GEEs: SAS Syntax and Examples

The **repeated** statement is used in **genmod** to fit GLMs by GEE. The format of the repeated statement is given by: **repeated subject=** *subject-effect*/**options;**, where *subject-effect* is a **class** variable for which observations having the same level of this effect are treated as repeated measures. The data set does not need to be sorted by subject. The working correlation structure is specified through the **type =** option with valid choices being AR or AR(1) for first order autoregressive, EXCH or CS for an exchangeable or compound symmetric structure, IND for an independence model, MDEP(m) for an m-dependent structure (with a value for m specified), and UN for unstructured. You can specify your own working correlation matrix with the **type=user** option.

Other popular options include: **corrw** gives the estimated working correlation matrix, **corrb** and **covb** give the model based and empirical correlation matrix and covariance matrix for the estimated regression effects. If you just wish the empirical summaries use **ecorrb** and **ecovb**, while **mcorrb** and **mcovb** give only the model based summaries. A **within=***within-subject* option is used to specify an ordering for unequally spaced repeated measures or repeated measures with missing time points. In the **mixed** procedure this was implemented by following the repeated statement with the class variable that specified the ordering. In **genmod**, this appears as an option.

The GEE analysis will include the standard ML and standard errors that are appropriate when all responses are independent, plus the GEE parameter estimates and empirical standard errors based on the specified working correlation structure. Model based standard errors are obtained using the **modelse** option on the repeated statement.

I will note that the **scale** and **noscale** options are specified with the model statement, but might be ignored with the GEE fit. For example, the first two examples below specify the noscale option, which is obeyed by the ML fit. In the logistic analysis, the GEE procedures are based on a fixed scale, while the Poisson GEE analysis estimates the scale parameter. In

principle, a scale parameter is not required with a GEE analysis, but SAS may insist that it be used. Also, I believe that the scale parameter in the GEE analysis may be estimated using the Pearson statistic based on the ML fit, rather than estimated iteratively, as described in the notes. This has no effect on the large sample distribution of the GEE estimates of the regression effects.

The next 3 sections contains analyses for 3 data sets. A careful analysis would include looking at residual plots, something that you should be familiar with for GLMs. Although the same ideas apply to GLMs fit by GEEs, I have neglected to include any residual analyses with the output. I will, however, note that fewer diagnostics are available for GEE fits than for ML fits.

Example 1: Binary Data (Respiratory Disorder Data)

The data for this example are used as one of the SAS online examples illustrating the genmod procedure. The data are available under "Gee Model for Binary Data" in the SAS/STAT Sample Program Library. The data were originally analyzed in Stokes, Davis, and Koch (1995) using a SAS macro to fit a GEE model.

Patients in each of two centers were randomly assigned to groups that received either an active treatment (A) for a respiratory disorder or a placebo (P). During treatment, respiratory status (coded here as 0=poor, 1=good) is determined for each of four visits (visit1 to visit4). The variables center (1 and 2), treatmnt (A and P), sex (M and F), and baseline (baseline respiratory status, 0 and 1) have two levels. The age variable (at time of entry into the study) is continuous. Complete data was available on 111 patients.

In the first SAS program, the data are in multivariate format, with 1 record per individual. The data step creates three 0-1 variables from treatment, center and sex. These were used in the SAS analysis and by Stokes et al. but are not really needed - you could use class variables in the genmod procedure instead. All 3 variable creation statements use comparison operators. For example, the statement **active = (treatmnt='A')** is equivalent

to **if treatmnt = 'A' then active = 1; else active = 0;**. Similarly, center2 is 1 for data collected from the second center, and 0 for data collected from center 1, whereas female is 1 for women and 0 for men. For plotting purposes, I created a variable actbycen = 10*active + center2 which has 4 distinct values, one for each of the four combinations of center and treatment: 0 and 1 for subjects receiving placebo in centers 1 and 2; 10 and 11 for subjects receiving the active treatment in centers 1 and 2.

There are two other aspects of the data step. The four statements of the form **visit=I; outcome=visitI; output;** transform the data from a multivariate format to a regression format with 4 records per subject (easier than using the transpose procedure but destroys original data format). The response variable is now labeled **outcome**. The **keep** statement indicates which of the original and created variables are kept in the data set for later use. Only a small subset of the input data and the created data set are provided with this handout.

After printing the data set, I used the means procedure to compute the mean outcome (which is just the sample proportion with a good outcome) by time, for each combination of sex, center, and treatment (the latter 2 captured in the variable actbycen). These are printed, and then plotted by sex, giving two plots of the sample proportions over time.

```
options ls = 79;
data resp;
     keep id active center center2 female age baseline visit outcome actbycen;
     input center id treatmnt $ sex $ age baseline visit1-visit4;
     active  = (treatmnt='A');
     center2 = (center=2);
     female  = (sex='F');
     actbycen = 10*active + center2;
     visit=1;  outcome=visit1;  output;
     visit=2;  outcome=visit2;  output;
     visit=3;  outcome=visit3;  output;
     visit=4;  outcome=visit4;  output;
     cards;
  1  1 P M 46 0 0 0 0 0
  1  2 P M 28 0 0 0 0 0
  1  3 A M 23 1 1 1 1 1
  1  4 P M 44 1 1 1 1 0

  2  1 P F 39 0 0 0 0 0
  2  2 A M 25 0 0 1 1 1
  2  3 A M 58 1 1 1 1 1
  2  4 P F 51 1 1 0 1 1
;
```

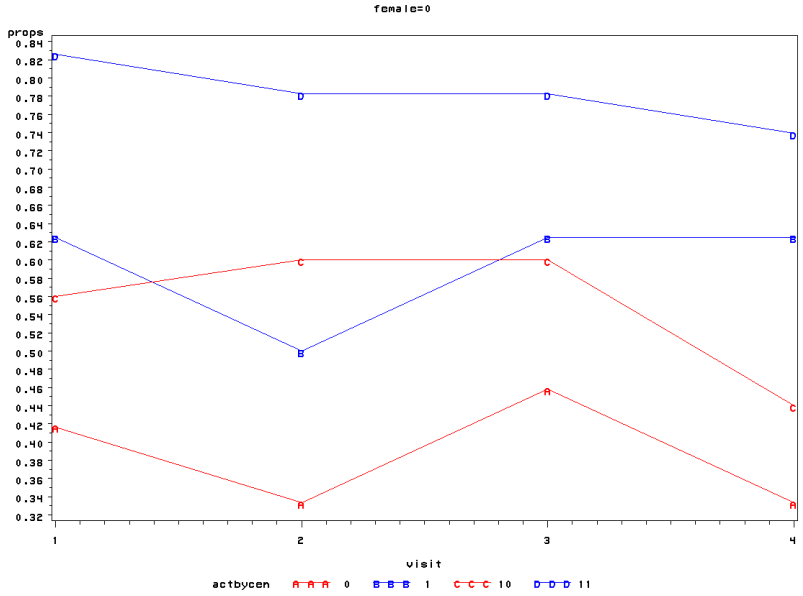
```
proc print data=resp;
  var id active center2 female age baseline visit outcome actbycen;

proc sort data=resp out=resp;
  by female actbycen visit;

proc means noprint data = resp;
  by female actbycen visit;
  var outcome;
  output out=d5 mean=props;

proc print data = d5;

symbol1 color=red   interpol=join value=a   height=1;
symbol2 color=blue  interpol=join value=b   height=1;
symbol3 color=red   interpol=join value=c   height=1;
symbol4 color=blue  interpol=join value=d   height=1;

proc gplot;
  by female;
  plot props*visit = actbycen;
run;
```

