Agenda: #62F AP Stats Ch19C

Next class 15min easy quiz

Objectives :

1) Learn what two-sided alternative hypothesis

2) Learn 1-PropZTest on TI-Calc

CW1) (20 min)

Reading Page 500~510

we stop one by one and answer questions.

CW2) (20min) 1-PropZTest on TI-Calc

3) Conclusion 15min Quiz is about easy

"box-whisker plot"

Alternative Alternatives

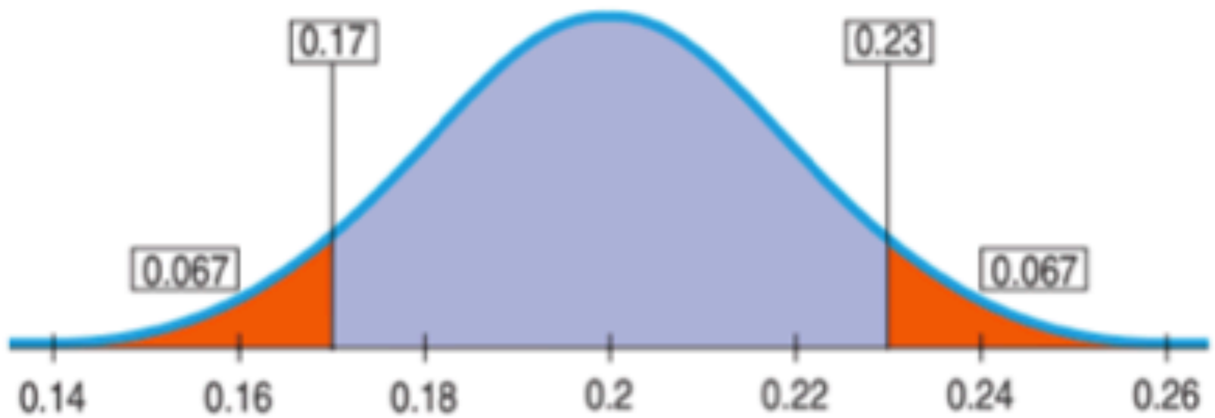
Tests on the ingot data can be viewed in two different ways. We know the old cracking rate is 20%, so the null hypothesis is

$$H_0: p = 0.20.$$

But we have a choice of alternative hypotheses. A metallurgist working for the company might be interested in *any* change in the cracking rate due to the new process. Even if the rate got worse, she might learn something useful from it. In that case, she's interested in possible changes on both sides of the null hypothesis. So she would write her alternative hypothesis as

$$H_A: p \neq 0.20.$$

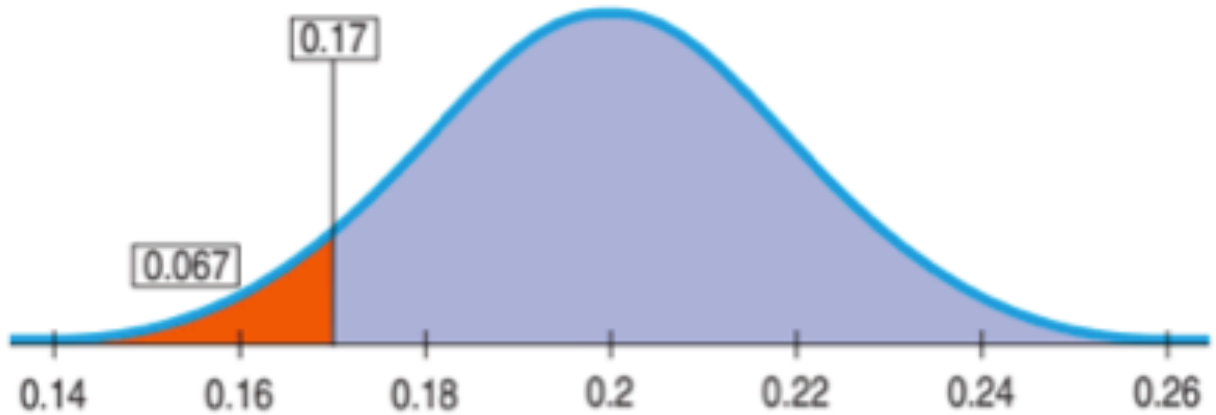
An alternative hypothesis such as this is known as a **two-sided alternative**[3] because we are equally interested in deviations on either side of the null hypothesis value. For two-sided alternatives, the P-value is the probability of deviating in *either* direction from the null hypothesis value.



