SABRE

EXERCISE 2 Longitudinal Logit ANSWERS

File: WEMP2.DAT

THE VARIABLES

case	individual identifier
femp	wife's employment status; 1=employed, 0=unemployed
mune	husband's employment status; 1=unemployed, 0=employed
time	calendar time (year-1975)
und1	children aged < 1 year old; 1=yes, 0=no
und5	children aged 1 - 5 years old; 1=yes, 0=no
age	mother's age

1. Declare the variables; read the data; specify the response variable.
You should have read in 1580 observations.

2. To keep a log of your work.

<S> outfile out2

3. Fit a cross-sectional (i.e. pooled) logistic regression model with mune as an explanatory variable.

<S> lfit int mune

4. Note the deviance and degrees of freedom for this model.

1970 with 1578 df.

5. Fit a longitudinal logistic regression model with mune as an explanatory variable.

<S> fit int mune

6. Note the deviance and the degrees of freedom of this model.

1401 with 1575 df.

7. Is the longitudinal model an improvement?

Yes. A change in deviance of over 500 for 3 df.

8. Is there significant residual heterogeneity?

Yes. Scale = 1.84; s.e. = 0.31.

ESRC Longitudinal Data Analysis Workshop 2B;
11th November 2003 St Andrews University, Dr Vernon Gayle.

1

9.Compute a Z score, Wald test for the variable mune.

Z=5.95; Wald = 3; p<.001

10. Construct a 95% confidence interval for the estimate of mune. Could the true value of the parameter estimate be zero?

-2.81; -1.22; Zero is not within the 95% C.I.

11. What do these measures suggest about the effects of husband's employment status on a wife's labour market participation?

A husband being unemployed has a negative effect on women's employment status.

12. Now look at the effects of the other variables und1 und5 & age. Fit each variable in turn and note the change in deviance and degrees of freedom.

+ und1 deviance = 1332.45 df = 1574;
+ und5 deviance = 1216.30 df = 1573;
+ age   deviance = 1213.62 df =1572.

13. Which variables are significant?

und1 and und5.

14. Add the time variable to this model. Is it significant?

No.

15. What can we conclude?

Calendar year is not significant in this analysis.

16. Some theorists argue that the effects of age are not linear. Let us construct a variable called for age^2.

<S> tran agesq age ^ 2   or   <S> tran agesq   age * age

17. Add both age and agesq to the model (with int mune und1 und5). What might we conclude?

Both variables are significant. There is a support for the idea that age has a non-linear effect.

18. Could we fit just agesq in the model?

No. Part of the explanation is captured by age and part by age squared. Age squared on its own is not significant.

NOTE: Always be strongly guided by theory especially when fitting such effects. Also be cautious of their substantive interpretation.

19. Overall, what is our most appropriate model and why?

int mune und1 und5 age agesq

deviance of 1209.72 @ 1571 df.

20. What substantive conclusions can we draw from this our analysis of this dataset?

Husband's employment status, having young children and the age of the women all affect her chances of being unemployed.

ESRC Longitudinal Data Analysis Workshop 2B;
11th November 2003 St Andrews University, Dr Vernon Gayle.

3