The data for a binary approach

Contents

Introduction

An analysis of a data set where students are clustered within schools is used to illustrate features of random-effects analysis of clustered grouped-time survival data.

We focus on actual usage of tobacco products and on subsequent data collected from the respondents.

Schools were randomized to one of four study conditions: (a) a social-resistance classroom curriculum (CC); (b) a media (television) intervention (TV); (c) a combination of curriculum and TV conditions; and (d) a no-treatment control group. These conditions form a 2 x 2 factorial design of CC (yes or no) by TV (yes or no).

The outcome variable of interest in this chapter is the response the question "Have you ever tried a cigarette?". Students were assessed at 4 occasions:

- pre-intervention (January 1986, also referred to as Wave A)
- post-intervention (April 1986, *i.e.* Wave B)
- year follow-up (April 1987, *i.e.* Wave C)
- year follow-up (April 1988, *i.e.* Wave D)

As the intervention procedures were implemented following the pretest, we focus in the analyses to follow on the three post-intervention time points and include only those students who had not answered yes to this question at pretest. Of the original 1,600 respondents, 1,556 are included in the data considered here. Thus, our analysis examines the degree to which the intervention prevented or delayed students from initiating smoking experimentation. Because the intervention was also aimed at smoking cessation for individuals who had initiated smoking, here we are examining only a part of the intervention aims.

The first few lines of the LISREL spreadsheet **SMKBCD2.lsf** used in this section are shown below. Note that there is a maximum of 3 observations associated with each student – not all students have data at all 3 occasions.

SMKBCD2.LSF

	School	Class	Student	Event	TimeC	TimeD	SexM	CC	TV
1	193.0	193101.0	193101105.0	0.0	0.0	0.0	1.0	0.0	0.0
2	193.0	193101.0	193101105.0	0.0	1.0	0.0	1.0	0.0	0.0
3	193.0	193101.0	193101105.0	0.0	0.0	1.0	1.0	0.0	0.0
4	193.0	193101.0	193101108.0	0.0	0.0	0.0	0.0	0.0	0.0
5	193.0	193101.0	193101108.0	0.0	1.0	0.0	0.0	0.0	0.0
6	193.0	193101.0	193101108.0	0.0	0.0	1.0	0.0	0.0	0.0
7	193.0	193101.0	193101113.0	1.0	0.0	0.0	1.0	0.0	0.0
8	193.0	193101.0	193101117.0	1.0	0.0	0.0	1.0	0.0	0.0
9	193.0	193101.0	193101121.0	0.0	0.0	0.0	1.0	0.0	0.0
10	193.0	193101.0	193101121.0	0.0	1.0	0.0	1.0	0.0	0.0

The variables of interest are:

- School indicates the school a student is from.
- Class identifies the classroom to which a student belongs.
- Student represents the student identification number.
- Event indicates occurrence of the event (1 indicating "yes" and 0 "no.").
- TimeC is an indicator variable indicating the first follow-up occasion after the post-intervention measurement occasion. It assumes a value of 1 if a measurement was made at the first follow-up occasion, and 0 otherwise.
- TimeD is the indicator variable for the second follow-up occasion. It assumes a value of 1 if a measurement was made at the second follow-up occasion and 0 otherwise.
- SexM is an indicator variable for gender, with "1" indicating male respondents, and "0" female respondents.
- CC is a binary variable indicating whether a social-resistance classroom curriculum was introduced, with 0 indicating "no" and 1 "yes."
- TV is an indicator variable for the use of media (television) intervention, with a "1" indicating the use of media intervention, and "0" the absence thereof.

The post-intervention measurement, which is the first of the three measurement occasions in this data set, serves as the reference cell. In terms of the indicator variables TimeC and TimeD it would be a measurement for which TimeC = TimeD = 0.

SMKBCD2.LSF

	CCTV	SexTC	SexTD	CCTC	CCTD	TVTC	TVTD	CCTVTC	CCTVTD
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

In addition to these variables, **SMKBCD2.lsf** includes a number of interaction terms:

- CCTV was constructed by multiplying the variables TV and CC, and represents the CC by TV interaction.
- SexTC denotes the SexM by TimeC interaction.
- SexTD denotes the SexM by TimeD interaction.
- CCTC denotes the interaction between classroom curriculum intervention CC and TimeC.
- CCTD denotes the interaction between CC and TimeD.
- TVTC denotes the interaction between media intervention TV and TimeC.
- TVTD denotes the interaction between TV and TimeD.
- CCTVTC represents the interaction between the CC by TV interaction at the TimeC.
- CCTVTD represents the interaction between the CC by TV interaction at the TimeD.

In all, there were 1556 students included in the analysis of smoking initiation. Of these students, approximately 40% (n = 634) answered yes to the smoking question at one of the three post-intervention time points, while the other 60% (n = 922) either answered no at the last time point or were censored prior to the last time point.

Consider a level-2 model, with schools as the level-2 units. In general, for $i = 1, \ldots, N$ N level-2 units, containing $j = 1, \ldots, n_i$ level-1 units (subjects or multiple failure times) the concept of a censoring or event indicator can be expressed as follows. First, we assume that the assessment time takes on discrete positive values $t = 1, 2, \ldots, m$ representing time points

or intervals and that each *ij* unit is observed until time t_{ij}. The censor/event indicator δ_{ij} is coded depending on what happens at time t_{ij}:

- an event occurs $\left(t_{ij} = t \quad and \quad \delta_{ij} = 1\right)$
- the observation is censored $\left(t_{ij} = t \quad and \quad \delta_{ij} = 0\right)$

The term censoring is used when a unit is observed at t_{ij}, but not at $t_{ij} + 1$ (and we know that the event has not occurred up to time t_{ij}).