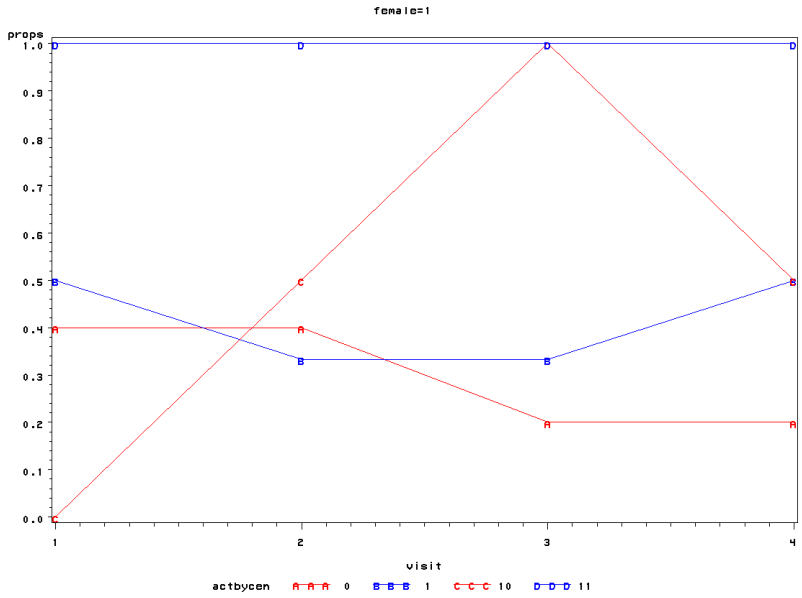
Obs	id	active	center2	female	age	baseline	visit	outcome	actbycen
1	1	0	0	0	46	0	1	0	0
2	1	0	0	0	46	0	2	0	0
3	1	0	0	0	46	0	3	0	0
4	1	0	0	0	46	0	4	0	0
9	3	1	0	0	23	1	1	1	10
10	3	1	0	0	23	1	2	1	10
11	3	1	0	0	23	1	3	1	10
12	3	1	0	0	23	1	4	1	10
225	1	0	1	1	39	0	1	0	1
226	1	0	1	1	39	0	2	0	1
227	1	0	1	1	39	0	3	0	1
228	1	0	1	1	39	0	4	0	1
229	2	1	1	0	25	0	1	0	11
230	2	1	1	0	25	0	2	1	11
231	2	1	1	0	25	0	3	1	11
232	2	1	1	0	25	0	4	1	11

Proportions:

Obs	female	actbycen	visit	_TYPE_	_FREQ_	props
1	0	0	1	0	24	0.41667
2	0	0	2	0	24	0.33333
3	0	0	3	0	24	0.45833
4	0	0	4	0	24	0.33333
5	0	1	1	0	16	0.62500
6	0	1	2	0	16	0.50000
7	0	1	3	0	16	0.62500
8	0	1	4	0	16	0.62500
9	0	10	1	0	25	0.56000
10	0	10	2	0	25	0.60000
11	0	10	3	0	25	0.60000
12	0	10	4	0	25	0.44000

13	0	11	1	0	23	0.82609
14	0	11	2	0	23	0.78261
15	0	11	3	0	23	0.78261
16	0	11	4	0	23	0.73913
17	1	0	1	0	5	0.40000
18	1	0	2	0	5	0.40000
19	1	0	3	0	5	0.20000
20	1	0	4	0	5	0.20000
21	1	1	1	0	12	0.50000
22	1	1	2	0	12	0.33333
23	1	1	3	0	12	0.33333
24	1	1	4	0	12	0.50000
25	1	10	1	0	2	0.00000
26	1	10	2	0	2	0.50000
27	1	10	3	0	2	1.00000
28	1	10	4	0	2	0.50000
29	1	11	1	0	4	1.00000
30	1	11	2	0	4	1.00000
31	1	11	3	0	4	1.00000
32	1	11	4	0	4	1.00000





The plots of the sample proportions suggest that there are some differences between centers and treatments, but differences between sexes are harder to assess in part because of the small number of women in the study - look at the frequencies reported in the listing of the sample proportions.

Stokes et al. consider a logistic regression model for these data. Let Y_{ij} be the respiratory status of patient i at the j^{th} visit, $j = 1, ..., 4$ and let p_{ij} be the corresponding probability of a good outcome for this patient at that visit: $\Pr(Y_{ij} = 1) = p_{ij}$. They assume

$$\log\left(\frac{p_{ij}}{1-p_{ij}}\right) = \beta_0 + \beta_1 \text{center2}_i + \beta_2 \text{active}_i + \beta_3 \text{female}_i + \beta_4 \text{age}_i + \beta_5 \text{baseline}_i,$$

where center2, active, female, and baseline are the binary variables defined above. The model does not include an effect for time, over and above adjusting for a subject's baseline level. The trajectory of the sample proportions do not show a strong time effect. It is reasonable to assume a priori that the 4 post-baseline responses within an individual are correlated.

I did two analyses. The first analysis specifies an exchangeable structure, while the second allows the working correlation matrix to be unstructured. Each analysis requests the estimated working correlation matrix (**corrw**) and model based standard errors (**modelse**) with the output. Both analyses fix the scale parameter at 1 (can binary data be overdispersed?). Each of the effects in the model is treated as a predictor. The same results would be obtained had I defined center2, active, sex and baseline as **class** variables, provided the baseline levels for the class variables correspond to 0 for the predictor variables (i.e. 0 for center2 is center=1; 0 for active is placebo group; 0 for female is male; 0 for baseline is 0 or poor).

Note that the **subject=** effect is **id** nested within **center**, so id and center must be class variables. This explains why Stokes et al. kept both center and center2 in the data set - they wished to use a center effect as a predictor in the model, but also needed to use it as class variable in the repeated statement. Also note that the **descend** option is used on the model statement, so we are modeling the probability of a good outcome (outcome=1) instead of a poor outcome (outcome=0).

Important features in the output:

1. The first analysis provides the complete output. The second analysis only includes the GEE output.

2. (First analysis) The Deviance and Pearson statistics can be unreliable when responses are correlated over time, whether a scale parameter is used or not.

3. (First analysis) The table titled 'Analysis Of Initial Parameter Estimates' gives the standard ML analysis ignoring the repeated measures.

4. (First analysis) The GEE output identifies that we requested an compound symmetric correlation matrix, and gives the estimated (common) correlation as .33. Two tables of parameter estimates and standard errors are given. The two sets of parameter

estimates are necessarily identical - they are just the GEE based estimates using the working correlation matrix. The two sets of standard errors are the empirical and model based estimates.

5. (First analysis) The ML and GEE based parameter estimates agree (but they will not in general agree nor do they agree if you add other effects to this model, such as a quadratic term in age). However, the ML based standard errors are noticeably smaller than the two GEE-based estimates, which are similar to each other. A result of this discrepancy is that the p-values for the regression effects are smaller in the ML based analysis that ignores the correlation (a common result with positively correlated responses). Notice that the p-value for age is about .03 in the ML analysis and about .17 in the GEE analyses.

6. The second analysis assumes an unstructured correlation matrix, but gives similar estimates and standard errors as the compound symmetric GEE analysis. Am implication here is that the choice of the correlation structure is not important nor is the decision on whether to use the empirical or model based procedures.

