The Bootstrap

confidence interval for regression coefficients

$(1-\alpha)\times 100\%$ $\quad CI = \hat{\beta} \pm t(n-k,\ 1-\frac{\alpha}{2}) \cdot SE(\hat{\beta})$

95% $\quad CI = \hat{\beta} \pm 2 \cdot SE(\hat{\beta})$

Assuming iid Normal errors ε_i

confidence interval for binomial proportion

Wald $CI = \hat{p} \pm 2\sqrt{\hat{p}(1-\hat{p})/n}$

Agresti $CI = \hat{p}' \pm 2\sqrt{\hat{p}'(1-\hat{p}')/(n+4)}$, $\quad \hat{p}' = \frac{S_n+2}{n+4}$

Assuming the binomial sum is approximately Normal.

Both of these CI assume Normality.
What if we don't want to assume normality? Or what if it's too complicated to analytically compute CI?

The Bootstrap is a technique to obtain standard errors or confidence intervals without assuming a distribution (i.e., it's a nonparametric technique to quantify uncertainty) by Resampling data.

bootstrap estimated binomial proportion:

input T = training dataset
for $b = 1, \dots, B$:

- $T^{(b)}$ = Re-sample m observations X_i with replacement
- $\hat{p}^{(b)}$ = estimated binomial proportion from $T^{(b)}$

Sort the $\hat{p}$'s so that $\hat{p}^{(1)} \le \dots \le \hat{p}^{(B)}$

Then $SE(\hat{p}) = sd(\hat{p}^{(1)}, \dots, \hat{p}^{(B)})$, 95% CI $= [\hat{p}^{(2.5^{th} \text{ quantile})}, \hat{p}^{(97.5^{th} \text{ quantile})}]$

bootstrap a regression parameter:

input T = training dataset
for $b = 1, \dots, B$:

- $T^{(b)}$ = Re-sample m observations (X_i, y_i) with replacement
- fit the Regression from $T^{(b)}$ and record the parameter estimate $\hat{\beta}^{(b)}$

Sort the $\hat{\beta}$'s so that $\hat{\beta}^{(1)} \le \dots \le \hat{\beta}^{(B)}$

Then $SE(\hat{\beta}) = sd(\hat{\beta}^{(1)}, \dots, \hat{\beta}^{(B)})$, 95% CI $= [\hat{\beta}^{(2.5^{th} \text{ quantile})}, \hat{\beta}^{(97.5^{th} \text{ quantile})}]$

Visualizing the bootstrap Resampling scheme:

$$
\begin{tikzcd}
 & T \arrow[dl] \arrow[d] \arrow[drr, "\text{resampling}"] & & \\
T^{(1)} \arrow[d] & T^{(2)} \arrow[d] & \cdots & T^{(b)} \arrow[d] \\
\hat{\beta}^{(1)} \arrow[dr] & \hat{\beta}^{(2)} \arrow[d] & \cdots & \hat{\beta}^{(b)} \arrow[dll] \\
 & SE(\hat{\beta}) \text{ and } \widehat{CI}(\beta) & &
\end{tikzcd}
$$

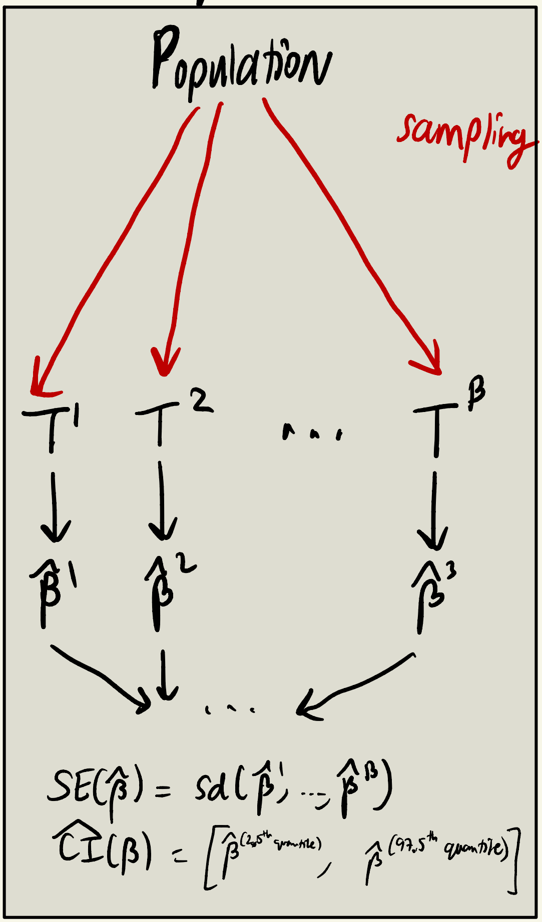
Why does this Resampling scheme work?