As mentioned previously, the dichotomous variable EVENT indicated the occurrence of an event. Occurrence of an event was recorded at three time points (WaveB, WaveC, and WaveD), though some subjects dropped out of the study and were not measured at all three time points. To model the time until the event as the outcome variable in a binary analysis of the data, person-time indicators are created (Singer & Willett, 1993). For this, the number of records for each person depends on the timing of the event or censoring for that person. For example, if there were two follow-up points, the two person-time indicators T1 and T2 would be coded as follows:

- T1 = 1: event occurred at T1 (or in interval between T0 and T1)
- T1 = 0: event did not occur at T1 (or in interval between T0 and T1) and T1 was the subject's last measured time point
- T1 = 0 and T2 = 1: event did not occur at T1 but did occur at T2 (or in the interval between T1 and T2)
- T1 = 0 and T2 = 0: individual was censored at T2 (the subject did not experience the event at either T1 or T2)

Note that for the first two scenarios above, subjects would contribute a single record in the data set (for the T1 indicator), whereas they would contribute two records (one for each person-time indicator T1 and T2) for the latter two scenarios. These indicators would represent the dependent variable in the analysis, akin to the variable named EVENT in our TVSFP data.

For this data, there were three follow-up occasions, and thus three person-time indicators are necessary to describe the occurrence of event/censoring. The three person-time indicators form the EVENT variable in the data set, and the timing of the event/censoring is represented by the two variables TimeC and TimeD in the data set. The coding of the person-time indicators (T1, T2, T3) that form the EVENT variable are given in Table 1.1.

Table 1.1: Three time points with censoring

Outcome	Up to 3 records per person
Censor at T1	T1 = 0
Event at T1	T1 = 1
Censor at T2	T1 = 0; T2 = 0
Event at T2	T1 = 0; T2 = 1
Censor at T3	T1 = 0; T2 = 0; T3 = 0
Event at T3	T1 = 0; T2 = 0; T3 = 1

Table 1.2: Coding of time and event indicators for binary TVSFP analysis

EVENT records	Time indicators		Outcome description
	TimeC	TimeD	
T1 =0	0	0	Censor at T1
T1 = 0 T2 = 0	0 1	0 0	No event at T1 Censor at T2
T1 = 0 T2 = 0 T3 = 0	0 1 0	0 0 1	No event at T1 No event at T2 Censor at T3
T1 =1	0	0	Event at T1
T1 = 0 T2 = 1	0 1	0 0	No event at T1 Event at T2
T1 = 0 T2 = 0 T3 = 1	0 1 0	0 0 1	No event at T1 No event at T2 Event at T3

Note that each person would contribute from one to three records in the data set depending on their outcome. For example, for the current data, the EVENT records and their corresponding time indicators are coded as shown in Table 1.2.

The breakdown of cigarette onset for gender and condition subgroups is presented in Table 1.3. Percentages given in the table are calculated relative to the totals for that subgroup at the time of response.

At Wave B (post-intervention time point; TimeC = 0 and TimeD = 0), 130 females (SexM = 0) and 156 males (SexM = 1) reported an event (Event = 1), while 105 females and 83 males were censored (Event = 0). These censored subjects did not experience the event at Wave B

and were not measured at subsequent waves. The total numbers of females and males that provided data at Wave B were 814 and 742 respectively. The totals at Wave C (TimeC =1) are only 579 and 503 females and males, respectively because the numbers of Wave B event and censored subjects are removed from the Wave C totals. For example, the total number of females at Wave C equals 814 (the number at Wave B) – 130 (females experiencing the event at Wave B) – 105 (censored females at Wave B) = 579. The male total of 503 is obtained in the same way. Of the 579 females, 117 experienced the event at Wave C and 154 were censored at Wave C. Similar calculations for Wave D (TimeD =1) yield the total of 308 females (= 579 – 117 – 154), where 79 females experienced the event and 229 did not and were censored at this last time point. Regarding the differences between males and females, it can be seen that the proportion of males who experienced the event is relatively similar across the three waves. Alternatively, females were initially lower than males (16% versus 21% at Wave B) but increasingly experienced the event across the waves. At the end, the total proportion of males who experienced the event is 41.5% (156 + 89 + 63 of 742), and similarly it is 40.0% for females (130 + 117 + 79 of 814). Thus, the initial gender difference is largely gone by the end of the study.

In terms of the invention groups, the differences do not appear to be very large. If anything, there is some suggestion that control subjects have lower rates of the event, but this difference is not striking.

Table 8:3: Onset of cigarette experimentation across three time points

	TimeB			TimeC			TimeD		
	with event	censored	total	with event	censored	total	with event	censored	total
Males	156 (21.0)	83 (11.2)	742	89 (17.7)	134 (26.6)	503	63 (22.5)	217 (77.5)	280
Females	130 (16.0)	105 (12.9)	814	117 (20.2)	154 (26.6)	579	79 (25.6)	229 (74.4)	308
Control	66 (16.5)	60 (15.0)	401	53 (19.3)	69 (25.1)	275	34 (22.2)	119 (77.8)	153
CC only	75 (19.1)	27 (6.9)	392	53 (18.3)	61 (21.0)	290	49 (27.8)	127 (72.2)	176
TV only	71 (17.3)	54 (13.2)	410	60 (21.1)	79 (27.7)	285	38 (26.0)	108 (74.0)	146
CC & TV	74 (21.0)	47 (13.3)	353	40 (17.2)	79 (34.1)	232	21 (18.6)	92 (81.4)	113

In terms of clustering, these 1556 students were from 28 schools with between 13 and 151 students per school ($\bar{n}$ = 56, S.D. = 38) Thus, the data are highly unbalanced with large variation in the number of clustered observations.