This analysis is consistent with many analyses that I have done in that the GEE methods gives similar point estimates but different standard errors than ML, except if you specify an independence structure which essentially gives the ML fit for independent responses.

```
proc genmod data=resp descend;
   class id center;
   model outcome=center2 active female age baseline/
         dist=bin link=logit noscale;
   repeated subject=id(center) /type=exch corrw modelse;

proc genmod data=resp descend;
   class id center;
   model outcome=center2 active female age baseline/
         dist=bin link=logit noscale;
   repeated subject=id(center) /type=un corrw modelse;

FIRST ANALYSIS:
```

The GENMOD Procedure

Model Information

Data Set	WORK.RESP
Distribution	Binomial
Link Function	Logit
Dependent Variable	outcome

Number of Observations Read	444
Number of Observations Used	444
Number of Events	248
Number of Trials	444

Class Level Information

Class	Levels	Values
id	56	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56
center	2	1 2

Response Profile

Ordered Value	outcome	Total Frequency
1	1	248
2	0	196

PROC GENMOD is modeling the probability that outcome='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	center2
Prm3	active
Prm4	female
Prm5	age
Prm6	baseline

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	438	487.4873	1.1130
Scaled Deviance	438	487.4873	1.1130
Pearson Chi-Square	438	444.7942	1.0155
Scaled Pearson X2	438	444.7942	1.0155
Log Likelihood		-243.7436	

Algorithm converged.

Analysis Of Initial Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-0.8561	0.3351	-1.5130	-0.1992	6.53	0.0106

center2	1	0.6495	0.2383	0.1825	1.1165	7.43	0.0064
active	1	1.2654	0.2350	0.8048	1.7259	28.99	<.0001
female	1	0.1368	0.2933	-0.4381	0.7117	0.22	0.6410
age	1	-0.0188	0.0088	-0.0360	-0.0015	4.53	0.0334
baseline	1	1.8457	0.2393	1.3768	2.3147	59.51	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

GEE Model Information

Correlation Structure	Exchangeable
Subject Effect	id(center) (111 levels)
Number of Clusters	111
Correlation Matrix Dimension	4
Maximum Cluster Size	4
Minimum Cluster Size	4

Algorithm converged.

Working Correlation Matrix

	Col1	Col2	Col3	Col4
Row1	1.0000	0.3270	0.3270	0.3270
Row2	0.3270	1.0000	0.3270	0.3270
Row3	0.3270	0.3270	1.0000	0.3270
Row4	0.3270	0.3270	0.3270	1.0000

Exchangeable Working Correlation

Correlation 0.3270345338

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	-0.8561	0.4564	-1.7506	0.0384	-1.88	0.0607
center2	0.6495	0.3532	-0.0428	1.3418	1.84	0.0660
active	1.2654	0.3467	0.5859	1.9448	3.65	0.0003
female	0.1368	0.4402	-0.7261	0.9996	0.31	0.7560
age	-0.0188	0.0130	-0.0442	0.0067	-1.45	0.1480
baseline	1.8457	0.3460	1.1676	2.5238	5.33	<.0001

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	-0.8561	0.4717	-1.7807	0.0684	-1.81	0.0695
center2	0.6495	0.3354	-0.0078	1.3068	1.94	0.0528
active	1.2654	0.3308	0.6171	1.9136	3.83	0.0001
female	0.1368	0.4129	-0.6724	0.9460	0.33	0.7404
age	-0.0188	0.0124	-0.0431	0.0056	-1.51	0.1306
baseline	1.8457	0.3368	1.1857	2.5058	5.48	<.0001
Scale	1.0000	.	.	.	.	.

NOTE: The scale parameter was held fixed.

SECOND ANALYSIS:

GEE Model Information

Correlation Structure	Unstructured
Subject Effect	id(center) (111 levels)
Number of Clusters	111
Correlation Matrix Dimension	4
Maximum Cluster Size	4
Minimum Cluster Size	4

Algorithm converged.

Working Correlation Matrix

	Col1	Col2	Col3	Col4
Row1	1.0000	0.3351	0.2140	0.2953
Row2	0.3351	1.0000	0.4429	0.3581
Row3	0.2140	0.4429	1.0000	0.3964
Row4	0.2953	0.3581	0.3964	1.0000

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	-0.8882	0.4568	-1.7835	0.0071	-1.94	0.0519
center2	0.6558	0.3512	-0.0326	1.3442	1.87	0.0619
active	1.2442	0.3455	0.5669	1.9214	3.60	0.0003
female	0.1128	0.4408	-0.7512	0.9768	0.26	0.7981
age	-0.0175	0.0129	-0.0427	0.0077	-1.36	0.1728
baseline	1.8981	0.3441	1.2237	2.5725	5.52	<.0001

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	-0.8882	0.4773	-1.8237	0.0473	-1.86	0.0628
center2	0.6558	0.3384	-0.0075	1.3191	1.94	0.0527
active	1.2442	0.3340	0.5895	1.8989	3.72	0.0002
female	0.1128	0.4169	-0.7044	0.9299	0.27	0.7868
age	-0.0175	0.0125	-0.0421	0.0070	-1.40	0.1619
baseline	1.8981	0.3410	1.2297	2.5665	5.57	<.0001
Scale	1.0000	.	.	.	.	.

NOTE: The scale parameter was held fixed.

As a follow up analysis one might consider omitting the sex effect, which is far from significant (p-value $>$ .50) in each of the GEE analyses. The output below was obtained after omitting the sex effect. Looking at the output, which assumes an exchangeable working

correlation structure, one might also consider whether the age effect might be safely omitted. If excluding age does not change the other estimates or standard errors by much, it may not matter whether age is ultimately excluded, at least if the focus is on the interpretation of the effect of individual predictors on outcome. Based on the model with four effects, the predicted probabilities from the GEE analysis satisfy:

$$\log\left(\frac{\hat{p}_{ij}}{1-\hat{p}_{ij}}\right) = -.85 + .67\text{center2}_i + 1.24\text{active}_i - .018\text{age}_i + 1.83\text{baseline}_i.$$

Recalling the definitions of center2 (1 if center = 2 and 0 if center =1), active (1 if active treatment, 0 if placebo), baseline (1 if good at baseline, 0 otherwise), we see that the predicted probability of a good response is higher for (a) center 2 than 1 (b) active treatment than placebo (c) younger subjects and (d) for subjects with good response at baseline (holding all other variables constant).

Given that the logit transform is the log-odds of a good outcome, we can interpret the regression coefficient for an effect in the model (and in general for any effect in a logistic regression model that does not appear as part of a higher order effect such as an interaction) as the change in log-odds associated with a unit increase in the predictor, holding all other variables constant. For example, the log-odds of a good outcome for someone that receives the active treatment is

$$\log\left(\frac{\hat{p}_{ij}}{1-\hat{p}_{ij}}\right) = -.85 + .67\text{center2}_i + 1.24 - .018\text{age}_i + 1.83\text{baseline}_i$$

for any fixed levels of center2, age and baseline, while the corresponding log-odds is

$$\log\left(\frac{\hat{p}_{ij}}{1-\hat{p}_{ij}}\right) = -.85 + .67\text{center2}_i - .018\text{age}_i + 1.83\text{baseline}_i$$

for someone that receives a placebo. The difference between the two is 1.24, the estimated treatment effect. Exponentiating the differences in log-odds ($\exp(1.22) = 3.46$) is the ratio of the odds of a good outcome for someone receiving the active treatment relative to the placebo. This is called an **adjusted odds-ratio** (adjusted for the other effects). A **crude**

odds ratio would be obtained from a model with no other effects but treatment. One can get a CI for the adjusted odds ratio by exponentiating the endpoints of the CI for the individual regression estimates, but I will leave that to you to do if you are interested.

```
proc genmod data=resp descend;
   class id center2;
   model outcome=center active age baseline /dist=bin link=logit noscale;
   repeated subject=id(center2) /type=exch corrw modelse;
```

The GENMOD Procedure

PROC GENMOD is modeling the probability that outcome='1'.

