

Errata: Design of Observational Studies

Page 78. In expressions (3.14), (3.15) and (3.18), replace each π_j by the corresponding odds $\omega_j = \pi_j / (1 - \pi_j)$. Thanks to Keith Goldfeld.

Typos

Page 50. Bottom of page, second sentence of the new section. This should read: “have to be part of the story.” Thanks to Jose Zubizarreta.

Page 52. On line 8, “the collection is \mathcal{K} contains” should read “the collection \mathcal{K} contains” Thanks to Sue Marcus.

Sensitivity analysis model when pairs are matched for observed covariates

The sensitivity analysis model (3.13) is quite general in its applicability [85, Chapter 4], but here its implications for matched pairs are developed [74]. Suppose that two subjects, k and ℓ , with the same observed covariates, $\mathbf{x}_k = \mathbf{x}_\ell$, are paired, with precisely the additional fact that one of them is treated and the other control, $Z_k + Z_\ell = 1$. Then in the representation (3.1), the chance that k is treated and ℓ is control is $\Pr(Z_k = 1, Z_\ell = 0 \mid r_{Tk}, r_{Ck}, \mathbf{x}_k, u_k, r_{T\ell}, r_{C\ell}, \mathbf{x}_\ell, u_\ell, Z_k + Z_\ell = 1)$

$$= \frac{\pi_k(1 - \pi_\ell)}{\pi_k(1 - \pi_\ell) + \pi_\ell(1 - \pi_k)}. \quad (3.14)$$

If in addition the sensitivity model (3.13) were true in (3.1), then simple algebra yields

$$\frac{1}{1 + \Gamma} \leq \frac{\pi_k(1 - \pi_\ell)}{\pi_k(1 - \pi_\ell) + \pi_\ell(1 - \pi_k)} \leq \frac{\Gamma}{1 + \Gamma}. \quad (3.15)$$

In words, the condition (3.13) becomes a new condition (3.15) on paired individuals where one is treated and the other control, $Z_k + Z_\ell = 1$. If $\Gamma = 1$, then all three terms in (3.15) equal $\frac{1}{2}$, as in the randomized experiment in Chapter 2. As $\Gamma \rightarrow \infty$, the lower bound in (3.13) tends to zero and the upper bound tends to one.

Instead of pairing just two individuals, k and ℓ , suppose we pair $2I$ distinct individuals of the L individuals in the population in just this way, insisting that within each pair the two subjects have the same observed covariates and different treatments. Renumber these paired subjects into I pairs of two subjects, $i = 1, 2, \dots, I$, $j = 1, 2$, so $\mathbf{x}_{i1} = \mathbf{x}_{i2}$, $Z_{i1} = 1 - Z_{i2}$ in each of the I pairs.¹⁴ If (3.1) and (3.13) are true, then the distribution of treatment assignments in the I pairs satisfies

$$Z_{i1}, i = 1, \dots, I \text{ are mutually independent,} \quad (3.16)$$

$$Z_{i2} = 1 - Z_{i1}, i = 1, \dots, I, \quad (3.17)$$

$$\frac{1}{1 + \Gamma} \leq \frac{\pi_k(1 - \pi_\ell)}{\pi_k(1 - \pi_\ell) + \pi_\ell(1 - \pi_k)} \leq \frac{\Gamma}{1 + \Gamma}, i = 1, \dots, I. \quad (3.18)$$

insist that $\pi_\ell = \Pr(Z_\ell = 1 \mid \mathbf{x}_\ell, u_\ell)$. Conversely, if (3.1) and (3.13) were true as they stand, then there is an unobserved covariate \tilde{u}_ℓ such that (3.1) and (3.13) are true with $\pi_\ell = \Pr(Z_\ell = 1 \mid \mathbf{x}_\ell, \tilde{u}_\ell)$; simply take $\tilde{u}_\ell = \pi_\ell = \Pr(Z_\ell = 1 \mid r_{T\ell}, r_{C\ell}, \mathbf{x}_\ell, u_\ell)$.

¹⁴ In a fussy technical sense, the numbering of pairs and people within pairs is supposed to convey nothing about these people, except that they were eligible to be paired, that is, they have the same observed covariates, different treatments, with $2I$ distinct people. Information about people is supposed to be recorded in variables that describe them, such as $Z, \mathbf{x}, u, r_T, r_C$, not in their position in the data set. You can't put your brother-in-law in the last pair just because of that remark he made last Thanksgiving; you have to code him in an explicit brother-in-law variable. Obviously, it is easy to make up subscripts that meet this fussy requirement: number the pairs at random, then number the people in a pair at random. The fussy technical point is that, in going from the L people in (3.1) to the $2I$ paired people, no information has been added and tucked away into the subject numbers — the criteria for pairs are precisely $\mathbf{x}_{i1} = \mathbf{x}_{i2}$, $Z_{i1} + Z_{i2} = 1$ with $2I$ distinct individuals.