MAT1372, Classwork24, Fall2025

7.1 One-sample means with t-distribution (Conti.)

12. **Hypothesis Testing For a Sigle Mean:** Once you've determined a one-mean hypothesis test is the correct procedure, there are four steps to completing the test:

choose find

Prepare. Identify parameter of interest, build hypothesis, α, $\bar{x}$, s, n

Check. Verify conditions to ensure $\bar{x}$ is nearly normal

Calculate. If the conditions hold, compute SE, compute T-score, find p-value

Conclude. Evaluate the hypothesis test by comparing p-value to α.

13. US runner getting faster or slower over time? (Example of Hypothesis testing for a single mean)

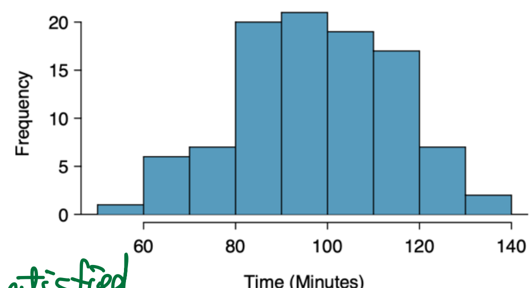
The average time for all runners who finished the Cherry Blossom Race in 2006 was 93.29 minutes (93 minutes 17 seconds). We want to determine using data from 100 participants in the 2017 Cherry Blossom Race whether runners in this race are getting faster or slower, versus the other possibility that there has been no change.

(a) What are appropriate hypotheses for this context?

H_0: The average time was the same for 2006 and 2017 $\mu = 93.29$

H_A: The average time for 2017 was different than that of 2006, $\mu \neq 93.29$

(b) The data come from a simple random sample of all participants, so the observations are independent. However, should we be worried about the normality condition?



$n = 100 > 30$, we should be concerned if there are particularly extrem outlier. The histogram of the data does not show any outlier. With both conditions satisfied we can proceed with a hypothesis test using t-distri.

(c) The sample mean and sample standard deviation of the sample of 100 runners from the 2017 Race are

μ 97.32 and s 16.98 minutes, respectively. Find the test statistic and p-value. What is your conclusion?

null value $(\mu_0) = 93.29$. , we choose $\alpha = 0.05$

To compute T-Score, we need $SE = \frac{\sigma}{\sqrt{n}} \approx \frac{s}{\sqrt{100}} = \frac{16.98}{\sqrt{100}} = 1.698$

Then the T-Score of 97.32: $T = \frac{\mu - \mu_0}{SE} = \frac{97.32 - 93.29}{1.698} = 2.37$

For $df = 100 - 1 = 99$, p-value = area of 2 tails = 0.02

p-value $< \alpha = 0.05 \Rightarrow$ reject H_0

7.2 Paired Data

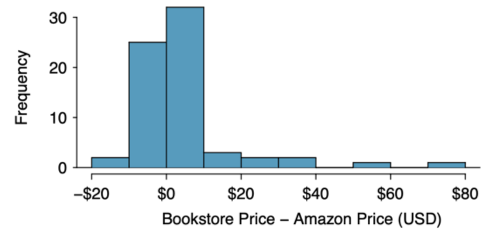
1. Paired Date.

Two sets of observations are paired if each observation in one set has a special correspondence or connection with exactly one observation in the other data set.

2. Example of Paired Data.

In the textbook's market, Amazon prices were, on average, lower than those of campus bookstores

	subject	course_number	bookstore	amazon	price_difference
1	American Indian Studies	M10	47.97	47.45	0.52
2	Anthropology	2	14.26	13.55	0.71
3	Arts and Architecture	10	13.50	12.53	0.97
⋮	⋮	⋮	⋮	⋮	⋮
68	Jewish Studies	M10	35.96	32.40	3.56



To analyze a paired data set, we simply analyze the differences by using the t- distribution techniques.

(a) Set up a hypothesis test to determine whether, on average, there is a difference between Amazon's prices for a book and the campus bookstore's prices. $\mu_B - \mu_A$

n_{diff}	$\bar{x}_{diff}$	s_{diff}
68	3.58	13.42

$\alpha = 0.05$

H_0: The prices are the same from 2 resources $\mu_{diff} = 0$

H_A: The prizes are difference, $\mu_{diff} \neq 0$

(b) Check the conditions for whether we can move forward with the test using the t-distribution.

Independence. The observations are based on a simple random sample, Yes. It is independent

Normality. $n = 68 > 30$, and none of the outliers are particularly extreme, so $\bar{x}_{diff}$ is nearly normal.

(c) Find the test statistic and p-value. What is your conclusion?

To compute the test compute the standard error associated with $\bar{x}_{\text{diff}}$ using the standard deviation of the differences (s_{diff}= 13.42) and the number of differences (n_{diff} = 68):

$$SE_{\bar{x}_{\text{diff}}} = \frac{s_{\text{diff}}}{\sqrt{n_{\text{diff}}}} = \frac{13.42}{\sqrt{68}} = 1.627$$

The test statistic is the T-score of $\bar{x}_{\text{diff}}$ under the null condition that the actual mean difference were 0:

$$T = \frac{\bar{x}_{\text{diff}} - 0}{\text{SE}_{\bar{x}_{\text{diff}}}} = \frac{3.58 - 0}{1.627} = 2.2$$

2% ~ 5%

with df = $n-1 = 67$. Using statistical software, we find the one-tail area of 0.0156. Doubling this area gives the p-value: 0.0312. Because the p-value is less than $\alpha = 0.05$, we reject the null hypothesis. (Amazon prices are, on average, lower than the campus bookstore prices.)