Sept 26, 2020
Marta Karas

Assume $X_1, \dots, X_{N_2}$ are iid $\sim N(\mu, \sigma^2)$.
Then random variable

$$\bar{X}_{N_2} = \left(\frac{1}{N_2} \sum_{i=1}^{N_2} X_i \right) \sim N\left(\mu, \frac{\sigma^2}{N_2}\right). \tag{1}$$

0.1 Scenario 1: I have sample size N_2 available

Then I can

- Estimate μ with

$$\hat{\mu} = \frac{1}{N_2} \sum_{i=1}^{N_2} x_i \tag{2}$$

- Estimate σ^2 with

$$\hat{\sigma}^2 = \frac{1}{N_2 - 1} \sum_{i=1}^{N_2} (x_i - \hat{\mu})^2 \tag{3}$$

  where:

  - $E(\hat{\sigma}^2) = \sigma^2$
  - $var(\hat{\sigma}^2) = \frac{2\sigma^4}{(N_2-1)}$

- Estimate $\sigma^2_{N_2} := \frac{\sigma^2}{N_2}$, variance of sample mean $\bar{X}_{N_2}$, with

$$\hat{\sigma}^2_{N_2} = \frac{\hat{\sigma}^2}{N_2} \tag{4}$$

  where:

  - $E(\hat{\sigma}^2_{N_2}) = E\left(\frac{\hat{\sigma}^2}{N_2}\right) = \frac{E(\hat{\sigma}^2)}{N_2} = \frac{\sigma^2}{N_2}$
  - $var(\hat{\sigma}^2_{N_2}) = var\left(\frac{\hat{\sigma}^2}{N_2}\right) = \left(\frac{1}{N_2}\right)^2 \cdot var(\hat{\sigma}^2) = \frac{2\sigma^4}{(N_2-1)N_2^2}$

0.2 Scenario 2: I have sample size N_1 available only, but want to determine the above quantities in case I have sample size N_2 ($N_2 > N_1$) available

I proceed as follows:

- I use upstrap: I generate $B = 10,000$ resamples of sample of size N_2, denote them $\mathbf{x}_1, \dots, \mathbf{x}_B$.
- For each resample, I generate its sample mean: $\bar{x}_1, \dots, \bar{x}_B$.
- I estimate $\sigma^2_{N_2}$, variance of sample mean $\bar{X}_{N_2}$, with upstrap estimator defined as sample variance of $\bar{x}_b$'s

$$\hat{\sigma}^{2;\mathbf{UP}}_{N_2} = \frac{1}{B-1} \sum_{b=1}^{B} \left[\bar{x}_b - \left(\frac{1}{B} \sum_{i=1}^{B} \bar{x}_i \right) \right]^2 \tag{5}$$

0.3 Desired properties of upstrap estimator

I would like to observe that

- The expected value of $\hat{\sigma}^{2;\mathbf{UP}}_{N_2}$ is the same as expected value of the estimator I use when I have sample size N_2 available, that is

$$E\left(\hat{\sigma}^{2;\mathbf{UP}}_{N_2}\right) = E\left(\hat{\sigma}^2_{N_2}\right) = \frac{\sigma^2}{N_2} \tag{6}$$

- The variance of $\hat{\sigma}^{2;\mathbf{UP}}_{N_2}$ is the same as variance of the estimator I use when I have sample size N_2 available, that is

$$var\left(\hat{\sigma}^{2;\mathbf{UP}}_{N_2}\right) = var\left(\hat{\sigma}^2_{N_2}\right) = \frac{\sigma^2}{N_2} = \frac{2\sigma^4}{\left(N_2 - 1\right) N_2^2} \tag{7}$$

0.4 Desired properties of upstrap estimator - checking

We check if the desired properties of upstrap estimator hold by:

- Repeating some big number of times, $R = 100,000$, the scenario 2 situation (using upstrap estimator of $\hat{\sigma}^{2;\mathbf{UP}}_{N_2}$ estimator)
- Compare the obtained distribution of $R = 100,000$ "draws" of $\hat{\sigma}^{2;\mathbf{UP}}_{N_2}$ with its desired distributional results we mention above.