Calculation Details for the Binomial Melding Test

Michael P. Fay

August 25, 2014

1 Overview

These notes give some calculation details for the binomial melding test. See Fay, Proschan and Brittain (2014) for the motivation and other melding test examples.

Suppose $X_i \sim Binomial(n_i, \theta_i)$ for two independent samples. We are interested in two-sample inferences on functions of the parameters, $\beta = g(\theta_1, \theta_2)$, such as the

difference: $\beta = g(\theta_1, \theta_2) = \theta_2 - \theta_1$,

ratio: $\beta = g(\theta_1, \theta_2) = \frac{\theta_2}{\theta_1}$, or

odds ratio: $\beta = g(\theta_1, \theta_2) = \frac{\theta_2(1-\theta_1)}{\theta_1(1-\theta_2)}$

The $100(1-\alpha)\%$ one-sided lower and upper one-sided melded confidence limits for $\beta = g(\theta_1, \theta_2)$ are

$$L_\beta(\mathbf{x}, 1-\alpha) = \text{the } \alpha\text{th quantile of } g\left\{U_{\theta_1}(\mathbf{x}_1, A), L_{\theta_2}(\mathbf{x}_2, B)\right\}, \tag{1}$$

and

$$U_\beta(\mathbf{x}, 1-\alpha) = \text{the } (1-\alpha)\text{th quantile of } g\left\{L_{\theta_1}(\mathbf{x}_1, A), U_{\theta_2}(\mathbf{x}_2, B)\right\}, \tag{2}$$

where $L_{\theta_i}(\mathbf{x}_i, q)$ and $U_{\theta_i}(\mathbf{x}_i, q)$ are the 100q% one-sided confidence limits for θ_i, and A and B are independent and uniform random variables. The one-sided confidence intervals can be combined to get two-sided intervals that are central ones. For example, a 95% central (and hence two-sided) confidence interval is $\left\{L_\beta(\mathbf{x}, 0.975), U_\beta(\mathbf{x}, 0.975)\right\}$.

For the binomial problem we use exact one-sided limits for $L_{\theta_i}(\mathbf{x}_i, q)$ and $U_{\theta_i}(\mathbf{x}_i, q)$. This means that

$$T_{Li} \equiv L_{\theta_i}(x_i, A) \sim Beta(x_i, n_i - x_i + 1)$$

and

$$T_{Ui} \equiv U_{\theta_i}(x_i, B) \sim Beta(x_i + 1, n_i - x_i).$$

except if $x_i = 0$ then $L_{\theta_i}(0, A)$ is a point mass at 0, and if $x_i = n_i$ then $U_{\theta_i}(n_i, B)$ is a point mass at 1.

For alternative is greater, we test

$$\begin{aligned} H_0 &: g(\theta_1, \theta_2) \leq \beta_0 \\ H_1 &: g(\theta_1, \theta_2) > \beta_0 \end{aligned}$$

Let $p_L(\beta_0)$ be the associated one-sided p-value, which is the solution to

$$L_\beta(\mathbf{x}, 1 - p_L(\beta_0)) = \beta_0.$$

For alternative is less, we test

$$\begin{aligned} H_0 &: g(\theta_1, \theta_2) \geq \beta_0 \\ H_1 &: g(\theta_1, \theta_2) < \beta_0 \end{aligned}$$

and we let $p_U(\beta_0)$ be the associated one-sided p-value, which is the solution to

$$U_\beta(\mathbf{x}, 1 - p_L(\beta_0)) = \beta_0.$$

Now we give the details for each of the three functions for g.

2 Difference

2.1 Lower Limit

When $x_2 > 0$ and $x_1 < n_1$ then we use numeric integration. For the difference, another way to define p_L is

$$p_L(\beta_0) \;=\; P_{A,B}\left[L_{\theta_2}(B) - U_{\theta_1}(A) \leq \beta_0\right] = P\left[T_{L2} \leq \beta_0 + T_{U1}\right] = \int_0^1 F_{L2}(t + \beta_0) f_{U1}(t)dt,$$

where F_{L2} is the cumulative distribution of T_{L2}, and f_{U1} is the density function of T_{U1}. Then to find $L_\beta(\mathbf{x}, 1 - \alpha)$ we use a root solving function and find the value of β_0 such that $p_L(\beta_0) = \alpha$.

