607
Bootstrap and Asymptotic Refinement

Jon Faust

http://e105.org/e607

October 13, 2016

- **So far**
  - We have described the bootstrap and explained that it seems like a reasonable idea.
  - We've shown it seems to work at least passably well in some contexts.
  - And we've gestured at some asymptotic expansion arguments that may provide theoretical underpinnings why the bootstrap would be better than conventional asymptotics.

- **This lecture**
  - In this lecture, I go a bit further.
  - I stop gesturing and move to full blown hand waving

    That is, I give a hand-wavy proof-ish

- **Readings**
  - Hansen's text probably would be the next more thorough (but still hand-wavy) discussion.
  - Next, I'd probably go to Peter Hall's book, 'The bootstrap and Edgeworth expansions'

- **Our standard setup**
  - Suppose $Y \sim P_\theta, \theta \in \Theta$
  - We have some statistic, ϕ, with distribution function $G_{\theta,T} = F^{\phi}_{\theta,T}$ when the data are generated by θ.

    We are naming it G just for cleaner notation.

  - True θ is θ^*.

▶ Edgeworth expansion of G

- Suppose that ϕ is asymptotically pivotal and has a valid 2-term Edgeworth expansion:

$$G_{\theta,T}(c) = \frac{g_0(c)}{T^0} + \frac{g_1(c,\theta)}{T^{1/2}} + o(T^{1/2})$$

  there is no θ in g_0 b/c the stat. is asymptotically pivotal.

- We want to konw $G_{\theta^*,T}(c)$
- Conventional (that is, first order) asymptotics uses only the first term in the above expansion and has an error equal to:

$$\frac{g_1(c,\theta^*)}{T^{1/2}} + o(T^{1/2})$$

  which is $o(T^0)$, due to that first term.
- In the bootstrap, we approximate $G_{\theta^*,T}(c)$ by $G_{\theta_b,T}$ which has expansion
- These two just differ by what parameter we are using, θ^* vs. θ_b.
- Lets write the two two-term expansions:

$$\begin{aligned} G_{\theta^*,T}(c) &= \frac{g_0(c)}{T^0} + \frac{g_1(c,\theta^*)}{T^{1/2}} + o(T^{-1/2}) \\ G_{\theta_b,T}(c) &= \frac{g_0(c)}{T^0} + \frac{g_1(c,\theta_b)}{T^{1/2}} + o(T^{-1/2}) \end{aligned}$$

- Call the first, the ideal 2-term expansion (since it uses the true θ^*).
- The ideal expasion has an error that is $o(T^{1/2})$

  Whereas conventional asymptotics has a larger error in the sense that it is $o(T^0)$.

- Subract the two expressions:

$$G_{\theta^*,T}(c) - G_{\theta_b,T}(c) = \frac{g_1(c,\theta^*) - g_1(c,\theta_b)}{T^{1/2}} + o(T^{-1/2})$$

- The g_0 term drops out because our statistic is asymptotically pivotal so θ doesn't affect this term.

  This is why we always emphasize that a necessary condition for the bootstrap providing an asymptotic refinement is that the statistic be asymptotically pivotal.

- The first term on the RHS is essentially the accuracy penalty you pay for using θ_b instead of the ideal θ^*.
- If that penalty term goes to zero even after being multiplies by $T^{1/2}$, then it is $o(T^{-1/2})$ and we can say that

$$G_{\theta^*,T}(c) - G_{\theta_b,T}(c) = o(T^{-1/2})$$

That is, the bootstrap approximation is as accurate as the ideal 2-term expansion—where 'as accurate' means it has an error that vanishes at the same rate with T.

- Of course, that penalty term is $o(T^{-1/2})$ so long as

$$g_1(c, \theta^*) - g_1(c, \theta_b) \to 0$$

with T in some relevant sense.

- Here comes the hand waving!
- Remember that we are choosing θ_b as a function of of the sample, and, hence, as a function of T. Thus we need:

$$g_1(c, \theta^*) - g_1(c, \theta_{b,T}) \to 0$$

- It should make sense to you that this condition will hold if $\theta_{b,T} \to \theta^*$ in some relevant sense (this condition is a bit like consistency, and would be consistency if that arrow were $\to_p$. We are not sure however, just what arrow that needs to be at this point. Therein lies the tricky bit.

▶ **The keys to refinement**

- 1. The statistic is asymptotically pivotal so that the first term in the expansion is the same under θ^* and θ_b.
- 2. $\theta_{b,T}$ is chosen cleverly so that it converges to θ^* so that the numerator of the second term under θ^* and θ_b is converging to zero with T.
- That second condition is essentially that θ_b is close enough to θ^* that you are indifferent between using the two in $g_1(c, \theta)$.

▶ **The hand-waving**

- Several steps of this require a good bit of work to finish.
- First, we need the convergence of $g_1(c, \theta_b)$ to $g_1(c, \theta^*)$ to be uniform in c.

We had not mentioned that recently so its worth emphasizing here.

- Second, our desire is to get results in the context of non-parametric, or semi-parametric problems. With θ infinite dimensional, we have to be careful saying things like θ_b converges to θ^*.
- Finally, in sketch just given it wasn't clear that in practice we choose θ_b as a function of a sample, so that it it stochastic. A real proof would have to take account of that.
- All these issues can be (have been) dealt with, delivering the desired result.

▶ **Bottom line, a pretty general statement**

- If your statistic is distributed $G_{\theta^*,T}$ and,
  - Has a valid 2-term Edgeworth expansion
  - And is asymptotically pivotal
  - And we can choose $\theta_{b,T}(Y)$ converging to θ^* appropriately
- Then the error in the bootstrap approximation converge to zero at the same rate as if we used the 'ideal' two-term Edgeworth expansion.

  That rate is: $o(T^{-1/2})$.

- Thus, the bootstrap provides a better approximation than conventional asymptotics in the sense of capturing one more term in the asymptotic expansion.
- Based on this sort of proof, you will often hear that 'the bootstrap provides an asymptotic refinement'

▶ **Comments**

- In any given sample size, the 2-term ideal approximation is not necessarily better than the 1-term expansion.

  So refinement may hurt you. But bootstrapping seems to be a very good idea in a wide range of case investigated by Monte Carlo.

- The proof shows that the bootstrap delivers the same quality (in the sense of rate of convergence) as the ideal 2-term expansion.
- In analytic examples, we can specify θ^* and compute how well the bootstrap does relative to using the ideal 2-term expansion.
- In a wide range of cases, the bootstrap does better than the ideal 2-term expansion.

  Thus, there is almost certainly (in my view) a stronger case to be made for the bootstrap.

- Peter Hall's book, 'The bootstrap and Edgeworth expansions' has a good discussion of this.
- Figure out that better case, and you will be very famous.