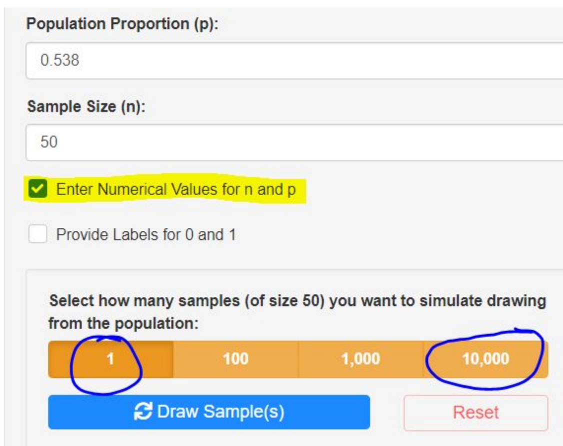
Show your work to receive full credit.

Problem 1

In the 2010 California gubernatorial election, 53.8% of all voters voted for Jerry Brown. We will use this information as the population proportion $p = 0.538$ and simulate exit polls using web apps (`https://istats.shinyapps.io/SampDist_Prop/`)

- Set the population proportion $p = 0.538$, the actual population proportion of voters who voted for Brown. (You should check "Enter Numerical Values for n and p" first. see the picture below.)
- In the box for the sample size n, type 50.
- Leave the default setting where we take 1 sample of this size (n=50)
- Click the "Draw sample" button once.



Answer the following :

a) Simulate drawing one sample of size 50. Let's treat this as a random sample of 50 people voters. What sample proportion $\bar{p}$ did you get? Did you expect to see the sample proportion to be exactly the same as the population proportion $p = 53.8\%$? Why / why not?

b) Keep the sample size n as 50 and population proportion p as 0.538, but now simulate drawing 40,000 samples of that size. (Click "10,000" circled in the picture above then click "Draw samples" four times). i) Submit the picture of the histogram of 40,000 sample proportions you generated in your HW answer. ii) Describe the shape, center and spread of this sampling distribution of sample proportion. (Hint: The app shows mean, standard deviation of sampling distribution).

c) Use a formula we learned from Ch 7 to verify the mean and standard deviation of sampling distribution from part b). (Note: It is okay to see a small difference (+/-0.0002) between the calculated values using formulas and values from simulation.)

d) Now change the sample size (n) to 200, keeping the population proportion 0.538. Simulation the exit poll at least 10,000 times. How did the sampling distribution of sample proportion change from part b)? Compare shape, mean, and spread when $n = 200$ vs $n = 50$.

Problem 2

Use the sampling distribution web app from Problem 1. This time use the population proportion $p = 0.97$, and $n = 100$. Simulate at least 10,000 samples. What is the shape of the sampling distribution of sample proportion? Is it approximately normal? If not, why not? (Hint: Check lecture notes page 99, Central Limit Theorem for $\hat{p}$)

Problem 3

Rafe was diagnosed with high blood pressure. He was able to keep his blood pressure in control for several month by taking blood pressure medicine. Rafe's blood pressure is monitored by taking three readings a day, in early morning, at midday, and in the evening.

a) During this period, the probability distribution of his systolic blood pressure reading had a mean of 130 ($\mu = 130$) and a standard deviation of 6 ($\sigma = 6$). If the successive observations behave like a random sample from this distribution, find the mean and standard deviation of the sampling distribution of the sample mean for the three observations each day.

b) Suppose that the probability distribution (population distribution) of his blood pressure reading is normal. What is the shape of the sampling distribution? Why?

c) Refer to part b. Find the probability that the sample mean exceeds 140, which is considered problematically high.

d) Find the probability that Rafe's systolic blood pressure reading from one randomly selected morning is greater than 140.

Problem 4

Refer to Problem 3 where the probability distribution of Rafe's systolic blood pressure reading had a mean of 130 ($\mu = 130$) and a standard deviation of 6 ($\sigma = 6$).
This time we are interested in sample mean from 21 observations.

a) If we repeat Problem 3 part c) to find $P(\bar{X} > 140)$ but with sample mean from 21 observations ($n = 21$) instead of 3 observations, how does your answer change? Increase, decrease or remain the same? Explain.

b) Use a 68 - 95- 99.7 rule to the sampling distribution of sample mean with (n=21) to find the interval that contains the middle 95% of the distribution. If we can't apply the 68-95-99.7 rule, explain why.

Problem 5

Based on data from the 2010 Major League Baseball season, X = number of home runs the San Francisco Giants hit in a game has the population mean of 1.0 and population standard deviation of 1.0.

a) Do you think X has a normal distribution? Why or why not?

b) Suppose that this year X has the same distribution. Report the shape, mean, and standard deviation of the sampling distribution of the mean number of home runs the team will hit in its 36 games.

c) Based on the answer to part b), find the probability that the sample mean number of home runs per game from 36 games will exceed 1.25.

Problem 6: Multiple Choice

Which of the following is *not* correct? **Briefly explain why it is incorrect.**
The standard deviation of a statistic describes :

(A) The standard deviation of the sampling distribution of that statistic

(B) The standard deviation of the individual observations in a sample data estimates

(C) The variability in the values of the statistic for repeated random samples of the same size

Problem 7: Multiple choice

Which of the following is correct? For each incorrect option, briefly explain why it is incorrect.
The sampling distribution of a sample mean for a random sample size of 100 describes :

(A) How sample means tend to vary from a random sample to sample of size 100.

(B) How observations tend to vary from person to person in a random sample of size 100

(C) How the data distribution looks like the population distribution when the sample size is larger than 30.

(D) How sample standard deviation (s) varies among samples of size 100.