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	center2
Prm3	active
Prm4	age
Prm5	baseline

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	439	487.7051	1.1109
Scaled Deviance	439	487.7051	1.1109
Pearson Chi-Square	439	444.7508	1.0131
Scaled Pearson X2	439	444.7508	1.0131
Log Likelihood		-243.8526	

Algorithm converged.

Analysis Of Initial Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-0.8516	0.3343	-1.5067	-0.1964	6.49	0.0108
center2	1	0.6689	0.2348	0.2087	1.1292	8.11	0.0044
active	1	1.2406	0.2286	0.7925	1.6886	29.45	<.0001
age	1	-0.0178	0.0086	-0.0346	-0.0010	4.32	0.0376
baseline	1	1.8336	0.2376	1.3680	2.2992	59.57	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

GEE Model Information

Correlation Structure	Exchangeable
Subject Effect	id(center) (111 levels)
Number of Clusters	111
Correlation Matrix Dimension	4
Maximum Cluster Size	4

```
          Minimum Cluster Size                                  4

  Algorithm converged.

                        Working Correlation Matrix
                     Col1         Col2         Col3         Col4
        Row1       1.0000       0.3261       0.3261       0.3261
        Row2       0.3261       1.0000       0.3261       0.3261
        Row3       0.3261       0.3261       1.0000       0.3261
        Row4       0.3261       0.3261       0.3261       1.0000

                      Exchangeable Working
                           Correlation

                      Correlation    0.3260978045

                   Analysis Of GEE Parameter Estimates
                    Empirical Standard Error Estimates
                          Standard   95% Confidence
      Parameter Estimate    Error        Limits            Z Pr > |Z|

      Intercept  -0.8516   0.4528  -1.7391   0.0359   -1.88   0.0600
      center2     0.6689   0.3398   0.0030   1.3349    1.97   0.0490
      active      1.2406   0.3238   0.6060   1.8751    3.83   0.0001
      age        -0.0178   0.0120  -0.0413   0.0057   -1.48   0.1378
      baseline    1.8336   0.3373   1.1725   2.4946    5.44   <.0001

                   Analysis Of GEE Parameter Estimates
                   Model-Based Standard Error Estimates
                          Standard   95% Confidence
      Parameter Estimate    Error        Limits            Z Pr > |Z|

      Intercept  -0.8516   0.4701  -1.7730   0.0699   -1.81   0.0701
      center2     0.6689   0.3303   0.0216   1.3163    2.03   0.0428
      active      1.2406   0.3215   0.6104   1.8708    3.86   0.0001
      age        -0.0178   0.0120  -0.0414   0.0058   -1.48   0.1394
      baseline    1.8336   0.3341   1.1787   2.4885    5.49   <.0001
      Scale       1.0000        .        .        .       .        .

  NOTE: The scale parameter was held fixed.
```

Example 2: Count Data (Epilepsy Seizure Data)

These data, from Thall and Vail (1990), concern the treatment of people suffering from epileptic seizure episodes. These data are also analyzed in Diggle, Liang, and Zeger's (1994) text on longitudinal data analysis. The data consist of the number of epileptic seizures in an eight-week baseline period, before any treatment, and in each of four two-week treatment

periods, in which patients received either a placebo or the drug Progabide in addition to other therapy.

A portion of the data is displayed the SAS program below. As with the previous analysis, the data for this example are one of the SAS online examples illustrating the genmod procedure. See the "Gee Model for Count Data, Exchangeable Correlation" in the SAS/STAT Sample Program Library for the complete data set.

Let Y_{ij} be the number of epileptic seizures for subject i in interval j, where t_j is the length of interval $j = 1, 2, ..., 5$. Here $t_1 = 8$ and the other t_j are 2. A reasonable model for these data would be to assume that the responses Y_{ij} have marginal Poisson distributions, but that the responses within an individual are correlated over time. An important aspect of this problem is that the observation length is 8 weeks for period 1 and 2 weeks for the next 4 periods. The standard way to incorporate the different length periods is to write the mean of the Poisson distribution in the form $\mathrm{E}(Y_{ij}) = \mu^*_{ij} = \mu_{ij} t_i$ where μ_{ij} is the mean response per unit time (a week). This mean per unit time is directly comparable across different periods, so we will be modeling the μ_{ij} as a function of predictor variables. With a log link $\log(\mu^*_{ij}) = \log(\mu_{ij}) + \log(t_i)$, so a log-linear model for the mean response per unit time will lead to a log-linear model for the μ^*_{ij} with an **offset** term - the $\log(t_i)$.

The original data set (**thall**) is in regression format with 1 record for each post-baseline count. The response is labeled **y**. The baseline count is included in each post-baseline record, in column **bline**. Other columns in the data set are individual **id**, **visit** number (1-4), treatment (**trt** with levels 0 for placebo and 1 for progabide), and **age** at baseline.

The desired model for these data includes the baseline as one of the longitudinal time points. The data needs to be transformed to include a new row for each subject with baseline levels. This transformation is performed in the **new** data step, so now **visit** has 5 levels (0-4, with 0 as baseline). The next data step (new2) creates a log time interval variable for use as an offset, and an indicator variable for whether the observation is for a baseline measurement or for a post-baseline visit. Patient 207, who had 152 seizures during baseline,

and averaged 75 seizures per subsequent two-week periods, is deleted as an outlier, as in the Diggle, Liang, and Zeger (1994) analysis. Several records from **new2** are provided with the output.

After creating the data set to be analyzed, I used the **means** procedure to compute the mean number of seizures per observation period by treatment. I then standardized by the interval length, printed the raw and standardized means, and then log-transformed the standardized means. The log of the sample mean per unit time provides information about how $\log(\mu_{ij}) = log(\mu^*_{ij}) - \log(t_i)$ might depend on time and treatment.

Looking at the data plots we see that the mean number of seizures per week is higher in the placebo group, that the mean appears to increase slightly from baseline in the placebo group, but decrease from baseline in the progabide group.

An **important** feature to recognize is that the unstandardized cell means provide a strong indication of **overdispersion**. If responses within each treatment-by-time combination have marginal Poisson distributions, then the unstandardized sample means should estimate both the population mean **and** the population variance. Thus, the estimated unstandardized cell means should be approximately the square of the estimated standard deviations. I am considering the unstandardized means because the standard deviations provided are from the raw data. Here, the standard deviations are approximately equal to the means, or larger. This clear indication of overdispersion may be due to heterogeneity - this analysis may ignore certain important covariates that produce the extra-Poisson variation.

```
data thall;
input id y visit trt bline age;
cards;
104 5 1  0 11 31
104 3 2  0 11 31
104 3 3  0 11 31
104 3 4  0 11 31

116  7 1 0 66 22
116 18 2 0 66 22
116  9 3 0 66 22
116 21 4 0 66 22

130 5 1  0 23 37
130 6 2  0 23 37
130 6 3  0 23 37
130 5 4  0 23 37
```

