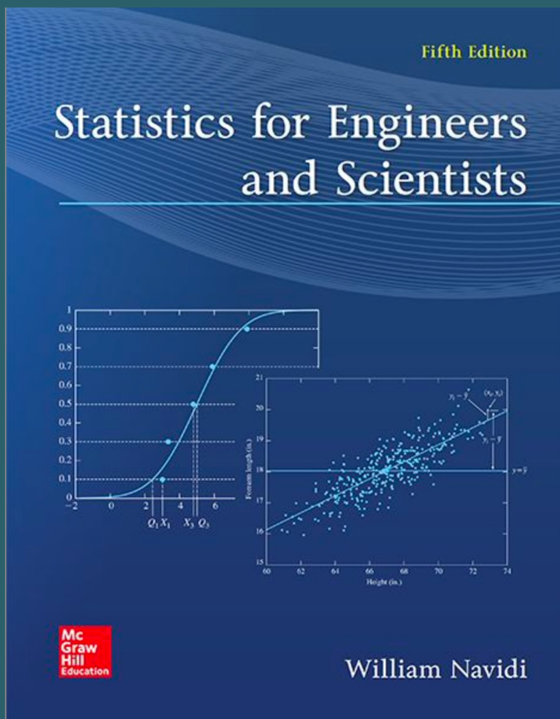


CHAPTER 6

Hypothesis Testing (part 3)

Ch. 6 - Hypothesis Testing

6-1 Large-Sample Tests for a Population Mean

6-2 Drawing Conclusions from the Results of Hypothesis Tests

6-3 Tests for a Population Proportion

6-4 Small-Sample Tests for a Population Mean

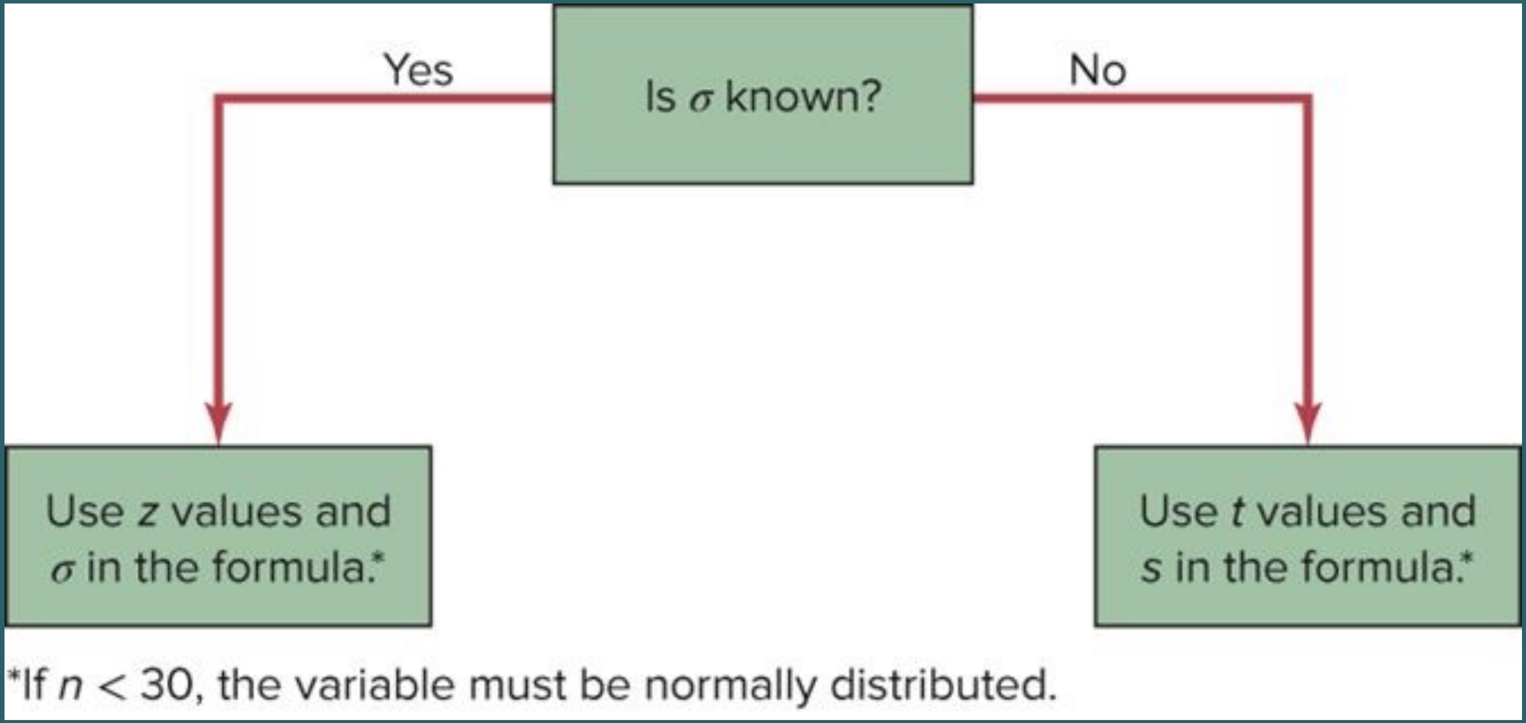
6-5 Large-Sample Tests for the Difference Between Two Means

