Lu Tian's exact meta-analysis method (Tian et al., 2009) as a special example under the general framework of combining CDs

Guang Yang, Dungang Liu, Minge Xie

Objective: We demonstrate that the meta-analysis method proposed in Tian et al. (2009) (hereafter LT's method) is a special case and can be covered by the general framework of combining confidence distributions (CDs) (cf., Singh et al., 2005).

1. Problem Setting (Tian et al., 2009):

- We are interested in constructing a $100(1-\alpha)\%$ one-sided CI $(a, +\infty)$ for Δ, a common parameter, based on all data from n independent studies.
- We assume that for each study we have confidence intervals for any arbitrary η-level.

  (The seond bullet suggests that we can obtain a CD for the parameter of interest from each study — let us denote it by $H_i(\Delta)$.)

2. LT's method:

- Let $J_{ij} = (a_{ij}, \infty)$ denote the η_j-level one-sided CI based on the ith study, for $i = 1, \ldots, n$ and $j = 1, \ldots, K$. Without loss of generality, we assume that $0 < \eta_1 < \ldots < \eta_K < 1$ and $a_{i1} \geq \ldots \geq a_{iK}$.
- The final combined interval for Δ is

$$\left\{ \Delta \middle| \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j \{\mathbf{I}(\Delta > a_{ij}) - \eta_j\} \geq d \right\} \tag{1}$$

  where $\mathbf{I}(\cdot)$ is the indicator function, w_i is a study-specific weight, $\widetilde{w}_j$ is a positive weight for η_j-level intervals, and d is a cut-off value chosen such that

$$\text{pr}\left\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j (B_{ij} - \eta_j) < d \right\} \leq \alpha. \tag{2}$$

  Here, $(B_{i1}, \ldots, B_{ik})', i = 1, \ldots, n$ are n independent random vectors, in which B_{ij}s are correlated Bernoulli variables with $B_{i1} \leq B_{i2} \leq \ldots \leq B_{iK}$ and $\text{pr}(B_{ij} = 1) = \eta_j$.
- Weights suggested in Tian et al. (2009): take the weight w_i to be the sample size for the ith study n_i; take $\widetilde{w}_j$ to be $\{\eta_j(1-\eta_j)\}^{-1}$.

3. An equivalent expression under the CD combination framework:

- Review of the general framework of CD combination (cf., Singh et al., 2005):
  - For any given function $g_c(u_1, \ldots, u_n)$ on $[0,1]^k \to \Re$, which is monotonic in each coordinate, we can have a combined CD for Δ by the recipe:

$$H^{(c)}(\Delta) = G_c\{g_c(H_1(\Delta), \ldots, H_n(\Delta))\}$$

Here, the function

$$G_c(t) = \text{pr}(g_c(U_1, \ldots, U_n) \le t) \tag{3}$$

is completely determined by the given g_c function, where $U_1, \ldots, U_n$ are independent U[0,1] random variable.

– A special choice of the function g_c is

$$g_c(u_1, \ldots, u_n) = w_1 a_0(u_1) + \ldots + w_n a_0(u_n)$$

where $a_0(\cdot)$ is any monotonically increasing function and w_i are the combining weights for each study.

- **<u>Connection to LT's method</u>**: Choose $a_0(u) = \sum_{j=1}^{K} \widetilde{w}_j(\mathbf{I}(u > 1 - \eta_j) - \eta_j)$, then we have the combined CD as

$$H^{(c)}(\Delta) = G_c \left\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j(\mathbf{I}(H_i(\Delta) > 1 - \eta_j) - \eta_j) \right\}. \tag{4}$$

<u>Our claim</u>: The inference based on the above CD is exactly the same as that obtained from LT's method.

- <u>Proof of the Claim</u>:

By its definition in (3), the function $G_c(\cdot)$ used in (4) is

$$\begin{aligned} G_c(t) &= \text{pr}\left\{ \sum_{i=1}^{n} w_i\, a_0(U_i) \le t \right\} \\ &= \text{pr}\left\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j(\mathbf{I}(U_i > 1 - \eta_j) - \eta_j) \le t \right\} \\ &= \text{pr}\left\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j(B_{ij} - \eta_j) \le t \right\} \end{aligned}$$

