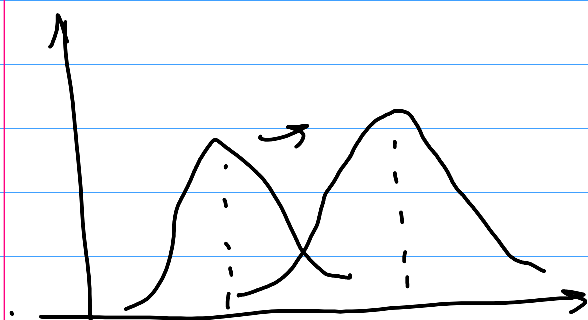
CUPED (Controlled experiments using pre-experiment data)

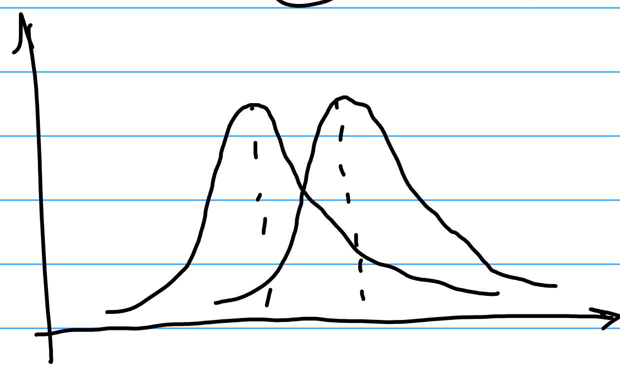
&

Diff-in-diff



✓

× ↓ CUPED



Naive approach : $\Delta = \bar{Y}_T - \bar{Y}_c$

Example : Δu after policy implementation ΔW.

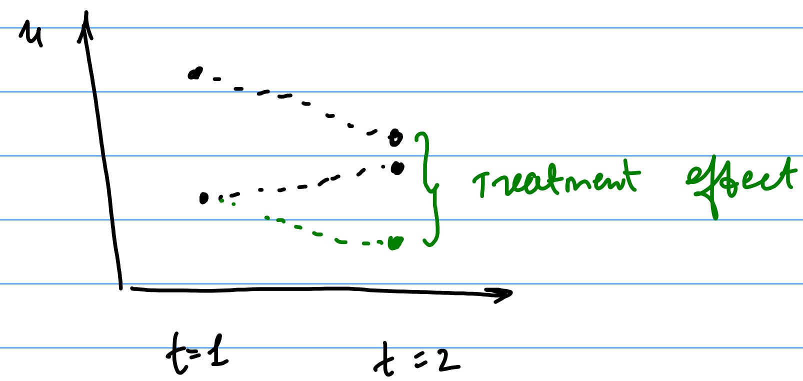
① $\Delta = \bar{u}_2 - \bar{u}_1$ other variates changing over time aren't controlled for

+

② $\Delta = \bar{u}_T - \bar{u}_c$ states can be different

"

Diff - in - diff



$$E(y \mid T, t=1) = \mu_1 + \alpha_T$$

$$E(y \mid T, t=2) = \mu_2 + \alpha_T + \delta$$

$\uparrow$ unknown TE

$$\Delta^T = \mu_2 - \mu_1 + \delta$$

$$E(y \mid T, t=1) = \mu_1 + \alpha_T$$

$$E(y \mid c, t=2) = \mu_2 + \alpha_c$$

$$\Delta^c = \mu_2 - \mu_1$$

$$\Delta^T - \Delta^c = \delta$$

"Diff-in-diff"

$$\Delta = \beta_2 + \delta \cdot D_T + \varepsilon_i$$

CUPED (Related to Frisch-Waugh)

y - explained data

X - covariate (pre-experiment)

$y \mid D, X$ $\hat{\theta}$

1) $y^c \mid X \Rightarrow \hat{y}^c$

2) $\hat{y}^2 \mid D$ $\hat{\theta}$

1) $y^{cuped} = y - \theta \cdot X$

$E[y] = E[\bar{y}] = E[\bar{y} - \theta X] + \theta E[X] =$

$= E[\bar{y} - \theta \bar{X}] + \theta E[X]$

$\hat{y} = \bar{y} - \theta \bar{X} + \theta E[X]$

$\theta = \dfrac{Cov(X,y)}{Var(X)}$ $\hat{\theta} = \dfrac{\hat{Cov}(X,y)}{\hat{Var}(X)}$

2) $\bar{y}^{cuped}_T - \bar{y}^{cuped}_c$

$Var(y^{cuped}) = (1 - \rho^2) Var(y) \leq Var(y)$

$y = \beta + \delta \cdot D_T + \theta \cdot X + \varepsilon_i$

variance exp.   var. ind.   var unexplained

Generalize: $y = \beta + \delta_1 D_{T1} + \delta_2 D_{T2} \dots + \theta_1 \cdot X_1 + \dots + \varepsilon_i$

Stratification:

$$\mathrm{Var}(\bar{y}) = \underbrace{\sum \frac{W_k}{n} \sigma_k^2}_{\text{within}} + \underbrace{\sum \frac{W_k}{n} (\mu_k - \mu)^2}_{\text{between}} \geq$$

$$\sum \frac{W_k}{n} \sigma_k^2 = \mathrm{Var}(\hat{y}_{strat})$$

= Cuped for X – categorical

$$X = \begin{cases} 1, & p_1 = w \\ 0, & p_2 = 1-w \end{cases}$$

$$\hat{y}_{STRAT} = w\bar{y}_1 + 1-w(\bar{y}_0)$$

$$\hat{y}_{cuped} = \bar{y} - \hat{\theta}\bar{X} + \hat{\theta} w$$

$$\hat{\theta} = \frac{\widehat{\mathrm{Cov}}(X,y)}{\widehat{\mathrm{Var}}(X)} = \bar{y}_1 - \bar{y}_0 \quad (\text{since } X\text{- dummy})$$

$$\hat{y}_{cuped} = \bar{y} - (\bar{y}_1 - \bar{y}_0)\bar{X} + (\bar{y}_1 - \bar{y}_0)w =$$

$$\bar{y} = \bar{X} \cdot \bar{y}_1 + (1-\bar{X})\bar{y}_0$$

$$= (1-\bar{X})y_0 + \bar{y}_0\bar{X} + (\bar{y}_1 - \bar{y}_0)w =$$

$$= w\bar{y}_1 + (1-w)\bar{y}_0 = \hat{y}_{STRAT}$$