6-7 Small-Sample Tests for the Difference Between Two Means

6-8 Tests with Paired Data

6-13 Power

Using *z* Test or *t* Test?



When *n* is large (generally **n>30**), *t*-distribution is very close to *z*-distribution, so it is acceptable to use z-distribution, even if σ is unknown.

Example: Using *z* Test or *t* Test?

In 1992, a study was done to determine whether 98.6 degrees Fahrenheit is an accurate mean body temperature for adults. The measures for a sample of 130 subjects, both male and female, are listed below.

$$\bar{X} = 98.25°$$

$$s = 0.73°$$

Test, at $\alpha = 0.05$ level of significance, the claim that the actual adult body temperature is 98.6 degrees Fahrenheit.

Example: Using *z* Test or *t* Test?

<u>Step 1</u>

State the hypotheses and identify the claim.

The claim is that the actual mean adult body temperature is 98.6^0 Fahrenheit.

So, this will be a two-tailed test.

Remember that the null hypothesis always contains the equal sign. The claim in this case is the null hypothesis.

This is a two-tailed test. The alternative hypothesis will be that the mean is not equal to 98.6^0 Fahrenheit.

H_o: $\mu = 98.6$ (claim)

H_1: $\mu \neq 98.6$

Example: Using *z* Test or *t* Test?

Step 2

Compute the test value

The population standard deviation is unknown. We will use the t-distribution.

$$\boldsymbol{t} = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} = \frac{98.25 - 98.6}{\frac{0.73}{\sqrt{130}}} = -5.47$$

Example: Using *z* Test or *t* Test?

<u>Step 3</u>

Find the P-value

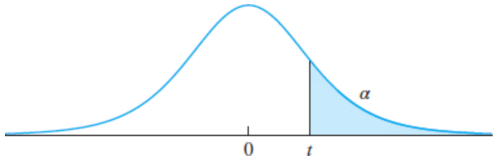
Our sample size is 130. So, the number of degrees of freedom would be 129.

Table A.3 does not contain a row for 129 degrees of freedom. We will go to the next smallest listed value which would be 120.

For two-tailed test, the P-value < 2*0.0005

So, the **P-value < 0.001** (using Table A.3)

TABLE A.3 Upper percentage points for the Student's *t* distribution



	α								
v	**0.40**	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**	**0.001**	**0.0005**
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	318.309	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.385	3.646
35	0.255	0.682	1.306	1.690	2.030	2.438	2.724	3.340	3.591
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.090	3.291

Example: Using *z* Test or *t* Test?

<u>Step 4</u>

Make the decision to reject or not reject the null hypothesis

P-value < 0.001 < $\alpha = 0.05$

Thus, our decision should be **to reject the null hypothesis**.

Example: Using *z* Test or *t* Test?

Step 5

Summarize the results

There is enough evidence to reject the claim (H_0) that the mean adult body temperature is 98.6 degrees Fahrenheit.

Example: Using *z* Test or *t* Test?

Alternative solution:

Step 2

Compute the test value

The sample is large (n=130). Thus, we can use the z-distribution and approximate σ with s.

$$\boldsymbol{z} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{98.25 - 98.6}{\frac{0.73}{\sqrt{130}}} = -5.47$$

Example: Using z Test or t Test?

