm4 | Prm5 | Prm6 | Prm7 | Prm8 | Prm9 | Prm10 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| treat | 1 | 1 | 1 | 0 | 0.25 | 0.25 | 0.25 | 0.25 | 492E-7 | 0 | 3.3198 |
| treat | 2 | 1 | 0 | 1 | 0.25 | 0.25 | 0.25 | 0.25 | 0 | 492E-7 | 3.3198 |

Least Squares Means

| Effect | treat | Estimate | Standard Error | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|
| treat | 0 | 1.8552 | 0.1047 | 1 | 313.92 | <.0001 |
| treat | 1 | 1.5084 | 0.1480 | 1 | 103.94 | <.0001 |

Differences of Least Squares Means

| Effect | treat | _treat | Estimate | Standard Error | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|---|
| treat | 0 | 1 | 0.3469 | 0.1798 | 1 | 3.72 | 0.0536 |

Contrast Estimate Results

| Label | Estimate | Standard Error | Alpha | Confidence Limits | |
|---|---|---|---|---|---|
| lsm trt diff at mean base | 0.3469 | 0.1798 | 0.05 | -0.0054 | 0.6992 |
| lsm trt diff at zero base | 2.1457 | 0.6601 | 0.05 | 0.8518 | 3.4395 |

Contrast Estimate Results

| Label | Chi-Square | Pr > ChiSq |
|---|---|---|
| lsm trt diff at mean base | 3.72 | 0.0536 |
| lsm trt diff at zero base | 10.56 | 0.0012 |

Contrast Results for GEE Analysis

| Contrast | DF | Chi-Square | Pr > ChiSq | Type |
|---|---|---|---|---|
| lbase slopes= | 1 | 3.64 | 0.0565 | Score |
| lsm trt diff at mean base | 1 | 3.74 | 0.0531 | Score |

Using the log(base)-average log(base), puts the estimate for treatment difference within the parameter space. Now the treatment difference at the mean log(base) value is approaching significance. We know however, that there are heterogeneous treatment slopes for log(base). We can investigate this further by going back to the original log(base) model and add estimate statements for a range of log(base) values from the original scale numbers of 10, 20, 30, and 50.

```
proc genmod data=epil;
class pat time treat;
model epis= treat time lbase(treat) lage / dist=p link=log type3;
repeated subject=pat/modelse type=cs corrw;
lsmeans treat/e diff;
contrast 'lbase slopes=' lbase(treat) 1 -1;
estimate 'lsm trt diff at 0 base' treat 1 -1;
estimate 'lsm trt diff at 10 base' treat 1 -1 lbase(treat) 2.303
-2.303;
estimate 'lsm trt diff at 20 base' treat 1 -1 lbase(treat) 2.9957
-2.9957;
estimate 'lsm trt diff at mean base' treat 1 -1 lbase(treat) 3.1542
-3.1542;
estimate 'lsm trt diff at 30 base' treat 1 -1 lbase(treat) 3.4011
-3.4011;
estimate 'lsm trt diff at 50 base' treat 1 -1 lbase(treat) 3.9120
-3.9120;
```

Contrast Estimate Results

| Label | Chi-Square | Pr > ChiSq |
|---|---|---|
| lsm trt diff at 0 base | 10.56 | 0.0012 |
| lsm trt diff at 10 base | 8.49 | 0.0036 |
| lsm trt diff at 20 base | 5.01 | 0.0252 |
| lsm trt diff at mean base | 3.72 | 0.0536 |
| lsm trt diff at 30 base | 1.62 | 0.2035 |
| lsm trt diff at 50 base | 0.28 | 0.5938 |

We can see the heterogeneous treatment slopes for baseline epiliptic seizures. As the number of baseline seizures increase, the treatment difference decreases.

Because the compound symmetry covariance pattern may not complex enough, the analyses was rerun using three additional covariance patterns: AR(1), Toeplitz, and the Unstructured. Note that for the next three, only the parameter estimate for Treatment 0 (which is the same as the difference between treatments) is presented for both the empirical and model-based analysis.

