# Chapter 9 – Sampling Distribution Models

We know the sample proportion of observed "successes," $\hat{p}$, is a random variable. We don't know what value we'll find until we actually take a sample. Likewise, we also can recognize that if a different sample were taken (involving different individuals or objects) we'd most likely get a different value. We need a way to model the distributions of sample statistics like $\hat{p}$ or the sample mean $\bar{x}$. This will form the basis for statistical inference to follow.

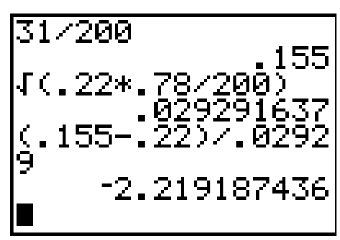
Intuitively, the mean for a statistic such as $\hat{p}$ (that estimates the population proportion *p*) should be close to *p* if we have a good (random and independent) sample. Similarly, $\bar{x}$ should be close to the population mean $\mu$. How close? That's measured by the standard deviation. The standard deviation of $\hat{p}$ is defined (actually mathematically proven) to be $\sigma(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{pq}{n}}$. It can be shown through simulation that if both $np \geq 10$ and $n(1-p) \geq 10$ that $\hat{p}$ will be approximately normally distributed. Similarly, we find (for large samples, or in the case of a normal population) that $\bar{x}$ has a normal distribution with mean $\mu$ and standard deviation $\sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}}$.

How can we make use of these facts? Ultimately for inference – making conclusions about what we believe is true about the actual population parameters. For now, we'll use these ideas for some probability questions.

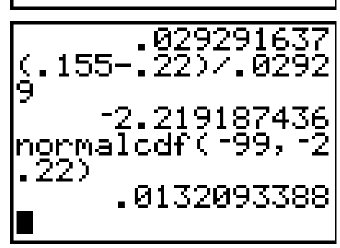
## USING SAMPLING DISTRIBUTIONS FOR A PROPORTION

The Centers for Disease Control report that 22% of 18-year-old women have a body mass index (BMI) of 25 or more. These values are associated with increased health risk. As part of a routine health check at a large college, the physical education department decided to try a self-report system, instead of having girls actually come in to be measured (height and weight since BMI = weight in kg/(height in m)$^2$, they asked 200 randomly selected female students to report their heights and weights. Only 31 had a BMI higher than 25. Should the physical education department continue to use the self-reporting system?

The sample fulfills the conditions necessary for doing the calculations needed. We had 31 females with BMI over 25, so there must have been 169 with BMI 25 or less (according to their reported heights and weights.) The females were randomly selected, and it's reasonable that they are less than 10% of all females at this "large college." First of all, the observed proportion of the randomly selected females is $\hat{p} = 31/200 = 15.5\%$. This seems low. What's the chance of getting this type of result (or something more extreme) if the national rate really holds at this university? If the national rate is true here, we should have $p = 0.22$. This is the mean for the normal distribution. The standard deviation will be $\sigma(\hat{p}) = \sqrt{\frac{.22(.78)}{200}} = 0.02929$.




To find the chance of our results (assuming, of course that the national model holds) we'll first compute the z-score for our observed result. Our observed 15.5% is about 2.2 standard deviations *below* the expected 22%. Is this unusual? Based on the 68-95-99.7 Rule, we can say it is. How unusual is it, exactly?



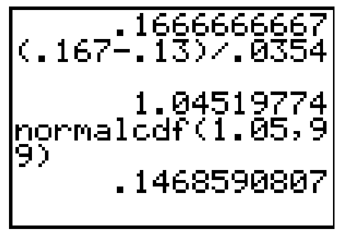

We're back to normal models, so we use `normalcdf` to find the probability of being 2.2 (or more) standard deviations below the mean. There is only about a 1.3% chance of getting 31 (or fewer) females having a BMI of more than 25 in a sample of 200, if the national rate holds here.

What should the college do about the self-reporting policy? Unless they've noticed that their female students are particularly thin, they most likely should scrap this system. Come on – how many women (or men for that matter) do you know that would *really* tell the truth about their height and weight?

## Another Example

If we suppose that 13% of the population is left-handed, what's the chance that a 200-seat auditorium with 15 "lefty desks" will be able to accommodate all the lefties in a large class? Suppose there are 90 students in the class. We'd think there should be plenty of seats of each type, wouldn't we?

First, consider the conditions. We can reasonable believe that 90 students in the class are less than 10% of the population of all students at the school and could be considered a random sample, "handedness" is independent among individuals (of course, assuming no twins, triplets, etc), and we expect more than 10 lefties as well as more than 10 righties since 0.13*90 = 11.7 and 0.87*90 = 78.3. We can proceed to answer the question of interest: what's the chance of more than 15 lefties in a class of 90 students?