Binary case: a 2-level model

In the binary case, the survival time of individual i at occasion j is treated as a set of dichotomous observations indicating whether or not an individual failed in each time unit until a person either experiences the event or is censored. Thus, each survival time is represented as a $t_{ij} \times 1$ vector of zeros for censored individuals, while for individuals experiencing the event the last element of this $t_{ij} \times 1$ vector of zeros is changed to a one. These multiple person-time indicators are then treated as distinct observations in a dichotomous regression model. In the case of clustered data, a random-effects dichotomous regression model is used. This method has been called the pooling of repeated observations method by Cupples (1985). It is particularly useful for handling time-dependent covariates and fitting nonproportional hazards models because the covariate values can change across each individuals' t_{ij} time points.

For this approach, define p_{ijt} to be the probability of failure in time interval *t*, conditional on survival prior to *t*:

$$p_{ijt} = \Pr\left[t_{ij} = t \mid t_{ij} \geq t \right]$$

Similarly, $1 - p_{ijt}$ is the probability of survival beyond time interval t, conditional on survival prior to *t*. The proportional hazards model is then written as

$$\log\left[-\log\left(1 - p_{ijt}\right) \right] = \alpha_{0t} + \mathbf{x}'_{ijt}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{v}_i$$

and the corresponding proportional odds model is

$$\log\left[p_{ijt} \big/ \left(1 - p_{ijt}\right) \right] = \alpha_{0t} + \mathbf{x}'_{ijt}\boldsymbol{\beta} + \mathbf{z}'_{ij}\mathbf{v}_i$$

where now the covariates **x** can vary across time and so are denoted as $\mathbf{x}_{ijt}$. Augmenting the model intercept, which we will denote α_{01}, the remaining intercept terms $\alpha_{0t}\ (t = 2, \ldots, m)$ are obtained by including as regressors $m-1$ time indicators representing deviations from the first time point. Because the covariate vector **x** now varies with *t*, this approach automatically allows for time-dependent covariates, and relaxing the proportional hazards assumption only involves including interactions of covariates with the $m-1$ time point dummy codes. It is further assumed that the random effects vector has a $\boldsymbol{N}\left(\mathbf{0}, \boldsymbol{\Phi}_{(2)}\right)$ distribution.

In the examples to follow, he type of intervention (CC and/or TV), the gender of the student and the interactions between gender and time (SexTC and SexTD) are included as fixed effects, along with indicators of the time of assessment (TimeC and TimeD).

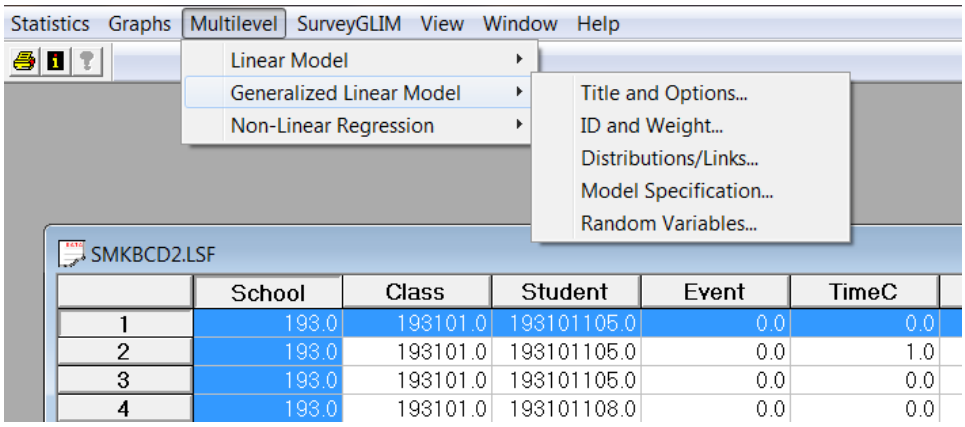
The first model fitted to the data will use the binary case and is of the form

$$\log\left[-\log\left(1-p_{ijt}\right)\right]=\alpha_{01}+(TimeC_{ij})\alpha_{02}+(TimeD_{ij})\alpha_{03}+(SexM_{ij})\beta_1+(CC_j)\beta_2 +(TV_j)\beta_3+v_{0i}.$$

In the current model specification, the baseline hazard is a function of the model intercept and the coefficients for the time indicators. Specifically, the baseline hazard estimate at the first time point equals the estimated model intercept, the baseline hazard estimate at the second time point is the sum of the model intercept and the estimated coefficient for the TimeC indicator, the baseline hazard at the third time point is the sum of the model intercept and the estimated coefficient for the TimeD indicator. Thus, two of these baseline hazard estimates involve sums of the estimated parameters.

