MGTECON 603 - Questions(Wk 2) (Instructor: Guido Imbens)

Wooyong Park

October 4, 2025

This is a document for the questions I had on the second week of the course.

1 Slides 3 Neyman

1.1 Variance Calculation in the Population

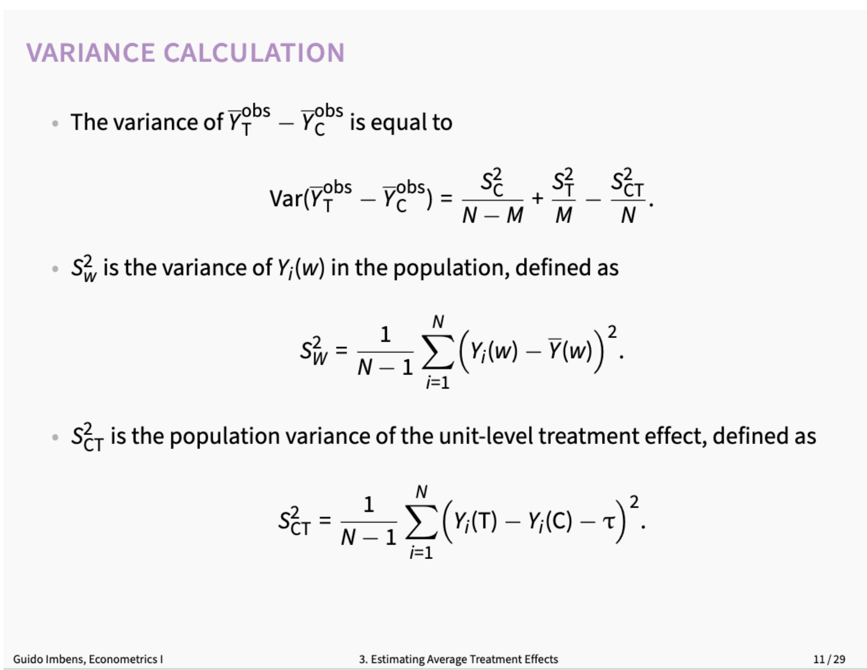

Figure 1.1: Variance Calculation in the Population

In the slide above, although the variance S_W^2 and S_{CT}^2 are defined in the population, the denominator is $N-1$ instead of N. This is contradictory to what I learned in my undergraduate statistics course, where the population variance is defined as $\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$.

Also, in this slide, I want to know how to derive the variance $V(\overline{Y}_T^{obs} - \overline{Y}_C^{obs}) = \frac{S_C^2}{N-M} + \frac{S_T^2}{M} - \frac{S_{CT}}{N}$; I have tried to derive it by myself, but I couldn't get the same result.

Sarah's answer: The DoF correction in the denominator makes S_W^2, S_{CT}^2 unbiased estimators of the 'super-population' variance.

1.2 Covariates and the third term of the variance

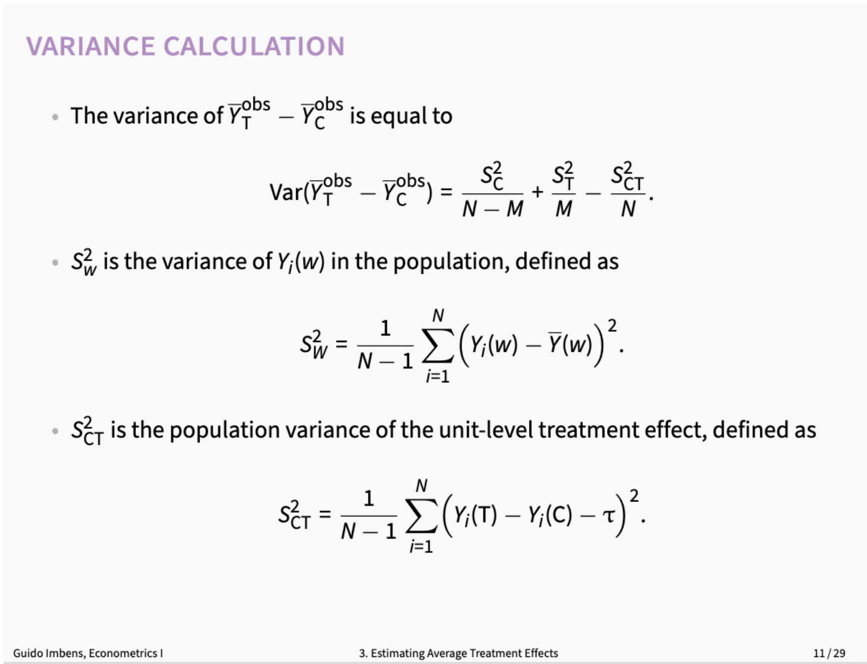

Figure 1.2: Covariates and the third term of the variance

In the slide above, the third term of the enitre variance of the difference-in-means, S_{CT}^2, is proportional to the heterogeneity of the treatment effects. In Monday's lecture, the professor said that having covariates can increase precision by reducing this term. However, I wonder if this is true for any kind of covariates. Giving this some thoughts, I think that if the covariates are colliders(i.e., $\{Y_i(0), Y_i(1)\} \perp W_i$ but $\{Y_i(0), Y_i(1)\} \perp W_i | X_i$), this might bias the results from the beginning. Thus, I wonder if there are conditions that the covariates should satisfy to reduce the third term without biasing the estimation.

2 Slides 4 OLS

2.1 Relationship between the linear model and the Potential Outcomes Framework

In a private question with the professor, I asked about the relationship between the linear model and the Potential Outcomes Framework. I wanted to know how the error term ϵ in the linear model is related to the potential outcomes $Y_i(0)$ and $Y_i(1)$. In most cases, I find that the unconfoundedness assumption, $\{Y_i(0), Y_i(1)\} \perp W_i | X_i$, corresponds to the assumption that the error term ϵ is uncorrelated with the treatment W_i when X_i is included in the model.

The professor answered that the error term can be interpreted as $\varepsilon_i = Y_i(0) - X_i'\beta$, where X_i' is the vector of covariates excluding the treatment indicator. In a simple linear regression without covariates, it would be $\varepsilon_i = Y_i(0) - \alpha$. In this case, regarding $Y_i(0)$ as constants or predetermined does not make sense, as ε_i is a random variable. Then, would the Neyman interpretation of the $Y_i(0)$ as constants or predetermined still hold?

Sarah's answer and a little bit of my interpretation: The super-population interpretation of randomized experiments, which follows from OLS, assumes that the stochasticity comes from re-sampling from a super-population, not from the permutation of the treatment. Thus, $Y_i(0)$ is a random variable sampled from a super-population, and ε_i inherits that randomness.