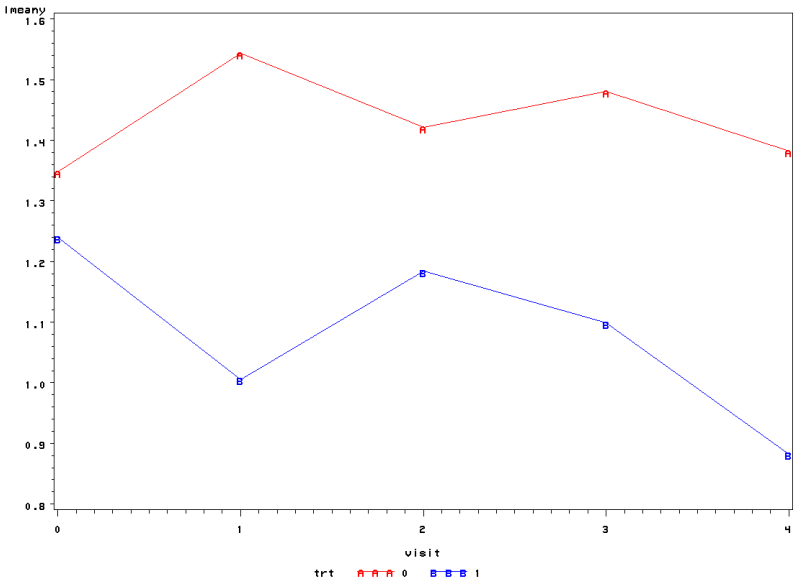
```
135 14 1 0 10 28
135 13 2 0 10 28
135  6 3 0 10 28
135  0 4 0 10 28

141 26 1 0 52 36
141 12 2 0 52 36
141  6 3 0 52 36
141 22 4 0 52 36
;

data new;
   set thall (drop=age);
   output;
   if visit=1 then do;
      y=bline; visit=0;
      output;
end;


proc sort;
   by id visit;

data new2;
   set new;
   if id ne 207;
   if visit=0 then do;
     x1=0; ltime=log(8);
     end;
     else do;
       x1=1; ltime=log(2);
     end;
 run;

proc print data=new2;

proc sort data=new2 out=new2;
  by trt visit;

proc means noprint data = new2;
  by trt visit;
  var y;
  output out=new3 mean=mean std=std;


data new4;
  set new3;
  if visit = 0 then meany = mean/8;
     else meany = mean/2;

proc print data = new4;

data new5;
  set new4;
  lmeany = log(meany);

symbol1 color=red   interpol=join value=a   height=1;
symbol2 color=blue  interpol=join value=b   height=1;

proc gplot data = new5;
  plot lmeany*visit = trt;

run;
```

TRANSFORMED DATA SET:

Obs	id	y	visit	trt	bline	x1	ltime
1	101	76	0	1	76	0	2.07944
2	101	11	1	1	76	1	0.69315
3	101	14	2	1	76	1	0.69315
4	101	9	3	1	76	1	0.69315
5	101	8	4	1	76	1	0.69315
6	102	38	0	1	38	0	2.07944
7	102	8	1	1	38	1	0.69315
8	102	7	2	1	38	1	0.69315
9	102	9	3	1	38	1	0.69315
10	102	4	4	1	38	1	0.69315
11	103	19	0	1	19	0	2.07944
12	103	0	1	1	19	1	0.69315
13	103	4	2	1	19	1	0.69315
14	103	3	3	1	19	1	0.69315
15	103	0	4	1	19	1	0.69315

MEANS PER WEEK

Obs	trt	visit	_TYPE_	_FREQ_	mean	std	meany
1	0	0	0	28	30.7857	26.1043	3.84821
2	0	1	0	28	9.3571	10.1369	4.67857
3	0	2	0	28	8.2857	8.1643	4.14286
4	0	3	0	28	8.7857	14.6726	4.39286
5	0	4	0	28	7.9643	7.6278	3.98214
6	1	0	0	30	27.6333	17.3811	3.45417
7	1	1	0	30	5.4667	5.7639	2.73333
8	1	2	0	30	6.5333	5.6062	3.26667
9	1	3	0	30	6.0000	7.3719	3.00000
10	1	4	0	30	4.8333	4.2838	2.41667



The original SAS analysis considered a simple two-factor model with interaction

$$\log(\mu_{ij}) = \beta_0 + \beta_1 \text{x1}_i + \beta_2 \text{trt}_i + \beta_3 \text{x1}_i * \text{trt}_i$$

or equivalently

$$\log(\mu^*_{ij}) = \beta_0 + \beta_1 \text{x1}_i + \beta_2 \text{trt}_i + \beta_3 \text{x1}_i * \text{trt}_i + \log(t_i).$$

Note that x1 is 0 for the baseline visit and 1 for subsequent visits, so the model includes a binary time effect. In genmod, the model is specified in terms of the Poisson response Y_{ij}, which has mean μ^*_{ij}. The model for μ^*_{ij} has an offset term with a fixed and known regression coefficient (1).

The interaction model includes 4 effects, one for each of the 4 combinations of 2 treatments and 2 times (baseline, post-baseline). The original SAS analysis treats x1 and trt as predictors rather than class variables (no big deal). The following table lists the mean structure as a function of time and treatment.

Treatment	Visit	$log(\mu_{ij})$
Placebo	Baseline	β_0
Placebo	1-4	$\beta_0 + \beta_1$
Progabide	Baseline	$\beta_0 + \beta_2$
Progabide	1-4	$\beta_0 + \beta_1 + \beta_2 + \beta_3$

Thus, β_0 is the mean (on log scale) at baseline for the baseline treatment (trt=0 is placebo); β_2 is the difference in means at baseline between the progabide and placebo groups - this is the main effect for treatment; β_1 is the change over time from baseline in the placebo group - this is the main effect for time; β_3 is the effect of interaction between time and treatment. If $\beta_3 = 0$ then the difference between baseline and post-baseline is the same (β_1) for both treatments, and the difference between placebo and progabide (β_2) is independent of time. Alternatively, if $\beta_3 < 0$ then the change from baseline in the placebo group (β_1) is greater than the change from baseline in the progabide group ($\beta_1 + \beta_3$) with the reverse implication if $\beta_3 > 0$. Thus, a value of $\beta_3 < 0$ indicates a reduction in the expected change in the seizure rate from baseline associated with the treatment.

Our first analysis fits this regression model, assuming an exchangeable working correlation structure for the marginal Poisson counts. I specified that the scale parameter should be fixed at 1, and I used the **offset=** option to specify the column containing the offset term. You should not include the offset term as a predictor in the model. If you do, a regression coefficient will be estimated for this effect. As in the logistic analyses, I request that SAS provide the model based standard errors and the estimated working correlation matrix.

Some important features of the output

1. The GEE analyses estimate a scale parameter, even though a noscale option was used in the model statement. I am not sure why, but I am thankful! The estimated scale is approximately 3.25, which is very different from 1. The large scale effect is consistent with the overdispersion observed in the table of means, so even if a non-longitudinal

analysis was appropriate (i.e. if correlation within an individual was small) it should include a correction for overdispersion!

2. As in earlier analyses, the ML and GEE based parameter estimates are similar, but the ML based standard errors are much smaller than either the model based or empirical standard errors from the GEE analysis. This is consistent with the large (.59) correlation between responses estimated under the exchangeable correlation model, and the large effect of overdispersion.

3. The usual ML analysis that ignores correlation suggests that the interaction between time and treatment is significant (p-value $< .0001$). Since the estimated coefficient for the interaction effect is negative, this suggests that change from baseline is lower in the treatment group than in the placebo group.

4. In the usual ML analysis, all effects are significant at the 5% level.

5. Using either the empirical or model based standard errors in the GEE analysis, no effects are significant at the 5% level - only the interaction is close to being significant.

6. The empirical and model based standard errors are similar, suggesting that the working correlation structure is adequate.

