Short calculation

Laus Wolsing Wullum (CGS519)

September 15, 2023

Contents

1 Stratification

We know that not taking stratification into account when modelling yields incorrect SE's. In (Bugni et al., 2018) they develop a method to take stratification into account for the t-test. This is also mentioned in the FDA covariate adjustment guidance document from May 2023. A generalisation of the results in (Bugni et al., 2018) is found in (Wang et al., 2023).

2 Simple example

To gain an intuition for the effect of modelling the stratified randomization we posit the simplest model possible and write up an explicit formula for the correction term in the asymptotic variance presented in (Wang et al., 2023). To this end, assume a binary endpoint Y , binary treatment A and a binary stratification variable such that $X = S$.
The target estimand is the marginal treatment effect

$$\Delta = \mathbb{E}[Y \mid A = 1] - \mathbb{E}[Y \mid A = 0].$$

The influence function for the marginal ATE in the non-parametric setting is easily found by example 3.4.3 in (Kennedy, 2022),

$$\mathbb{IF}(\Delta) = \frac{A}{\pi}(Y - \mathbb{E}[Y \mid A = 1]) - \frac{1-A}{1-\pi}(Y - \mathbb{E}[Y \mid A = 0]).$$

The asymptotic variance of $\hat{\Delta}$ under stratified randomization is given by equation 1 in (Wang et al., 2023),

$$V = \tilde{V} - V_{\text{strata}} = \tilde{V} - \frac{1}{\pi(1-\pi)}\mathbb{E}\left[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X)]^2\right].$$

where $\tilde{V}$ is the standard asymptotic variance under simple randomization given by $\mathbb{E}(\mathbb{IF}(\Delta)^2)$ and the last term is denoted the correction term.
We can now compute the correction factor

$$\mathbb{E}\left[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X)]^2\right]$$

since we already know that

$$\tilde{V} = \mathbb{E}(\mathbb{IF}(\Delta)^2) = \frac{1}{\pi}\mu_1(1-\mu_1) + \frac{1}{1-\pi}\mu_0(1-\mu_0).$$

where $\mu_1 = \mathbb{E}[Y \mid A=1]$ and $\mu_0 = \mathbb{E}[Y \mid A=0]$.

2.1 Calculation of correction term

The tower property gives

$$\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X] = \mathbb{E}[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X, A] \mid X]$$

such that the innermost expectation becomes

$$\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X, A] = (A-\pi)\left(\frac{A}{\pi}(\mathbb{E}[Y \mid X, A] - \mu_1) - \frac{1-A}{1-\pi}(\mathbb{E}[Y \mid X, A] - \mu_0)\right).$$

Writing out the expectations on the left hand side of the above utilizing that A are binary yields

$$\mathbb{E}[Y \mid X, A] = A\mathbb{E}[Y \mid X, A=1] + (1-A)\mathbb{E}[Y \mid X, A=0] = A\mu_{X,1} + (1-A)\mu_{X,0}.$$

Where $\mu_{X,a} = \mathbb{E}[Y \mid X, A=a]$. Inserting and simplifying gives

$$\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X, A] = \frac{1-\pi}{\pi}A(\mu_{X,1}-\mu_1) + \frac{\pi}{1-\pi}(1-A)(\mu_{X,0}-\mu_0)$$

By taking the expectation given X using both the tower property, as well as our assumed balance within strata, i.e, $\mathbb{E}[A \mid X] = \pi$, we get:

$$\begin{aligned}\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X] &= \mathbb{E}[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X, A] \mid X] \\ &= (1-\pi)(\mu_{X,1}-\mu_1) + \pi(\mu_{X,0}-\mu_0)\end{aligned}$$

Defining $\pi_{xa} = P(Y=1 \mid A=a, X=x)$ we reach the final expression by squaring and computing the expectation

$$\mathbb{E}\left[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X)]^2\right] = \mathbb{E}[((1-\pi)(\mu_{X,1}-\mu_1) + \pi(\mu_{X,0}-\mu_0))^2]$$

That is,

$$\begin{aligned}V_{\text{strata}} &= \frac{1}{\pi(1-\pi)}\mathbb{E}\left[\mathbb{E}[(A-\pi)\mathbb{IF}(\Delta) \mid X)]^2\right] \\ &= \frac{1-\pi}{\pi}\mathbb{E}[(\mu_{X,1}-\mu_1)^2] + \frac{\pi}{1-\pi}\mathbb{E}[(\mu_{X,0}-\mu_0)^2] + 2\mathbb{E}[(\mu_{X,1}-\mu_1)(\mu_{X,0}-\mu_0)]\end{aligned}$$

Finally

$$\begin{aligned}V_{\text{strata}} = &\frac{1-\pi}{\pi}\left(p_X\pi_{11}^2 + (1-p_X)\pi_{01}^2 - \mu_1^2\right) \\ &+\frac{\pi}{1-\pi}\left(p_X\pi_{10}^2 + (1-p_X)\pi_{00}^2 - \mu_0^2\right) \\ &\quad +2(p_X\pi_{11}\pi_{10} + (1-p_X)\pi_{01}\pi_{00})\end{aligned}$$

where $p_X = P(X=1)$.

