Fixed and Random Effects

When analysing data from a study in which you have a nominal predictor variable (or an ordinal predictor which you are treating as nominal), you must decide if that variable will be treated in the analysis as a fixed effect or a random effect. The calculations needed to carry out the analysis in the two cases are different, although generally if you are using software for your calculations, you will not need to worry about this, but will rather only need to correctly label the variable when you set up the analysis.

The difference between fixed and random effects depends on how many levels (possible outcomes) of the variable occur in your sample and what levels of the variable you want to be able to draw conclusions about.

There are three standard cases:

1. You have sampled all possible levels of the variable. For instance, the variable is "gender" and you have included both males and females in your study. In this case, gender is a fixed effect.

2. You have only sampled some levels of your variable, but you only care about these levels in your conclusions. For instance, your variable is "make of car". You have included Toyotas, Hondas and Hyundais in your study, and are only interested in comparing these three makes in your conclusions. In this case, make of car is a fixed effect.

3. You have only sampled some levels of your variable, but you would like to be able to make conclusions about all levels of the variable. For instance, your variable is "primary school attended" and you have chosen students for your study from a sample of 100 primary schools around the UK, but you want to be able to draw conclusions about students at any UK primary school. In this case, primary school attended is a random effect.

Note that if you have only sampled a few levels of the variable out of several thousand possible, you are not likely to get reliable results treating it as a random effect. Some thought is thus required at the planning stage of your study to determine how you will want to treat the variable and how you need to arrange your sampling to suit this. If you already have data, you need to make a judgement if you have sampled enough levels to justify treating the variable as a random effect.

There is one other situation in which you might choose to treat a nominal predictor variable as a random effect. This situation is a bit more statistically advanced than the three above. This is the situation in which you do not have a very large sample size, and are more interested in other predictors than in the precise effect of each level of the particular nominal predictor in question. In this case, by treating the nominal predictor as a random effect, you can get a better estimate of the effects of other predictors than if you treated it as a fixed

effect. This is due to the form of the models used for random versus fixed effects and the issue of deciding how many parameters you can afford in your model given a sample size that is not large. A random effect variable requires fewer parameters than a fixed effect variable with several levels. There is more information about this topic under the glossary entry “parameters”.