Example 2: A mixed model with count data

The following example is from SAS System for Mixed Models. The experiment is a split-plot with different management methods as the whole-plot treatment factor and different seed mixes as the split-plot treatment factor. The whole-plots were arranged in randomized complete blocks. The measurement was the number of plants of a given species per experimental unit. There are 7 management types (TRT), 4 blocks (BLK), 4 types of seed mix (MIX) and the response is the number of plants (COUNT). The data are located in the splitplotcount.dat dataset. Because it is count data, we will assume a Poisson distribution. The model that we will be fitting is

$$\log(\mu_{ijk}) = a + r_i + \tau_j + (r\tau)_{ij} + \delta_k + (\tau\delta)_{jk}$$

where

μ_{ijk}	is the conditional mean count given the random effects
a	is the intercept
r_i	is the BLK effect
τ_j	is the TRT effect
$(r\tau)_{ij}$	is the BLK*TRT (whole-plot error) effect
δ_k	is the MIX effect
$(\tau\delta)_{jk}$	is the TRT*MIX interaction

$r_i \sim N(0,\sigma^2_R)$

$(r\tau)_{ij} \sim N(0,\sigma^2_{RT})$

The following SAS code reads in the data and prints out the first 19 observations

```
options ps=66 ls=80;
filename sp 'f:\kathy mixed model\splitplotcount.dat';
data spd;
infile sp;
input trt blk mix count;
run;

data select;
set spd;
if _n_<20;
proc print;
run;
```

The following is the printout of the first observations. There are 112 observations (7*4*4).

```
                  Obs     trt     blk     mix     count
```

```
 1      1      1      1      24
 2      1      1      2      12
 3      1      1      3       8
 4      1      1      4      13
 5      1      2      1       9
 6      1      2      2       9
 7      1      2      3       9
 8      1      2      4      18
 9      1      3      1      12
10      1      3      2       8
11      1      3      3      44
12      1      3      4       0
13      1      4      1       8
14      1      4      2      12
15      1      4      3      25
16      1      4      4       0
17      2      1      1      11
18      2      1      2      32
19      2      1      3      12
```

The GLIMMIX procedure is used for this analysis.

```
proc glimmix data=spd;
class trt blk mix;
model count=trt mix trt*mix/dist=poisson link=log ddfm=satterth;
lsmeans trt mix/diff;
random blk blk*trt;
run;
```

The class statement specifies the classification effects in the model. Because we are using the Poisson distribution, the model dependent variable (left hand side) is the count information. The independent variables on the right side of the equation include the fixed effects, which are TRT, MIX, and the TRT*MIX interaction effect. The random statement specifies the random effects (BLK, and whole plot error BLK*TRT) that are included in the model.

The first information produced by this program is the model information.

```
                    Model Information

     Data Set                      WORK.SPD
     Response Variable             count
     Response Distribution         Poisson
     Link Function                 Log
     Variance Function             Default
     Variance Matrix               Not blocked
     Estimation Technique          Residual PL
     Degrees of Freedom Method     Satterthwaite
```

The model information is correct for the model specified. The next output is the class level information and the dimension information for the design matrices (X for fixed effects and Z for random effects). The number of covariance parameters is 2 (BLK and BLK*TRT). The columns in X are 40, which is 1 for the intercept, + 7 for TRT, +4 for MIX, + 7*4=28 for TRT*MIX. The columns in Z are 32, which is 4 for the BLK, +4*7=28 for BLK*TRT.

```
                 Class Level Information

         Class    Levels    Values

         trt           7    1 2 3 4 5 6 7
         blk           4    1 2 3 4
         mix           4    1 2 3 4


      Number of Observations Read         112
      Number of Observations Used         112


                        Dimensions

          G-side Cov. Parameters         2
          Columns in X                  40
          Columns in Z                  32
          Subjects (Blocks in V)         1
          Max Obs per Subject          112
```

