Statistics of Sample Distribution of Proportions

Expected Value / Expectation / Mean

$$\mathrm{E}(\hat{P}) = \bar{x} = \mathrm{E}\left(\frac{X}{n}\right) = \frac{1}{n}\mathrm{E}(X), \qquad \text{since } \mathrm{E}(X) = np, \qquad \mathrm{E}(\hat{P}) = \frac{1}{n}(np) = p$$

Mean and Sample Size

As the sample size increases the expected value remains constant.

Variance

$$\mathrm{var}(\hat{P}) = s^2 = \mathrm{Var}\left(\frac{X}{n}\right) = \left(\frac{1}{n}\right)^2 \mathrm{Var}(X)$$

$$\text{Since } \mathrm{Var}(X) = np(1-p), \qquad \mathrm{var}(\hat{P}) = \frac{1}{n^2}\big(np(1-p)\big) = \frac{p(1-p)}{n}$$

Variance and Sample Size

As the sample size increases, the variance decreases.

2016 Sample Exam 2 Question 3aiii / 2013 Exam 2 Question 2 (Modified)

It has been found that the probability that any member of FullyFit will complete a set of exercises called S in less than three minutes is $\frac{5}{8}$. This is independent of any other member. A random sample of 20 FullyFit members is taken. For a sample of 20 members, let X be the random variable that represents the number of members who complete S in less than three minutes. For samples of 20 members, $\hat{P}$ is the random variable of the distribution of sample proportions of people who complete S in less than three minutes. The expected value and variance of $\hat{P}$ are

$$\mathrm{E}(\hat{P}) = \frac{5}{8}, \qquad \mathrm{var}(\hat{P}) = \frac{\frac{5}{8}\left(1-\frac{5}{8}\right)}{20} = \frac{\frac{5}{8}\times\frac{3}{8}}{20} = \frac{\frac{15}{64}}{20} = \frac{3}{256}$$

Standard Deviation / Standard Error

The standard error is the standard deviation of the sample distribution of the proportion.

$$\mathrm{sd}(\hat{P}) = s = \sqrt{\mathrm{var}(\hat{P})} = \sqrt{\frac{p(1-p)}{n}}$$

Standard Deviation and Sample Size

As the sample size increases, the standard deviation decreases.

2016 Sample Exam 1 Question 7a

A student performs an experiment in which a computer is used to simulate drawing a random sample of size n from a large population. The proportion of the population with the characteristic of interest to the student is p. Let the random variable $\hat{P}$ represent the sample proportion observed in the experiment. If $p = \frac{1}{5}$, the smallest integer value of the sample size such that the standard deviation of $\hat{P}$ is less than or equal to $\frac{1}{100}$ is

$$\mathrm{sd}(\hat{P}) = \sqrt{\frac{\frac{1}{5}\left(1-\frac{1}{5}\right)}{n}} = \sqrt{\frac{\frac{1}{5}\times\frac{4}{5}}{n}} = \frac{\frac{2}{5}}{\sqrt{n}}, \qquad \frac{\frac{2}{5}}{\sqrt{n}} \leq \frac{1}{100} \Rightarrow \frac{1}{\sqrt{n}} \leq \frac{1}{1600} \Rightarrow \sqrt{n} \geq 1600 \Rightarrow n \geq 400.$$

$\therefore$ the smallest integer value of the sample size is 400.