Applet #2 Activities: Instructor Guide

Learners studying introductory statistics usually find concepts relating to the sampling distribution of a statistic difficult to understand. One reason for this is that how a statistic varies over multiple samples is not easy to convey in a textbook or traditional lecture. A variety of methods have been proposed for aiding learning about sampling distributions, one being on–line simulation–based tools, or *applets*, that permit visualisation of multiple samples from a distribution and how their statistics vary. Research suggests that how students engage with such applets can impact the learning obtained, as without structure to their interactions with a visualisation tool learners may miss important concepts.

Prof. Mike Whitlock (in the Department of Zoology, University of British Columbia) has developed a suite of applets for use in introductory courses. The second of these, "Central Limit Theorem", allows the learner to explore how the sample mean behaves when sampling from an arbitrary distribution. The accompanying activity aims to assist the learner in their engagement with the applet, helping to focus attention on key learning goals. The activity may be used as an in–class activity (say as part of a lab–based tutorial) or outside class as a homework assignment. The activity has been tested on over two hundred students at the University of British Columbia. There follows an instructor guide to the activity.

The activity preamble suggests that the learner has previously engaged with the first applet, though that is not strictly necessary and the wording may be revised when there is no prior exposure to the applet on sampling from a Normal distribution. Initially the applet presents the "Coffee" example, and the activity explains this scenario and requires the learner to describe the distribution from which they may sample (sometimes referred to as the *parent distribution*). The applet includes a tutorial, and the learner is obliged to work through that in part 2. The learner is asked to reflect on what they learned from the tutorial and write down what are perceived as important points. If nothing else, this ensures the learner works through the tutorial and pays some attention.

Parts 3 - 5 were motivated by student responses in a focus group conducted to evaluate the applet. To attempt to measure any learning gains from the applet, students were posed some concept–based questions before and after engaging with the applet. One such question involved taking a sample of size 4 from a non–Normal, but not very skewed, distribution, and

predicting the shape of the distribution of 1000 sample means. Students in the focus group claimed they answered this incorrectly after working through the applet's tutorial, where in the "Coffee" example the sampling distribution based on $n = 4$ is skewed and not Normal–like as in the concept question. Instructors may feel this relates to a subtle point, as in reality when taking samples as small as $n = 4$ one can say little about the shape of the parent distribution and so could not safely apply the Central Limit Theorem. Such instructors may wish to omit these parts. However experience tells us students tend to urge instructors to provide a rule for when n is large enough to use the Normal approximation; there is no rule as such, and these activities aim to enable learners to appreciate the interplay between sample size and the shape of the parent distribution.

In part 3, the learner observes than when sampling from a very skewed distribution the sample mean is not Normal for small n. However in part 2, although sampling from a non–Normal distribution, small sample sizes do nevertheless yield sampling distributions for the mean that are approximately Normal. The learner is asked to identify in part 5 that it is skewness that plays an important role in the convergence to normality.

Part 6 requests the learner create their own parent distribution via the "Custom" mode. In beta–testing an earlier form of the applet some students had difficulty with this feature, and it is hoped that the pop–up instructions that were subsequently added are helpful (and further testing suggested they are). Note the customised distribution is re-scaled when the sample size is changed, though that has no material impact on what the learner sees. Here the learner is asked to create a very skewed distribution and incrementally increase the sample size until the sampling distribution of the mean looks roughly Normal. There is no "right answer" as such.

Finally part 7 asks the learner to fix the sample size at 30 and attempt to make a parent distribution for which the sampling distribution of the sample mean is not Normal–like. This is close to impossible using the custom tool, something that is suggested to the learner (if they had not appreciated the fact already) by the "yes or no" format of the answer expected. Instructors may opt to use just one of parts 6 and 7, as both parts target the same concept.