AR(1):

Working Correlation Matrix

| | Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|---|
| Row1 | 1.0000 | 0.4759 | 0.2265 | 0.1078 |
| Row2 | 0.4759 | 1.0000 | 0.4759 | 0.2265 |
| Row3 | 0.2265 | 0.4759 | 1.0000 | 0.4759 |
| Row4 | 0.1078 | 0.2265 | 0.4759 | 1.0000 |

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| treat | 0 | 2.3759 | 0.6404 | 1.1207 | 3.6311 | 3.71 | 0.0002 |

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| treat | 0 | 2.3759 | 0.7233 | 0.9583 | 3.7935 | 3.28 | 0.0010 |

For the AR(1) analysis, the empirical standard error is smaller than the Model-based standard error. It appears that the AR(1) covariance pattern may fit the data slightly better than the compound symmetry. However, the small difference in the empirical and model-based standard error may indicate that the Poisson may not be the correct distribution.

Toeplitz:

Working Correlation Matrix

| | Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|---|
| Row1 | 1.0000 | 0.4771 | 0.0000 | 0.0000 |
| Row2 | 0.4771 | 1.0000 | 0.4771 | 0.0000 |
| Row3 | 0.0000 | 0.4771 | 1.0000 | 0.4771 |
| Row4 | 0.0000 | 0.0000 | 0.4771 | 1.0000 |

Notice that the off diagonal values greater than 1 apart are all zero, which doesn't seem quite right. You should always look at the SAS LOG when running analyses to make sure that the analyses did not have any problems. The SAS LOG for this analysis had the following notes, indicating that there was a problem with the correlation matrix becoming singular:

```
NOTE: The working correlation has been ridged with a maximum value of 0.3603775516 to
avoid singularity.
NOTE: The working correlation has been ridged with a maximum value of 0.3130979867 to
avoid singularity.
NOTE: The working correlation has been ridged with a maximum value of 0.3927971448 to
avoid singularity.
.
.
NOTE: The working correlation has been ridged with a maximum value of 0.4993798592 to
avoid singularity.
```

When I ran the exact same model as presented in the book using the Toeplitz covariance pattern, I ended up with the same problem. I am not sure why the analysis worked for the authors of the book, but doesn't work for me. Therefore, the results for the Toeplitz covariance pattern are suspect and won't be considered further.

Unstructured or general:

Working Correlation Matrix

| | Col1 | Col2 | Col3 | Col4 |
|---|---|---|---|---|
| Row1 | 1.0000 | 0.3149 | 0.2853 | 0.1707 |
| Row2 | 0.3149 | 1.0000 | 0.7431 | 0.3969 |
| Row3 | 0.2853 | 0.7431 | 1.0000 | 0.5199 |
| Row4 | 0.1707 | 0.3969 | 0.5199 | 1.0000 |

Empirical Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| treat | 0 | 2.4124 | 0.6503 | 1.1379 | 3.6869 | 3.71 | 0.0002 |

Model-Based Standard Error Estimates

| Parameter | | Estimate | Standard Error | 95% Confidence Limits | | Z | Pr > \|Z\| |
|---|---|---|---|---|---|---|---|
| treat | 0 | 2.4124 | 0.7515 | 0.9396 | 3.8852 | 3.21 | 0.0013 |

The results for the unstructured are fairly similar to the AR(1) and CS. Unlike using PROC MIXED for repeated measures, there are no quasi-likelihood or information criteria values outputted, so it is not possible to compare the models statistically. Notice that the empirical standard errors for all three models are similar. The empirical estimates reflect the different covariance between treatment groups, so are similar whatever model is fitted. Because of the slight differences in the emprical and model-based standard erros, it is possible that the Poisson distribution may not be appropriate. This may be due to the small number of very large frequencies that were noted on the figures, which could produce outlying residuals. Even though the Poisson model may not be appropriate, we will investigate the treatment differences, ignoring the significant differences in slopes over log(baseline) for different treatments. We will use the model-based results from the unstructured covariance pattern to look at relative rates and 95% confidence intervals.

The estimate of treatment difference is 2.4124. This gives us a relative rate of

seizure rate on placebo/seizure rate on active = exp(2.4124)=11.18.

We can get the confidence interval by exponentiating the 95% confidence limits from the output:

exp(.9396)=2.5589
exp(3.8852)=48.68

**Analysis using a categorical model**

The textbook continues on analyzing these data using a categorical mixed model. The categorize the response into 4 categories; 0, 1-3, 4-10, and 11+. They then used a special SAS macro written by Lipsitz et al. (1994) to fit the categorical model with compound symmetry, Toeplitz and general covariance patterns. They were only able to achieve convergence with the compound symmetry covariance pattern. The results that they present are on a model that does not include the log(age) covariate, so are not comparable to the analyses that have been presented above. If there is time at the end of the semester, we will revisit categorical mixed models.