

BOOTSTRAP FOR SAMPLE MEAN

(1) Basic References:

- Horowitz, J. (1998), Journal of Human Resources.

(2) Bootstrap for sample mean

- $\{x_1, \dots, x_T\}$ is a random sample from a population with μ .
- Wish to estimate μ and test $H_0: \mu = 5$ Vs $H_a: \mu \neq 5$.

Asymptotic Procedure:

- Use $\bar{X} = T^{-1} \sum_{t=1}^T X_t$ as a point estimator of μ .
- Let $s_X^2 = T^{-1} \sum_{t=1}^T (X_t - \bar{X})^2$ which is a consistent estimator of σ^2 .
- Estimated $\text{var}(\bar{X}) = s_X^2 / T$.
- Under $H_0: \mu = 5$, $t = \left| \frac{\bar{X} - 5}{\sqrt{s_X^2 / T}} \right| \approx N(0,1)$.
- If you choose α (significance level) = 5%, $|z_{\alpha/2}| = 1.96$. Reject H_0 if $|t| > |z_{\alpha/2}|$.

Motivation of Bootstrap:

- The above method is relevant for large T . But, if T is small, \bar{X} may deviate a lot from μ and the test may generate wrong inferences.
- Wish to correct the possible bias in \bar{X} and find the better critical value.

Bootstrap Procedure:

- Treat $S \equiv \{x_1, \dots, x_T\}$ like a population with $f^*(x_t) = 1/T$.
- Draw a random sample of size T from S : say,
 - $S^{[1]} \equiv \{x_1^{[1]}, \dots, x_T^{[1]}\}$ (called a bootstrap sample)
 - How can we do this? (***) Explained later.
- Compute $\bar{X}^{[1]} = T^{-1} \sum_{t=1}^T X_t^{[1]}$ and $s_X^{2[1]} = T^{-1} \sum_{t=1}^T (X_t^{[1]} - \bar{X}^{[1]})^2$, and

$$|t^{[1]}| = \left| \frac{\bar{X}^{[1]} - \bar{X}}{\sqrt{s_X^{2[1]} / T}} \right|.$$

- Draw another random sample of size T from S : say
 - $S^{[2]} \equiv \{x_1^{[2]}, \dots, x_T^{[2]}\}$ (called a bootstrap sample).
 - Compute $\bar{X}^{[2]} = T^{-1} \sum_{t=1}^T X_t^{[2]}$ and $s_X^{2[2]} = T^{-1} \sum_{t=1}^T (X_t^{[2]} - \bar{X}^{[2]})^2$, and

$$|t^{[2]}| = \left| \frac{\bar{X}^{[2]} - \bar{X}}{\sqrt{s_X^{2[2]} / T}} \right|.$$

- Repeat this procedure b times:
 - Greater b is always better.
 - My guess is that $b = 1000$ would be enough.
- Define $\hat{B} = b^{-1} \sum_{i=1}^b \bar{X}^{[i]} - \bar{X}$ (Bootstrap estimate of bias in \bar{X}).
- $\bar{X}_B = \bar{X} - \hat{B}$ (Bias-corrected bootstrap point estimate of μ).
- Order $|t^{[i]}|$ in the ascending order (let $\{|t^{(1)}|, \dots, |t^{(b)}|\}$ be this ordered set), and choose $|t_{B,\alpha}| \equiv$ the $(1-\alpha)$ quantile of $\{|t^{(1)}|, \dots, |t^{(b)}|\}$.
- If $|t| > |t_{B,\alpha}|$, reject H_0 .

Digression --- Answer to (***)

- Let $X = [x_1, \dots, x_T]'$.
- Let $[[x_t]] =$ the smallest integer which is bigger than x_t .
- Generate a $T \times 1$ vector of uniform numbers, $U = [u_1, \dots, u_T]'$, where $0 < u_t < 1$ (in GAUSS, `rndu(tt,1)`).
- Let $UU = [uu_1, \dots, uu_T]' = [[[T \times u_1]], \dots, [[T \times u_T]]]'$.

Note that $[[T \times u_t]]$ is an integer between 1 and T .

Some numbers in UU can be repeated.

- In GAUSS, $uu = \text{ceil}(tt * \text{rndu}(tt,1))$.
- Choose $x_1^{[1]} = uu_1'$ th element of the original sample S .

$x_2^{[1]} = uu_2'$ th element of S , etc.

- In GAUSS, `x[uu,.]`.

- EX: $x = \begin{pmatrix} 1.5 \\ 2.5 \\ 3.5 \\ 4.5 \end{pmatrix}; uu = \begin{pmatrix} 1 \\ 3 \\ 2 \\ 2 \end{pmatrix} \rightarrow x[uu,.] = \begin{pmatrix} 1.5 \\ 3.5 \\ 2.5 \\ 2.5 \end{pmatrix}.$

End of Digression

Question:

- Why bootstraps generate better statistical inferences?
 - For technical details, see Hall and Horowitz (1996, Econometrica).
 - The bootstrap test is asymptotically equivalent to the usual asymptotic test. It can be shown that

$$E^*(X_t^{[1]}) = \bar{X} \text{ and } \text{var}^*(X_t^{[1]}) = s_X^2.$$

$$E^*(\bar{X}^{[i]}) = \bar{X} \text{ and } E^*(s_X^{2[i]}) \approx s_X^2.$$

Accordingly, by CLT, $\{t^{[1]}, \dots, t^{[b]}\}$ will be roughly distributed with $N(0,1)$ if T is large. Thus, the $(1-\alpha)$ quantile point of $\{|t^{[1]}|, \dots, |t^{[b]}|\}$ would be close to $|z_{\alpha/2}|$.

- As T increases, the gain by bootstrapping decreases.

(3) In the program **bmeanmon.prg** and **bmean.prg**:

- Uses t^2 which is asymptotically $\chi^2(1)$.
- Accordingly, use $\{t^{[1]2}, \dots, t^{[b]2}\}$ to compute bootstrap critical values.