**Analysis of a Binary Outcome with SPSS**

These instructions were included in the original Tips and Tricks with SPSS. I have made minor updates (7 June 2016). They relate to logistic regression when there is a single measurement per subject. For repeated measurements see "SPSS generalized mixed models.docx".

Import the dataset "binary outcomes.xls" to work your way through these instructions. I generated the data with the binary-outcomes spreadsheet in the Excel workbook accompanying the article on Understanding stats via simulations. The dependent is Event, with values 0 and 1 in control and exptal subjects defined by the variable Group. The predictor is some subject characteristic X, which I have recoded into a variable XM0SDp5 ("X with mean 0 and SD .5"), as per the next point. The variable ModeledP is the "true" probability that the event will occur for each subject, modeled from population values that are shown in the binary-outcomes spreadsheet. You can ignore this variable.

Transform any numeric covariates by making new covariates with a mean of zero and a standard deviation of 0.5. Do it in Excel, not SPSS. This strategy is necessary to make any sense of the output. If there are certain values of the covariates (other than the means) at which you are interested in evaluating proportion ratio, create more new variables from the covariates so that the new variables have a mean of zero at the values of interest. Make sure they have SDs of 0.5, too.

1. From the menu bar, select Analyze/Regression/Binary Logistic…
2. In the Logistic Regression window that opens, select the dependent (an outcome variable having only two levels 0 and 1, female and male, etc.).
   1. Select each covariate (predictor), then select those you want to interact by control-clicking on them, then click the >a\*b> button.
   2. If you have any categorical covariates, click on Categorical, click on the variable, choose Indicator (it should be the default), click on the Reference Category (e.g., if the levels are "control" and "exptal", control comes before exptal alphabetically, so if you want exptal-control, click on First), **then don't forget to click Change**. Click Continue.
3. Back in the Logistic Regression window, click Options, select CI for exp(B) and make it 90%.
   1. Select Display/At last step. Keep the default Include constant in model. Click Continue.
   2. Click OK.
4. Scroll to the bottom (ignore everything else). Exp(B) has the odds ratios for each effect, and the odds (not odds ratio) for the constant.
   1. You now have to jump through various hoops to convert odds ratios into relative proportions, which are more meaningful. You use the fact that odds = proportion/(1-proportion), or equivalently proportion = odds/(1+odds).
   2. The odds ratios for categorical covariates are sort-of controlled for other covariates: that is, they are the odds ratios when there are equal numbers of subjects on the other levels of any other categorical covariates, **but when the numeric covariates are zero. This is not what is usually meant by controlling for a numeric covariate.** In fact, what you have here is the solution to the model, not what SPSS should provide, the so-called estimated marginal means or least-squares means. So make sure you have already transformed the numeric covariates to have a mean of zero (or to have a value of zero corresponding to any other values of the covariates(s) at which you want to evaluate the relative proportion).
   3. The value for Constant is the odds with the categorical covariates at their reference levels and numeric covariates at zero. The proportion for subjects at this level or these levels is odds/(1+odds) = Exp(B)/(1+Exp(B). You can do the same thing to the confidence limits, if you want them. To get the proportion at another level, multiply the Exp(B) for the constant by the Exp(B) for the other level. You now have the odds for the new level, so convert it to a proportion using the same odds/(1+odds) formula. Sorry, now you *can't* do the same thing with the confidence limits. To get those, you will have to rerun the analysis and choose a new reference level.
   4. Get the relative proportion for the new level vs the reference level by dividing one proportion by the other. The confidence limits for this proportion ratio have to come from my spreadsheet "Confidence limits and clinical chances. Use the panel for log-normally distributed effects, and insert the relative proportion and the p value (the value in the Sig. column).
   5. To convert the odds ratios for the numeric covariates to relative proportions, you can't just use Exp(B)/(1+Exp(B). Instead, you have to first convert the odds ratio of the covariate into a proportion for someone who already has a certain proportion, a reference proportion. The most obvious levels are those of the categorical covariates. So, you multiply the odds for that level by the odds ratio for the covariate, then convert the resulting odds into a proportion using the odds/(1+odds) formula. Its confidence limits come from the p value for the numeric covariate. Note that you have evaluated the effect of two SDs of the covariate. The smallest worthwhile relative proportion for the effect of two SDs of a covariate is the same as what you would consider to be the smallest worthwhile effect between two levels of a categorical predictor, 1.11 and its inverse, 0.9. If the proportions are small (<10%), you can interpret the odds ratio directly as a relative proportion.