1 Randomized Trials

1.1 Causal Effect

Difference between the treatment group and the control group. First we can construct a dummy variable D_i to indicate whether the i-th component is in the treatment group or the control group:

$$D_i = \begin{cases} 1, & \text{if the } i\text{-th component is in the treatment group} \\ 0, & \text{if the } i\text{-th component is in the control group} \end{cases}$$

$$\text{Diffenence in group means} = Avg_n[Y_{1i} \mid D_i = 1] - Avg_n[Y_{0i} \mid D_i = 0]$$

But this equation is not what we are looking for. (Why? Not the comparasion between apples and apples.) We use to denote the effect of the treatment, then we have:

$$Y_{1i} = Y_{0i} + \kappa$$

$$\begin{aligned} \text{Diffenence in group means} &= Avg_n[Y_{1i} \mid D_i = 1] - Avg_n[Y_{0i} \mid D_i = 0] \\ &= \kappa + Avg_n[Y_{0i} \mid D_i = 1] - Avg_n[Y_{0i} \mid D_i = 0] \\ &= \kappa + \{Avg_n[Y_{0i} \mid D_i = 1] - Avg_n[Y_{0i} \mid D_i = 0]\} \end{aligned}$$

According to this, the difference between the two groups can be written as:

$$\boxed{\text{Diffenence in group means} = \text{Average causal effect} + \text{Section bias}}$$

1.2 Fundamental Knowledge of Inference

- **Unbiasedness of the sample mean:** $E[\bar{Y}] = E[Y_i]$
- **Variance of the sample mean:** $Var[\bar{Y}] = \frac{\sigma_Y^2}{n}$
- **the Estimated standard error:** $\hat{SE}(\bar{Y}) = \frac{S(Y_i)}{\sqrt{n}}$
- **the Estimated standard error of the difference in means:** $\hat{SE}(\bar{Y}_1 - \bar{Y}_0) = \sigma_Y \sqrt{\frac{1}{n_1} + \frac{1}{n_0}}$
- **the t-statistic for the difference in means:** $t(\mu) = \frac{\bar{Y}_1 - \bar{Y}_0 - \mu}{\hat{SE}(\bar{Y}_1 - \bar{Y}_0)}$