How we would ideally obtain Standard Errors and Confidence Intervals (which is impossible):

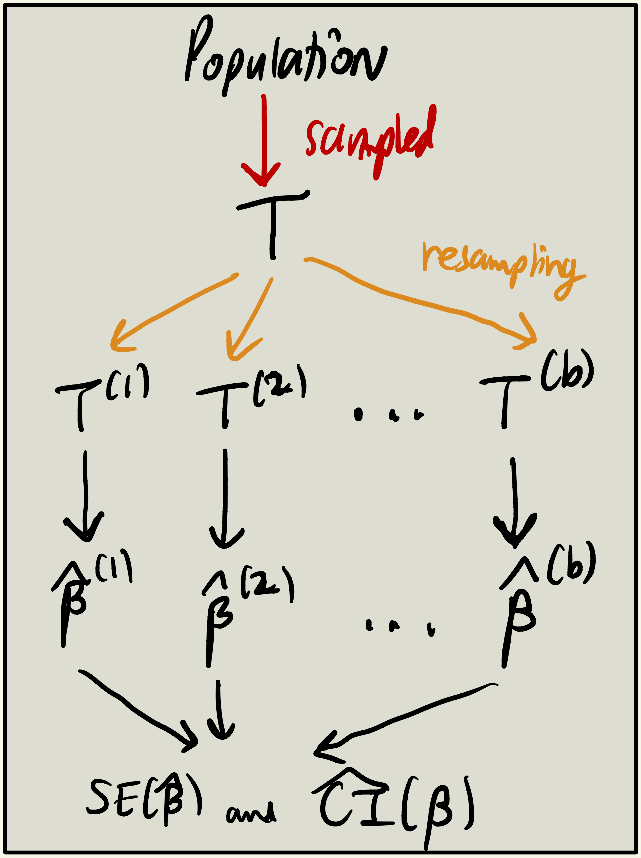
$$
\begin{tikzcd}
 & \text{Population} \arrow[dl] \arrow[d] \arrow[drr, "\text{sampling}"] & & \\
T^{1} \arrow[d] & T^{2} \arrow[d] & \cdots & T^{\beta} \arrow[d] \\
\hat{\beta}^{1} \arrow[dr] & \hat{\beta}^{2} \arrow[d] & & \hat{\beta}^{3} \arrow[dll] \\
 & \cdots & &
\end{tikzcd}
$$

$$SE(\hat{\beta}) = sd(\hat{\beta}^1, \dots, \hat{\beta}^B), \quad \widehat{CI}(\beta) = \left[\hat{\beta}^{(2.5^{th}\text{ quantile})}, \ \hat{\beta}^{(97.5^{th}\text{ quantile})}\right]$$

Ideal uncertainty quantification:



Bootstrapped uncertainty quantification:



Bootstrapping works because the training dataset T is itself a sample from the population; thus resampling from T mimics sampling from the population!

Drawbacks of the bootstrap

- doesn't always work
  e.g. works well for SE of means but not extrema (like max or min)

- computationally intensive

- can underestimate uncertainty: CI often not wide enough

  (see Brill "Analytics, Have Some Humility)

  ↳ can calibrate the bootstrap to acheive desired coverage

Class field trip to WaWa to buy M&Ms!