Chapter 6 in Hoffman Plus Some Notes from Allison

1. Poisson Regression with proc genmod

Poisson Distribution Assumptions:

- Mean = Variance
- Events are independent

Not independent then variance > mean. Known as overdispersion.

If mean > variance then underdispersion.

Mean = Variance known as equidispersion.

Because the Poisson model uses a log link function, the parameter estimates represent the expected change in the log scale. If you calculate $100(e^{\hat{\beta}}-1)$, you obtain the percent change in the expected number of events with each one-unit increase in the predictor variable.

For example, if $e^{\hat{\beta}}=0.80$, then a one-unit increase in X yields a 20% decrease in the estimated mean.

Aim is to predict expected counts for various groups.

Suggest you start with a full model then remove the nonsignificant factors one at a time, starting with the least-significant one (**width**), to reduce your model. See Allison for SAS coding examples and discussion.

2. Negative Binomial Regression with proc genmod

Negative binomial model is more appropriate than Poisson when events are not independent. See Allison for SAS coding examples and discussion.

3. Zero Inflated Poisson

In case there are a lot of zero counts the poisson regression model will give a poor fit. So, the zero inflated poisson model gives special treatment to the records with zero count. See Allison for SAS coding examples and discussion.