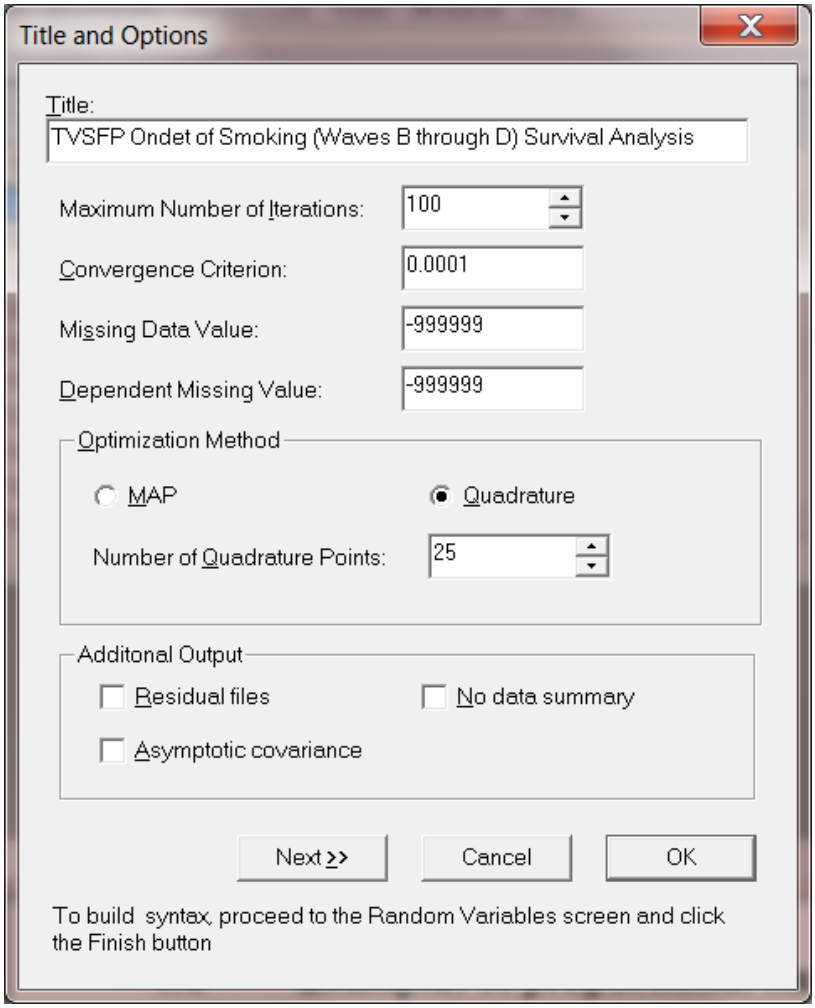
Setting up the analysis

Start by opening the file **SMKBCD2.LSF** from the **Multilevel Generalized Linear Model Examples** folder selecting the **Multilevel**, **Generalized Linear** option from the main menu bar as shown below.



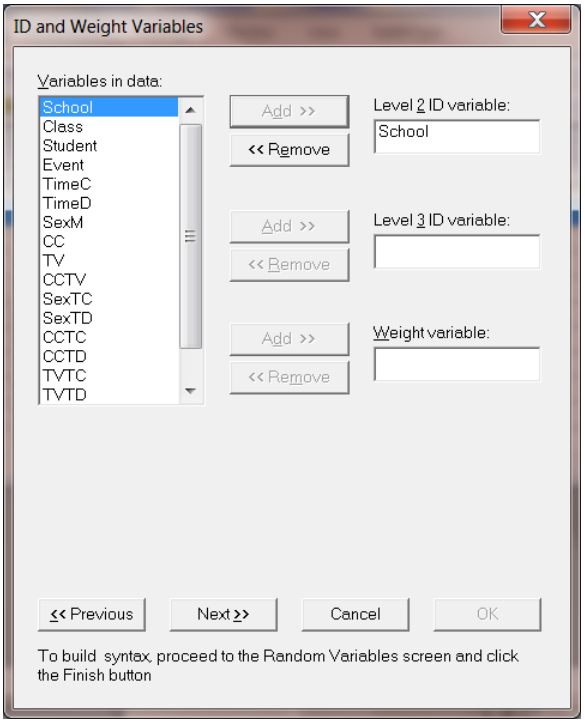
	School	Class	Student	Event	TimeC
1	193.0	193101.0	193101105.0	0.0	0.0
2	193.0	193101.0	193101105.0	0.0	1.0
3	193.0	193101.0	193101105.0	0.0	0.0
4	193.0	193101.0	193101108.0	0.0	0.0

Enter (optional) a title in the **Title and Options…** dialog.

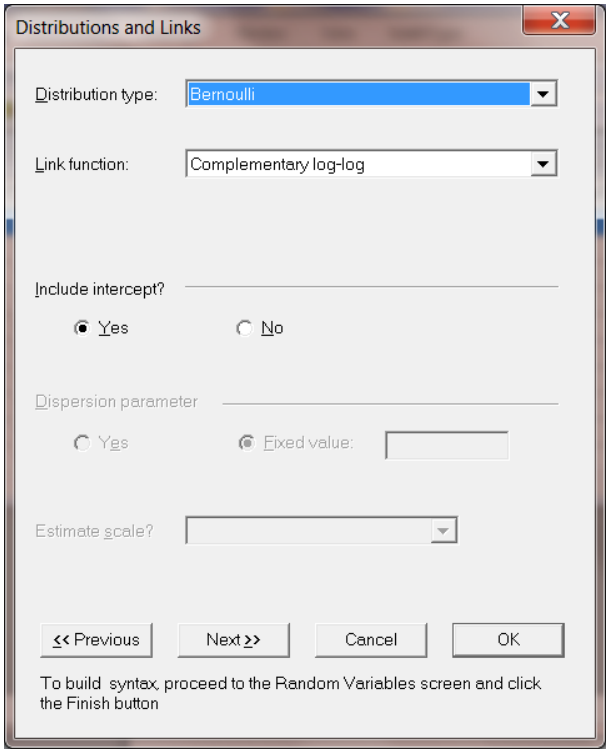


Change the number of quadrature points to 25, then click the **Next** button to obtain the **ID and Weight…** dialog.

