Mid-p Adjustment for exact2x2: Computational Details

Michael P. Fay

August 2, 2020

Overview

These notes give details on how the mid-p adjustment is done. Section 1 describes the mid-p adjustment as it is done for the exact2x2 and uncondExact2x2 functions. Section 2 describes the mid-p adjustment as implemented in the binomMeld.test function.

1 Usual Mid-p Adjustment for Two Binomial Distributions

The following is how the usual mid-p adjustment is done (for example in the exact2x2 and uncondExact2x2 functions). The mid-p value has a long history (see e.g., Lancaster, 1961 or the list of references in Hirji 2006, p. 50).

Let $\mathbf{X} = [X_1, X_2]$ with $X_a \sim Binom(n_a, \theta_a)$ for $a = 1, 2$. Suppose we are interested in $\beta = b(\theta)$, where $b(\theta)$ is some function of θ_1 and θ_2. Common examples are the difference, $\beta_d = \theta_2 - \theta_1$, the ratio, $\beta_r = \theta_2/\theta_1$, and the odds ratio, $\beta_{or} = \{\theta_2(1-\theta_1)\} / \{\theta_1(1-\theta_2)\}$.

Let $T(\mathbf{X})$ be some test statistic, where larger values are most extreme with respect to the null hypothesis. Let Θ_0 be the set of all possible values of $[\theta_1, \theta_2]$ under the null hypothesis. Then a valid (i.e., exact) p-value is

$$p(\mathbf{x}, \Theta_0) = \sup_{\theta \in \Theta_0} Pr_\theta \left[T(\mathbf{X}) \geq T(\mathbf{x}) \right].$$

These exact p-values are necessarily conservative because for most $\theta \in \Theta_0$ we have $Pr_\theta \left[p(\mathbf{X}, \Theta_0) \leq \alpha \right] < \alpha$. A less conservative approach, *but one that is no longer valid (i.e., no longer exact)*, is to use a mid-p value. The mid-p value is

$$p_{mid}(\mathbf{x}, \Theta_0) = \sup_{\theta \in \Theta_0} \left\{ Pr_\theta \left[T(\mathbf{X}) > T(\mathbf{x}) \right] + \frac{1}{2} Pr_\theta \left[T(\mathbf{X}) = T(\mathbf{x}) \right] \right\}.$$

It is convenient to write Θ_0 in terms of β. For example,

$$\Theta_0 = \{\theta : b(\theta) = \beta_0\}$$

For this example, instead of writing the null hypothesis as $H_0 : \theta \in \Theta_0$, we write it in terms of $\beta = b(\theta)$ as $H_0 : \beta = \beta_0$. We are generally interested in three classes of hypotheses: two-sided hypotheses,

$$\begin{aligned} H_0 : &\quad \beta = \beta_0 \\ H_1 : &\quad \beta \neq \beta_0 \end{aligned}$$

or one of the one-sided hypotheses,

Alternative is Less	Alternative is Greater
$H_0 : \beta \geq \beta_0$	$H_0 : \beta \leq \beta_0$
$H_1 : \beta < \beta_0$	$H_1 : \beta > \beta_0$.

Let $p_{ts}(\mathbf{x}, \beta_0)$ be the p-value for testing the two-sided hypotheses, let $p_U(\mathbf{x}, \beta_0)$ be the p-value for testing $H_0 : \beta \geq \beta_0$, and $p_L(\mathbf{x}, \beta_0)$ be the p-value for testing $H_0 : \beta \leq \beta_0$.