7. The three estimate commands, coupled with the estimate for the intercept, provide estimates and standard errors for the 4 distinct log-means corresponding to the 4 combinations of time and treatment. The point estimates are fairly similar, which should not be too surprising given the data plot. The estimate command always used the empirical covariance matrix to compute standard errors.

```
proc genmod data=new2;
   class id;
   model y=x1 trt x1*trt/noscale dist=poi offset=ltime;
   repeated subject=id/corrw type=exch modelse;
   estimate 'placebo post baseline' Intercept 1 x1 1;
   estimate 'prog at baseline'      Intercept 1 trt 1;
```

```
   estimate 'prog post baseline'    Intercept 1 x1 1
                                    trt 1 x1*trt 1;
run;
```

The GENMOD Procedure

Model Information

Data Set	WORK.NEW2
Distribution	Poisson
Link Function	Log
Dependent Variable	y
Offset Variable	ltime

Number of Observations Read	290
Number of Observations Used	290

Class Level Information

Class	Levels	Values
id	58	101 102 103 104 106 107 108 110 111 112 113 114 116 117 118 121 122 123 124 126 128 129 130 135 137 139 141 143 145 147 201 202 203 204 205 206 208 209 210 211 213 214 215 217 218 219 220 221 222 225 226 227 228 230 232 234 236 238

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	x1
Prm3	trt
Prm4	x1*trt

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	286	2413.0245	8.4371
Scaled Deviance	286	2413.0245	8.4371
Pearson Chi-Square	286	3015.1555	10.5425
Scaled Pearson X2	286	3015.1555	10.5425
Log Likelihood		5631.7547	

Algorithm converged.

Analysis Of Initial Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	1.3476	0.0341	1.2809	1.4144	1565.44	<.0001
x1	1	0.1108	0.0469	0.0189	0.2027	5.58	0.0181
trt	1	-0.1080	0.0486	-0.2034	-0.0127	4.93	0.0264
x1*trt	1	-0.3016	0.0697	-0.4383	-0.1649	18.70	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

NOTE: The scale parameter was held fixed.

GEE Model Information

Correlation Structure	Exchangeable
Subject Effect	id (58 levels)
Number of Clusters	58

Correlation Matrix Dimension	5
Maximum Cluster Size	5
Minimum Cluster Size	5

Algorithm converged.

Working Correlation Matrix

	Col1	Col2	Col3	Col4	Col5
Row1	1.0000	0.5941	0.5941	0.5941	0.5941
Row2	0.5941	1.0000	0.5941	0.5941	0.5941
Row3	0.5941	0.5941	1.0000	0.5941	0.5941
Row4	0.5941	0.5941	0.5941	1.0000	0.5941
Row5	0.5941	0.5941	0.5941	0.5941	1.0000

Exchangeable Working Correlation

Correlation	0.5941485833

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	1.3476	0.1574	1.0392	1.6560	8.56	<.0001
x1	0.1108	0.1161	-0.1168	0.3383	0.95	0.3399
trt	-0.1080	0.1937	-0.4876	0.2716	-0.56	0.5770
x1*trt	-0.3016	0.1712	-0.6371	0.0339	-1.76	0.0781

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

Parameter	Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept	1.3476	0.1106	1.1309	1.5644	12.19	<.0001
x1	0.1108	0.1233	-0.1308	0.3524	0.90	0.3687
trt	-0.1080	0.1579	-0.4176	0.2015	-0.68	0.4940
x1*trt	-0.3016	0.1936	-0.6811	0.0779	-1.56	0.1193
Scale	3.2469	.	.	.	.	

NOTE: The scale parameter for GEE estimation was computed as the square root of the normalized Pearson's chi-square.

Contrast Estimate Results

Label	Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square
placebo post baseline	1.4584	0.1896	0.05	1.0869	1.8300	59.19
prog at baseline	1.2396	0.1129	0.05	1.0183	1.4609	120.53
prog post baseline	1.0488	0.1577	0.05	0.7396	1.3580	44.20

Contrast Estimate Results

Label	Pr > ChiSq
placebo post baseline	<.0001
prog at baseline	<.0001
prog post baseline	<.0001

As a follow-up analysis, one might delete the interaction term (which is marginally significant in the GEE analyses) and check whether either main effect is significant. I considered an alternative approach. I treated x1 and trt as class variables, and fit a log linear model with a single effect x1*trt (i.e. no main effects), plus an intercept and the offset. As above, I assumed an exchangeable working correlation structure. The log-linear model is equivalent to a one-way ANOVA with 4 treatment groups. The type3 option provides a Wald test (accounting for the correlation) for the significance of the x1*trt effect. If this effect is eliminated from the model, then the population means per week are identical across the four groups. If included, the population means per week are different. Thus the p-value for the Wald test is checking whether the means are different. The p-value is large, which indicates that once the correlation is taken into account, that effects of time and treatment and differences in baseline levels between treatment groups observed in the data plots are possibly just a reflection of random variation.

```
proc genmod data=new2;
   class id x1 trt/desc;
   model y= x1*trt/noscale d=poisson offset=ltime type3;
   repeated subject=id/corrw type=exch modelse;
```

The GENMOD Procedure

Model Information

Data Set	WORK.NEW2
Distribution	Poisson
Link Function	Log
Dependent Variable	y
Offset Variable	ltime

Number of Observations Read	290
Number of Observations Used	290

Class Level Information

Class	Levels	Values
id	58	238 236 234 232 230 228 227 226 225 222 221 220 219 218 217 215 214 213 211 210 209 208 206 205 204 203 202 201 147 145 143 141 139 137 135 130 129 128 126 124 123 122 121 118 117 116 114 113 112 111 110 108 107 106 104 103 102 101
x1	2	1 0
trt	2	1 0

Parameter Information

Parameter	Effect	x1	trt
Prm1	Intercept		

Prm2		x1*trt	1	1
Prm3		x1*trt	1	0
Prm4		x1*trt	0	1
Prm5		x1*trt	0	0

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	286	2413.0245	8.4371
Scaled Deviance	286	2413.0245	8.4371
Pearson Chi-Square	286	3015.1555	10.5425
Scaled Pearson X2	286	3015.1555	10.5425
Log Likelihood		5631.7547	

Algorithm converged.

Analysis Of Initial Parameter Estimates

Parameter			DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square
Intercept			1	1.3476	0.0341	1.2809	1.4144	1565.44
x1*trt	1	1	1	-0.2988	0.0512	-0.3992	-0.1985	34.08
x1*trt	1	0	1	0.1108	0.0469	0.0189	0.2027	5.58
x1*trt	0	1	1	-0.1080	0.0486	-0.2034	-0.0127	4.93
x1*trt	0	0	0	0.0000	0.0000	0.0000	0.0000	.
Scale			0	1.0000	0.0000	1.0000	1.0000	

Analysis Of Initial Parameter Estimates

Parameter			Pr > ChiSq
Intercept			<.0001
x1*trt	1	1	<.0001
x1*trt	1	0	0.0181
x1*trt	0	1	0.0264
x1*trt	0	0	.
Scale			

NOTE: The scale parameter was held fixed.

GEE Model Information

Correlation Structure	Exchangeable
Subject Effect	id (58 levels)
Number of Clusters	58
Correlation Matrix Dimension	5
Maximum Cluster Size	5
Minimum Cluster Size	5

Algorithm converged.

