

Probability and Statistics - Open and Free

- [SIGN IN]

Syllabus | Outline | Help | More

Unit 3 :: Probability

This course is not led by an instructor

Random Variables

Sampling Distributions

Module 9 / Behavior of Sample Proportion (3 of 3)

165

Apply the sampling distribution of the sample proportion (when appropriate). In particular, be able to identify unusual samples from a given population.

If a sampling distribution is normally shaped, then we can apply the Standard Deviation Rule and use z-scores to determine probabilities. Let's look at some examples.

A random sample of 100 students is taken from the population of all part-time students in the United States, for which the overall proportion of females is .6.

(a) There is a 95% chance that the sample proportion (\hat{p}) falls between what two values? First note that the distribution of \hat{p} has the mean $p = .6$, standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.6(1-0.6)}{100}} = 0.05$, and a shape that is close to normal, since $np = 100(.6) = 60$ and $n(1-p) = 100(.4) = 40$ are both greater than 10. The Standard Deviation Rule applies: the probability is approximately .95 that \hat{p} falls within 2 standard deviations of the mean, that is, between $0.6 - 2(.05)$ and $0.6 + 2(.05)$. There is roughly a 95% chance that \hat{p} falls in the interval (.5, .7).

(b) What is the probability that sample proportion \hat{p} is less than or equal to .56?

To find $P(\hat{p} \leq 0.56)$, we standardize .56 to $z = (.56 - .60) / .05 = -.80$:

$$P(\hat{p} \leq 0.56) = P(Z \leq -0.80) = 0.2119$$

To see the impact of the sample size on these probability calculations, consider the following variation of our example.

A random sample of **2,500** students is taken from the population of all part-time students in the United States, for which the overall proportion of females is .6.

(a) There is a 95% chance that the sample proportion (\hat{p}) falls between what two values? First note that the distribution of \hat{p} has the mean $p = .6$, standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.6(1-0.6)}{2500}} = 0.01$, and a shape that is close to normal, since $np = 2500(.6) = 1500$ and $n(1-p) = 2500(.4) = 1000$ are both greater than 10. The standard deviation rule applies: the probability is approximately .95 that \hat{p} falls within 2 standard deviations of the mean, that is, between $0.6 - 2(.01)$ and $0.6 + 2(.01)$. There is roughly a 95% chance that \hat{p} falls in the interval (.58, .62).

(b) What is the probability that sample proportion \hat{p} is less than .56?

To find $P(\hat{p} \leq 0.56)$, we standardize .56 to $Z = (.56 - .60) / .01 = -4.00$.

$$P(\hat{p} \leq 0.56) = P(Z \leq -4.0) = 0, \text{ approximately.}$$

Comment

As long as the sample is truly random, the distribution of \hat{p} is centered at p , no matter what size sample has been taken. Larger samples have less spread. Specifically, when we multiplied the sample size by 25, increasing it from 100 to 2,500, the standard deviation was reduced to 1/5 of the original standard deviation. Sample proportion strays less from population proportion .6 when the sample is larger: it tends to fall anywhere between .5 and .7 for samples of size 100, whereas it tends to fall between .58 and .62 for samples of size 2,500. It is not so improbable to take a value as low as .56 for samples of 100 (probability is more than 20%) but it is almost impossible to take a value as low as .56 for samples of 2,500 (probability is virtually zero).

The purpose of this next activity is to give guided practice in finding the sampling distribution of the sample proportion

(\hat{p}), and use it to draw conclusions about what values of \hat{p} we are most likely to get.

Learn By Doing

The proportion of left-handed people in the general population is about 0.1. Suppose a random sample of 225 people is observed.

What is the sampling distribution of the sample proportion (\hat{p})? In other words, what can we say about the behavior of the different possible values of the sample proportion that we can get when we take such a sample? (Note: normal approximation is valid because $.1(225) = 22.5$ and $.9(225) = 202.5$ are both more than 10.)

Submit and Compare

Since the sample proportion has a normal distribution, its values follow the Standard Deviation Rule. What interval is almost certain (probability .997) to contain the sample proportion of left-handed people?

Submit and Compare

In a sample of 225 people, would it be unusual to find that 40 people in the sample are left-handed?

Submit and Compare

Find the approximate probability of at least 27 in 225 (proportion .12) being left-handed. In other words, what is $P(\hat{p} \geq 0.12)$?

Guidance: Note that 0.12 is exactly 1 standard deviation (0.02) above the mean (0.1). Now use the Standard Deviation Rule.

Submit and Compare

Theoretical Comment (Optional)

The above results for the distribution of sample proportion \hat{p} are directly related to the results already obtained for the distribution of sample count X in a binomial experiment. Remember that X had mean np , standard deviation $\sqrt{np(1-p)}$, and a shape that allowed for normal approximations as long as both np and $n(1-p)$ were at least 10. Since sample proportion is $\hat{p} = \frac{X}{n}$, we could derive the mean and standard deviation of \hat{p} by applying the Rules for Means and Variances: $\mu_{\hat{p}} = \mu_{\frac{X}{n}} = \frac{1}{n}\mu_x = \frac{1}{n}(np) = p$ and $\sigma_{\hat{p}}^2 = \sigma_{\frac{X}{n}}^2 = \frac{1}{n^2}\sigma_x^2 = \frac{1}{n^2}(np(1-p)) = \frac{1}{n}p(1-p)$ so $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. The requirements that np and $n(1-p)$ be at least 10 are the same, whether we are focusing on the distribution of sample count or the distribution of sample proportion. After all, the shape of \hat{p} is the same as the shape of X : the scale of the horizontal axis is just uniformly divided by n .

165



Unless otherwise noted this work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-nc-sa/3.0/).

[Open Learning Initiative](#)