Then we can create $100(1-\alpha)\%$ confidence regions as the set of β_0 value that fail to reject the associated null hypothesis. For example,

$$C_{ts}(\mathbf{x}, 1-\alpha) \quad = \quad \{\beta : p_{ts}(\mathbf{x}, \beta) > \alpha\}$$

gives a "two-sided" confidence region. The region may not be an interval if the p-value function is not unimodal. This problem occurs with Fisher's exact test (the Fisher-Irwin version, or 'minlike' version). For central confidence regions we take the union of the one-sided confidence regions, in other words,

$$C_c(\mathbf{x}, 1-\alpha) = C_L(\mathbf{x}, 1-\alpha/2) \cup C_U(\mathbf{x}, 1-\alpha/2),$$

where C_L and C_U are the one-sided confidence regions,

$$C_L(\mathbf{x}, 1-\alpha/2) \quad = \quad \{\beta : p_L(\mathbf{x}, \beta) > \alpha/2\}$$

and

$$C_U(\mathbf{x}, 1-\alpha/2) \quad = \quad \{\beta : p_U(\mathbf{x}, \beta) > \alpha/2\}\,.$$

If the regions are intervals, and we let $L(\mathbf{x}, 1-\alpha/2) = \min C_L(\mathbf{x}, 1-\alpha/2)$ and $U(\mathbf{x}, 1-\alpha/2) = \max C_U(\mathbf{x}, 1-\alpha/2)$, then the central interval is

$$C_c(\mathbf{x}, 1-\alpha) \quad = \quad \{L(\mathbf{x}, 1-\alpha/2), U(\mathbf{x}, 1-\alpha/2)\}\,.$$

For the mid-p confidence regions, we replace the p-values with the mid-p values.

2 Mid-p Modifications with Binomial Melding

For a single binomial response, the mid p-value and associated central confidence interval can be represented using confidence distribution random variables. Suppose that the exact central 100(1-α) percent binomial confidence interval for a single binomial random variable (i.e., the default in binom.test) is $(L(1-\alpha/2), U(1-\alpha/2))$. Then the lower and upper confidence distribution random variables are respectively, $W_L = L(A_1)$ and $W_U = U(A_2)$, where A_1 and A_2 and independent uniform random variables. Let B be an independent Bernoulli random variable with parameter 1/2. Then the 95 percent central mid-p confidence interval for the binomial parameter is the middle 95 percent of the distribution of $W = B * W_L + (1-B) * W_U$. This is shown in the appendix of Fay and Brittain (2016).

The way the midp=TRUE option is done in binomMeld.test is to replace the upper and lower confidence distribution random variables in the usual melding equations, with the "mid-p" confidence distribution random variable (CD-RV) analogous to W for each group. For example if the lower and upper CD-RVs for group 1 are W_{1L} and W_{1U}, then the mid-p CD-RV is $W_1 = B_1 * W_{1L} + (1 - B_1) * W_{1U}$, where B_1 is a Bernoulli random variable with parameter $1/2$. The mid-p CD-RV W_2 is defined analogously. It is fairly simple to program a Monte Carlo estimate of the "mid" p-value and associated confidence interval. Let $g(\theta_1, \theta_2)$ be the parameter of interest (e.g., $g(\theta_1, \theta_2) = \theta_2 - \theta_1$ for parmtype="difference"). The one-sided p-values are the proportion of times that $g(W_1, W_2)$ is $\leq$ nullparm (for alternative="greater") or $\geq$ nullparm (for alternative="less"). The confidence intervals just use the appropriate quantiles of the Monte Carlo values of $g(W1, W_2)$.

When nmc=0, we estimate the one-sided p-values with numeric integration. Conceptually, the usual melded p-value might be, for example when alternative="greater" and nullparm= β_0:

$$Pr[g(W_{1U}, W_{2L}) \leq \beta_0] \;=\; \int Pr[g(W_1, w_2) \leq \beta_0 | W_2 = w_2] Pr[W_2 = w_2]$$

where W_{1U} is the upper confidence distribution random variable (CD-RV) for group 1 and W_{2L} is the lower CD-RV for group 2. These CD-RVs are beta distributions (see Fay, Proschan, and Brittain, 2015). For the mid-p version, we use

