Kennesaw State University
College of Science and Mathmatics
Department of Mathematics and Statistics

Adolescent Smoking and Development of Long Term Habits: A Longitudinal Analysis in SAS®

Elizabeth Leslie, MSAS
Kennesaw State University, Kennesaw GA
Supervised by Bradley Barney, Ph.D.

SAS GLOBAL FORUM

BACKGROUND

This study was an investigation into the impact of early adolescent smoking on adult smoking habits of National Longitudinal Survey of Youth 1997 Participants[1] over the course of 13 years. Research shows that adolescents who experiment with smoking are more likely to remain smokers and that even low-rate smokers find it extremely difficult to quit smoking. The longitudinal data allows for examination of the effects of time on the number of cigarettes smoked, as well as the effects of other factors over time, such as starting age.

WHAT IS LONGITUDINAL ANALYSIS?

Longitudinal analysis is the study of repeated measurements of the same objects over time. In this case, the measure was reported number of cigarettes smoked in one month and it was repeatedly taken for 13 years.

THE DATA STRUCTURE: Converting from Person-Level to Person Period

Person-Level Data Structure Sample

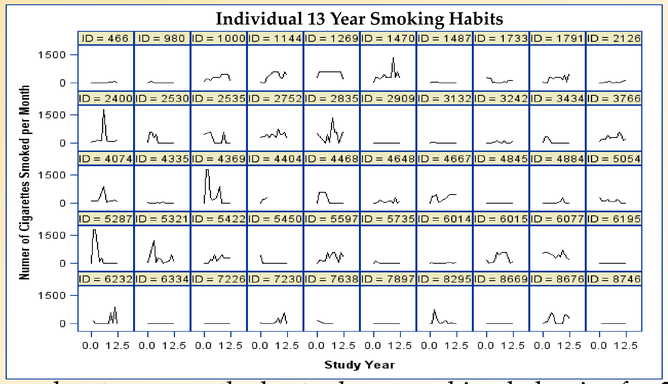
ID	Cigarettes Smoked '97	Cigarettes Smoked '98	Cigarettes Smoked '99	Cigarettes Smoked '00
80	120	0	150	150
86	450	150	14	120
102	40	75	300	360
103	210	390	200	270

Person-Period Data Structure Sample

ID	time	cigarettes
80	0	120
80	1	0
80	2	150
80	3	150
86	0	450
86	1	150
86	2	14
86	3	120
102	0	40
102	1	75
102	2	300
102	3	360
103	0	210
103	1	390
103	2	200
103	3	270

Why?
Transforming the data from person-level (1,212 observations, 74 variables) to person-period (13,544 observations, 11 variables) allows for the creation of a time variable that is flexible enough to examine various time structures as well as generating a quantified effect measure. The late-teen smoking habits were of particular interest, so for ease of interpretation time was centered at zero for year 1997

TEMPORAL TRENDS



These exploratory growth charts show smoking behavior for 25 of the 1,212 participants. Even in this small sample, there is a lot of variability between individual smoking behaviors which must be considered when building an appropriate model.

THE MIXED MODEL

A linear mixed model was fitted to the data in order to accurately model the inter-individual variability. The general form is:

$$Y_i = X_i\beta + Z_i\gamma_i + \varepsilon_i$$

- Y is the response vector
- X is the fixed design matrix
- Z is the random design matrix
- β and γ are the unknown fixed and random effects
- ε is the unknown random error

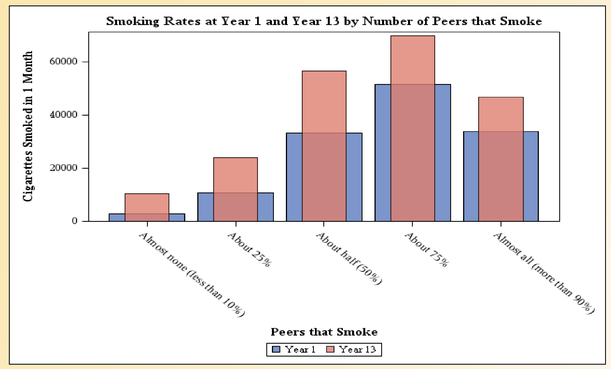
COVARIANCE STRUCTURE

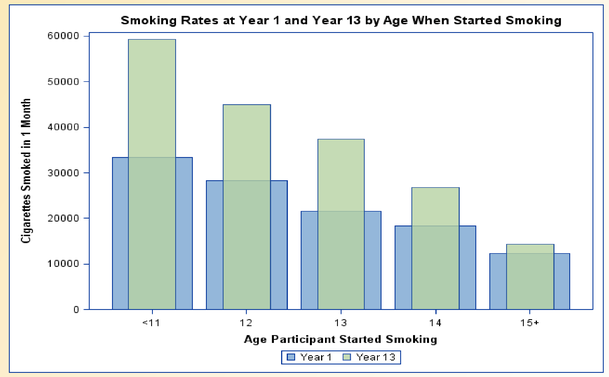
Why is covariance structure so important?

We want the variance of y: $V = ZGZ' + R$. If the structure is wrong then the model is untrustworthy. Using a chi-squared distribution to test nested covariance structures (simpler vs. more complex structures), the best covariance fit for this data was the Unstructured Structure where there are no restrictions, other than the universal constraints that all variances must be non-negative and all correlations must be between -1 and 1.

$$\begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ \sigma_{12} & \sigma_{22} & \sigma_{32} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}$$

RESULTS





The effect of peer pressure and effect of starting at a younger age clearly influence the amount smoked into adulthood. Significant effects in the model include

- Time (as a factor)
- Number of peers that smoke
- Number of peers that do drugs
- Sex
- Race
- Age when started
- Interactions: age when started and time, age when started and sex

CONCLUSIONS

Adolescent smoking habits increase over time, and sex and age when started smoking have an effect on number of cigarettes smoked. Additionally, age when started smoking impacts each sex differently, the number of cigarettes smoked varies by race, and the number of cigarettes increases as the number of peers who smoke and does drugs increases.

1.Bureau of Labor Statistics, U.S. Department of Labor. National Longitudinal Survey of Youth 1997 cohort, 1997-2009 (rounds 1-13) [computer file]. Produced by the National Opinion Research Center, the University of Chicago and distributed by the Center for Human Resource Research, The Ohio State University. Columbus, OH: 2009.
Special thanks to Dr. B for his patience and quick e-mail response times!!