Working Correlation Matrix

	Col1	Col2	Col3	Col4	Col5
Row1	1.0000	0.5941	0.5941	0.5941	0.5941
Row2	0.5941	1.0000	0.5941	0.5941	0.5941
Row3	0.5941	0.5941	1.0000	0.5941	0.5941
Row4	0.5941	0.5941	0.5941	1.0000	0.5941
Row5	0.5941	0.5941	0.5941	0.5941	1.0000

Exchangeable Working Correlation

Correlation	0.5941485833

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			1.3476	0.1574	1.0392	1.6560	8.56	<.0001
x1*trt	1	1	-0.2988	0.2228	-0.7355	0.1379	-1.34	0.1799
x1*trt	1	0	0.1108	0.1161	-0.1168	0.3383	0.95	0.3399
x1*trt	0	1	-0.1080	0.1937	-0.4876	0.2716	-0.56	0.5770
x1*trt	0	0	0.0000	0.0000	0.0000	0.0000	.	.

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

Parameter			Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept			1.3476	0.1106	1.1309	1.5644	12.19	<.0001
x1*trt	1	1	-0.2988	0.2346	-0.7587	0.1610	-1.27	0.2028
x1*trt	1	0	0.1108	0.1233	-0.1308	0.3524	0.90	0.3687
x1*trt	0	1	-0.1080	0.1579	-0.4176	0.2015	-0.68	0.4940
x1*trt	0	0	0.0000	0.0000	0.0000	0.0000	.	.
Scale			3.2469	.	.	.	.	.

NOTE: The scale parameter for GEE estimation was computed as the square root of the normalized Pearson's chi-square.

Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
x1*trt	3	3.67	0.2997

Example 3: Normal responses (Rat Weight Gains)

GEEs are sometimes used to analyze longitudinal data with normal responses. GEEs require the specification of a marginal model, so general forms of random effects can not be handled with this framework. Further, genmod does not allow different correlation structures for different groups, a feature of mixed that we used extensively. A nice feature of the GEE analysis is that valid inferences are obtained using the empirical covariance matrix even when the working correlation matrix is incorrectly specified. ML or REML based inferences using the mixed procedure hold only if the variance structure is correctly specified.

I will use data on rats to contrast a ML based analysis for a longitudinal regression model with a GEE analysis. The data were from a study of weight gain where investigators randomly assigned rats to one of 3 treatments: treatment 1 was a control (no additive);

treatments 2 and 3 consisted of two different additives (Thiouracil and Thyroxin respectively) to the rats drinking water. Weight was measured at baseline (week 0) and at weeks 1, 2, 3, and 4. The first SAS program enters the data in multivariate format, then transforms it to regression format. A subset of the data is presented so we can see the result of the **transpose** procedure. The means and standard deviations for each group by week, and a trajectory plot of the means by group, is provided with the output.

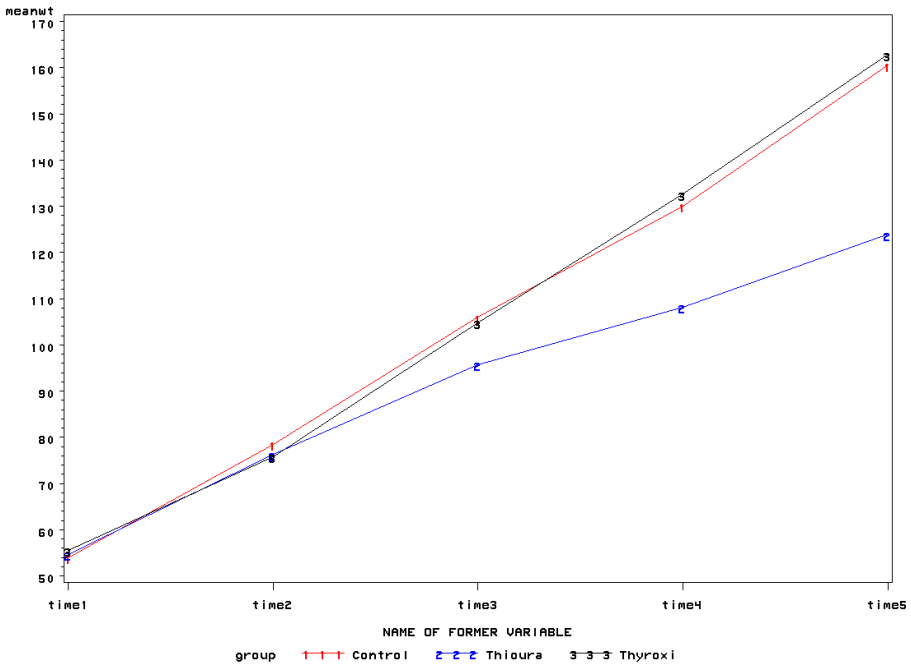
```
options ls=75;
data d1;
input ID treat time1-time5;
if treat = 1 then group = 'Control';
if treat = 2 then group = 'Thioura';
if treat = 3 then group = 'Thyroxi';
datalines;
           1       1      57     86     114     139     172
           2       1      60     93     123     146     177
           3       1      52     77     111     144     185
           4       1      49     67     100     129     164
           5       1      56     81     104     121     151
           6       1      46     70     102     131     153
           7       1      51     71      94     110     141
           8       1      63     91     112     130     154
           9       1      49     67      90     112     140
          10       1      57     82     110     139     169
          11       2      61     86     109     120     129
          12       2      59     80     101     111     122
          13       2      53     79     100     106     133
          14       2      59     88     100     111     122
          15       2      51     75     101     123     140
          16       2      51     75      92     100     119
          17       2      56     78      95     103     108
          18       2      58     69      93     114     138
          19       2      46     61      78      90     107
          20       2      53     72      89     104     122
          21       3      59     85     121     146     181
          22       3      54     71      90     110     138
          23       3      56     75     108     151     189
          24       3      59     85     116     148     177
          25       3      57     72      97     120     144
          26       3      52     73      97     116     140
          27       3      52     70     105     138     171
;
proc sort;
  by id treat group;

proc transpose out=d3 name=time prefix=weight;
  var time1-time5;
  by id treat group;

proc print;

proc sort data=d3 out = d4;
  by group time;

proc means data = d4 noprint;
  by group time;
  var weight1;
  output out=d5 mean=meanwt std=stdwt;

proc print data = d5;
```

```
symbol1 color=red   interpol=join value=1   height=1;
symbol2 color=blue  interpol=join value=2   height=1;
symbol3 color=black interpol=join value=3   height=1;

proc gplot data = d5;
  plot meanwt*time = group;

run;
```

Obs	ID	treat	group	time	weight1
1	1	1	Control	time1	57
2	1	1	Control	time2	86
3	1	1	Control	time3	114
4	1	1	Control	time4	139
5	1	1	Control	time5	172
51	11	2	Thioura	time1	61
52	11	2	Thioura	time2	86
53	11	2	Thioura	time3	109
54	11	2	Thioura	time4	120
55	11	2	Thioura	time5	129
106	22	3	Thyroxi	time1	54
107	22	3	Thyroxi	time2	71
108	22	3	Thyroxi	time3	90
109	22	3	Thyroxi	time4	110
110	22	3	Thyroxi	time5	138

Obs	group	time	_TYPE_	_FREQ_	meanwt	stdwt
1	Control	time1	0	10	54.000	5.4365
2	Control	time2	0	10	78.500	9.6408
3	Control	time3	0	10	106.000	9.9219
4	Control	time4	0	10	130.100	12.5649
5	Control	time5	0	10	160.600	15.1965
6	Thioura	time1	0	10	54.700	4.6916
7	Thioura	time2	0	10	76.300	7.9169
8	Thioura	time3	0	10	95.800	8.4958
9	Thioura	time4	0	10	108.200	9.7502
10	Thioura	time5	0	10	124.000	11.2546
11	Thyroxi	time1	0	7	55.571	2.9921
12	Thyroxi	time2	0	7	75.857	6.4402
13	Thyroxi	time3	0	7	104.857	11.0970
14	Thyroxi	time4	0	7	132.714	16.9776
15	Thyroxi	time5	0	7	162.857	21.5053

As a next step in the analysis, I assumed that the mean weight was a quadratic function of time. The trajectories of the means appear to be either linear or quadratic in each group, so a quadratic model would appear to be adequate. In the data step below, I created numerical variables for week and the square of week. I then fit (by REML) a quadratic to each group using a SGP, assuming that the correlation between responses follows an AR(1) model.