But management is really interested only in *lowering* the cracking rate below 20%. The scientific value of knowing how to *increase* the cracking rate may not appeal to them. The only alternative of interest to them is that the cracking rate *decreases*. So, they would write their alternative hypothesis as

$$H_A: p < 0.20.$$

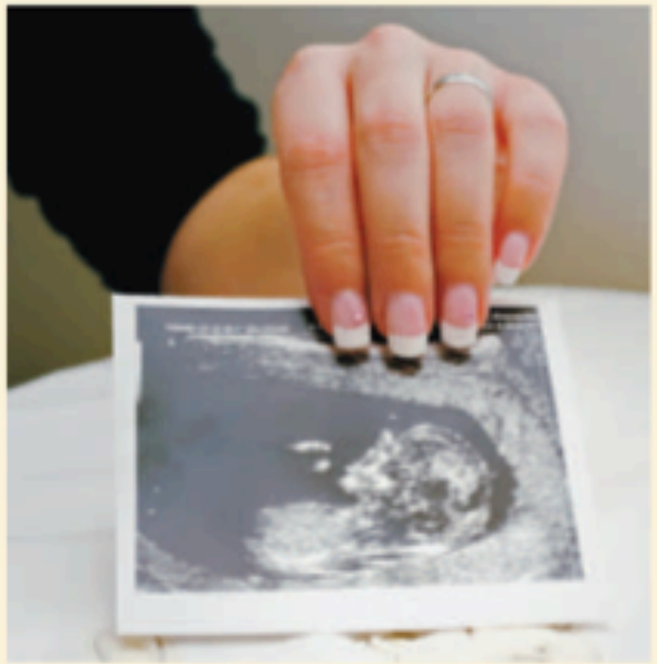
An alternative hypothesis that focuses on deviations from the null hypothesis value in only one direction is called a **one-sided alternative**.



[3]It is also called a **two-tailed alternative**, because the probabilities we care about are found in both tails of the sampling distribution.

Two-sided alternative (Two-tailed alternative)	An alternative hypothesis is two-sided ($H_A: p \neq p_0$) when we are interested in deviations in *either* direction away from the hypothesized parameter value. (p. 500)
One-sided alternative (One-tailed alternative)	An alternative hypothesis is one-sided (e.g., $H_A: p > p_0$ or $H_A: p < p_0$) when we are interested in deviations in *only one* direction away from the hypothesized parameter value. (p. 500)

Step-by-Step Example TESTING A HYPOTHESIS



Advances in medical care such as prenatal ultrasound examination now make it possible to determine a child's sex early in a pregnancy. There is a fear that in some cultures some parents may use this technology to select the sex of their children. A study from Punjab, India (E. E. Booth, M. Verma, and R. S. Beri, "Fetal Sex Determination in Infants in Punjab, India: Correlations and Implications," *BMJ* 309 [12 November 1994]: 1259–1261), reports that, in 1993, in one hospital, 56.9% of the 550 live births that year were boys. It's a medical fact that male babies are slightly more common than female babies. The study's authors report a baseline for this region of 51.7% male live births.

Question: Is there evidence that the proportion of male births is different for this hospital?

I want to know whether the proportion of male births in this hospital is different from the established baseline of 51.7%. The data are the recorded sexes of the 550 live births from a hospital in Punjab, India, in 1993, collected for a study on fetal sex determination. The parameter of interest, p, is the proportion of male births:

$$H_0: p = 0.517$$
$$H_A: p \neq 0.517$$

✓ **Independence Assumption:** There is no reason to think that the sex of one baby can affect the sex of other babies, so births can reasonably be assumed to be independent with regard to the sex of the child.

✓ **Randomization Condition:** The 550 live births are not a random sample, so I must be cautious about any general conclusions. I hope that this is a representative year, and I think that the births at this hospital may be typical of this area of India.

✓ **10% Condition:** I would like to be able to make statements about births at similar hospitals in India. These 550 births are fewer than 10% of all of those births.