3 Adjusting for stratification indicators

When setting up an RCT we often need to prespecify an analysis plan and deciding baseline covariates we want to adjust for in our analysis. This is not a trivial task, but when the trial uses a stratified randomization procedure we can look towards corollary 1 in (Wang et al., 2023). This corollary shows that if we adjust for indicators for the randomization strata and treatment-by-randomization strata interaction terms, then the correction term in the variance becomes zero, i.e., $V = \tilde{V}$. This matches our intuition, since including information of the stratification in the model should express the same information as not including this information, but using the correction term in the variance.

We show that this is indeed the case in the setting where we target the marginal ATE, but still include information from covariates $X = S$. By using this estimator, we can gain extra efficiency in the estimation of the average treatment effect Δ since $A \perp\!\!\!\perp X$. The independence of treatment and baseline covariates is given by the simple randomization. From example 5.4 in (Tsiatis, 2006) we know the efficient influence function for this case:

$$
\begin{aligned}
\mathbb{EIF}(\Delta) &= \mathbb{IF}(\Delta) - \Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}) \\
&= \mathbb{IF}(\Delta) - \Big(\frac{A-\pi}{\pi}[\mu_{X,1} - \mu_1] + \frac{A-\pi}{1-\pi}[\mu_{X,0} - \mu_0]\Big) \\
&= \Big(\frac{A}{\pi}Y - \frac{A-\pi}{\pi}\mathbb{E}[Y \mid A=1, X]\Big) - \Big(\frac{1-A}{1-\pi}Y + \frac{A-\pi}{1-\pi}\mathbb{E}[Y \mid A=0, X]\Big) - \Delta.
\end{aligned}
$$

Where $\mathcal{J}$ is the tangent space of semiparametric models in this setup.
Corollary 1 in (Wang et al., 2023) tells us that the correction term should be zero. To see this in practice, let us compute the asymptotic variance of $\hat{\Delta}$ in this case while utilizing that

$$(\mathbb{IF}(\Delta) - \Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}) \perp \mathbb{IF}(\Delta)$$

Pythagoras now yields:

$$
\begin{aligned}
\tilde{V}_{\text{eff}} &= \mathbb{E}[\mathbb{EIF}(\Delta)^2] \\
&= \mathbb{E}[(\mathbb{IF}(\Delta) - \Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}))^2] \\
&= \left\|\mathbb{IF}(\Delta) - \Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp})\right\|^2 \\
&= \|\mathbb{IF}(\Delta)\|^2 - \left\|\Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp})\right\|^2 \\
&= \mathbb{E}[\mathbb{IF}(\Delta)^2] - \mathbb{E}[(\Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}))^2] \\
&= \tilde{V} - \mathbb{E}[(\Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}))^2]
\end{aligned}
$$

We want to show that $V_{\text{strata}_{\text{eff}}} = 0$ or equally

$$\mathbb{E}[(\Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}))^2] = V_{\text{strata}}.$$

This is easily seen, as

$$
\begin{aligned}
\mathbb{E}[(\Pi(\mathbb{IF}(\Delta) \mid \mathcal{J}^{\perp}))^2] &= \mathbb{E}\left[\Big(\frac{A-\pi}{\pi}[\mu_{X,1} - \mu_1] + \frac{A-\pi}{1-\pi}[\mu_{X,0} - \mu_0]\Big)^2\right] \\
= \frac{1-\pi}{\pi}\mathbb{E}[(\mu_{X,1} - \mu_1)^2] &+ \frac{\pi}{1-\pi}\mathbb{E}[(\mu_{X,0} - \mu_0)^2] + 2\mathbb{E}[(\mu_{X,1} - \mu_1)(\mu_{X,0} - \mu_0)] = V_{\text{strata}}
\end{aligned}
$$

Which is exactly what we wanted.

References

Bugni, F. A., Canay, I. A., & Shaikh, A. M. (2018). Inference under covariate-adaptive randomization. *Journal of the American Statistical Association*, *113*(524), 1784–1796.

Kennedy, E. H. (2022). Semiparametric doubly robust targeted double machine learning: A review. *arXiv preprint arXiv:2203.06469*.

Tsiatis, A. A. (2006). *Semiparametric theory and missing data*. Springer.

Wang, B., Susukida, R., Mojtabai, R., Amin-Esmaeili, M., & Rosenblum, M. (2023). Model-robust inference for clinical trials that improve precision by stratified randomization and covariate adjustment. *Journal of the American Statistical Association*, *118*(542), 1152–1163.