My Blog 1 (Donglin Zeng)

I was asked to write an "opinion" blog which I have never done before. So to be cautious, I looked up word "opinion" in dictionary. Well, it says that an opinion is " a view or judgment formed about something, not necessarily based on fact or knowledge." The latter makes me quite relieve. It seems that I won't be held responsible if I say something irritating audience–you just attributed it to my ignorance.

I want to revisit an old issue on how to define a causal effect. Today, it is well known that a causal effect of a treatment is the expected difference between two counterfactual outcomes from all individuals, say $Y_i^{(1)}$ and $Y_i^{(-1)}$, where $Y_i^{(a)}$ is the outcome if subject i were to receive treatment ($a = 1$) or placebo ($a = -1$). It is also well accepted that a randomized trial is an ideal way to evaluate this effect.

Now here comes something often puzzling myself. First, since subject i will always receive one particular treatment, $Y_i^{(1)}$ and $Y_i^{(-1)}$ cannot be available at the same time. Thus, it makes no sense to talk about the difference of two variables who cannot co-exist in reality. Second, I always have suspicion about why the causal effect should be the difference. Why not ratio, $Y_i^{(1)}/Y_i^{(-1)}$, or the sign of the difference? The ratio has been quite popular in other estimation (odds ratio, risk ratio, hazard ratio) while the sign can be used if $Y_i^{(1)}$ or $Y_i^{(-1)}$ is more qualitative measurements? Clearly, even if we trust that a randomized trial can provide a valid estimator for the distribution of $Y^{(a)}$ by using data from treatment group a so we can compare the mean difference, there is no way that it can estimate the mean of the ratio or the sign comparing the two counterfactual outcomes. In other words, choosing the mean difference (or other differences based on the marginal distribution of $Y_i^{(a)}$) as the definition of the causal effect is more or less for reasons of estimability, or simply mathematical convenience.

We are stuck with these two problems if we stick to the original counterfactual outcomes. To avoid the non-coexistance of $Y_i^{(a)}$, it seems that we have to define the causal effects purely based on sensible variables. Let U_i denote ALL pre-treatment variables, either observed or not observed. For illustration, we can assume U to be discrete. Then a sensible causal effect due to treatment (we use A to denote it) may be defined as

$$\delta = E_i \left\{ E_j \left[g(Y_i, Y_j) | A_j = -1, U_j = U_i \right] \Big| A_i = 1 \right\},$$

where E_j is the expectation with respect Y_j given (A_j, U_j), and $g(x, y)$ is the causality quantification in the second point above. For example, $g(x, y) = x - y, x/y$ or $\text{sign}(x - y)$, depending on whether the causality quantification is the difference, ratio or sign respectively. In other words, for each subject i in the treatment group, we calculate the causal quantity related to this subject using all the matched subjects in the placebo group who have the same pre-treatment variable values; we then take the average of this quantity over the treatment group.

In reality, there is no way to observe ALL pre-treatment variables U. Instead, we only collect a subset of U, denoted by X. Using more crude X, we can only estimate the conditional distributions of Y's given A and X. Then the question is under what assumptions we may still estimate δ. Clearly, for any $U = u$ leading to $X = x$, if we assume

$$E[g(y, Y_j) | A_j = -1, U_j = u] \text{ is the same so equal to } E[g(y, Y_j) | A_j = -1, X_j = x],$$

then

$$\delta = E_i \left\{ E_j \left[g(Y_i, Y_j) | A_j = -1, X_j = X_i \right] \Big| A_i = 1 \right\}.$$

In other words, if U is independent of Y given X in the placebo group, then δ is estimable. As a remark, in a randomized trial setting, if $g(x, y) = x - y$, then we can easily show that δ is the usual causal effect.

Can we find any alternative assumptions to estimate δ?