Sample Distribution of Proportions

Sample Proportion

The ratio of successful outcomes to the sample size. For a sample of fixed size n, the sample proportion is

$$\hat{p} = \frac{\text{number of successful outcomes in the sample}}{\text{sample size}} = \frac{x}{n}$$

The sample proportion, $\hat{p}$, can be used as an estimate of the population proportion, p.

Sample Proportion as a Random Variable $\hat{P}$

The proportion of a population or distribution X is a fixed value. The sample proportion, however, varies between samples with the same size. Therefore, we can consider the sample proportion to be a random variable $\hat{P} = \frac{X}{n}$, where X is a binomial random variable associated with the number of items that have a particular characteristic. $\hat{P}$ can take on the values $\left\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right\}$.

Example VCAA 2017 NHT Exam 1 Question 6ai

At a large sporting arena there are a number of food outlets, including a cafe. The cafe employs five men and four women. Four of these people are rostered at random to work each day. Let $\hat{P}$ represent the sample proportion of men rostered to work on a particular day.

The possible values that $\hat{P}$ can take are $\left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$.

Determining Probabilities with Sample Proportions

Use $\hat{P} = \frac{X}{n}$ to convert the problem back to a binomial distribution with $\Pr(\text{success}) = p$ and n trials

$$\Pr(a < \hat{P} < b) = \Pr(na < X < nb) \qquad \Pr(a \leq \hat{P} \leq b) = \Pr(na \leq X \leq nb)$$

$$\Pr(\hat{P} > a) = \Pr(X > na) \qquad \Pr(\hat{P} \geq a) = \Pr(X \geq na)$$

Since X takes on integers but $\hat{P}$ takes on real numbers, na and nb may not be integers, so na can be rounded up and nb can be rounded down such that $\Pr(na < X < nb) = \Pr(\lceil na \rceil \leq X \leq \lfloor nb \rfloor)$

Example

The probability that a sample of 10 items taken from a distribution with a population proportion $\frac{1}{3}$ with a proportion between $5 \pm \sqrt{2}$ is

$$\Pr\left(5 - \sqrt{2} < \hat{P} < 5 + \sqrt{2}\right) = \Pr(3.586 < X < 6.414) = \Pr(4 \leq X \leq 6) = 0.421$$

Example VCAA 2016 Sample Exam 2 Question 3av

It has been found that the probability that any member of FullyFit will complete **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.

$$\hat{P} = \frac{X}{20}, \qquad X \sim \text{Bi}\left(20, \frac{5}{8}\right)$$

$$\Pr\left(\hat{P} \geq \frac{3}{4} \middle| \hat{P} \geq \frac{5}{8}\right) = \Pr(X \geq 15 | X \geq 12.5) = \frac{\Pr(X \geq 15)}{\Pr(X \geq 12.5)} = \frac{0.1788}{0.5079} = 0.352$$