When $x_2 = 0$ and $x_1 < n_1$ then

$$\begin{aligned} L_\beta(\mathbf{x}, 1 - \alpha) &= \alpha\text{th quantile of } \{0 - T_{U1}\} \\ &= -1 \text{ times the } (1 - \alpha)\text{th quantile of } T_{U1} \\ &= -F_{U1}^{-1}(1 - \alpha) \end{aligned}$$

where $F_{U1}^{-1}(1-\alpha) = qbeta(1-\alpha, x_1+1, n_1-x_1)$.

The p-value is the p that solves $L_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{aligned} -F_{U1}^{-1}(1-p_L(\beta_0)) &= \beta_0 \\ \Rightarrow \quad 1-p_L(\beta_0) &= F_{U1}(-\beta_0) \\ \Rightarrow \quad p_L(\beta_0) &= 1-F_{U1}(-\beta_0) \end{aligned}$$

When $x_2 > 0$ and $x_1 = n_1$ then

$$\begin{aligned} L_\beta(\mathbf{x}, 1-\alpha) &= \alpha\text{th quantile of } \{T_{L2}-1\} \\ &= F_{L2}^{-1}(\alpha) - 1 \end{aligned}$$

where $F_{L2}^{-1}(\alpha) = qbeta(\alpha, x_2, n_2-x_2+1)$.

The p-value is the p that solves $L_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{aligned} F_{L2}^{-1}(p_L(\beta_0)) - 1 &= \beta_0 \\ \Rightarrow \quad p_L(\beta_0) &= F_{U1}(1+\beta_0) \end{aligned}$$

When $x_2 = 0$ and $x_1 = n_1$ then $L_\beta(\mathbf{x}, 1-\alpha) = -1$ for all α. So $p_L(\beta_0) = 1$ for all β_0.

2.2 Upper Limit

When $x_2 < n_2$ and $x_1 > 0$ then we use numeric integration. For the difference, another way to define p_U is

$$\begin{aligned} p_U(\beta_0) &= P_{A,B}\left[U_{\theta_2}(B) - L_{\theta_1}(A) \geq \beta_0\right] = P_{A,B}\left[-T_{U2} + T_{L1} \leq -\beta_0\right] \\ &= P\left[T_{L1} \leq T_{U2} - \beta_0\right] = \int_0^1 F_{L1}(t-\beta_0) f_{U2}(t)\,dt, \end{aligned}$$

Then to find $U_\beta(\mathbf{x}, 1-\alpha)$ we use a root solving function and find the value of β_0 such that $p_U(\beta_0) = \alpha$.

When $x_2 = n_2$ and $x_1 > 0$ then

$$\begin{aligned} U_\beta(\mathbf{x}, 1-\alpha) &= (1-\alpha)\text{th quantile of } \{1-T_{L1}\} \\ &= 1 \text{ plus the } (1-\alpha)\text{th quantile of } -T_{L1} \\ &= 1 \text{ minus the } \alpha\text{th quantile of } T_{L1} \\ &= 1 - F_{L1}^{-1}(\alpha) \end{aligned}$$

where $F_{L1}^{-1}(\alpha) = qbeta(\alpha, x_1, n_1 - x_1 + 1)$.

The p-value is the p that solves $U_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{array}{rcl} & 1 - F_{L1}^{-1}(p_L(\beta_0)) & = \beta_0 \\ \Rightarrow & p_L(\beta_0) & = F_{L1}(1-\beta_0) \end{array}$$

When $x_2 < n_2$ and $x_1 = 0$ then

$$\begin{aligned} U_\beta(\mathbf{x}, 1-\alpha) &= 1-\alpha\text{th quantile of } \{T_{U2} - 0\} \\ &= F_{U2}^{-1}(1-\alpha) \end{aligned}$$

where $F_{U2}^{-1}(1-\alpha) = qbeta(1-\alpha, x_2+1, n_2 - x_2)$.

