Bootstrap Hypothesis Testing

To illustrate the technique, consider the case of two independent samples:

- Observed Sample 1 of size n: $\{x_{obs,1}, x_{obs,2}, ..., x_{obs,n}\} \Rightarrow \overline{x}_{obs}$
- Observed Sample 2 of size m: $\{y_{obs,1}, y_{obs,2}, ..., y_{obs,m}\} \Rightarrow \overline{y}_{obs}$
- Observed sample mean difference: $t^*_{obs} = \overline{x}^*_{obs} - \overline{y}^*_{obs}$

Hypotheses and Alpha Level

- H_0: Both samples are from the same population
- H_1: Both samples are NOT from the same population and $\mu_x > \mu_y$
- $\alpha = 0.05$

Bootstrap Procedure

- Step 0: Merge the two observed samples into one sample of $(n + m)$ observations
- Step 1: Draw a *bootstrap* sample of $(n + m)$ observations *with replacement* from the merged sample.
- Step 2: Calculate the mean of the first n observations and call it $\overline{x}^*$, calculate the mean of the remaining m observations and call it $\overline{y}^*$, and finally, evaluate the test statistic:

$$t^* = \overline{x}^* - \overline{y}^* \tag{1}$$

- Step 3: Repeat Step 1 and Step 2 for B (e.g., 3000) times and obtain B values of the test statistic.
- Step 4: The desired p-value is then estimated as

$$p - value \cong \frac{number\ of\ times\{t^* > t^*_{obs}\}}{B} \tag{2}$$

Reject H_0 if $p - value < \alpha$ and retain H_0 otherwise.

Points to Note

- No assumption of normality
- As such, the sampling distribution (i.e., distribution of all possible t^*'s) does *not* generally follow a t-distribution.