The next information presented is the information on the optimization. Because we cannot get a closed form solution for the variance components, we need to iteratively solve the equations until convergence is reached. The default optimization technique used in PROC GLIMMIX is the Dual Quasi-Newton. We are trying to estimate two variance components BLK and BLK*TRT which are in the G matrix. This analysis took five iterations to reach convergence. The output also has a message that the Estimated G matrix is not positive definite. This indicates that one of the variance components that was being estimated had a negative estimate. SAS automatically sets negative estimates to 0.

```
               Optimization Information

  Optimization Technique        Dual Quasi-Newton
  Parameters in Optimization    2
```

```
          Lower Boundaries              2
          Upper Boundaries              0
          Fixed Effects                 Profiled
          Starting From                 Data


                              Iteration History

                                   Objective                          Max
Iteration   Restarts Subiterations   Function        Change       Gradient

    0          0              4   472.45598993     2.00000000    44.26195
    1          0              4   594.57559733     0.59355256     50.0345
    2          0              2   624.66008319     0.06096869    50.37533
    3          0              2   626.86889196     0.00056067    50.36349
    4          0              1   626.88648763     0.00000153    50.36334
    5          0              0   626.88649572     0.00000000    50.36334

                Convergence criterion (PCONV=1.11022E-8) satisfied.

                    Estimated G matrix is not positive definite.
```

The next information presented is the goodness of fit. The only fit statistics that we will focus on is the Pearson Chi-Square/DF. This gives us an estimate of the overdispersion parameter. This parameter is telling us that there is quite a bit of overdispersion and that the statistical tests will be liberal.

```
                                Fit Statistics

                  -2 Res Log Pseudo-Likelihood        626.89
                  Pseudo-AIC  (smaller is better)     628.89
                  Pseudo-AICC (smaller is better)     628.94
                  Pseudo-BIC  (smaller is better)     628.27
                  Pseudo-CAIC (smaller is better)     629.27
                  Pseudo-HQIC (smaller is better)     627.54
                  Pearson Chi-Square                  634.08
                  Pearson Chi-Square / DF               7.55
```

The covariance parameter estimates shows us that the variance estimate for BLK was the parameter that was set to zero. Remember that the TRT effect is tested over the TRT*BLK variance estimate.

```
                        Covariance Parameter Estimates

                        Cov                        Standard
                        Parm        Estimate          Error

                        blk         5.93E-19              .
                        trt*blk       0.1652        0.05596
```

The Type III tests for the fixed effects are presented next. Note that the denominator degrees of freedom are 1, which does not make sense. This is because we asked for the Satterthwaite degrees of freedom which are based on the variance being fixed. SAS doesn't know how to handle the degrees of freedom calculation when you are not estimating overdispersion, so the degrees of freedom are nonsense. If we ran the same model without the ddfm=satterth, the denominator degrees of freedom for MIX and TRT*MIX become 63 and the Pr > F for both of these become <.0001.

```
                     Type III Tests of Fixed Effects

                          Num        Den
           Effect          DF         DF     F Value     Pr > F
           trt              6      20.26        3.64     0.0130
           mix              3          1       14.82     0.1882
           trt*mix         18          1       10.48     0.2391
```

There is a significant treatment effect. Even though both MIX and TRT*MIX have large F values, neither of them are statistically significant. This is because the probabilities have taken into account the overdispersion.

