Ling 236 Homework #2 Answers

Sue has counted the available Kaurna sentences and has this data:

	Voice	
Aspect	Active	Passive
NonPerfective	5	0
Perfective	1	2

1. What are the maximum likelihood probability estimates for a nonperfective sentences to be active/passive? What are the maximum likelihood probability estimates for a perfective sentences to be active/passive?

   Answer: $\hat{P}(\text{active}|\text{nonperfective}) = 1$; $\hat{P}(\text{passive}|\text{nonperfective}) = 0$; $\hat{P}(\text{active}|\text{perfective}) = 1/3$; $\hat{P}(\text{passive}|\text{perfective}) = 2/3$.

2. Do a chi-square test (of homogeneity) to determine whether there seems to be an association between aspect and voice in this data, and give the result.

   Answer: $X^2 = 4.44$, which gives significance at a $p = 0.035$, i.e., $p < 0.05$ level, so to the extent that the assumptions of the chi-square test are met, this result is moderately strong evidence of an association between aspect and voice in Kaurna.

3. Is this test appropriate here? Why or why not?

   Answer: No. Even if the sampling, etc., is all valid, the chi-square distribution test is only a good approximation to the probability of this data being generated by chance under conditions of homogeneity if the sample size is sufficiently large. Ideally, we would like the expected value of all squares to be above 5; minimally we would like the average expected value to be above 5 and no cell to have an expected value below 1. Here, though, the data is so limited that no cell has an expected value above 5, and one cell has an expected value below 1.

4. State in words what one might be testing doing a Fisher's Exact Test on the same data (there's more than one way to phrase it). (I mean not just "are voice and aspect associated?", but a more detailed description.)

   Answer: If the probabilities of generating active and passive and perfective and nonperfective are as shown by the marginal totals, what is the probability that, just by chance, the distribution of items in the cells would come out as or more extreme (further from independence) than the observed values.

5. Do a Fisher's Exact Test on this data. Can one conclude that there is an association between aspect and voice in Kaurna?

   Answer: For the top left cell, say, – any cell could be chosen – it's value must be between 3 and 5. Using the hypergeometric distribution the chance of it taking on each of these values is: $P(3) = 0.357$; $P(4) = 0.536$; and $P(5) = 0.107$. So, we certainly can't reject the null hypothesis of independence between the two factors at the $p = 0.05$ level. [This result shows that the chi-square test was indeed a bad approximation of the true probability of generating equally or more extreme data – it underestimated the true probability by a factor of 3.]

6. What if Sue discovered on a manuscript 3 new examples of perfective sentences, 2 of which were passive. So the data now looks like:

	Voice	
Aspect	Active	Passive
NonPerfective	5	0
Perfective	2	4

Does this new data change what one can conclude about the association between aspect and voice in Kaurna?

Answer: Doing the top left cell again, one gets the following table:

1	0.01515
2	0.18182
3	0.45455
4	0.30303
5	0.04545

Under a 1-tailed test, this would just squeak in significance at the $p < 0.05$ level. But, the original question was simply "is there an association between aspect and voice", which could only naturally be interpreted as a question that deserves a 2-tailed test. So, we should take the sum of the probability of all extreme values, and the chance of having got 1 active non-perfective would have been a more extreme violation of non-association in the opposite direction, and so it should be counted as well. Thus the probability of an equally or more extreme association arising by chance is $p = 0.06$, and so the data still fails to achieve significance at the $p < 0.05$ level (even though it is a bit more suggestive of a trend than before...).

7. What assumptions does one have to make about this data sample for any of the above to be valid?

   Answer: One needs to assume that these sentences are representative of Kaurna as a whole (that is, randomly sampled from Kaurna, or at least randomly sampled on one of the dimensions). In particular, one needs to assume that the sentences are independent of each other: for example, if the sentences came from one text, textual cohesion tends to produce a strong effect where passives follow passives and probably the same is true for aspect. This could badly bias the results.