The variable School, which defines the units within which students are nested, is selected as the Level-2 ID from the **Level-2 IDs** drop-down list box.



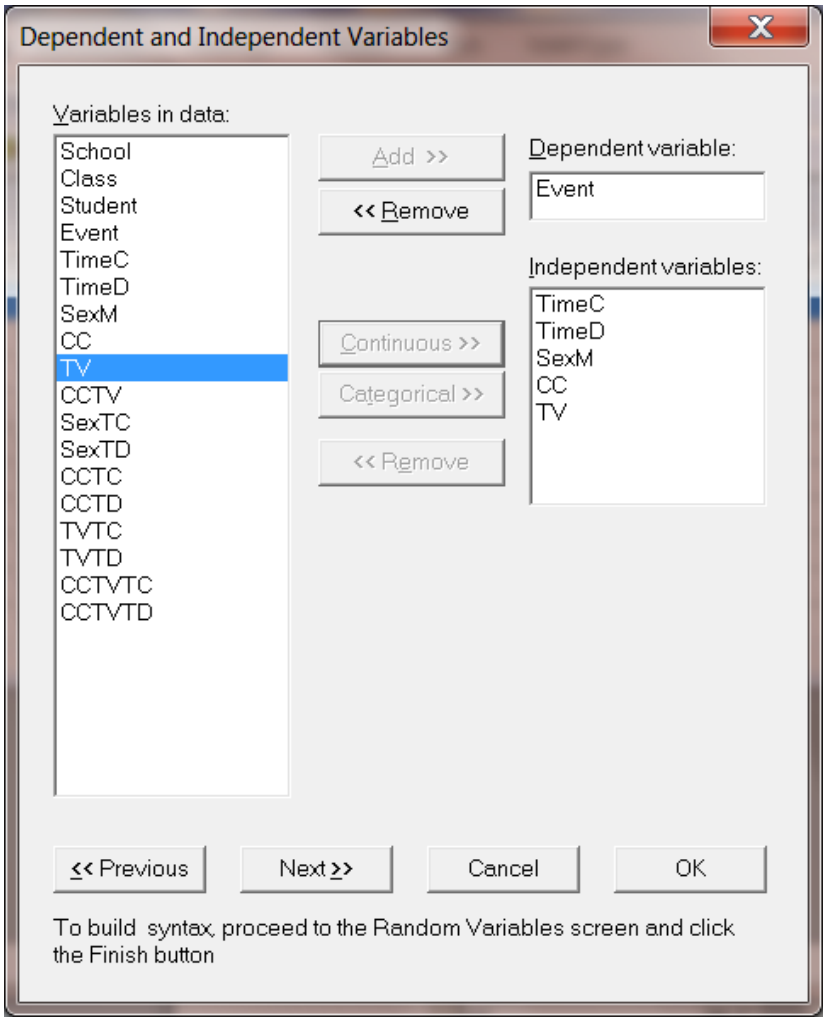
Next, proceed to the **Distributions and Links** dialog. Select Bernoulli as the **Distribution type** and Complementary log-log as the **Link function** as shown below.