The p-value is the p that solves $U_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{array}{rcl} & F_{U2}^{-1}(1 - p_L(\beta_0)) & = \beta_0 \\ \Rightarrow & p_L(\beta_0) & = 1 - F_{U2}(\beta_0) \end{array}$$

When $x_2 = n_2$ and $x_1 = 0$ then $U_\beta(\mathbf{x}, 1-\alpha) = 1$ for all α. So $p_U(\beta_0) = 1$ for all β_0.

3 Ratio

3.1 Lower Limit

When $x_2 > 0$ and $x_1 < n_1$ then we use numeric integration:

$$p_L(\beta_0) = P_{A,B}\left[\frac{L_{\theta_2}(B)}{U_{\theta_1}(A)} \leq \beta_0\right] = P\left[T_{L2} \leq \beta_0 T_{U1}\right] = \int_0^1 F_{L2}(t\beta_0) f_{U1}(t) dt,$$

where F_{L2} is the cumulative distribution of T_{L2}, and f_{U1} is the density function of T_{U1}. Then to find $L_\beta(\mathbf{x}, 1-\alpha)$ we use a root solving function and find the value of β_0 such that $p_L(\beta_0) = \alpha$.

When $x_2 = 0$ and $x_1 < n_1$ then

$$L_\beta(\mathbf{x}, 1-\alpha) = \alpha\text{th quantile of } \left\{\frac{0}{T_{U1}}\right\} = 0 \text{ for all } \alpha$$

So $p_L = 1$.

When $x_2 > 0$ and $x_1 = n_1$ then

$$\begin{aligned} L_\beta(\mathbf{x}, 1-\alpha) &= \alpha\text{th quantile of } \{T_{L2}\} \\ &= F_{L2}^{-1}(\alpha) \end{aligned}$$

where $F_{L2}^{-1}(\alpha) = qbeta(\alpha, x_2, n_2 - x_2 + 1)$.

The p-value is the p that solves $L_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{aligned} & F_{L2}^{-1}(p_L(\beta_0)) &= \beta_0 \\ \Rightarrow \quad & p_L(\beta_0) &= F_{L2}(\beta_0) \end{aligned}$$

When $x_2 = 0$ and $x_1 = n_1$ then $L_\beta(\mathbf{x}, 1-\alpha) = 0$ for all α. So $p_L(\beta_0) = 1$ for all β_0.

3.2 Upper Limit

When $x_2 < n_2$ and $x_1 > 0$ then we use numeric integration:

$$\begin{aligned} p_U(\beta_0) &= P_{A,B}\left[\frac{U_{\theta_2}(B)}{L_{\theta_1}(A)} \geq \beta_0\right] = P_{A,B}\left[T_{U2} \geq T_{L1}\beta_0\right] \\ &= P\left[T_{L1} \leq \frac{T_{U2}}{\beta_0}\right] = \int_0^1 F_{L1}(\frac{t}{\beta_0}) f_{U2}(t)dt, \end{aligned}$$

Then to find $U_\beta(\mathbf{x}, 1-\alpha)$ we use a root solving function and find the value of β_0 such that $p_U(\beta_0) = \alpha$.

When $x_2 = n_2$ and $x_1 > 0$ then

$$\begin{aligned} U_\beta(\mathbf{x}, 1-\alpha) &= (1-\alpha)\text{th quantile of } \left\{\frac{1}{T_{L1}}\right\} \\ &= 1 \text{ over the } \alpha\text{th quantile of } T_{L1} \\ &= \frac{1}{F_{L1}^{-1}(\alpha)} \end{aligned}$$

where $F_{L1}^{-1}(\alpha) = qbeta(\alpha, x_1, n_1 - x_1 + 1)$.

