Pearson & likelihood ratio test statistics

I will now continue looking at a goodness-of-fit test statistic for Poisson regression so we would like to test if a particular model that we are assuming really does fit the data or if we may want to extend our model and include maybe more covariance in our model.

We first of all look at another example so I'll first of all talk that through and then I'll be discussing goodness of fit test statistic so the Pearson chi-squared test and the log likelihood ratio test. So we've got the example of a recall of stressful events for example 3 and basically we would like to first of all start with something simple again a simple Poisson model without any covariates but then extend this to an example with co-variance.

So basically participants from a randomized study were asked if they can recall any stressful events over the last in fact 18 months and if yes in which months when did this particular stressful event happen and we then wanted to look at the number of a stressful event that people were able to recall and look at their distribution according to these 18 months so we had a 147 stressful events recorded in total. The H0 hypotheses is that first of all we started something very simple and started with some big sort of conservative attitude conservative assumption that these events are uniformly distributed over time. That basically means that H0 follows the equiprobable model that all these probabilities across months are the same so we have an event occurrence of 1 /18 so particular event can occur in any given months so it would be a percentage of 0.055 or in terms of percentage terms it's a 5.5% so we would expect about just over 5% of all events to happen / months. Looking at the actual count data at the actual data that was recorded so we've got first of all months ranging from 1 to 18 and then we've got the actual count data so the actual number of events that were recorded per month and then the percentage that relates to the actual count variable. So we can already see that some counts are significantly actually higher than 5.5 percent and some are actually a bit lower than 5.5% so just looking at the data we may already conclude that there is some divergence between or discrepancy between the observed values and what we would expect to see based on our equiprobable model.

Let's look at the evaluation of the Poisson model doing this more formally. So we will be using the Pearson chi-squared test and the deviance or also called the log likelihood ratio test for Poisson regression. Basically both are goodness of fit test statistics and basically compares two models one for the current model the model that we have attained so in this particular case the equiprobable model and then we compare this model with the so-called saturated model i.e. the model that is a yeah larger that is the saturated model that is the model that fits the data perfectly and that explains all of the viability. That basically means we are comparing observed and expected frequencies.

Looking at the Pearson and the log likelihood ratio statistics, basically we say that if H0 is true that means that if the equiprobable model actually holds then we would expect overall the distribution of 147 times 1/18 so that we would expect frequency of just over 8 / months basically.

So we basically have on parameter model that we would like to estimate and that would be just over eight events per month so we can compare the count for the observed counts with the expected count per month and we can see that obviously some counts are quite a bit higher than maybe 8 and some counts per month are quite a bit lower than 8 that we we would expect. Looking at the Pearson chi-square test it allows us basically to compare the observed and expected frequencies and you may have come across the Pearson chi-squared test and testing associations between two categorical variables so it's the same principle effectively here we are trying to compare observed and expected frequencies and basically it allows us to look at the sum of the standardized residuals in squared terms and we can calculate this particular statistics for for our example. So basically you just have to plug in the numbers for each cell basically so we've got 18 cells in total and for this particular example recall of stressful events we will have the chi-square test statistic are 45.4 and we now need to compare this to the value from a chi square distribution so basically the assumption is that if H0 is true if indeed the equiprobable model holds then this test statistic will follow the chi-square distribution so we can compare it with the distribution from the chi square tables for example so for that we need to have the degrees of freedom which is defined as the number of cells - the number of model parameters which is the number of cells as 18 so C- the model parameters here in this particular case for the simple model is just 1 because we have got only alpha that we need to estimate. So we've got the chi-square test statistic of 45.4 with basically 17 degrees of freedom so 18 - 1 at the 5% significance level. Looking at the chi-squared table value we've got 27.6 from the table and also we have got that basically associates with the p value of really rather small 0.001 so that means the the value we would then reject H0 based on those characteristics.

Conclusion is that there is strong evidence that the equiprobable model does not fit the data so just looking at the actual data from our table we've already looked at observed and expected frequencies and these already saw some discrepancy but exactly how significant the discrepancy is the discrepancy between observed and expected values we can then formally test this for example here with the Pearson chi-squared test and we concluded that the difference is rather large so it is not just due to chance and that the data does not follow the equiprobable model. So we have to do probably something else to improve this particular model.

Likewise looking at the log likelihood ratio test statistics for Poisson regression we can now also use this test statistics so again compare observed and expected frequencies so basically the formula here gives the log likelihood ratio test statistic and you can plug in the numbers the observed and expected frequencies and that is again a measure of the fit of the model, so the goodness of fit test statistics. And again similarly to before if H0 is true that actually this particular log likelihood test statistics would follow a chi square distribution, so again we need to define the degrees of freedom which again is the number of cells minus the number of model parameters so again it's 17 for example and the log likelihood ratio test statistics for our example is 50.8 if you plug in the numbers on 17 degrees of freedom so we've got a p-value of less than 0.001 again and you would again project H0, so basically in the same way as with the Pearson chi-squared test. So the conclusion is again the same there is strong evidence that the equiprobable model that does not fit the data. So basically now need to use this information to take this forward and to think about possibly fitting a more sophisticated model maybe including another co-variate to allow for differences between numbers of months. Just a couple of remarks about the Pearson chi-square test and the log likelihood ratio test. They're basically asymptotically equivalent. So basically they are relying on the large sample and you would expect them to be giving very similar results. And if they're not similar this could just simply be an indication that the large sample approximation doesn't actually hold. And also just to know that for fixed degrees of freedom so when an increase for larger samples that the distribution of the Pearson chi-squared tests usually converges to the actual chi square distribution and also it does it more quickly than the log likelihood ratio test.

And also to note that the chi-squared approximation is usually relatively poor or not appropriate if any of the expected cells are less than 5.