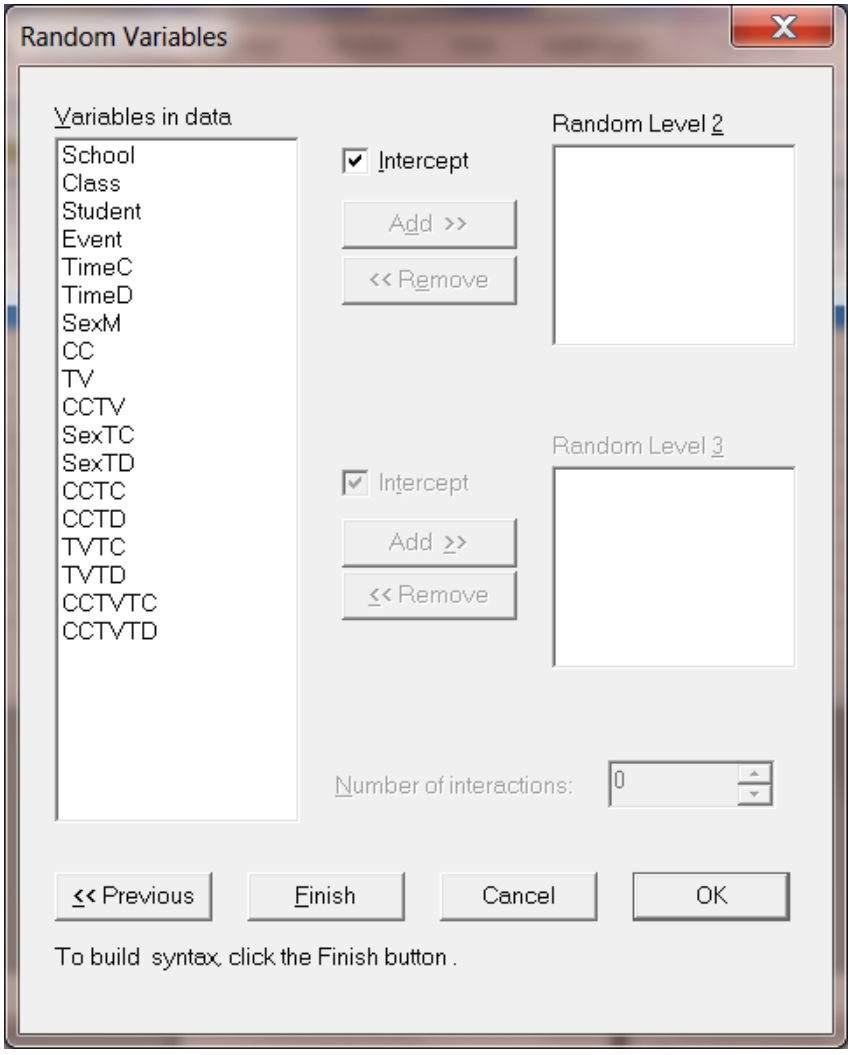


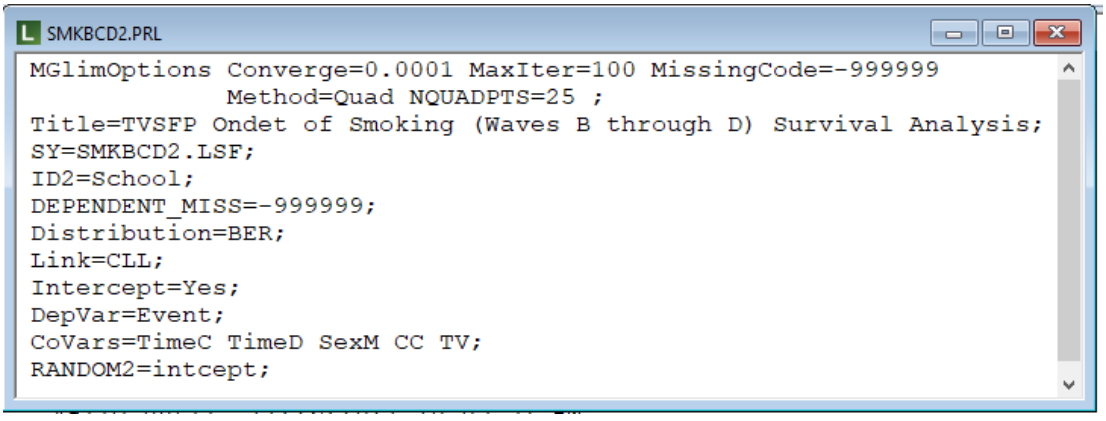
Click the **Include intercept?** Radio button and then click the **Next** button to obtain the **Dependent and Independent Variables** dialog.

Select the **binary** dependent (outcome) variable Event from the **Variables in data** list. Once done, TimeC, TimeD, SexM, CC, and TV are selected as the predictors (independent variables) of the fixed part of the model as shown below.

Finally, the **Random Variables** dialog is selected and Intercept is selected as the only random variable. To produce a syntax file, click the **Finish** button on the **Random Variables** dialog.







Next, click the **Run Prelis** icon on the main menu bar to run the analysis as shown below.

Discussion of results

Data summary

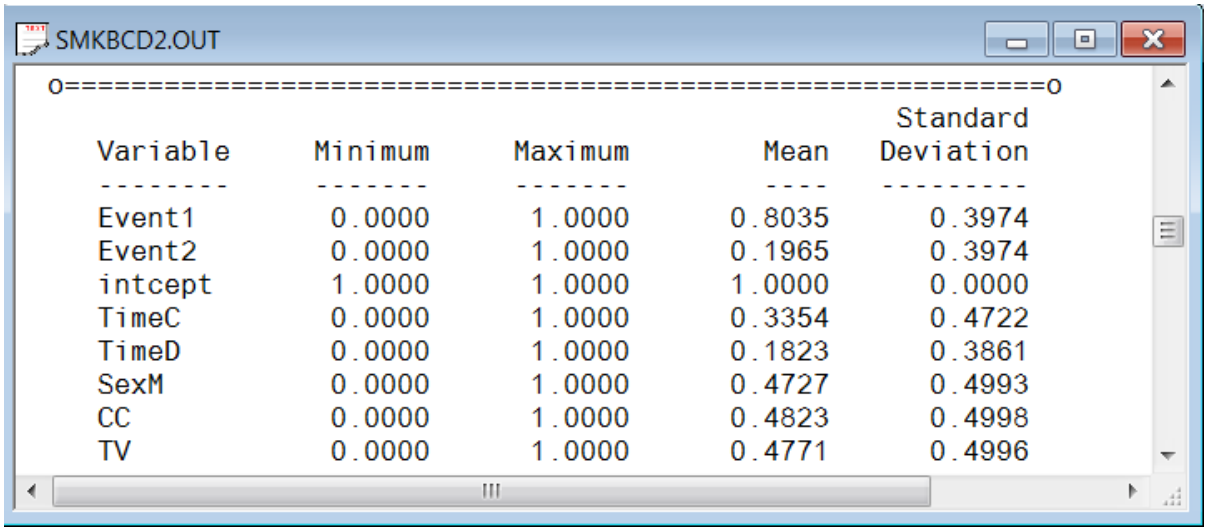
The portion of the output file shown below indicates that there are 28 schools. Nested within these level-2 units are 3226 measurements (note: this is not equal; to the number of students because of the creation of person-time indicators in this binary version of the survival analysis model). A summary of the number of level-1 observations per level-2 unit is also given.

