The purpose of this piece is to explain the derivation of Eq. 19 to Eq. 21 on p. 6265 of Xie (2013).

Xie, Y. 2013. “Population Heterogeneity and Causal Inference.” *Proceedings of the National Academy of Sciences* 110(16):6262–68.

All equations with numbers are from the original article.

This holds because ITE can be conceptualized as LATE in the IV framework hence can be estimated by a Wald estimator in the form of Eq. 17. In order to understand this, imagine the dataset being expanded so that each data point is represented twice, at time and , respectively.

|  |  |  |  |
| --- | --- | --- | --- |
| Potential outcome | Potential outcome | Treatment status | T (IV) |
|  |  | 0 | 1 |
|  |  | 0 | 1 |
|  |  | 1 | 1 |
|  |  | 0 | 2 |
|  |  | 1 | 2 |
|  |  | 1 | 2 |

If the potential outcomes are time-constant, then ignorability of T ( & ) is trivially fulfilled. And T only affects the observed mean via treatment expansion. Hence the Wald estimator gives a valid estimate of LATE/ITE.

Taking limit for Eq. 17 and get Eq. 18,

Now I prove Eq. 19.

The last equality is by the substitution rule of integration.

As is an antiderivative of , by fundamental theorem of calculus, we get

Because of the way is defined (the cumulative proportion), . Thus,

, which is the Wald estimator for the LATE:

As is universally true, we get

, which is just the treatment effect on the treated (TT) at time t.

Eq. 20 and Eq. 21 can be proved very similarly. For Eq. 21, simply note that

When , all units are treated, and when, no unit is treated. Hence, , which is the ATE.