The least-squares means are presented next.

```
                        trt Least Squares Means

                             Standard
     trt     Estimate           Error         DF     t Value     Pr > |t|

     1         2.4973          0.2164      22.41       11.54       <.0001
     2         2.8825          0.2129      20.91       13.54       <.0001
     3         2.8279          0.2125      20.83       13.31       <.0001
     4         3.0979          0.2106      20.08       14.71       <.0001
     5         3.4404          0.2084      19.27       16.51       <.0001
     6         3.2579          0.2100      19.85       15.51       <.0001
     7         3.6928          0.2073      18.88       17.81       <.0001
```

```
                     Differences of trt Least Squares Means

                                      Standard
   trt     _trt     Estimate          Error        DF     t Value     Pr > |t|

   1       2         -0.3852       0.3036    21.65       -1.27       0.2180
   1       3         -0.3307       0.3033    21.61       -1.09       0.2877
   1       4         -0.6006       0.3020    21.23       -1.99       0.0597
   1       5         -0.9431       0.3004    20.81       -3.14       0.0050
   1       6         -0.7606       0.3016    21.11       -2.52       0.0198
   1       7         -1.1955       0.2997    20.61       -3.99       0.0007
   2       3         0.05451       0.3008    20.87        0.18       0.8580
   2       4         -0.2155       0.2995    20.49       -0.72       0.4800
   2       5         -0.5579       0.2979    20.08       -1.87       0.0757
   2       6         -0.3755       0.2991    20.38       -1.26       0.2235
   2       7         -0.8104       0.2972    19.88       -2.73       0.0130
   3       4         -0.2700       0.2992    20.45       -0.90       0.3774
   3       5         -0.6124       0.2977    20.04       -2.06       0.0529
   3       6         -0.4300       0.2988    20.34       -1.44       0.1654
   3       7         -0.8649       0.2969    19.84       -2.91       0.0087
   4       5         -0.3425       0.2963    19.67       -1.16       0.2615
   4       6         -0.1600       0.2974    19.97       -0.54       0.5966
   4       7         -0.5949       0.2955    19.47       -2.01       0.0581
   5       6          0.1825       0.2959    19.56        0.62       0.5445
   5       7         -0.2524       0.2940    19.07       -0.86       0.4012
   6       7         -0.4349       0.2951    19.36       -1.47       0.1567
```

There are several pairs of treatments that are different from each other. Again, these values are reported on the log scale, so would need to be exponentiated in order to interpret them on the odds scale. We will do that in a later analysis. The least-squares means for MIX are presented next. When we looked at the overall affect of MIX earlier, it did not have a significant affect on count. However, when we look at the differences in the least-squares means for mixed, almost all of the comparisons are significantly different. This is because these tests did not take into account the overdispersion. The next analysis we'll do will take into account the overdispersion.

```
                           mix Least Squares Means
                            Standard
          mix     Estimate       Error        DF     t Value     Pr > |t|

          1        3.2716     0.08608    27.45       38.01       <.0001
          2        3.1716     0.08726       29       36.35       <.0001
          3        3.0641     0.08708    28.76       35.19       <.0001
          4        2.8907     0.08955    32.16       32.28       <.0001
                     Differences of mix Least Squares Means

                                      Standard
    mix     _mix     Estimate          Error         DF     t Value     Pr > |t|

    1       2         0.09999       0.05602        84        1.79       0.0779
    1       3          0.2075       0.05574        84        3.72       0.0004
```

```
    1       4         0.3809     0.05951        84        6.40       <.0001
    2       3         0.1075     0.05755        84        1.87       0.0653
    2       4         0.2809     0.06122        84        4.59       <.0001
    3       4         0.1734     0.06096        84        2.84       0.0056
```

Accounting for overdispersion

The following PROC GLIMMIX code accounts for overdispersion.

```
proc glimmix data=spd;
class trt blk mix;
model count=trt mix trt*mix/dist=poisson link=log ddfm=satterth;
lsmeans trt mix/diff;
ods output diffs=diff;
random blk blk*trt;
random _residual_;
run;
```

The code is the same as the previous model, except that we now have an additional random statement. This fits a polynomial Poisson regression model with overdispersion where the variance function $v(\mu)$ is replaced by $\phi v(\mu)$. There is also an ODS output statement. This is for outputting the differences from the LSMEANS in order to calculate the confidence intervals. The code for this will be presented later.

The model and class information are the same as for the previous model. However the output on the dimensions and optimization have changed.

```
                                   Dimensions

                        G-side Cov. Parameters          2
                        R-side Cov. Parameters          1
                        Columns in X                   40
                        Columns in Z                   32
                        Subjects (Blocks in V)          1
                        Max Obs per Subject           112
```

The dimension has the addition of the R-side covariance parameter, which is the overdispersion parameter. The optimization information has the additional line concerning the residual variance.

```
                              Optimization Information

                    Optimization Technique        Dual Quasi-Newton
                    Parameters in Optimization    2
                    Lower Boundaries              2
                    Upper Boundaries              0
                    Fixed Effects                 Profiled
                    Residual Variance             Profiled
                    Starting From                 Data
```

