

Longitudinal Homework 6

Tim Vigers

12 November 2019

1. Model planning

a. Software

i. Medication use

In order to adapt a generalized linear model (GzLM) for serially correlated data, I would use a generalized estimating equation (GEE). This could be fit using PROC GENMOD in SAS, with a Poisson distribution since this is count data. There is also a package called “geepack” in R that can fit GEEs, but I’ve never used it and don’t know much about it, so PROC GENMOD is probably safer.

ii. FEV1

I think you can probably get away with a normal theory model for FEV1 data, unless it’s really skewed. If so, I would use either PROC MIXED in SAS or gls() in R. If normal theory models won’t work, then I would use PROC GENMOD or geeglm() like above, to model the outcome with a non-normal distribution.

b. Data

i. Medication use

Because we need a GEE for this outcome, I would set up the data so that each subject has a row for every day during the relevant timeframe. On days without an albuterol count the outcome would be filled in as missing (NA in R), but the temporal spacing would be equal between rows within subject.

The SAS code would look something like:

```
proc genmod data=albuterol;
class id friday;
model albuterol_use =
    friday ln_mmax_pm25 temperature pressure humidity / solution dist=poisson;
repeated subject=id / TYPE = AR(1);
run;
```

c. Between-subject differences

i. Medication use

In order to fit a Poisson model with a random intercept for subject, I would use PROC GLIMMIX in SAS. PROC NLMIXED would also work, but I find it a little more confusing. The default in PROC GLIMMIX is to approximate the true likelihood using Laplace’s method, but you can also specify method = quad to use Gaussian quadrature. This can also be done using glmer() in R.

What are the drawbacks?

ii. FEV1

Again assuming that the normal theory model is okay for FEV1, I would use PROC MIXED in SAS or lme() in R (or possibly lmer() to be consistent with the model above). These models would be fit using ML or REML rather than Gaussian quadrature.

Gaussian quad vs. ML?

for an AR(1) correlation structure in the repeated measures, but does not allow for the incorporation of random effects. GEEs can be fit (using the same data structure as above) with:

##d. Random intercept and serial correlation

i. Medication use

Correlated count data like this probably requires a generalized linear mixed model (GzLMM) where the outcome is modeled as Poisson-distributed, a random intercept for subject, and with an AR(1) or spatial power correlation for repeated measures. I think the best way to do this in SAS is to use PROC GLIMMIX, and in R glmmPQL() should work.

There isn't a REPEATED statement in PROC GLIMMIX, so you need to add another random effect with the "residual" and the correlation structure:

```
proc glimmix data=albuterol;
model albuterol_use =
    friday ln_mmax_pm25 temperature pressure humidity / solution distribution=poisson;
random intercept / subject=id;
random _residual_ / subject=id type=ar(1);
run;
```

The R code would be something like:

```
glmmPQL(fixed = albuterol_use ~ friday + ln_mmax_pm25 + temperature + pressure + humidity,
        random = ~1|id, family = "poisson", correlation = corAR1(),
        data = df)
```

ii. FEV1

These normal models could be fit using either PROC MIXED in SAS or lme() in R. I don't know how to include the AR(1) structure in lmer(), but I'm sure it's possible.

The SAS code would be something like:

```
proc mixed data=albuterol;
class id friday;
model fev1 =
    friday ln_mmax_pm25 temperature pressure humidity / solution;
random intercept / subject=id;
repeated / type=AR(1) subject=id;
run;
```

And the R code something like:

```
lme(fev1 ~ friday + ln_mmax_pm25 + temperature + pressure + humidity,
    random = ~1|id, correlation = corAR1(), data = df, na.action = na.omit)
```