g. Another justification for $\chi^2$
   i. Each of the summands $X_j$ is approximately $N(m\pi_j, m\pi_j(1-\pi_j))$
   ii. Then $(X_j - m\pi_j)^2 / [m\pi_j(1-\pi_j)] \sim \chi^2$
   approximately
   iii. Sum should be $\sim \chi^2$
      • Except that summands are dependent
      • Compensate by Dropping 1 - $\pi_j$ from denominator and adjusting df.
   F: 13.6

13. Tests of Equality of Proportions
   a. $X_j \sim \text{Bin}(\pi_j, m_j)$ for $j = 1, \ldots, k$, all independent
   b. $H_0 : \pi_1 = \ldots = \pi_k$ vs. $H_A : \pi_i \neq \pi_j$ for some $i, j$.
   c. Approach 1: Generalized Likelihood Ratio
      i. $L(\pi_1, \ldots, \pi_k) = \prod_{j=1}^k \left( \frac{m_j}{\hat{\pi}_j} \right)^{X_j} (1 - \hat{\pi}_j)^{m_j - X_j}$
      ii. $\hat{\pi}_j = (\sum_{j=1}^k X_j) / (\sum_{j=1}^k m_j)$, $\hat{\pi}_j = X_j / m_j$
      iii. $\Lambda = \frac{\prod_{j=1}^k \left( \frac{m_j}{\hat{\pi}_j} \right)^{X_j} (1 - \hat{\pi}_j)^{m_j - X_j}}{\prod_{j=1}^k \left( \frac{X_j}{\hat{\pi}_j} \right)^{X_j} (1 - \hat{\pi}_j)^{m_j - X_j}}$
   iv. No simplification possible
   v. Take $-2 \log(\Lambda)$
   vi. Approximate by $\chi^2$
   d. Approach 2: Approximate $-2 \log(\Lambda)$ using Taylor series as before:
      $\approx \sum_{j=1}^k \left( \frac{X_j - m_j\hat{\pi}}{m_j\hat{\pi}} \right)^2 + \sum_{j=1}^k \frac{(m_j - X_j - m_j(1 - \hat{\pi}))^2}{m_j(1 - \hat{\pi})}$
   $= \sum_{j=1}^k \left( \frac{X_j - m_j\hat{\pi}}{m_j\hat{\pi}} \right)^2$
   and compare to $\chi^2_{k-1}$
   e. Call $m_j\hat{\pi}_j$ and $m_j(1 - \hat{\pi}_j)$ the expected number of successes and failures resp.
   f. Describe statistic as sum of observed minus expected, squared, divided by expected, for successes and failures, and all $k$ categories.

Lecture 23

14. Tests of Equality of Proportions – Multinomial Case
   a. $X_j \sim \text{Mul}(\pi_j, m_j)$ for $j = 1, \ldots, k$, all independent
   b. $H_0 : \pi_1 = \ldots = \pi_k$ vs. $H_A : \pi_i \neq \pi_j$ for some $i, j$.
   c. Generalized Likelihood Ratio
      i. $L(\pi_1, \ldots, \pi_k) = \prod_{j=1}^k \prod_{i=1}^l \pi_{ji}^{x_{ji}}$
      • Multinomial coefficients omitted since they cancel
      ii. $\hat{\pi}_{ij} = (\sum_{l=1}^k X_{li}) / (\sum_{l=1}^k m_l)$, $\hat{\pi}_{ji} = X_{ji} / m_j$
      iii. $\Lambda = \prod_{j=1}^k \prod_{i=1}^l \left( \frac{\hat{\pi}_{ji}^{x_{ji}}}{\hat{\pi}_{ji}} \right)^{X_{ji}}$
      iv. No simplification possible
      v. Take $-2 \log(\Lambda)$
      vi. Approximate by $\chi^2$
      • Degrees of freedom are $(k - 1)(J - 1)$
   vii. Test statistic and null distribution are same if you swap rows and columns