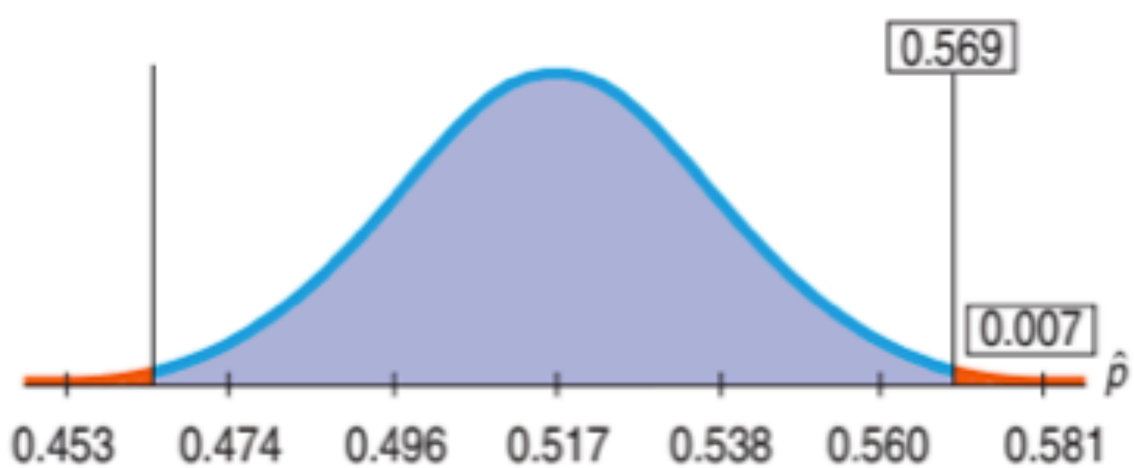
The null model is a Normal distribution with a mean of 0.517 and a standard deviation of

$$SD(\hat{p}) = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{(0.517)(1 - 0.517)}{550}}$$
$$= 0.0213.$$

The observed proportion, $\hat{p}$, is 0.569, so

$$z = \frac{\hat{p} - p_0}{SD(\hat{p})} = \frac{0.569 - 0.517}{0.0213} = 2.44.$$

The sample proportion lies 2.44 standard deviations above the mean.



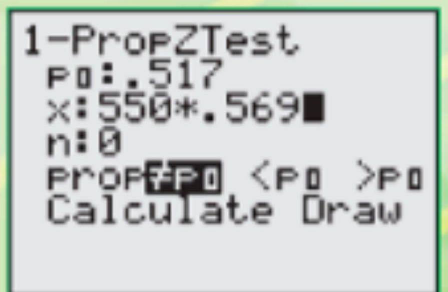
$$P = 2P(z > 2.44) = 2(0.0073) = 0.0146$$

The P-value of 0.0146 says that if the true proportion of male babies were still at 51.7%, then an observed proportion as different as 56.9% male babies would occur at random only about 15 times in 1000. With a P-value this small, I reject H_0. This is strong evidence that the proportion of boys is not equal to the baseline for the region. It appears that the proportion of boys may be larger.

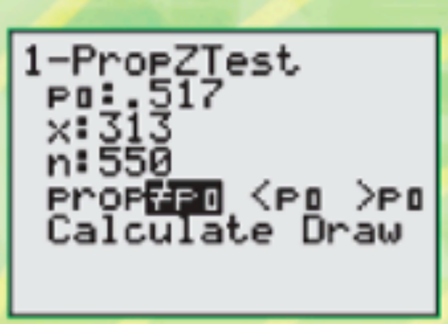
TI Tips TESTING A HYPOTHESIS

By now probably nothing surprises you about your calculator. Of course it can help you with the mechanics of a hypothesis test. But that's not much. It cannot write the correct hypotheses, check the appropriate conditions, interpret the results, or state a conclusion. You still have to do the tough stuff!

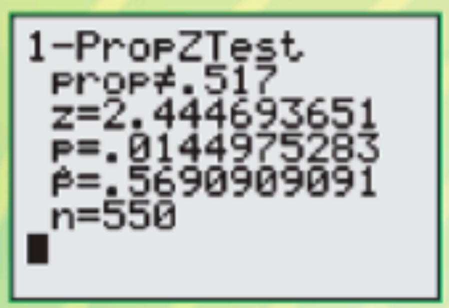
Let's do the mechanics of the Step-By-Step example about the post-ultrasound male birthrate. Based on historical evidence, we hypothesized that 51.7% of babies would be males, but one year at one hospital the rate was 56.9% among 550 births.



- Go to the `STAT TESTS` menu. Scroll down and select `1-PropZTest`.
- Specify the hypothesized proportion p_0
- Enter `x`, the observed number of males. Since you don't know the actual count, enter `550*.569` there and then round the resulting `312.95` off to a whole number.