$$\begin{aligned} Pr[g(W_1, W_2) \leq \beta_0] \;=\; & \frac{1}{4}\int Pr[g(W_{1L}, w_2) \leq \beta_0 | W_{2L} = w] Pr[W_{2L} = w] + \\ & \frac{1}{4}\int Pr[g(W_{1L}, w_2) \leq \beta_0 | W_{2U} = w] Pr[W_{2U} = w] + \\ & \frac{1}{4}\int Pr[g(W_{1U}, w_2) \leq \beta_0 | W_{2L} = w] Pr[W_{2L} = w] + \\ & \frac{1}{4}\int Pr[g(W_{1U}, w_2) \leq \beta_0 | W_{2U} = w] Pr[W_{2U} = w]. \end{aligned}$$

The integration simplifies for special cases (e.g., when $x_1 = 0$), and in other case we just use the integrate function. For the confidence intervals we solve for the β_0 values such that the p-values equal either α (for one-sided alternatives) or $\alpha/2$ (for two-sided alternatives), where alpha=1-conf.level. If there is no β_0 value that solves that, we set the confidence limit to the appropriate extreme.

It is known that the p-values that match the melded confidence intervals for two independent binomial observations exactly equal the one-sided Fisher's exact p-values (see Fay, et al, 2015). For example,

```
> x1<-6
> n1<-12
> x2<-15
> n2<- 17
> exact2x2(matrix(c(x2,n2-x2,x1,n1-x1),2,2), tsmethod="central", midp=FALSE)

        Central Fisher's Exact Test
```

```
data:  matrix(c(x2, n2 - x2, x1, n1 - x1), 2, 2)
p-value = 0.06506
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  0.9119249 89.4167455
sample estimates:
odds ratio
  6.924704

> binomMeld.test(x1,n1,x2,n2, parmtype="oddsratio",  midp=FALSE)

        melded binomial test for oddsratio

data:  sample 1:(6/12), sample 2:(15/17)
proportion 1 = 0.5, proportion 2 = 0.88235, p-value = 0.06506
alternative hypothesis: true oddsratio is not equal to 1
95 percent confidence interval:
   0.909023 106.265540
sample estimates:
odds ratio {p2(1-p1)}/{p1(1-p2)}
                             7.5
```

Note, the confidence intervals for the two methods are not equal.

This does not necessarily mean that the midp versions give equivalent p-values:

```
> x1<-6
> n1<-12
> x2<-15
> n2<- 17
> exact2x2(matrix(c(x2,n2-x2,x1,n1-x1),2,2), tsmethod="central", midp=TRUE)

        Central Fisher's Exact Test (mid-p version)

data:  matrix(c(x2, n2 - x2, x1, n1 - x1), 2, 2)
p-value = 0.03578
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
  1.12685 62.05021
sample estimates:
odds ratio
  6.924704

> binomMeld.test(x1,n1,x2,n2, parmtype="oddsratio",  midp=TRUE)

        melded binomial test for oddsratio, mid-p version
```

```
data:  sample 1:(6/12), sample 2:(15/17)
proportion 1 = 0.5, proportion 2 = 0.88235, p-value = 0.02899
alternative hypothesis: true oddsratio is not equal to 1
95 percent confidence interval:
  1.214721 66.148301
sample estimates:
odds ratio {p2(1-p1)}/{p1(1-p2)}
                             7.5
```

References

Fay, MP, and Brittain, EH (2016). "Finite sample pointwise confidence intervals for a survival distribution with right-censored data." *Statistics in Medicine* 35: 2726-2740.

Fay, MP, Proschan, MA, and Brittain, E (2015). "Combining One-Sample Confidence Procedures for Inference in the Two-Sample Case." *Biometrics* 71: 146-156.

hirji, K. F. (2006). *Exact Analysis of Discrete Data.* Chapman and Hal/CRC, New York.

Lancaster, H.O. (1961). Significance Tests in Discrete Distributions. *Journal of the American Statistical Association.* 56: 223-234.