The iteration history shows that more rounds were needed before convergence was reached. The more parameters that are being estimated, the more iterations that are needed for convergence.

```
                         Iteration History

                                        Objective                  Max
Iteration  Restarts  Subiterations       Function        Change  Gradient

  0          0              5     178.13148738     2.00000000  880.3884
  1          0              3     206.97129676     0.36514487  855.9358
  2          0              4     211.65787334     0.08370848  874.6871
  3          0              2     211.74945361     0.00518381  876.5665
  4          0              2       211.748509     0.00014994  876.6344
  5          0              1     211.74839278     0.00000325  876.6363
  6          0              1     211.74839003     0.00000051  876.6365
  7          0              1      211.7483896     0.00000043  876.6364
  8          0              0     211.74838996     0.00000000  876.6364

              Convergence criterion (PCONV=1.11022E-8) satisfied.
```

The estimated G matrix is still not positive definite.

```
                  Estimated G matrix is not positive definite.

                              Fit Statistics

                  -2 Res Log Pseudo-Likelihood        211.75
                  Pseudo-AIC  (smaller is better)     215.75
                  Pseudo-AICC (smaller is better)     215.90
                  Pseudo-BIC  (smaller is better)     214.52
                  Pseudo-CAIC (smaller is better)     216.52
                  Pseudo-HQIC (smaller is better)     213.05
                  Pearson Chi-Square                  805.57
                  Pearson Chi-Square / DF               9.59
```

Note that the Pearson Chi-Square/DF statistic in this model is larger than for the model without accounting for overdispersion. In the model without accounting for the overdispersion, a portion of it went into the BLK*TRT variance component estimate, so the estimate was smaller. When we look at the variance parameter estimates, we see that the BLK*TRT estimate is smaller than the (0.1652) that was estimated in the earlier model. The “Residual (VC)” parameter is the overdispersion parameter estimate.

```
                    Covariance Parameter Estimates

                                                Standard
                     Cov Parm         Estimate       Error

                     blk                     0           .
                     trt*blk           0.04877     0.04729
                     Residual (VC)      9.5901      1.7028
```

Next are the Type III tests of fixed effects. The TRT effect probability in the previous model was 0.03. Remember that the TRT effect is tested over the TRT*BLK variance component. The decrease in the size of that variance component estimate allows for the greater precision.

```
                 Type III Tests of Fixed Effects

                        Num       Den
        Effect           DF        DF    F Value    Pr > F

        trt               6     22.61       3.80    0.0092
        mix               3     63.44       1.55    0.2115
        trt*mix          18     63.44       1.09    0.3802
```

The treatment least-squares means and differences between least-squares means are presented next. These results are fairly similar to the results that we got with the model without overdispersion. Those treatments that were significantly different before are still different, and those that were not are still not.

```
                         trt Least Squares Means

                             Standard
          trt     Estimate      Error       DF    t Value    Pr > |t|

          1         2.5062     0.2538    55.67       9.87      <.0001
          2         2.9855     0.2089    30.04      14.29      <.0001
          3         2.8682     0.2155     33.9      13.31      <.0001
          4         3.1401     0.1974    24.48      15.90      <.0001
          5         3.4728     0.1766    15.64      19.66      <.0001
          6         3.3502     0.1862    19.16      17.99      <.0001
          7         3.7016     0.1672    12.52      22.14      <.0001
```

```
                 Differences of trt Least Squares Means

                                 Standard
 trt     _trt     Estimate        Error       DF     t Value     Pr > |t|

 1       2         -0.4793       0.3287    43.43       -1.46       0.1520
 1       3         -0.3620       0.3330    45.33       -1.09       0.2827
 1       4         -0.6339       0.3216    40.67       -1.97       0.0555
 1       5         -0.9665       0.3092    35.72       -3.13       0.0035
 1       6         -0.8440       0.3148    37.73       -2.68       0.0108
 1       7         -1.1954       0.3040    33.74       -3.93       0.0004
 2       3          0.1173       0.3001    31.96        0.39       0.6986
 2       4         -0.1546       0.2874    27.24       -0.54       0.5951
 2       5         -0.4873       0.2735     22.5       -1.78       0.0884
 2       6         -0.3647       0.2798    24.43       -1.30       0.2046
 2       7         -0.7161       0.2676    20.68       -2.68       0.0143
 3       4         -0.2719       0.2923    29.13       -0.93       0.3599
 3       5         -0.6045       0.2786     24.3       -2.17       0.0400
 3       6         -0.4820       0.2848    26.27       -1.69       0.1024
 3       7         -0.8333       0.2728    22.43       -3.06       0.0057
 4       5         -0.3327       0.2649    19.89       -1.26       0.2237
 4       6         -0.2102       0.2714    21.76       -0.77       0.4470
 4       7         -0.5615       0.2587    18.13       -2.17       0.0435
 5       6          0.1225       0.2566    17.37        0.48       0.6390
 5       7         -0.2288       0.2432    14.04       -0.94       0.3627
 6       7         -0.3513       0.2502    15.71       -1.40       0.1798
```

