XII. Latent Class Analysis

A. Setup

1. One unobserved categorical variable $W$

2. $L \geq 2$ observed categorical variables $X_j$

3. $P[\mathbf{X} = \mathbf{x} | W = w] = \prod_{l=1}^{L} P[X_l = x_l | W = w]$

4. Generally fewer parameters than saturated model.

5. Example: 4 manifest variables, each with two levels
   a. 2 class model has $2 \times 4 + 1 = 9$ parameters.
   b. Corresponding saturated model has 15 parameters.
   c. 1 class model has 4 parameters

6. Model for $I \times J \times \cdots \times L$ table with $k$ classes can be made a submodel of the model with $k - 1$ classes by
   a. setting class probability to 0 or 1: 1 restriction.
   b. setting two sets of conditional probabilities equal:
      $$(I - 1) + (J - 1) + \ldots + (K - 1)$$ restrictions
   c. Varying dimension of submodel makes it irregular, and standard $\chi^2$ likelihood ratio techniques don’t work.