<u>Step 3</u>

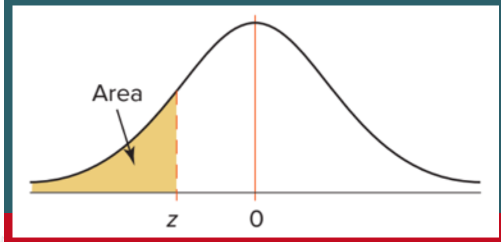
Find the P-value

Table A.2 does not contain $z = -5.47$

Thus, we can only conclude that the
P-value < 2*0.003 = 0.006

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641



Example: Using *z* Test or *t* Test?

Step 4

Make the decision to reject or not reject the null hypothesis

P-value < 0.006 < $\alpha = 0.05$

Thus, our decision should be **to reject the null hypothesis**.

Example: Using z Test or t Test?

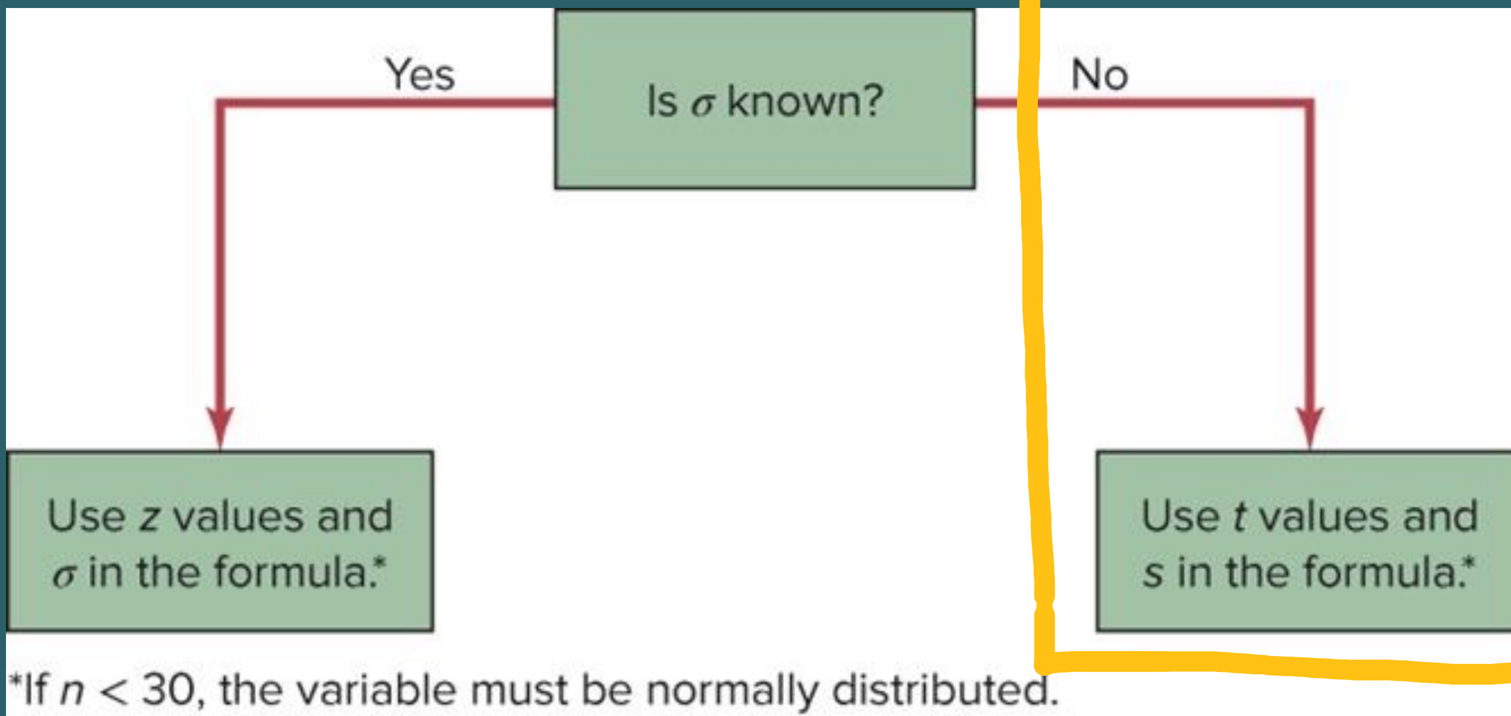
Step 5

Summarize the results

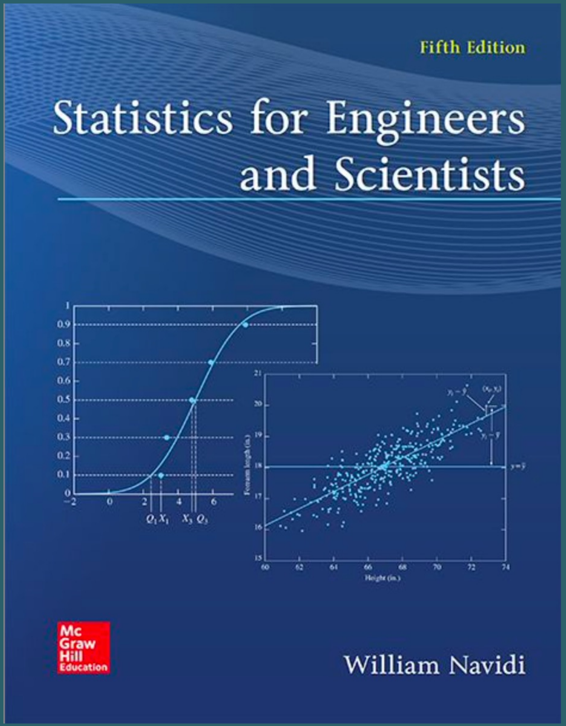
There is enough evidence to reject the claim (H_0) that the mean adult body temperature is 98.6 degrees Fahrenheit.

Using *z* Test or *t* Test?

PREFFERED!



When *n* is large (generally **n>30**), *t*-distribution is very close to *z*-distribution, so it is acceptable to use z-distribution, even if σ is unknown.



6-5 Large-Sample Tests for the Difference Between Two Means

Introduction

- With the *z*- and *t*-tests a sample mean can be compared to a specific population mean to determine whether the null hypothesis (H_0) should be rejected.
