Section 7.1: Sampling Distribution for Proportions

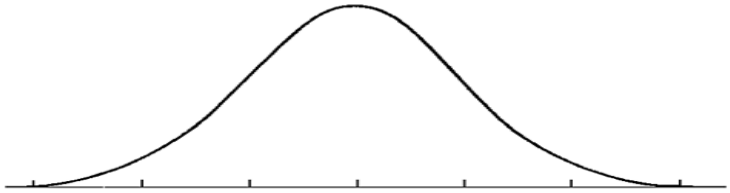
1) Harley-Davidson motorcycles make up 14% of all motorcycles registered in the United States. You plan to interview a simple random sample of 500 motorcycle owners and ask each if he/she owns a Harley.

   a. What is the population of interest?

   b. What is the sample of interest?

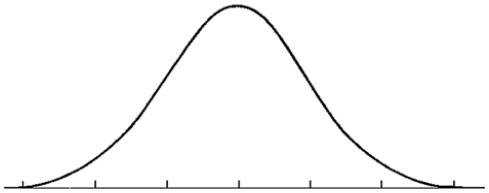
   c. Is 14% (i.e., 0.14) a parameter or a statistic?

   d. If we imagine *repeatedly* sampling (randomly) 500 motorcycle owners and computing a $\hat{p}$ value each time, and then making a histogram of *all* the $\hat{p}$ values, the graph will look normally distributed. Label the horizontal axis of the idealized histogram below, being sure to identify both the mean and at least two standard deviations from the mean in both directions.

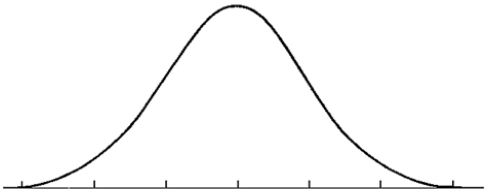


$\hat{p}$ distribution

   e. Is your sample likely to contain less than 15% Harley owners? (That is, what is $Prob(\hat{p} < 0.15)$ in the distribution above on the left?)

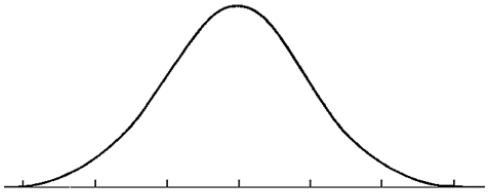


$\hat{p}$ distribution

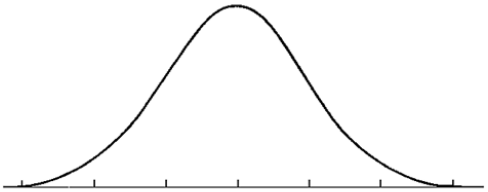


z-score distribution

   f. Is your sample likely to contain 20% or more who own Harleys? (That is, what is $Prob(\hat{p} > 0.20)$ in the distribution above on the left?)



$\hat{p}$ distribution



z -score distribution

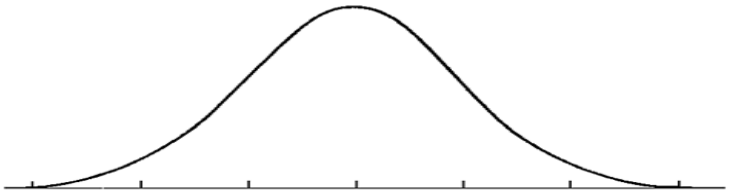
2) The College Alcohol Study interviewed a simple random sample of 14,941 college students about their drinking habits. Let's assume that, if it were possible to measure it, 50% of *all* college students would answer yes to the question, "Do you 'drink to get drunk' at least once in a while?"

a. What is the population of interest?

b. What is the sample of interest?

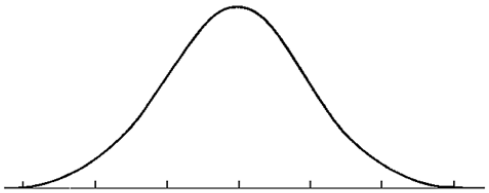
c. Is 50% a parameter or a statistic?

d. If we imagine *repeatedly* sampling (randomly) 14941 college students and computing a $\hat{p}$ value each time, and then making a histogram of *all* the $\hat{p}$ values, the graph will look normally distributed. Label the horizontal axis of the idealized histogram below, being sure to identify both the mean and at least two standard deviations from the mean in both directions.

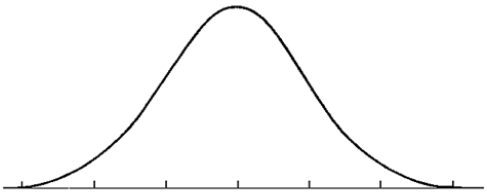


$\hat{p}$ distribution

e. In reality, we are only going to ask one random sample of 14941 college students and compute a single value of $\hat{p}$. What is the probability that this one value of $\hat{p}$ is between 49% and 51%? This is the probability that between 49% and 51% of your sample will answer yes to the question, "Do you 'drink to get drunk' at least once in a while?"

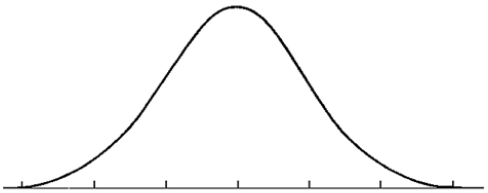


$\hat{p}$ distribution

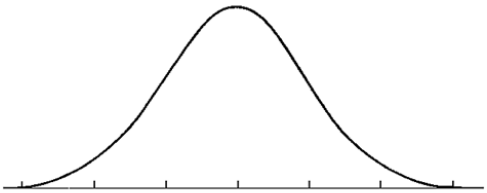


z -score distribution

f. Suppose that the random sample of college students only involves 100 students. What is the probability that between 49% and 51% of the students in your sample will say that they get drunk at least once in a while?



$\hat{p}$ distribution



z -score distribution