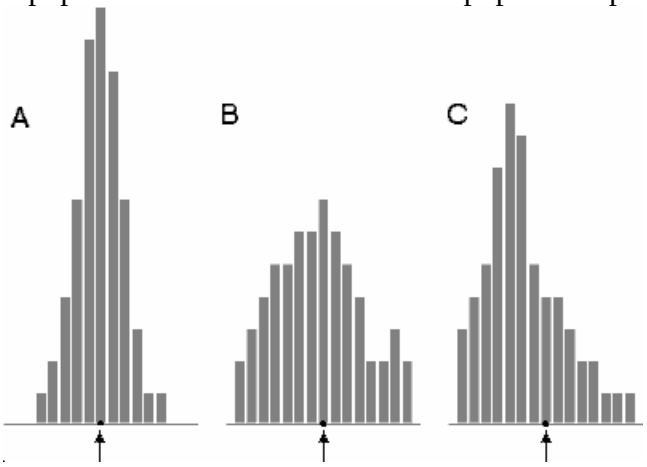
In items 1–3, classify each underlined number as a parameter or statistic. Give the appropriate notation for each.

1. Forty-two percent of today's 15-year-old girls will get pregnant in their teens.

2. A 1993 survey conducted by the *Richmond Times-Dispatch* one week before Election Day asked voters which candidate for the state's Attorney General they would vote for. Thirty-seven percent of the respondents said they would vote for the Democratic candidate. On Election Day, 41% actually voted for the Democratic candidate.

3. The National Center for Health Statistics reports that the mean systolic blood pressure for males 35 to 44 years of age is 128 and the standard deviation is 15. The medical director of a large company looks at the medical records of 72 executives in this age group and finds that the mean systolic blood pressure for these executives is 126.07.

Below are histograms of the values taken by three sample statistics in several hundred samples from the same population. The true value of the population parameter is marked on each histogram.



4. Which statistic has the largest bias among these three? Justify your answer.

5. Which statistic has the lowest variability among these three?

6. Based on the performance of the three statistics in many samples, which is preferred as an estimate of the parameter? Why?

An opinion poll asks, "Are you afraid to go outside at night within a mile of your home because of crime?" Suppose that the proportion of all adults who would say "Yes" to this question is $p = 0.4$.

1. Use the partial table of random digits below to simulate the result of an SRS of 20 adults. Be sure to explain clearly which digits you used to represent each of "Yes" and "No." Write directly on or above the table so that I can follow the results of your simulation. What proportion of your 20 responses were "Yes"?

6 8 4 1 7	3 5 0 1 3	1 5 5 2 9	7 2 7 6 5	8 5 0 8 9	5 7 0 6 7
8 2 7 3 9	5 7 8 9 0	2 0 8 0 7	4 7 5 1 1	8 1 6 7 6	5 5 3 0 0
6 0 9 4 0	7 2 0 2 4	1 7 8 6 8	2 4 9 4 3	6 1 7 9 0	9 0 6 5 6
3 6 0 0 9	1 9 3 6 5	1 5 4 1 2	3 9 6 3 8	8 5 4 5 3	4 6 8 1 6
3 8 4 4 8	4 8 7 8 9	1 8 3 3 8	2 4 6 9 7	3 9 3 6 4	4 2 0 0 6

2. Repeat problem 1 using the next consecutive lines of the digits table with one line per SRS until you have simulated the results of 5 SRSs of size 20 from the same population. Compute the proportion of "Yes" responses in each sample. These are the values of the statistic $\hat{p}$ in 5 samples. Find the mean of your 5 values of $\hat{p}$. Is it close to p?

3. The sampling distribution of $\hat{p}$ is the distribution of $\hat{p}$ from all possible SRSs of size 20 from this population. What would be the mean of this distribution?

4. If the population proportion changed to $p = 0.5$, what would then be the mean of the sampling distribution?