This is exactly the lef-hand-side of the inequality (2). Thus, the one-sided $100(1-\alpha)\%$CI for Δ by LT's method in (1) can be re-written as

$$\begin{aligned} &\left\{ \Delta \middle| \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j\{\mathbf{I}(\Delta > a_{ij}) - \eta_j\} \ge G_c^{-1}(\alpha) \right\} \\ = &\left\{ \Delta \middle| G_c\Big\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j\{\mathbf{I}(\Delta > a_{ij}) - \eta_j\} \Big\} \ge \alpha \right\} \\ = &\left\{ \Delta \middle| G_c\Big\{ \sum_{i=1}^{n} w_i \sum_{j=1}^{K} \widetilde{w}_j\{\mathbf{I}(H_i(\Delta) > 1 - \eta_j) - \eta_j\} \Big\} \ge \alpha \right\} \end{aligned}$$

$$= \left\{ \Delta \middle| H^{(c)}(\Delta) \geq \alpha \right\}.$$

Here, the second equation holds because $a_{ij} = H_i^{-1}(1-\eta_j)$ and the third equation holds because of (4). Note that $\{\Delta | H^{(c)}(\Delta) \geq \alpha\} = \{\Delta | \Delta \geq {H^{(c)}}^{-1}(\alpha)\}$ is exactly the one-sided $100(1-\alpha)\%$ CI for Δ from the combined CD. The claim is proved.

4. The case when $K \to \infty$ (The choice of K is a little bit ad hoc and we should perhaps let $K \to \infty$)

- Assume $0 < \eta_1 < \ldots < \eta_K < 1$ and they are equally spaced (i.e., $\eta_{j+1} - \eta_j \equiv \delta_K$ for all $j = 1, \ldots, K-1$). When the weight $\widetilde{w}_j = \{\eta_j(1-\eta_j)\}^{-1}$, we have $a_*(u) = \lim_{K\to\infty} \delta_K a_0(u) = \lim_{K\to\infty} \delta_K \sum_{j=1}^{K} \{\mathbf{I}(u > 1-\eta_j) - \eta_j\}/\{\eta_j(1-\eta_j)\} = \int_0^1 \{\mathbf{I}(u > 1-\eta) - \eta\}/\{\eta(1-\eta)\} d\eta = \ldots = \ln\left(\frac{u}{1-u}\right)$. Thus, an "equivalent" combined CD without relying on K is

$$H^{(c)}(\Delta) = G_c \left\{ \sum_{i=1}^{n} w_i \ln\left(\frac{H_i(\Delta)}{1 - H_i(\Delta)} \right) \right\},$$

where the choice of g_c is $g_c(u_1, \ldots, u_n) = w_1 a_*(u_1) + \ldots + w_n a_*(u_n) = w_1 \ln\left(\frac{u_1}{1-u_1}\right) + \ldots + w_n \ln\left(\frac{u_n}{1-u_n}\right)$ and $G_c(t) = \mathrm{pr}\left\{ w_1 \ln\left(\frac{U_1}{1-U_1}\right) + \ldots + w_n \ln\left(\frac{U_n}{1-U_n}\right) \leq t \right\}$.

5. Numerical Results

- We applied the LT method and the CD methods to the rosiglitazone data set to construct 95% CIs for the risk difference between rosiglitazone and the control arm with respect to mortality. The results are as follows:

	LT method	CD method-1	CD method-2
95% CI	$(-0.23, 0.13)\%$	$(-0.23, 0.13)\%$	$(-0.24, 0.14)\%$

  - Note: In the LT method and the CD method-1, we let $\{\eta_j\}$ be equally spaced levels from 0.1 to 0.95. The numerical result in LT's method is obtained by LT's on-line code. In the CD method-1, we take $a_0(u) = \sum_{j=1}^{K} \widetilde{w}_j (\mathbf{I}(u > 1-\eta_j) - \eta_j)$ in the CD combining recipe. The CD method-2 refers to the method above letting $K \to \infty$.

- Remark: The combined 95% CIs are similar to each other, and the small differences are due to the simulation method for generating the null distributions.

References

Singh, K., Xie, M., and Strawderman, W. E. (2005). Combining information from independent sources through confidence distributions. *The Annals of Statistics*, 33(1):159–183.

Tian, L., Cai, T., Pfeffer, M. A., Piankov, N., Cremieux, P.-Y., and Wei, L. (2009). Exact and efficient inference procedure for meta-analysis and its application to the analysis of independent 2× 2 tables with all available data but without artificial continuity correction. *Biostatistics*, 10(2):275–281.