STAT 770 Sep. 21 Lecture Part B
Examples for 2-way Tables

Reading for this lecture:

Same as part A: Chapter 2 in Agresti, plus review of Delta Method in Sec. 16.1.

Here we look at a variety of data formats and CIs for unknown parameters in 2x2 tables. The parameters are mostly the ones used to formulate tests, so that LRTs can be compared with Wald-type tests. The data example throughout is:

$$N_{11} = 47,\ N_{01} = 40,\ N_{10} = 33,\ N_{00} = 50$$

Delta Methods for RR and OR Estimates

In slides 2-3, n_z fixed and $N_{z1} \sim \text{Binom}(n_z, \pi_z)$ indep., $z = 0, 1$

RR case: $b = \binom{\pi_1}{\pi_0} \mapsto \beta = \binom{\pi_1/\pi_0}{\pi_0}, \quad \nabla_{\mathbf{b}}\beta_1 = \binom{1/\pi_0}{-\pi_1/\pi_0^2}$

$\widehat{\pi}_z = N_{z1}/n_z \sim \mathcal{N}\Big(\pi_z, n_z^{-1}\pi_z(1-\pi_z)\Big)$ **independent**

$$\text{a.var}(\widehat{\beta}_1) = \left(\frac{1}{\pi_0}\right)^2 (1, -\beta_1) \begin{bmatrix} \pi_1(1-\pi_1)/n_1 & 0 \\ 0 & \pi_0(1-\pi_0)/n_0 \end{bmatrix} \binom{1}{-\beta_1}$$

So can read off a.var and use approx. $\mathcal{N}$ to construct (Wald-type) CIs.

logOR: $\mathbf{b} = \binom{\pi_1}{\pi_0} \mapsto \beta = \left(\log\left\{\frac{\pi_1(1-\pi_0)}{(1-\pi_1)\pi_0}\right\}, \, \pi_0 \right), \qquad \frac{\partial \beta_1}{\partial b_z} = \frac{1}{\pi_z(1-\pi_z)}$

$\widehat{\pi}_z = N_{z1}/n_z \sim \mathcal{N}\Big(\pi_z, \, n_z^{-1}\pi_z(1-\pi_z)\Big)$ **independent**

$$\text{a.var}(\widehat{\beta}_1) = \sum_{z=0}^{1} \frac{1}{(\pi_z(1-\pi_z))^2} \pi_z(1-\pi_z) \frac{n}{n_z} = \sum_{z=0}^{1} \frac{n}{n_z \pi_z (1-\pi_z)}$$

Again read off a.var and use approx. $\mathcal{N}$ to construct (Wald-type) CIs.

Note. **a.var** is the variance for the limiting dist'n for $\sqrt{n}(\widehat{\beta} - \beta)$

OR CIs in 2×2 Table

First plug into the **avar** formula for $N_{11} = 47$, $N_{01} = 40$, $N_{10} = 33$, $N_{00} = 50$, $n_1 = 80$, $n_0 = 90, n = 170$

avar $= 16.4185$, CI $= \log\left\{\frac{47*50}{33*40}\right\} \pm 1.96\sqrt{\frac{16.42}{170}} = (-0.032,\ 1.186)$

How to do LRT inverted interval in this case?

We would express π_1 as function of β_1, π_0, put into likelihood for general β_1, and maximize in π_0 for fixed β_1. Start by

$$\frac{\pi_1}{1-\pi_1} = e^{\beta_1}\frac{\pi_0}{1-\pi_0} \Rightarrow \pi_1 = e^{\beta_1}\pi_0 \Big/ \Big(1 - \pi_0 + e^{\beta_1}\pi_0\Big)$$

$$L(\beta_1, \pi_0) = \pi_1^{N_{11}}(1-\pi_1)^{N_{10}}\pi_0^{N_{01}}(1-\pi_0)^{N_{00}}$$

LRT interval will be $\{\beta_1 :\ \mathrm{LRT}(\beta_1, \widehat{\pi}_{r,0}) \le \chi^2_{1,.05}\}$

Multinomial 2×2 **Tables**

Data $N_{zx},\ z = 0, 1,\ x = 0, 1,$ jointly Multinomial $(n,\ \{\pi_{zx}\})$

Parameter $\beta = (\log\left\{\frac{\pi_{11}\pi_{00}}{\pi_{10}\pi_{01}}\right\},\ \pi_{+1},\ \pi_{1+})\ ,\quad \beta_1 =$ Log Odds Ratio

$$\widehat{\beta}_1 = \log\left\{\frac{N_{11}N_{00}}{N_{10}N_{01}}\right\}$$ log of cross product

$$\nabla'\beta_1 = \left(\frac{1}{\pi_{11}} - \frac{1}{\pi_{00}},\ \frac{-1}{\pi_{01}} - \frac{1}{\pi_{00}},\ \frac{-1}{\pi_{10}} - \frac{1}{\pi_{00}}\right)$$ use in Δ Method

$$\text{a.var}(\widehat{\beta}_1) = \nabla'\beta_1\left[\texttt{diag}(\{\pi_{zx}\}) - (\{\pi_{zx}\})(\{\pi_{zx}\})'\right]\nabla'\beta_1 = \sum_{z,x=0}^{1} 1/\pi_{zx}$$

Contrast Wald CI using this a.var vs LRT test-based CI

Odds Ratio CIs in Multinomial 2×2 Table

Now with same data, plug into **a.var**

$$= 170(\tfrac{1}{47} + \tfrac{1}{33} + \tfrac{1}{40} + \tfrac{1}{50}) = 16.418 \qquad \text{same as before.}$$

So the Wald-type Confidence Interval for logOR is the same as before $\quad = (-0.032, 1.186)$

LRT test inversion requires expressions for all π_{zx} as functions of β_1, π_{+1}, π_{1+} substituted in likelihood $\prod_{z,x} \pi_{zx}^{N_{zx}}$ and maximized over π_{+1}, π_{1+}

LRT interval will be $\quad \{\beta_1 : \text{LRT}(\beta_1, \widehat{\pi}_{r,+1}, \widehat{\pi}_{r,1+}) \le \chi^2_{1,.05}\}$

Not the same as previous LRT interval.

Remaining Mode of Condtioning: fixing Marginals

This is the basis for the **Fisher exact test** as in Problem **(2)** on HW2.

LRT intervals are also possible using the Extended Hypergeometric derived in HW1 problem **(C).(b)**.