SMKBCD2.OUT

```
o=============================================================o
| TVSFP Ondet of Smoking (Waves B through D) Survival Analysis |
|                                                             |
o=============================================================o


            Model and Data Descriptions

   Sampling Distribution                    = Bernoulli
   Link Function                            = Complementary Log-Log (CLL)
   PROB(Success)= 1.0-EXP[-EXP(ETA)]

   Number of Level-2 Units                    28
   Number of Level-1 Units                    3226
   Number of Level-1 Units per Level-2 Unit =
    51   163    57    30    59   153    34    42    62    56    42    64
    86    99    54    46   175    82   123   200   155   264   117   194
   202   150   172   294
```

Descriptive statistics

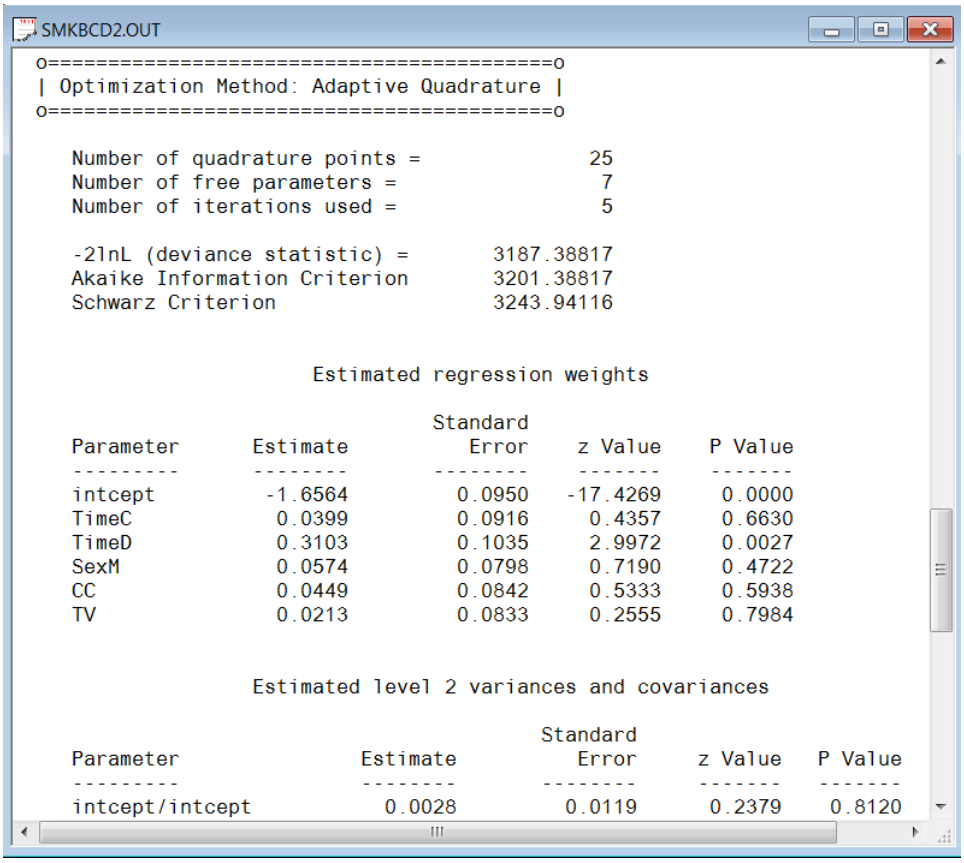
This is followed by descriptive statistics for all the variables. Except for the intercept term, the variables are all dichotomous. The proportions of subjects assigned a value of 0 or 1 are 0.80347 and 0.19653 respectively. In approximately 20% of the person-time indicators, an event occurred.



Variable	Minimum	Maximum	Mean	Standard Deviation
Event1	0.0000	1.0000	0.8035	0.3974
Event2	0.0000	1.0000	0.1965	0.3974
intcept	1.0000	1.0000	1.0000	0.0000
TimeC	0.0000	1.0000	0.3354	0.4722
TimeD	0.0000	1.0000	0.1823	0.3861
SexM	0.0000	1.0000	0.4727	0.4993
CC	0.0000	1.0000	0.4823	0.4998
TV	0.0000	1.0000	0.4771	0.4996

Fixed effects estimates

Parameter estimates are given in the next part of the output. The effect of SexM is positive and indicates that boys have a slightly, but non-significant, increased hazard (*i.e.*, a shorter time to the first occurrence), relative to girls. The coefficients associated with the TimeD indicator variable is significant at a 5% level. In contrast, the corresponding TimeC coefficient is not significant. These indicate that the baseline hazard does not significantly change between Waves B and C, however there is significant change between Waves B and D as relatively more students experiment with smoking at Wave D. Finally, the effects of the intervention variables CC and TV are not seen to be statistically significant, though the direction of their effects is positive (*i.e.*, increased hazard relative to the control group).



