Recall... Bootstrap C.I.'s for $\theta = \gamma (F)$ based on

$$T_n = \frac{1}{n} \sum_{r=1}^{n} (X_r, X_{r+1}, \ldots, X_{r+n})$$

and bootstrapped $T_{n_1}^* \leq T_{n_2}^* \leq \cdots \leq T_{n_B}^*$

There are several versions of improvements on basic bootstrap percentile C.I.'s - one is "Bias corrected Percentile C.I.'s".

Get

$$T_{n_1}^* \leq \cdots \leq T_{n_B}^*$$

Use

$$\left[ T_{\alpha_1 (B+1)}^*, T_{\alpha_2 (B+1)}^* \right]$$

where we will agree to round $\alpha_1 (B+1)$ down and $\alpha_2 (B+1)$ up, where

$$\alpha_1 = \Phi \left( \frac{Z_0 + \widehat{\Delta} + Z_{upper}}{\sqrt{\widehat{\Delta}^2 + Z_{upper}^2}} \right) \quad \alpha_2 = 1 - \Phi \left( \frac{Z_0 + \widehat{\Delta} + Z_{upper}}{\sqrt{\widehat{\Delta}^2 + Z_{upper}^2}} \right)$$
For $\Phi$ the std normal cdf

$z_{\alpha/2} = \text{upper \ 95\% pt of std normal dsn (for confidence level 1-\alpha)}$

$\hat{\Delta}_n = \Phi^{-1}(\text{the fraction of } T_{n_1}^* \text{ that are smaller than } \hat{T}_n)$

$\hat{\Delta}_n$ is a measure of "median bias" of $\hat{T}_n$ in "normal units."

$T_{nj} = T_n$ computed dropping the $j$th observation from the data set.

$\bar{T}_{nj} = \frac{1}{n} \sum_{j=1}^{n} T_{nj}$

$\hat{\Delta} = \frac{\sum_{i=1}^{n} (\bar{T}_{nj} - \bar{T}_{nj})^3}{6 \left( \sum_{i=1}^{n} (\bar{T}_{nj} - \bar{T}_{nj})^2 \right)^{3/2}}$

This is called an "estimated acceleration factor." +

An alternative method called the ABC (approximate bias-corrected percentile bootstrap) method that approximates the previous one analytically - both are implemented in R.

Folklore says that the bias-correction improves coverage probability (in the sense of making it match the nominal one).
What about bootstrapping in more complicated (than the 1-sample problem) contexts? — e.g. how might the basic "resampling" idea be used in nonlinear regression? —

There are lots of possibilities here — one is to resample data vectors

\[(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\]

to make bootstrap samples — another approach is to "bootstrap the residuals." —

see Efron & Tibshirani / Koehler

---

**Generalized Linear Models**

generalize to other (besides normal) distributions for responses in a "LM" setting

basic problem is that we'd like to use

\[Z \sim \mathcal{B} \]

for modeling things like Poisson means for \(y_i\) or Binomial success probabilities for \(y_i\) — but this can be negative or outside \((0,1)\) —
don't try to model $\lambda$ (Poisson case) or $p$ (Binomial case) but instead functions of them taking values across all of $\mathbb{R}$

e.g. in the Poisson case I might model

$$\ln \lambda_i = x_i' \beta$$  (in Poisson case)

$$\ln \frac{p_i}{1 - p_i} = x_i' \beta$$  (in Binomial case)

Q: How far can I push this (in terms of types of
dens in response and in terms of functions of
parameters that I might choose to model as $x_i' \beta$)?

A: We can use exponential families and "reasonable"
functions - See Faraway's 2nd book -