2. Methods based on inverting the score statistic

Let $\Delta = p_1 - p_2$ be the difference in proportions, with MLE $\hat{\Delta} = \hat{p}_1 - \hat{p}_2$. The score statistic for testing $H_0 : \Delta = \Delta_0$ as given by Mee (1984) equals

$$\begin{aligned} S_{\mathrm{M}}(\Delta_0) &= (\hat{\Delta} - \Delta_0)^2 \,/\, \widetilde{\mathrm{Var}}_{\Delta_0}\,, \quad \text{where} \\ \widetilde{\mathrm{Var}}_{\Delta_0} &= \frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2}\,. \end{aligned}$$

Here, $\tilde{p}_2$ is the restricted MLE of p_2, resulting from maximizing the product binomial likelihood under the null hypothesis H_0 (i.e., replacing p_1 with $p_2 + \Delta_0$ and maximizing w.r.t. p_2) and $\tilde{p}_1 = \tilde{p}_2 + \Delta_0$. A closed form solution for $\tilde{p}_2$ is given in Nurminen (1986). Note that if $\Delta_0 = 0$, $\tilde{p}_2 = \tilde{p}_1 = (y_1 + y_2)/(n_1 + n_2)$, the pooled proportion. Then, $E\left[\widetilde{\mathrm{Var}}_{\Delta_0}\right] = \frac{n_1+n_2-1}{n_1+n_2}\mathrm{Var}[\hat{\Delta}]$, showing that $\widetilde{\mathrm{Var}}_{\Delta_0}$ is a slightly biased estimator of $\mathrm{Var}[\hat{\Delta}]$, the variance of the numerator of the score statistic. This lead Miettinen and Nurminen (1985) to propose a bias-corrected score statistic (even when $\Delta_0 \neq 0$)

$$\begin{aligned} S_{\mathrm{MN}}(\Delta_0) &= (\hat{\Delta} - \Delta_0)^2 \,/\, \widetilde{\mathrm{Var}}_0\,, \quad \text{where} \\ \widetilde{\mathrm{Var}}_0 &= \left(\frac{n_1+n_2}{n_1+n_2-1}\right)\left(\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2}\right), \end{aligned}$$

with $\tilde{p}_2$ and $\tilde{p}_1$ the same restricted MLEs as in $S_{\mathrm{M}}(\Delta_0)$. For large n_1 or n_2, the two statistics are virtually identical ($S_{\mathrm{MN}}(\Delta_0) = \frac{n_1+n_2-1}{n_1+n_2} S_{\mathrm{M}}(\Delta_0)$), but for small n_1 and n_2 Newcombe and Nurminen (2011) found that incorporating this bias adjustment is very effective in bringing the level of the test close to the nominal one and that $S_{\mathrm{MN}}(\Delta_0)$ is preferred over $S_{\mathrm{M}}(\Delta_0)$ for all Δ_0.

2.1. Asymptotic score interval

We get an asymptotic $100(1-\alpha)\%$ confidence interval by inverting the score test. Inverting $S_{\mathrm{MN}}(\Delta_0)$ by finding all values for $\Delta_0 \in [-1, 1]$ for which

$S_{\mathrm{MN}}(\Delta_0) \leq \chi_{1-\alpha}$, where $\chi_{1-\alpha}$ is the $1-\alpha$ quantile of the Chi-square distribution (with $df = 1$) is not possible in closed form. However, the two endpoints of the interval (i.e., the two solutions to $S_{\mathrm{MN}}(\Delta_0) - \chi_{1-\alpha} = 0$ in terms of Δ_0) can easily be obtained by numerical methods such as interval halving or some other root finding algorithm. The Web-App mentioned in the introduction provides this interval as well as the one obtained by inverting $S_{\mathrm{M}}(\Delta_0)$. Chapter 3 shows that the interval obtained by inverting $S_{\mathrm{MN}}(\Delta_0)$ has excellent coverage probabilities. In addition, for score inference no artificial adjustments, such as adding pseudo observations are needed when $y_i = 0$ or $n_i, i = 1, 2$.

2.2. Exact interval based on the score statistic

Rather than using the asymptotic Chi-square distribution we can find the exact P-value of $S_{\mathrm{MN}}(\Delta_0)$, given by

$$p_{\mathrm{exact}}(\Delta_0) = \sup_{p_2} \sum_{y_1=0}^{n_1} \sum_{y_2=0}^{n_2} I(y_1, y_2, \Delta_0) \times \mathrm{bin}(y_1, n_1, p_2 + \Delta_0) \times \mathrm{bin}(y_2, n_2, p_2),$$

where the supremum is taken over the permissable range $(\max\{0, -\Delta_0\}, \min\{1, 1-\Delta_0\})$ for p_2 and $\mathrm{bin}(y, n, p)$ stands for the binomial probability of y successes in n trials with success probability p. Here, $I(y_1, y_2, \Delta_0)$ is a function equal to 1 when $S_{\mathrm{MN}}(\Delta_0)$ computed with that y_1 and y_2 is larger than the actually observed value of $S_{\mathrm{MN}}(\Delta_0)$, equal to $1/2$ if $S_{\mathrm{MN}}(\Delta_0)$ is exactly equal to the observed value and equal to 0 otherwise. This uses the mid-P approach in the calculation of the exact P-value, which does not guarantee a coverage probability above the nominal level but reduces conservativeness. (For a guaranteed coverage probability, choose $I(y_1, y_2, \Delta_0)$ equal to 1 when $S_{\mathrm{MN}}(\Delta_0)$ is *as large* or larger than the observed value.) When finding the supremum over the entire permissable range is too computationally expensive, one can restrict the region to a 99.99% confidence interval for p_2. However, since we use exact computations of coverage and expected length here, we are already constructing a fine grid of p_2 values.