Thoughts on teaching an SBI course for a large class

Matthew Beckman

8/24/2017

What this is & what this isn't

This post is intended share some pragmatic thoughts for *teaching* SBI in a large class, and not necessarily converting your curriculum to the SBI framework. A number of suggestions on the latter have been published in this blog and elsewhere. Besides, my colleague–Kari Lock Morgan–had already done a remarkable job accomplishing that feat in the course to be described before I arrived. What follows are simply remarks about rubber-meets-the-road strategies from teaching an SBI course with 225 students to either capitalize on large class size or at least help navigate some logistical challenges that surface with increased enrollment.

A familiar activity scaled for a large class

While teaching smaller classes, I adopted a popular illustration with M&M's (sometimes Skittles) to introduce bootstrapping. I didn't invent or perfect the activity, but here's a summary: Each student gets a fun size bag of M&M's, calculates the proportion of blue, and then marks the result on a class dotplot in front of the room. We can now have a conversation about a sampling distribution under the assumption that each fun size bag represents a random sample of M&M's. We emphasize the point that each mark on the dotplot represents a statistic calculated from an actual sample in the room, and perhaps point out a few (say, extremes) and identify the responsible students to emphasize the point.

Students then sample with replacement from their own bag to build a bootstrap distribution. Students do this a few times by hand, and then introduce software to speed things up. We emphasize that each dot in the bootstrap distribution is now a proportion of blue M&M's out of 16 draws with replacement from their own fun size bag (assuming 16 is the sample size contained). We highlight similarities and differences between the sampling distribution and bootstrap distributions. For example, bootstrap distributions will generally be centered in different places, but still wind up with a useful estimate of standard error.

I tend to get a lot of mileage out of this activity as the semester progresses. When I sense that students are losing sight of the fundamentals, we periodically back up and talk about the M&M's again to get back on solid footing.

The thought of scaling this for 225 students was daunting at first, but since I like having this example in my back pocket later in the semester it was worth a shot. To start, I basically buy the entire stock of fun-size M&M's available at a super-store in town. We'll have plenty of M&Ms in the room to make our point, so I also supplement with an alternative of some kind for those who don't want/can't eat M&M's.

The student dotplot is really the main hurdle here. With 225 students in fixed seating, a human wave filing down the aisles to mark the chalk board is a non-starter so we use Google Sheets or Forms. Students use a smart phone to access a shortened link (e.g. tiny.cc) or scan a QR code displayed on the front screen to access the spreadsheet/form and enter their result from their seats. After a quick filter, I cut and paste the data into software, and make our class dotplot. From start to finish, this method in a large class might even be faster than the manual approach I had used in smaller classes.

A few themes...

Of course this is just one specific activity, but there are some themes here that apply to lots of activies.

Try it anyway. The "worst case scenarios" that I concoct when I imagine attempting an activity developed for small classes never really seem to happen. I probably fear that the mild noise, mess, and chaos of smaller

classes will escalate into a completely wasted class period with so many students. Thankfully, that hasn't happened (. . . yet?) largely because a little technology can neutralize the few inflection points where things would most likely derail (e.g. 225 students physically marking the chalkboard).

Redeem the smart phone! Don't underestimate the tiny computers that most of your students already bring to class. QR codes–those pixelated square codes you scan–and shortened URLs (e.g. tiny.cc; bit.ly) can direct students to an applet, Google Sheet (Form, Doc, etc), or anything else on the web right from their seats. Students can even share with a neighbor or look over the shoulder in front of them–particularly in stadium seat lecture halls. As another example, my class uses the Lock5 text and the accompanying StatKey software works great on a smart phone. I can introduce authentic data analysis tasks during "lecture" for students to tackle in pairs; one runs StatKey, the other takes notes.

The large enrollment actually has some perks. Since the class is large, class data sets are that much closer to Asymptopia. I'm sure I'm not the only one who's attempted to use a class plot to lay the groundwork for the CLT and wound up with a skewed or bimodal-looking mess that looks nothing like the bell-shape I had in mind. . . Furthermore, rare events and oddities that lead to interesting discussions almost always show up. Someone will enter the count rather than proportion, and someone else usually observes something rare just by chance. Since I'm the gatekeeper for the class plot, I decide whether there's time to discuss data entry errors or if we need to filter them out and plow ahead in the interest of time. Also, a student with no blue M&Ms, for example, isn't hypothetical or a mistake. Rare events in the tails of the distribution really do happen!

Keep an eye toward the big picture. You'll rarely need 100% cooperation to make your point. For example, in the M&M activity some students eat the candy before we start, others won't get to the spreadsheet in time, and some didn't get M&M's to begin with! Even if only 50 students fully participate ($< 25\%$ in my case), the illustration still serves it's purpose and benefits the whole class.