When we look at the MIX results, they are different from the results from the previous model. The estimates for the differences are the same. However the standard error estimates are larger. The probabilities for the differences are in line with the overall MIX effect probability.

```
                         mix Least Squares Means

                             Standard
          mix     Estimate      Error       DF    t Value    Pr > |t|

          1         3.3185     0.1258    81.35      26.39      <.0001
          2         3.2185     0.1334    83.21      24.13      <.0001
          3         3.1110     0.1322    83.02      23.53      <.0001
```

```
         4        2.9376      0.1472    83.96      19.96      <.0001


                        Differences of mix Least Squares Means

                                          Standard
   mix      _mix     Estimate          Error       DF     t Value     Pr > |t|

   1        2         0.09999       0.1735    63.44       0.58       0.5664
   1        3          0.2075       0.1726    63.44       1.20       0.2338
   1        4          0.3809       0.1843    63.44       2.07       0.0429
   2        3          0.1075       0.1782    63.44       0.60       0.5485
   2        4          0.2809       0.1896    63.44       1.48       0.1434
   3        4          0.1734       0.1888    63.44       0.92       0.3619
```

Ignoring the Poisson distribution and treating the data as normally distributed

Before we had software available for analyzing generalized linear mixed models, many researchers treated the data as having an underlying normal distribution and relied on the robustness of the normal distribution. The following code analyzes the data in this way.

```
proc mixed data=spd;
class trt blk mix;
model count=trt mix trt*mix/ddfm=satterth;
lsmeans trt mix/diff;
random blk blk*trt;
run;
```

Only some of the results will be presented. Just as with the GLMM, the BLK variance parameter was so small that the estimate was set to zero.

```
                      Covariance Parameter
                                Estimates

                             Cov Parm      Estimate
                             blk                  0
                             trt*blk        18.6070
                             Residual        315.49
```

The Type III treatment effects are fairly similar to the GLMM.

```
                         Type 3 Tests of Fixed Effects

                                 Num      Den
                 Effect           DF       DF    F Value    Pr > F

                 trt               6       21       4.33    0.0054
                 mix               3       63       2.44    0.0729
                 trt*mix          18       63       1.10    0.3763
```

