Modelling seasonal effects in Longitudinal Studies

Target: Statistics in Med/ Jasa applications

Motivating Example and research question

How does TC vary over seasons of the year, and what factors relate to this variability. In practice this question is very difficult to answer.

i. subjects have staggered entry
ii. there may be sequence effects (1st interview gives highest results)
iii. there are errors in variables and interviewer effects for independent variables
iv. the pattern we wish to capture is not necessarily linear, nor a simple sine curve.
v. There may be heterogeneous variances.

Description of Season’s study

Choices in Modelling seasonal effects

Model months
Model discrete seasons (simple discussion)
Use sine/cosine models (ed has a program)*
With 5 measures per subject, how many terms can be added:

Ie. $A_i \sin(\frac{t}{P_i} + \lambda_i)$ for I=1,…,4, where $P_i = \frac{365}{i}$ (could be evaluated via simulation)

How important is equal spacing in the power of the analysis?
Power for adding additional sine and cosine terms.
Use polar coordinates

Use polynomial models (maybe just mention)

Covariates: How to include

Cross-sectional covariates (no real problem)
Time dependent covariates (time = known)( adding the mean of the time dependent covariates)
Errors in variables (using BLUPs as a screener) (Interviewer effects, day of the week)
When controlling for day of the week (or interviewer effects), should the effect be treated as a random effect, or a fixed effect? (I don’t know, but we could investigate with a small scale simulation).
Bias in other effects due to inclusion of covariates with error. (Like Carroll, Stafanski, and Ruppert (errors in variables) and the paper by Ed at the Biometrics meetings).

Is it always better to add a covariable (even if known) to the model? Will predicted values always have smaller MSE if a covariable (even though weakly related) is included in the model).

From the Mixed models for Energy intake, there is different day to day residual variability between men and women. What effect does this have after using BLUP estimates to reduce daily variability. How important is it.
A different relationship may occur between subjects (for the outcome and a time dependent variable based on the mean (of the time dependent variable for a subject), as opposed to the effect within subjects. This requires inclusion of the mean of the independent variable in the model.

Variance structures
   - Compound symmetry
   - Heteroskedastic errors
   - Auto-correlation
   - How to choose?
   - How important is the choice? (simulation)

Presentation of results
   - Seasonal plots with CI.
   - Plots with random effects for amplitude and phase
   - Estimating amplitude and phase variances. (deriving these variances for a cosine expansion). (How good are the variance estimates based on this approximation? )