9.3 Significance Tests for a Population Mean

Learning Objectives:
1. State and check the Random, 10%, and Normal/Large Sample conditions for performing a significance test about a population mean.
2. Perform a significance test about a population mean.
3. Use a confidence interval to draw conclusion for a two-sided test about a population parameter.
4. Perform a significance test about a mean difference using paired data.

Vocabulary: one sample *t* test for a population mean, paired data,

Read 574–579
What are the three conditions for conducting a significance test for a population mean? How are these different than the conditions for calculating a confidence interval for a population mean?

What test statistic do we use when testing a population mean? Is the formula on the formula sheet?

How do you calculate *P*-values using the *t* distributions?

Read 579–582

Alternate Example: *Short Subs*

Abby and Raquel like to eat sub sandwiches. However, they noticed that the lengths of the "6-inch sub" sandwiches they get at their favorite restaurant seemed shorter than the advertised length. To investigate, they randomly selected 24 different times during the next month and ordered a "6-inch" sub. Here are the actual lengths of each of the 24 sandwiches (in inches):

4.50	4.75	4.75	5.00	5.00	5.00	5.50	5.50
5.50	5.50	5.50	5.50	5.75	5.75	5.75	6.00
6.00	6.00	6.00	6.00	6.50	6.75	6.75	7.00

(a) Do these data provide convincing evidence at the $\alpha = 0.10$ level that the sandwiches at this restaurant are shorter than advertised, on average?

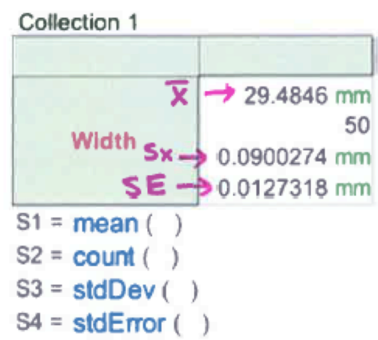
(b) Given your conclusion in part (a), which kind of mistake—a Type I or a Type II error—could you have made? Explain what this mistake would mean in context.

HW page 573 (54–58 even), page 595 (65, 69, 73)

9.3 Two-sided tests for μ

Read 582–583
Can you use your calculator for the Do step? Are there any drawbacks?

Read 583–586
Alternate Example: *Don't break the ice*
In the children's game Don't Break the Ice, small plastic ice cubes are squeezed into a square frame. Each child takes turns tapping out a cube of "ice" with a plastic hammer, hoping that the remaining cubes don't collapse. For the game to work correctly, the cubes must be big enough so that they hold each other in place in the plastic frame but not so big that they are too difficult to tap out. The machine that produces the plastic cubes is designed to make cubes that are 29.5 millimeters (mm) wide, but the actual width varies a little. To ensure that the machine is working well, a supervisor inspects a random sample of 50 cubes every hour and measures their width. The Fathom output summarizes the data from a sample taken during one hour.



(a) Interpret the standard deviation and the standard error provided by the computer output.

(b) Do these data give convincing evidence that the mean width of cubes produced this hour is not 29.5 mm? Use a significance test with α = 0.05 to find out.

(c) Calculate a 95% confidence interval for μ. Does your interval support your decision from (c)?

HW page 597 (75, 77, 79, 83)

9.3 Paired Data and Using Tests Wisely

Read 586–589

Alternate Example: *Is the express lane faster?*

For their second semester project in AP Statistics, Libby and Kathryn decided to investigate which line was faster in the supermarket: the express lane or the regular lane. To collect their data, they randomly selected 15 times during a week, went to the same store, and bought the same item. However, one of them used the express lane and the other used a regular lane. To decide which lane each of them would use, they flipped a coin. If it was heads, Libby used the express lane and Kathryn used the regular lane. If it was tails, Libby used the regular lane and Kathryn used the express lane. They entered their randomly assigned lanes at the same time, and each recorded the time in seconds it took them to complete the transaction. Carry out a test to see if there is convincing evidence that the express lane is faster.

Time in express lane (seconds)	Time in regular lane (seconds)
337	342
226	472
502	456
408	529
151	181
284	339
150	229
357	263
349	332
257	352
321	341
383	397
565	694
363	324
85	127

Read 592–593
What is the difference between statistical and practical significance?

What is the problem of multiple tests?

Suppose that 20 significance tests were conducted and in each case the null hypothesis was true. What is the probability that we avoid a Type I error in all 20 tests?

HW page 588 (85–93 odd, 95–102)