Five-start Ratings simulation

In this lab handout we will carry out a simulation using our textbook student resource website.

- Click the following link to textbook student resource website. `https://media.pearsoncmg.com/ph/esm/esm_agresti_ast4e_17/cw/ast4e_student.html`)
- Click "Sampling distribution of a sample mean (discrete population distribution)" under Chapter 7.
- By default, "Amazon Star Rating (Fitbit wristband)" is selected.

1. (i) Describe the population distribution. (Shape, center, spread).
2. For a randomly selected customer, what is the probability s/he rates the product with at least 4 stars?

   Now suppose that a researcher wants to reach out a sample of 2 customers and ask them for their rating. To simulate this result, select sample size (n) 2 and draw 1 sample by clicking "Draw Sample(s)".

3. What is the result of your simulated sample (Data distribution)? (e.g. 5 and 5). What is your sample mean and sample standard deviation? write them (result, sample mean, standard deviation) down on a piece of paper (we will use it in Part 6 below.) Compare the results with your neighbors. Do your neighbors have the same result as yours?
4. Draw a few more samples of sample size n=2. Observe how data distribution and sampling distribution change.
5. Simulate at least 20,000 samples of size n=2. What is the mean and standard deviation of **sampling distribution of sample mean**? What is the shape of the sampling distribution? Is it approximately normal? Compare the result with your neighbors.
6. In class, we learned :

   - standard deviation of probability distribution (population standard deviation) from Ch 6 $\sigma = \sqrt{\sum (x-\mu)^2 P(X=x)}$
   - sample standard deviation $s = \sqrt{\frac{\sum (x-\bar{x})^2}{(n-1)}}$ from Ch 2
   - standard deviation of sample proportion $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$ from Ch 7
   - standard deviation of sample mean $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ from Ch 7.

   Which formula do we use to obtain the standard deviation from Problem 3? Which formula do you use to obtain the standard deviation from Problem 5? **Repeat Problem 5 but this time use** $n = 30$.
7. Observe how mean and standard deviation of sample distribution change as sample size changes from $n = 2$ to $n = 30$. Does the shape of sampling distribution change?
8. Suppose the researcher surveyed a random sample of 30 customers and found the sample mean rating is 4. Do you think this value of sample mean is usually large? Why or why not?
9. How many standard deviation is this observed sample mean ($\bar{x} = 4$) away from the center of the sampling distribution?

10. What is the probability a random sample of 30 customers has a sample mean larger than 4?

    You may select any different population distribution (fair die, unfair die, or continuous distribution at `https://istats.shinyapps.io/sampdist_cont/`) and repeat.