The p-value is the p that solves $U_\beta(\mathbf{x}, 1-p) = \beta_0$, or

$$\begin{aligned} & \tfrac{1}{F_{L1}^{-1}(p_L(\beta_0))} &= \beta_0 \\ \Rightarrow \quad & p_L(\beta_0) &= F_{L1}\left(\frac{1}{\beta_0}\right) \end{aligned}$$

When $x_2 < n_2$ and $x_1 = 0$ then

$$U_\beta(\mathbf{x}, 1-\alpha) \quad = \quad 1-\alpha\text{th quantile of } \left\{\frac{T_{U2}}{0}\right\} = \infty \text{ for all } \alpha$$

So the $p_U = 1$.

When $x_2 = n_2$ and $x_1 = 0$ then $U_\beta(\mathbf{x}, 1-\alpha) = \infty$ for all α. So $p_U(\beta_0) = 1$ for all β_0.

4 Odds Ratio

4.1 Lower Limit

When $x_2 > 0$ and $x_1 < n_1$ then we use numeric integration:

$$\begin{aligned} p_L(\beta_0) &= P_{A,B}\left[\frac{L_{\theta_2}(B)(1-U_{\theta_1}(A))}{(1-L_{\theta_2}(B))U_{\theta_1}(A)} \leq \beta_0\right] \\ &= P\left[T_{L2} \leq \frac{\beta_0 T_{U1}}{1-T_{U1}+\beta_0 T_{U1}}\right] \\ &= \int_0^1 F_{L2}(\frac{\beta_0 t}{1-t+\beta_0 t}) f_{U1}(t)dt, \end{aligned}$$

where F_{L2} is the cumulative distribution of T_{L2}, and f_{U1} is the density function of T_{U1}. Then to find $L_\beta(\mathbf{x}, 1-\alpha)$ we use a root solving function and find the value of β_0 such that $p_L(\beta_0) = \alpha$.

When $x_2 = 0$ and $x_1 < n_1$ then

$$L_\beta(\mathbf{x}, 1-\alpha) \quad = \quad \alpha\text{th quantile of } \left\{\frac{0}{T_{U1}}\right\} = 0 \text{ for all } \alpha$$

So $p_L = 1$.

When $x_2 > 0$ and $x_1 = n_1$ then

$$L_\beta(\mathbf{x}, 1-\alpha) \quad = \quad \alpha\text{th quantile of } \left\{\frac{0}{(1-T_{L2})}\right\} = 0 \text{ for all } \alpha$$

So $p_L = 1$.

When $x_2 = 0$ and $x_1 = n_1$ then $L_\beta(\mathbf{x}, 1-\alpha) = 0$ for all α. So $p_L(\beta_0) = 1$ for all β_0.

4.2 Upper Limit

When $x_2 < n_2$ and $x_1 > 0$ then we use numeric integration:

$$\begin{aligned}
p_U(\beta_0) &= P_{A,B}\left[\frac{U_{\theta_2}(B)(1-L_{\theta_1}(A))}{L_{\theta_1}(A)(1-U_{\theta_2}(B))} \geq \beta_0\right] \\
&= P\left[T_{U2}(1-T_{L1}) \geq \beta_0 T_{L1}(1-T_{U2})\right] \\
&= P\left[T_{L1} \leq \frac{T_{U2}}{T_{U2}+\beta_0-\beta_0 T_{U2}}\right] \\
&= \int_0^1 F_{L1}(\frac{t}{t+\beta_0-\beta_0 t}) f_{U2}(t)dt,
\end{aligned}$$

Then to find $U_\beta(\mathbf{x}, 1-\alpha)$ we use a root solving function and find the value of β_0 such that $p_U(\beta_0) = \alpha$.

When ($x_2 = n_2$ and $x_1 > 0$) or ($x_2 < n_2$ and $x_1 = 0$) or ($x_2 = n_2$ and $x_1 = 0$) then

$$U_\beta(\mathbf{x}, 1-\alpha) = \infty \text{ for all } \alpha$$

So $p_U = 1$.