Math 142: Monday, Nov 18th, Day 25 Notes

Agenda for Monday, Nov 14th 2024

Coming up

- Student Hours Tuesday 1-2:15pm and Wednesday 10-11:15 am in C162
- Study session Exam 3 Tuesday 3-4:30pm in C168
- HW 17 and 18 should be done before the exam
- Exam 3 in class on Wednesday

- One job of a statistician is to make statistical inferences about populations based on samples taken from the population.
- In the last chapter, we used confidence intervals to estimate an unknown population parameter.
- In this chapter we will use statistical inference to make a decision about a parameter; that is, given two **competing** claims about an unknown parameter, we will decide which of them is more plausible.
- The method for doing this is called **hypothesis testing**.

population (parameter)

μ = mean

p = proportion

σ = st. dev.

sample (statistic)

$\bar{x}$ = sample mean

$\hat{p}$ = sample proportion

s = st. dev of sample

Example 1.

The ACME chemical corporation makes a fabric cleaner; the company claims that this product successfully removes 90% of all stains. A consumer organization questions this claim and decides to test the product. They select a random sample of 100 stained garments and apply the product to each. They find that the stain remover successfully cleaned only 78 of them. Does this sample data provide evidence that the company is overstating the effectiveness of their product? That is, does the sample data provide evidence that the success rate for the fabric cleaner is less than 90%? $\hat{p} = .78$

We let **p denote the success rate for the cleaner** – that is, this is the proportion of all stains in the population that will be removed by the product. Now we see that we have two competing claims: read H o or H null hypothesis

- claim • ACME claims that $p = 0.9$ H_0 = the null hypothesis
- claim • The consumer organization claims that $p < 0.9$ H_a = the alternative hypothesis

Which one of these claims is true?

- Note that if the company's claim is true, then we would expect to get about 90 successes in a sample of 100.
- Let us give the company the benefit of the doubt.
- Let us assume that the success rate really is 90% and calculate the probability of getting a sample of 100 garments with a success rate of 78% or less (a sample proportion less than 78% would also support the consumer organization's claim).
- If n = 100 and p = .90, then the sampling proportion for $\hat{p}$ is approximately normal with

$P(\hat{p} < .78) = ?\ .0000317$

$P(\bar{x} < 7ft)$ – sample of n

$P(x < 7ft)$ – one person

Recall: $\bar{x} \sim N\left(\mu_{\bar{x}} = \mu, \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\right)$

$\hat{p} \sim N\left(\mu_{\hat{p}} = p, \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}\right)$

$P(\hat{p} < .78) = \text{normalcdf}\left(-1E99, .78, .9, \sqrt{\frac{.9(1-.9)}{100}}\right)$

- Using the calculator, we see that the probability we want is:

$$P(\hat{p} < 0.78) = 0.0000317$$

- So, if ACME's claim really is true, then getting 78 or fewer successes in a sample of 100 garments would be highly unlikely (only about 3 times for every 100,000 samples).

- This means that there are two possibilities: The success rate really is 90%, and a very rare event occurred, **or** else the assumption that $p = .90$ is not correct. Given how small the probability we calculated was, the second option seems to be more plausible.

- In other words, the sample data is not compatible with ACME's claim about the success rate of their product; so we would reject the assumption that the success rate is 90%. Thus, this sample data provides evidence that ACME was overstating the success rate of their product.

we reject the H_0, the null hypothesis

- The line of reasoning we used here is known as **hypothesis testing**, and it is a widely used method of statistical inference.

- In general, a hypothesis test will always consist of two contradictory hypotheses or statements, a decision based on sample data, and a conclusion. Note that in all of our problems, we will be given either the sample data or the summary statistics. Reviewing our example, we see that a hypothesis test will involve the following five steps:

1. Set up two contradictory hypotheses, which we call the null hypothesis and the alternative hypothesis.

   H_o : $=$, $\leq$, $\geq$

   H_a : $\neq$, $>$, $<$

2. Determine the correct sampling distribution to perform the calculations.

   p or μ

3. Assuming that the null hypothesis is true, we calculate the probability of getting sample data like the data we have actually observed.

4. If this probability is sufficiently small, we reject the null hypothesis.

5. Interpret the decision to write a meaningful conclusion; i.e. interpret the decision to answer the original question.