- Specify the sample size.
- Since this is a two-tailed test, indicate that you want to see if the observed proportion is significantly different (≠) from what was hypothesized.
- `Calculate` the result.



Okay, the rest is up to you. The calculator reports a z-score of 2.445 and a P-value of 0.0145. Such a small P-value indicates that this higher rate of male births is unlikely to be just sampling error. Be careful how you state your conclusion.

P-Values and Decisions: What to Tell About a Hypothesis Test

TELL MORE

Don't We Want to Reject the Null? Often the folks who collect the data or perform the experiment hope to reject the null. (They hope the new drug is better than the placebo, or the new ad campaign is better than the old one.) But when we practice Statistics, we can't allow that hope to affect our decision. The essential attitude for a hypothesis tester is skepticism. Until we become convinced otherwise, we cling to the null's assertion that there's nothing unusual, no effect, no difference, etc. As in a jury trial, the burden of proof rests with the alternative hypothesis—innocent until proven guilty. When you test a hypothesis, you must act as judge and jury, but you are not the prosecutor.

Hypothesis tests are particularly useful when we must make a decision. Is the defendant guilty or not? Should we choose print advertising or television? The absolute nature of the hypothesis test decision, however, makes some people (including the authors) uneasy. Whenever possible, it's a good idea to report a confidence interval for the parameter of interest as well.

How small should the P-value be to reject the null hypothesis? A jury needs enough evidence to show the defendant guilty "beyond a reasonable doubt." How does that translate to P-values? The answer is that there is no good, universal answer. How small the P-value has to be to reject the null hypothesis is highly context-dependent. When we're screening for a disease and want to be sure we treat all those who are sick, we may be willing to reject the null hypothesis of no disease with a P-value as large as 0.10. That would mean that 10% of the healthy people would be treated as sick and subjected to further testing. We might rather treat (or recommend further testing for) the occasional healthy person than fail to treat someone who was really sick. But a long-standing hypothesis, believed by many to be true, needs stronger evidence (and a correspondingly small P-value) to reject it.

See if you require the same P-value to reject each of the following null hypotheses:

- A renowned musicologist claims that she can distinguish between the works of Mozart and Haydn simply by hearing a randomly selected 20 seconds of music from any work by either composer. What's the null hypothesis? If she's just guessing, she'll get 50% of the pieces correct, on average. So our null hypothesis is that *p* equals 50%. If she's for real, she'll get more than 50% correct. Now, we present her with 10 pieces of Mozart or Haydn chosen at random. She gets 9 out of 10 correct. It turns out that the P-value associated with that result is 0.011. (In other words, if you tried to just guess, you'd get at least 9 out of 10 correct only about 1% of the time.) What would *you* conclude? Most people would probably reject the null hypothesis and be convinced that she has some ability to do as she claims. Why? Because the P-value is small and we don't have any particular reason to doubt the alternative.

- On the other hand, imagine a student who bets that he can make a flipped coin land the way he wants just by thinking hard. To test him, we flip a fair coin 10 times. Suppose he gets 9 out of 10 right. This also has a P-value of 0.011. Are you willing now to reject this null hypothesis? Are you convinced that he's not just lucky? What amount of evidence *would* convince you? We require more evidence if rejecting the null hypothesis would contradict long-standing beliefs or other scientific results. Of course, with sufficient evidence we would revise our opinions (and scientific theories). That's how science makes progress.

- A researcher claims that the proportion of college students who hold part-time jobs now is higher than the proportion known to hold such jobs a decade ago. You might be willing to believe the claim (and reject the null hypothesis of no change) with a P-value of 0.05.
- An engineer claims that even though there were several problems with the rivets holding the wing on an airplane in their fleet, they've retested the proportion of faulty rivets and now the P-value is small enough to reject the null hypothesis that the proportion is the same. What P-value would be small enough to get you to fly on that plane?

Your conclusion about any null hypothesis should always be accompanied by the P-value of the test. Don't just declare the null hypothesis rejected or not rejected. Report the P-value to show the strength of the evidence against the hypothesis and the effect size. This will let each reader decide whether or not to reject the null hypothesis and whether or not to consider the result important if it is statistically significant.

To complete your analysis, follow your test with a confidence interval for the parameter of interest, to report the size of the effect.