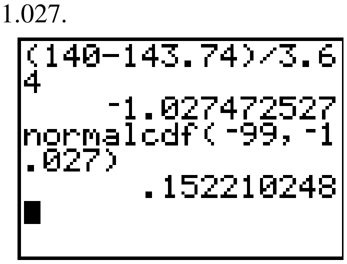

Since $\mu = p = 0.13$, we also have $\sigma(\hat{p}) = \sqrt{\frac{.13(.87)}{90}} = 0.0354$. Now we can do the `normalcdf` calculation to find the chance of more than 15 left-handed students, which means we'd see a sample proportion in excess of $\hat{p} = 15/90 = 0.167$. For our particular sampling distribution, 0.167 has a *z*-score of $z = \frac{.167 - .13}{.0354} = 1.05$. We see that there is about a 14.7% chance of not having enough lefty desks for the class.

## SAMPLING DISTRIBUTIONS FOR A SAMPLE MEAN

The Central Limit Theorem says that as sample sizes get "large," the distribution of the sample mean $\bar{x}$ becomes normally distributed. (This is also what really happens to the sample proportion $\hat{p}$ if we consider it an average of 0's (for failures) and 1's (for successes.) There are various "rules of thumb" for how large is a large enough sample, but the real answer is that it depends on the shape of the parent (population) distribution. If the distribution is unimodal and fairly symmetric, even very small samples will give means that look normally distributed.

Did the women at the college "shave" their weights in the example above? The CDC reports that 18-year-old women in the U.S. have a mean weight of 143.74 lb with standard deviation of 51.54 lb. The 200 women in the sample reported an average weight of 140 lb. Is this unusually low, or might this just be random sampling variation at work?

We considered some of the conditions above (we can believe these women are less than 10% of the women at the "large college", and they were randomly selected.) Now, since 200 is a pretty large sample, and weights are most likely unimodal and fairly symmetric, we can also believe that the sample mean $\bar{x}$ will have a normal distribution. If these women follow the national pattern, they should have μ = 143.74 lb. The standard deviation of their average reported weights is not σ = 51.41 lb, however, but $\sigma(\bar{x}) = 51.41/\sqrt{200} = 3.64$. We find the *z*-score on this distribution that corresponds to an observed average of 140 is (140-143.74)/3.64 = -1.027.

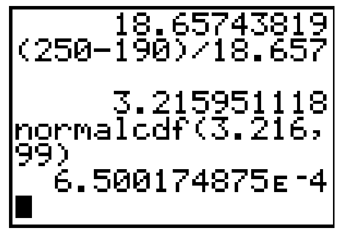



Doing the normal calculation gives a 15.2% chance of observing an average reported weight of 140 lb or less in this sample. That's not overwhelming evidence of weight shaving. Their results have a good chance of occurring by randomness.

## Another Example

The Centers for Disease Control and Prevention reports that the mean weight of adult men in the U.S. is 190 lb with standard deviation 59 lb. If an elevator has a limit of 10 people or 2500 lb, what's the chance that 10 men will overload it?

We can reasonably assume the 10 men are a random sample from the population, and that they represent less than 10% of the population of men who might use this elevator. It's also reasonable to believe their weights are independent of one another. Also, since weights are unimodal and fairly symmetric, it's reasonable that the average for these 10 men will have a normal distribution by the Central Limit Theorem.




We must first realize that for the elevator to be overloaded (have total weight more than 2500 lb), the 10 men must average more than 250 lb each. If these men follow the national model, they will have $\mu = \mu(\bar{x}) = 190$ lb with $\sigma(\bar{x}) = 59/\sqrt{10} = 18.657$ lb. The $z$-score on this distribution that corresponds to an observed average of 250 is (250-190)/18.657 = 3.216. The probability of these 10 men averaging more than 250 lb is 0.00065. We have a very small chance of overloading the elevator if people follow the 10 person maximum.

## WHAT CAN GO WRONG?

### My Results Don't Match! (Rounding Errors)

Most "errors" are the result of rounding in intermediate steps of the problem. Rounding the standard deviations, the z-scores, etc can have an impact on the final answers. This author's advice is to carry more decimal places than is really necessary and do all rounding at the end of a series of calculations. Your instructor may have different rules. If in doubt, ASK!

### My Results Don't Match! (Population vs. Sample)

The other typical error made by students in working with sampling distributions is failure to recognize the difference between dealing with one observation (or realization of a process) and a sample mean. This means that the error is usually related to having forgotten to divide the population standard deviation by $\sqrt{n}$ .