CIs in the Mediation R Package

Tim Vigers

1/12/2021

Notation

- Decompose a statistical model into two parts:

$$Y_i \sim f(\theta_i, \alpha)$$

and

$$\theta_i = g(X_i, \beta)$$

.

- For a normal linear regression $Y_i \sim N(\mu_i, \sigma^2)$, $\mu_i = X_i\beta$, and $\alpha = Var(Y_i) = \sigma^2$.

Quasi-Bayesian Approximation (King et al., 2000)

1. Get point estimates and covariance from the appropriate model. "Stack" the point estimates in a vector $\hat{\gamma} = vec(\hat{\beta}, \hat{\alpha})$ and call the covariance matrix $\hat{V}(\hat{\gamma})$.

2. The CLT (assuming a large enough sample size and bounded variance) allows us to draw from a multivariate normal distribution assuming:

$$\hat{\gamma} \sim N(\hat{\gamma}, \hat{V}(\hat{\gamma}))$$

Quasi-Bayesian Approximation (King et al., 2000)

3. Draw one instance of the vector from the multivariate normal distribution:

$$\tilde{\gamma} = \mathit{vec}(\tilde{\beta}, \tilde{\alpha})$$

Repeat step 3 $M = 1000$ times (by default, but can be increased).

4. Calculate CIs by sorting the draws from lowest to highest and taking the 25th and 976th values (if $M = 1000$).

The Bootstrap Approach

1. Assuming n data points $x_1, ..., x_n$, draw a re-sample of size n with replacement.
2. Compute the statistic of interest.
3. Repeat steps 1 and 2 $M = 1000$ times (again, this can be increased).

Bootstrap CIs

- A common approach to calculating bootstrap CIs is using the percentile values, like in the King et al. approach.
- However, these CIs can be too narrow, particularly for a small n.
- One relatively simple alternative is to adjust which percentiles are used based on the estimation of acceleration and bias correction factors (BCa intervals).

BCa Intervals

- Instead of using α^{th} and $(1-\alpha)^{th}$ percentiles, BCa uses an interval based on different percentiles (α_1 and α_2).
- These are calculated using the formula:

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z(\alpha)}{1 - a(\hat{z}_0 + z(\alpha))}\right)$$

$$\alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z(1-\alpha)}{1 - a(\hat{z}_0 + z(1-\alpha))}\right)$$

Where $\hat{z}_0$ is the bias correction factor and a is the acceleration parameter. So, if $\hat{z}_0 = a = 0$, this is the same as the usual percentile approach.

BCa Intervals

- $\hat{z}_0$ is the difference between the median of the bootstrap statistics and $\hat{\theta}$.
- The acceleration factor a is calculated using a jackknife approach. Let $\hat{\theta}_{(i)}$ be the estimate of θ after taking out x_i and $\hat{\theta}_{(\cdot)} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{(i)}$ be the average of the jackknife estimates.

BCa Intervals

- Then a is calculated as:

$$a = \frac{\sum_{i=1}^{n}(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^3}{6(\sum_{i=1}^{n}(\hat{\theta}_{(\cdot)} - \hat{\theta}_{(i)})^2)^{3/2}}$$

- This "estimates the rate of change of the standard error of $\hat{\theta}$ with respect to the true parameter θ."[4]
- BCa intervals are range preserving and transformation-invariant, but not particularly intuitive.

References

1. King G, Tomz M, Wittenberg J. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. Am J Polit Sci. 2000;44(2):347. doi:10.2307/2669316
2. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for Causal Mediation Analysis. J Stat Softw. 2014;59(1):1-38. doi:10.18637/jss.v059.i05
3. DiCiccio TJ, Efron B, Hall P, et al. Bootstrap confidence intervals. Stat Sci. 1996;11(3):189-228. doi:10.1214/ss/1032280214
4. Helwig, Nathaniel E. Bootstrap Confidence Intervals. Jan. 2017. http://users.stat.umn.edu/~helwig/notes/bootci-Notes.pdf