Stat 312: Lecture 27
Testing for Independence

Moo K. Chung
mchung@stat.wisc.edu

May 6, 2003

Concepts

1. Suppose that each individual in a population is classified with respect to two different *factors* into a two-way contingency table of size $I \times J$.

2. The null hypothesis of interest would be testing if the two factors are independent:

$$H_0 : p_{ij} = p_{i\bullet} p_{\bullet j} \text{ for all } i, j,$$

where $p_{i\bullet} = \sum_{j=1}^{J} p_{ij}$ and $p_{\bullet j} = \sum_{i=1}^{J} p_{ij}$.

3. Under H_0, the expected number of element $\mathbb{E}N_{ij} = np_{ij} = np_{i\bullet}p_{\bullet j}$. Since the marginal probabilities are unknown, we estimate them by $\hat{p}_{i\bullet} = n_{i\bullet}/n$, the sample proportion for category i for factor 1 and $\hat{p}_{\bullet j} = n_{\bullet j}/n$, the sample proportion for category j for factor 2. So $\mathbb{E}N_{ij} = n_{i\bullet}n_{\bullet j}/n$.

4. Test statistic

$$\chi^2 = \sum_{i,j} \frac{(n_{ij} - \mathbb{E}N_{ij})^2}{\mathbb{E}N_{ij}} \sim \chi^2_{(I-1)(J-1)}.$$

This is the test statistic that has been used in testing homogeneity!

In-class problems

Excercise 14.41. Test if the parental use of alcohol and drugs is independent of the student usage of marijuana for 445 college students.

parent\student	never	occasional	regular
neither	141	54	40
one	68	44	51
both	17	11	19

Self-study problems

Example 14.14.
Example. Show that the two hypotheses $H_0 : p_{ij} = p_{i\bullet}p_{\bullet j}$ for all i, j and $H_0 : p_{1j} = p_{2j} = \cdots = p_{Ij}$ for all j are equivalent. Hence testing for homogeneity is equivalent to testing independence.
The End of Lectures. ra...ra...ra... Thank you all. In the next class, I will solve some sample final exams.