When we look at the differences in the TRT and MIX LSMeans, they are also similar in the probabilities.

```
                  Differences of Least Squares Means
```

```
                                                 Standard
 Effect   trt   mix   _trt   _mix   Estimate       Error    DF   t Value   Pr > |t|

 trt      1           2               -7.6875     6.9814    21    -1.10     0.2833
 trt      1           3               -4.6250     6.9814    21    -0.66     0.5149
 trt      1           4              -10.8750     6.9814    21    -1.56     0.1342
 trt      1           5              -20.0000     6.9814    21    -2.86     0.0093
 trt      1           6              -18.2500     6.9814    21    -2.61     0.0162
 trt      1           7              -29.8750     6.9814    21    -4.28     0.0003
 trt      2           3                3.0625     6.9814    21     0.44     0.6654
 trt      2           4               -3.1875     6.9814    21    -0.46     0.6527
 trt      2           5              -12.3125     6.9814    21    -1.76     0.0923
 trt      2           6              -10.5625     6.9814    21    -1.51     0.1452
 trt      2           7              -22.1875     6.9814    21    -3.18     0.0045
 trt      3           4               -6.2500     6.9814    21    -0.90     0.3808
 trt      3           5              -15.3750     6.9814    21    -2.20     0.0390
 trt      3           6              -13.6250     6.9814    21    -1.95     0.0644
 trt      3           7              -25.2500     6.9814    21    -3.62     0.0016
 trt      4           5               -9.1250     6.9814    21    -1.31     0.2053
 trt      4           6               -7.3750     6.9814    21    -1.06     0.3028
 trt      4           7              -19.0000     6.9814    21    -2.72     0.0128
 trt      5           6                1.7500     6.9814    21     0.25     0.8045
trt      5           7                -9.8750     6.9814    21    -1.4      0.1719
trt      6           7               -11.6250     6.9814    21    -1.67     0.1107
mix            1            2         2.4643     4.7471    63     0.52     0.6055
mix            1            3         8.3571     4.7471    63     1.76     0.0832
mix            1            4        11.3929     4.7471    63     2.40     0.0194
mix            2            3         5.8929     4.7471    63     1.24     0.2191
mix            2            4         8.9286     4.7471    63     1.88     0.0646
mix           3           4          3.0357     4.7471    63     0.64     0.5248
```

Going back to the PROC GLIMMIX model accounting for overdispersion, the following code calculates the confidence intervals and then exponentiates them to put them on an interpretable scale.

```
  data cl_hat;
  set diff;
  if effect='trt';
  tp=(tinv(.975,df));
  cl=exp(estimate-tp*stderr);
  cu=exp(estimate+tp*stderr);
  phat=exp(estimate);

  proc print data=cl_hat noobs;
  var trt _trt estimate stderr df probt tp phat cl cu;
run;
```

```
trt  _trt  Estimate    StdErr     DF   Probt      tp       phat       cl        cu

 1    2     -0.4793    0.3287  43.43  0.1520  2.01611  0.61922  0.31916  1.20139
 1    3     -0.3620    0.3330  45.33  0.2827  2.01370  0.69628  0.35610  1.36140
 1    4     -0.6339    0.3216  40.67  0.0555  2.02004  0.53054  0.27707  1.01588
 1    5     -0.9665    0.3092  35.72  0.0035  2.02865  0.38040  0.20314  0.71233
```

```
1     6      -0.8440    0.3148  37.73  0.0108  2.02487  0.42998  0.22731  0.81335
1     7      -1.1954    0.3040  33.74  0.0004  2.03281  0.30260  0.16312  0.56132
2     3       0.1173    0.3001  31.96  0.6986  2.03704  1.12443  0.61013  2.07226
2     4      -0.1546    0.2874  27.24  0.5951  2.05100  0.85677  0.47516  1.54488
2     5      -0.4873    0.2735   22.5  0.0884  2.07120  0.61431  0.34861  1.08254
2     6      -0.3647    0.2798  24.43  0.2046  2.06198  0.69438  0.38996  1.23646
2     7      -0.7161    0.2676  20.68  0.0143  2.08159  0.48867  0.27998  0.85291
3     4      -0.2719    0.2923  29.13  0.3599  2.04483  0.76196  0.41915  1.38513
3     5      -0.6045    0.2786   24.3  0.0400  2.06254  0.54633  0.30752  0.97059
3     6      -0.4820    0.2848  26.27  0.1024  2.05451  0.61754  0.34400  1.10860
3     7      -0.8333    0.2728  22.43  0.0057  2.07155  0.43459  0.24700  0.76467
4     5      -0.3327    0.2649  19.89  0.2237  2.08673  0.71701  0.41252  1.24623
4     6      -0.2102    0.2714  21.76  0.4470  2.07521  0.81046  0.46147  1.42338
4     7      -0.5615    0.2587  18.13  0.0435  2.09981  0.57036  0.33129  0.98197
5     6       0.1225    0.2566  17.37  0.6390  2.10642  1.13034  0.65832  1.94078
5     7      -0.2288    0.2432  14.04  0.3627  2.14416  0.79548  0.47223  1.33999
 6     7      -0.3513    0.2502  15.71  0.1798  2.12311  0.70375  0.41369  1.19719
```