Delta Method

Central Limit Theorem

* $X_1, \dots, X_n \sim_{iid} F$ with finite μ and $\sigma^2 > 0$
* $\dfrac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow[n \to \infty]{} \mathcal{N}(0,1)$

Common Application:

Confidence Interval for μ: $\bar{X}_n \pm z_{\alpha/2} \dfrac{\sigma}{\sqrt{n}}$

Drawback: Assumption of iid of $X_1, \dots, X_n$

Delta Method

For simplicity assume that $\sigma = 1$ and consider $T_n \triangleq T(X_1, \dots, X_n)$ - some statistic defined on a sample, θ - constant such that

$$\sqrt{n}(T_n - \theta) \to \mathcal{N}(0,1)$$

$$\Rightarrow$$

$$T_n - \theta = O(1/\sqrt{n})$$

Consider $\phi: \mathbb{R} \to \mathbb{R}$ - continious transformation

Using Taylor expansion we have

$$\phi(T_n) - \phi(\theta) = \phi'(\theta)(T_n - \theta) + O(T_n - \theta)^2$$

$$\Rightarrow$$

$$\sqrt{n}\{\phi(T_n) - \phi(\theta)\} \to \mathcal{N}(0, \phi'(\theta)^2)$$

Remark

Delta Method allows us to extend asymptotic normality to any continous transformation $\phi(T_n)$

Infering Percent Changes

* $X_1, \dots, X_n \sim_{iid} F_X$ (with mean μ_x and $\sigma_x^2 > 0$)
  from Control Group

  $Y_1, \dots, Y_n \sim_{iid} F_Y$ (with μ_y and $\sigma_y^2 > 0$)
  from Experimental Group

* We would like to build CI for $\boxed{\hat{\Delta}\% = \dfrac{\bar{Y} - \bar{X}}{\bar{X}}}$

* Application of DM:

$$T_n = (\bar{Y}, \bar{X}), \quad \theta = (\mu_y, \mu_x), \quad \phi(y, x) = y/x$$

using $\underline{\phi(T_n) - \phi(\theta) \approx \nabla\phi(\theta)(T_n - \theta)}$ we have

$$\frac{\bar{Y}}{\bar{X}} - \frac{\mu_y}{\mu_x} \approx \frac{1}{\mu_x}(\bar{Y} - \mu_y) - \frac{\mu_y}{\mu_x^2}(\bar{X} - \mu_x)$$

$$\forall\, i = \overline{1, n} \quad W_i = \frac{Y_i}{\mu_x} - \mu_y \frac{X_i}{\mu_x^2} \Rightarrow \frac{\bar{Y}}{\bar{X}} - \frac{\mu_y}{\mu_x} \approx \underbrace{\sum_{i=1}^{n} W_i \cdot \frac{1}{n}}_{\substack{\| \\ \bar{W}_n \text{ - average of iid random variables}}}$$

! $\bar{W}_n$ - average of iid random variables

CI for $\hat{\Delta}\%$

$$\underbrace{\frac{\bar{Y}}{\bar{X}} - 1}_{\text{point estimate}} \pm \frac{z_{\alpha/2}}{\sqrt{n}\,\bar{X}} \underbrace{\sqrt{S_y^2 - 2\frac{\bar{Y}}{\bar{X}} S_{xy} + \frac{\bar{Y}^2}{\bar{X}^2} S_x^2}}_{Var(\hat{\Delta}\%)}$$

Proof is pretty straight Forward ∎

Remark

By applying DM we can build CI for ratio-of-averages (of iid random variables)

Decoding Cluster Randomization

* Key concepts in AB test are
  * randomization unit (for example, user)
  * analysis unit (for example, view)

* Analytics is straightforward when randomization and analysis units agree. (When randomizing by user and computing average revenue per user)

* In practice randomization unit is a cluster of analysis units

* WLOG we will consider only treatment group with K clusters, $i = 1, \dots, K$

  N_i – number of analysis units in cluster i.

  Y_{ij} $(j = \overline{1, N_i})$ – analysis unit observations

  $$\overline{Y} = \sum_{ij} Y_{ij} / \sum_i N_i .$$

Lets rewrite this expression as

$$\overline{Y} = \frac{\sum_i \sum_j Y_{ij}}{\sum_i N_i} = \frac{\sum_i S_i / K}{\sum_i N_i / K} = \frac{\overline{S}_K}{\overline{N}_K} \sim \mathcal{N}(\cdot, \cdot)$$

ratio-of-averages (of iid random vars)

Remark

Thus we can apply DM to compute $Var(\overline{Y})$

$$Var(\overline{Y}) \approx \frac{1}{K \mu_N^2} \left(\sigma_S^2 - 2 \frac{\mu_S}{\mu_N} \sigma_{SN} + \frac{\mu_S^2}{\mu_N^2} \sigma_N^2 \right)$$

Hypothesis Testing

* Suppose we measured $\bar{Y}$ and $\bar{X}$ (revenue per view) for treatment and control groups.
* We would like to test H_0: there is no significant difference between $\bar{Y}$ and $\bar{X}$
  * $H_0: \mu_X = \mu_Y$

    $H_1: \mu_X \neq \mu_Y$
  * $T(X^n, Y^m) = \bar{Y} - \bar{X}$

    under the H_0 $\quad T(X^n, Y^m) \sim \mathcal{N}(0, \delta^2)$

    where $\delta^2 = Var(\bar{Y} - \bar{X}) = Var(\bar{Y}) + Var(\bar{X})$
  * We can test H_0 in two ways:
    1) Build CI for $\mu_Y - \mu_X$:

       $$\bar{Y} - \bar{X} \pm z_{\alpha/2}\left[Var(\bar{Y}) + Var(\bar{X})\right]$$

       and check $0 \in CI$.

       If $0 \in CI$ then we can not reject H_0
    2) Compute p-value of $T(x^n, y^m)$ and compare it against α.

       If $p \leq \alpha$ then we reject H_0 in favour of H_1

Efficient Variance Estimation for Quantile Metrics

TODO