We can read off the coefficients for the quadratics from the parameter estimates table to produce the estimated mean weight for each group:

$$\begin{aligned} PredWt &= 54.01 + 23.73Week + 0.73Week^2 \text{ for Placebo(1)} \\ &= 54.72 + 22.22Week - 1.23Week^2 \text{ for Thioura(2)} \\ &= 55.54 + 21.15Week + 1.42Week^2 \text{ for Thyroxi(3)} \end{aligned}$$

```
data e1;
  set d3;
  if time='time1' then week  = 0 ;
  if time='time2' then week  = 1;
  if time='time3' then week  = 2;
  if time='time4' then week  = 3;
```

```
  if time='time5' then week  = 4;
  week2 = week*week;
proc print;
proc mixed data = e1;
 class id time group;
 model weight1 = group group*week group*week2/noint solution;
 repeated/sub=id type=ar(1);
run;
```

The Mixed Procedure

Model Information

Data Set	WORK.E1
Dependent Variable	weight1
Covariance Structure	Autoregressive
Subject Effect	ID
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Between-Within

Class Level Information

Class	Levels	Values
ID	27	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
time	5	time1 time2 time3 time4 time5
group	3	Control Thioura Thyroxi

Dimensions

Covariance Parameters	2
Columns in X	9
Columns in Z	0
Subjects	27
Max Obs Per Subject	5

Number of Observations

Number of Observations Read	135
Number of Observations Used	135
Number of Observations Not Used	0

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	995.70298708	
1	2	863.26604654	0.00057188
2	1	863.07445135	0.00000544
3	1	863.07272336	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
AR(1)	ID	0.8647

Residual	129.00

Fit Statistics

-2 Res Log Likelihood	863.1
AIC (smaller is better)	867.1
AICC (smaller is better)	867.2
BIC (smaller is better)	869.7

Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
1	132.63	<.0001

Solution for Fixed Effects

Effect	group	Estimate	Standard Error	DF	t Value	Pr > |t|
group	Control	54.0127	3.5917	24	15.04	<.0001
group	Thioura	54.7211	3.5917	24	15.24	<.0001
group	Thyroxi	55.5449	4.2929	24	12.94	<.0001
week*group	Control	23.7299	1.9065	102	12.45	<.0001
week*group	Thioura	22.2204	1.9065	102	11.66	<.0001
week*group	Thyroxi	21.1540	2.2787	102	9.28	<.0001
week2*group	Control	0.7283	0.4275	102	1.70	0.0915
week2*group	Thioura	-1.2266	0.4275	102	-2.87	0.0050
week2*group	Thyroxi	1.4201	0.5109	102	2.78	0.0065

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
group	3	24	208.56	<.0001
week*group	3	102	125.65	<.0001
week2*group	3	102	6.29	0.0006

I refit the model using GEE with an AR(1) working correlation matrix. I excluded the Initial Parameter Estimates table, which you should realize gives the ML estimates and standard errors assuming response within mice are independent. This is not the same analysis obtained from the mixed procedure gives ML (or REML) estimates assuming the correlation matrix is AR(1).

Important points to notice

1. The regression parameter estimates and model based standard errors from the GEE analysis are similar to the REML analysis given above.

2. The estimated AR(1) correlation and scale from the GEE analysis are .84 and 10.81. The scale is the estimated standard deviation so the estimated variance is $10.81^2 = 116.96$. The estimated AR(1) correlation and variance from the REML analysis are .86 and 129. The two sets of estimates are similar.

3. The empirical standard errors are smaller than the model based standard errors so inferences about regression effects based on the robust covariance matrix yield smaller p-values (on individual effects on Type 3 tests compared to REML). One general concern here is that the sample size is small - only 27 mice total in 3 groups. There is some evidence that the estimated empirical covariance matrix is biased in small samples, and that it tends to give estimates that are too small.

```
proc genmod data = e1;
 class id time group;
 model weight1 = group group*week group*week2/
       noint dist=normal link=identity type3;
 repeated subject=id/type=ar(1) modelse corrw;
```

The GENMOD Procedure

Model Information

Data Set	WORK.E1
Distribution	Normal
Link Function	Identity
Dependent Variable	weight1
Number of Observations Read	135
Number of Observations Used	135

GEE Model Information

Correlation Structure	AR(1)
Subject Effect	ID (27 levels)
Number of Clusters	27
Correlation Matrix Dimension	5
Maximum Cluster Size	5
Minimum Cluster Size	5

Algorithm converged.

Working Correlation Matrix

	Col1	Col2	Col3	Col4	Col5
Row1	1.0000	0.8365	0.6998	0.5854	0.4897
Row2	0.8365	1.0000	0.8365	0.6998	0.5854
Row3	0.6998	0.8365	1.0000	0.8365	0.6998
Row4	0.5854	0.6998	0.8365	1.0000	0.8365

Row5	0.4897	0.5854	0.6998	0.8365	1.0000

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.0000	0.0000	0.0000	0.0000	.	.
group	Control	54.0177	1.6436	50.7962	57.2391	32.87	<.0001
group	Thioura	54.7297	1.4048	51.9764	57.4829	38.96	<.0001
group	Thyroxi	55.5336	1.0467	53.4822	57.5851	53.06	<.0001
week*group	Control	23.7303	1.4635	20.8618	26.5987	16.21	<.0001
week*group	Thioura	22.2189	1.6057	19.0717	25.3661	13.84	<.0001
week*group	Thyroxi	21.1621	1.7371	17.7575	24.5667	12.18	<.0001
week2*group	Control	0.7275	0.3424	0.0564	1.3986	2.12	0.0336
week2*group	Thioura	-1.2274	0.4626	-2.1341	-0.3206	-2.65	0.0080
week2*group	Thyroxi	1.4194	0.3807	0.6733	2.1655	3.73	0.0002

Analysis Of GEE Parameter Estimates
Model-Based Standard Error Estimates

Parameter		Estimate	Standard Error	95% Confidence Limits		Z	Pr > |Z|
Intercept		0.0000	0.0000	0.0000	0.0000	.	.
group	Control	54.0177	3.4200	47.3146	60.7207	15.79	<.0001
group	Thioura	54.7297	3.4200	48.0266	61.4327	16.00	<.0001
group	Thyroxi	55.5336	4.0877	47.5219	63.5453	13.59	<.0001
week*group	Control	23.7303	1.9907	19.8286	27.6319	11.92	<.0001
week*group	Thioura	22.2189	1.9907	18.3172	26.1205	11.16	<.0001
week*group	Thyroxi	21.1621	2.3793	16.4987	25.8255	8.89	<.0001
week2*group	Control	0.7275	0.4484	-0.1513	1.6064	1.62	0.1047
week2*group	Thioura	-1.2274	0.4484	-2.1062	-0.3485	-2.74	0.0062
week2*group	Thyroxi	1.4194	0.5359	0.3690	2.4698	2.65	0.0081
Scale		10.8152	.	.	.	.	.

NOTE: The scale parameter for GEE estimation was computed as the square root of the normalized Pearson's chi-square.

Score Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
group	3	26.83	<.0001
week*group	3	25.83	<.0001
week2*group	3	11.90	0.0077