Intraclass correlation (ICC)

The last part of the output contains an estimate of the intracluster correlation. This estimate indicates a very modest school effect, and we also note that the random effect variance term is not significant. From this, we conclude that the time until the occurrence of an event does not vary significantly across schools. However, from a design point of view, because schools were randomized to the intervention conditions in this study, one can argue that the clustering attributable to schools is an important part of the model regardless of its significance.

```
Calculation of the intracluster correlation
------------------------------------------
residual variance = pi*pi / 6 (assumed)
cluster  variance =     0.0028

intracluster correlation =     0.0028 / (    0.0028 + (pi*pi/6)) = 0.002
```

Population Average Estimates

Parameter	Estimate	Standard Error	z Value	P Value
intcept	-1.6553	0.0942	-17.5721	0.0000
TimeC	0.0399	0.0916	0.4357	0.6630
TimeD	0.3102	0.1034	2.9984	0.0027
SexM	0.0574	0.0798	0.7190	0.4721
CC	0.0449	0.0842	0.5334	0.5938
TV	0.0213	0.0833	0.2554	0.7984

Interpreting the output

Estimated unit-specific probabilities

We now use the estimated coefficients from the fitted model

$$\begin{aligned}\log\left[-\log(1-p_{ijt})\right] &= \hat{\alpha}_{01} + (TimeC_{ij})\hat{\alpha}_{02} + (TimeD_{ij})\hat{\alpha}_{03} + (SexM_{ij})\hat{\beta}_1 + (CC_j)\hat{\beta}_2 + (TV_j)\hat{\beta}_3 \\ &= -1.6564 + (TimeC_{ij})0.0399 + (TimeD_{ij})0.3103 + (SexM_{ij})0.0574 \\ &\quad + (CC_j)0.0449 + (TV_j)0.0213\end{aligned}$$

and the inverse cumulative log-log link function

$$P(z) = 1 - \exp[-\exp(z)]$$

to calculate the probability of Event = 1 at various time points and for different covariate values.

At the first time point (Wave B), $TimeC_{ij} = TimeD_{ij} = 0$, and thus the relevant part of the fitted model (see above) is

$$\begin{aligned}\log\left[-\log(1-p_{ijt})\right] &= \hat{\alpha}_{01} + (SexM_{ij})\hat{\beta}_1 + (CC_j)\hat{\beta}_2 + (TV_j)\hat{\beta}_3 \\ &= -1.6564 + (SexM_{ij})0.0574 + (CC_j)0.0449 + (TV_j)0.0213\end{aligned}$$

For female students (SexM = 0) from the control group (CC = TV = 0) the probability of smoking experimentation (Event = 1) at the point of post-intervention can be expressed as

$$P(Event = 1 at WaveB,\ female) = 1 - \exp\left[-\exp(-1.6564)\right]$$
$$= 0.1737.$$

For male students in the control group adding the intercept with the SexM estimate together yields $z = -1.6564 + 0.0574 = -1.599$, and so

$$P(Event = 1 at WaveB, male) = 1 - \exp\left[-\exp(-1.599)\right]$$
$$= .1830.$$

Results for all groups are summarized in Table 1.5. The probability of smoking experimentation at the time of post-intervention is larger for males than for females. The results also indicate an increased probability of failure with an increase of time. In the current model, it is assumed that the ratio of the estimated hazards over time will be constant for two individuals with the same values on the covariates. To check whether the effect of gender is dependent on time, and thus to check on the proportional hazards assumption, interactions with time indicators should be included in the model.

Table 1.5: Unit-specific probabilities for groups

Gender	CC	TV	WaveB (TimeC = 0, TimeD = 0)	WaveC (TimeC = 1, TimeD = 0)	WaveD (TimeC = 0, TimeD = 1)
Female	0	0	0.1737	0.1801	0.2291
	1	0	0.1809	0.1876	0.2383
	0	1	0.1771	0.1836	0.2335
	1	1	0.1844	0.1912	0.2428
Male	0	0	0.1830	0.1897	0.2409
	1	0	0.1905	0.1975	0.2505
	0	1	0.1865	0.1933	0.2454
	1	1	0.1942	0.2012	0.2551

Table 1.6 shows the differences between the estimated unit-specific probabilities and the observed proportions for each of the 24 subgroups formed by crossing all predictors currently in the model.

Looking at the direction of the differences, we note that for females all the estimated probabilities are larger in size than the observed ratios at WaveB, but consistently lower than the observed ratios at the next two time points, with the exception of the situation where

TimeD = CC = TV = 1. It seems as if the model is overestimating the probabilities of failure at the first time point, but underestimating probabilities at the last time of measurement. However, the pattern for males is almost the opposite. At the first wave, only one estimated probability is larger than the observed proportion, at WaveC this is true for 2 of the four cells, and at WaveD for three of the four cells.

Table 1.6: Differences between unit-specific probabilities and observed proportions

Gender	CC	TV	Difference at WaveB (estimated – observed)	Difference at WaveC (estimated – observed)	Difference at WaveD (estimated – observed)
Female	0	0	0.0227	–0.0419	–0.0179
	1	0	0.0149	0.0016	–0.0117
	0	1	0.0091	–0.0174	–0.0875
	1	1	0.0204	–0.0058	0.0568
Male	0	0	–0.0150	–0.0073	–0.0361
	1	0	0.0165	–0.0025	0.0625
	0	1	–0.0275	0.0303	0.0064
	1	1	–0.0678	0.0613	0.0710

This trend could be the result of a gender effect (which we know to be non-significant in the current model) or from an interaction between gender and time. While only TimeD had a significant estimated coefficient, this apparent trend leads us to conclude that testing of the assumption of proportional hazards is appropriate. Specifically, the interaction between gender and the time of measurement will be explored.