Practical 8: Generalized estimating equations

Data

- The `Beta-blocker trial`
  The data come from a randomised, double blind placebo trial to establish the efficacy of beta-blockers for reducing blood pressure in patients with abnormally high blood pressure, or *hypertension*.
  We are using a modified version of the original data held in the file called `beta_blocker_gee.dta`.
  The variables are:

```
------------------------------------------------------------------------------
id              patient identifier
time            time of visit
treat           treatment: 1=active; 0=placebo
pre             pre randomisation diastolic BP in mmHg (centered)
dbp             diastolic blood pressure post randomisation
first5          Indicator that data are complete
c_pre           pre randomisation diastolic BP in mmHg (centered)
------------------------------------------------------------------------------
```

Questions

1. Load the Beta-Blocker dataset and examine whether the two treatment groups have similar pre-treatment DBP values and also whether they are equally affected by missingness with:

   ```
   . tabstat  pre if time==1,by(treat) s(count mean sd) c(s) nototal
   . tab first5 treat, col nokey
   ```

   (Note that `nototal` in the first command removes the overall summary and `nokey` in the second one simplifies the output.)

2. Examine the number of available observations over time separately by treatment. Plot the DBP average profiles of those without complete data, separately by treatment with:

```
. xtset id time
. xtline dbp if treat==0 & first5==0, overlay legend(off)  ///
          saving(plot1,replace) title(Placebo)
. xtline dbp if treat==1 & first5==0, overlay legend(off) ////
          saving(plot2,replace)  title(Active)
. graph combine plot1.gph plot2.gph, title(With missing values)
```

Do the same for those without missing values.

Describe the profiles (for example as follows) and comment on the pattern of missingness.

```
gen dbp_act = dbp if treat==1
gen dbp_pl = dbp if treat ==0
tabstat dbp_act dbp_pl, by(time) s(count mean sd) c(s) nototal
```

3. For simplicity we will now restrict the analyses to patients with complete data.

   Fit a random intercept model to these data, controlling for the pre-treatment value of DBP, with:

   ```
   . mixed dbp time treat c_pre if first5==1 ||id: , var  ml
   ```

   Interpret the results.

4. Now use GEEs to estimate the marginal effect of treatment and time, controlling for pre-treatment DBP. First use an identity matrix:

   ```
   . xtset id time
   . xtgee dbp time treat c_pre  if first5==1, cor(indep)
   ```

   Interpret the estimated parameters.

5. Try a different working matrix:

   ```
   . xtgee dbp time treat c_pre  if first5==1, cor(exch)
   ```

   Compare the results with those found when adopting a different working matrix and also when fitting a random intercept model.

6. Refit both models but now requiring robust standard errors:

   ```
   . xtgee dbp time treat c_pre  if first5==1, cor(indep) vce(robust)
   . xtgee dbp time treat c_pre  if first5==1 , cor(exch) vce(robust)
   ```

   Do the estimates change?

7. Repeat questions 3-6 using all observations. Do these equalities still apply?