STATS305B – Lecture 3

Jonathon Taylor
Scribed by Michael Howes

10/03/22

Contents

1 Associations in ordinal data

Last lecture we defined the odds ratio

$$\theta = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}},$$

which measured association in a 2×2 table. For $I \times J$ tables, we also defined the local odds ratios

$$\theta_{ij} = \frac{\pi_{ij}\pi_{i+1,j+1}}{\pi_{i,j+1}\pi_{i+1,j}},$$

for $i = 1, \ldots, I-1$ and $j = 1, \ldots, J-1$. The local odds ratios are a collection of $(I-1)(J-1)$ parameters that describe the associations in a $I \times J$ table. If X or Y are ordinal variables rather than categorical variables, then we can use fewer parameters to describe the association. Consider the following table:

	Career satisfaction		
Age	1	2	3
< 30	34	53	88
$30 - 50$	80	174	304
> 50	29	75	172

In this example both X and Y are ordinal and we would like to know does a higher X (age) result in a higher Y (career satisfaction). To answer this question we can use a single parameter of interested instead of analyzing the four local odd ratios. This is done by studying the probabilities of concordance and discordance.

Note that we can put a partial order on the values (i, j) where $(i, j) \succ (h, k)$ if $i < h$ and $j < k$. In the previous table we would have

$$(> 50, 3) \succ (30 - 50, 2), (30 - 50, 1), (< 30, 2), (< 30, 1).$$

We would not have $(> 50, 3) \succ (> 50, 2)$ or $(> 50, 3) \succ (< 30, 3)$ since we do not allow for either X or Y to tie. A pair of observations (X_1, Y_1) and (X_2, Y_2) are said to be *concordant* if $(X_1, Y_1) \succ (X_2, Y_2)$ or $(X_2, Y_2) \succ (X_1, Y_1)$. In our example a concordant pair is a pair of individuals where one individual is strictly old, and the older one has strictly higher career satisfaction. A pair (X_1, Y_1) and (X_2, Y_2) are *discordant* if $(X_1, Y_2) \succ (X_2, Y_1)$ or $(X_2, Y_1) \succ (X_1, Y_1)$. Thus, a discordant pair is again a pair where one individual is strictly older, but now the older individual has strictly lower career satisfaction. We can then define

$$\pi_c = \mathbb{P}(\text{drawing a concordant pair}) = 2\sum_{ij} \pi_{ij} \sum_{h>i,k>j} \pi_{hk},$$

and

$$\pi_d = \mathbb{P}(\text{drawing a discordant pair}) = 2\sum_{ij} \pi_{ij} \sum_{h>i,k<j} \pi_{hk}.$$

The parameter

$$\gamma = \frac{\pi_c - \pi_d}{\pi_c + \pi_d},$$

is a measure of the ordinal association between X and Y. Under the null that there is no ordinal association, $\gamma = 0$. We can estimate γ by

$$\widehat{\gamma} = \frac{C - D}{C + D},$$

where

$$C = 2\sum_{ij} n_{ij} \sum_{h>i,k>j} n_{hk} \text{ and } D = 2\sum_{ij} n_{ij} \sum_{h>i,k<j} n_{hk}.$$

In our example,