Weighted 2-level models

1. The data

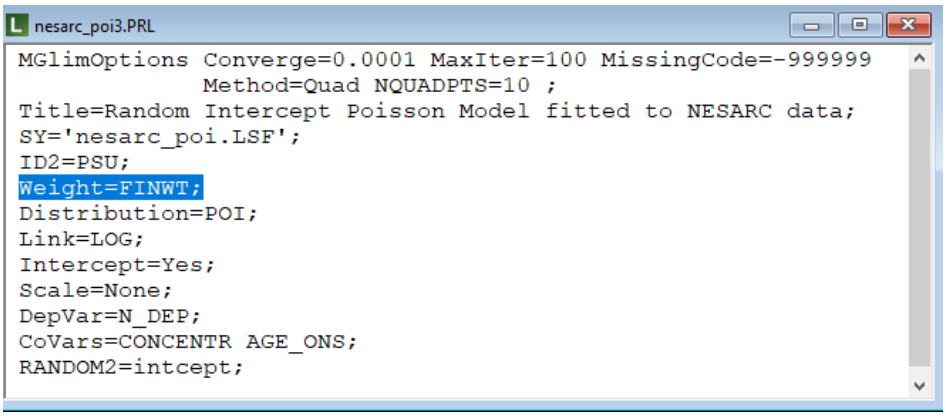
The sampling frame of many multistage surveys frequently entails selection of units with known, but unequal, selection probabilities. This situation is the result of a number of design factors, of which the cost of doing the survey is an important consideration. When this is the case, it is appropriate to weigh observations in order to produce unbiased estimates of the population parameters.

The variable FINWT represents the NESARC weights sample results used to form national-level estimates. The final weight is the product of the NESARC base weight and other individual weighting factors. In this section, we explore the effect of inclusion of the weights on the results obtained in the section describing the unweighted Poisson model.

2. Setting up the analysis

The models remain the same, and we again use **nesarc_poi.lsf** from the **Multilevel Generalized Linear Model Example** folder.with only the selection of the weight variable on the **Advanced** tab of the **Model Specification** screen to be added. Below, we show how this is done in the case of the Poisson distribution model.

Open the syntax file for the Poisson distribution model (**nesarc_poi1.prl**). Save the syntax file file as **nesarc_poi3.prl** and add Weight = FINWT below the **ID2** paragraph as shown below.



Save the file and run the analysis.

3. Discussion of results

Results for this analysis are reported in the table below. The results from the unweighted Poisson distribution model are included in order to facilitate evaluation of the impact of the weights on the results.

Comparison of results for weighted and unweighted Poisson models

Parameter	Unweighted model		Weighted model	
	Estimate	Standard error	Estimate	Standard error
intcept	0.7982	0.0641	0.7225	0.0660
CONCENTR	0.2922	0.0510	0.3055	0.0532
AGE_ONS	–0.0165	0.0012	–0.0156	0.0013
Level-2 variance	0.1347	0.0184	0.1378	0.1089

Results for the two models are very similar, and interpretation of the results of both models would lead to the same conclusions, both in terms of significance and in terms of the expected number of depression episodes. However, this is more the exception than the rule – users are cautioned to use weight variables whenever they are available in order to prevent skewed or biased results that may occur when weights are excluded in the analysis of a disproportionally drawn sample.