- There are, however, many instances when researchers wish to compare two sample means, using experimental and control groups.
- For example, *two different brands of fertilizer might be tested to see whether one is better than the other for growing plants*.

Introduction…

- In the comparison of two means, the same basic steps for hypothesis testing shown in the previous lectures with the *z* or *t* test are used.

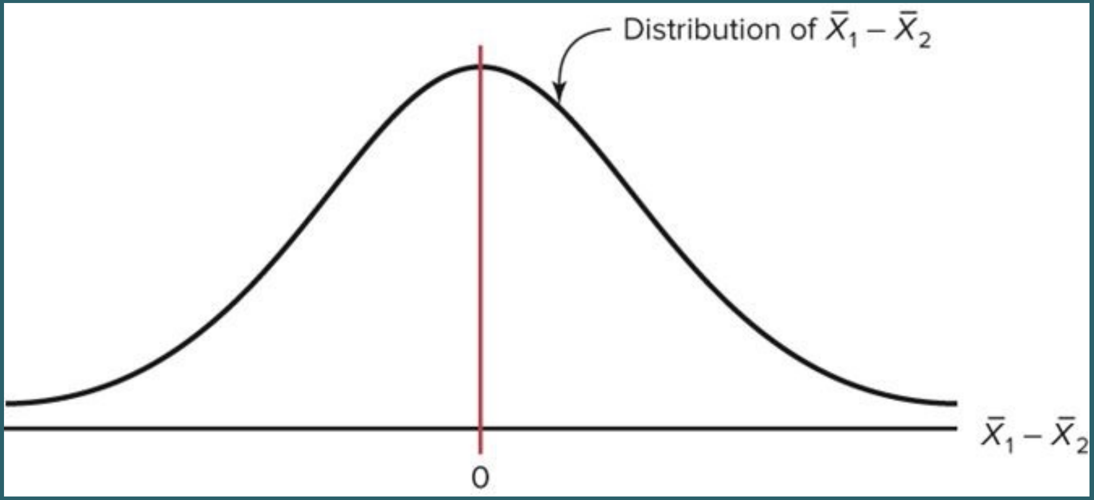
- When comparing two means by using the *t* test, the researcher must decide if the two samples are *independent* or *dependent*.

- Typical hypotheses

$$\boxed{\begin{aligned} H_0&: \mu_1 = \mu_2 \\ H_1&: \mu_1 \neq \mu_2 \end{aligned}} \quad \text{or} \quad \begin{aligned} H_0&: \mu_1 - \mu_2 = 0 \\ H_1&: \mu_1 - \mu_2 \neq 0 \end{aligned}$$

Introduction…

- The theory behind testing the difference between two means is based on selecting pairs of samples and comparing the means of the pairs.
- The population means do not need to be known.



Differences of Means of Pairs of Samples

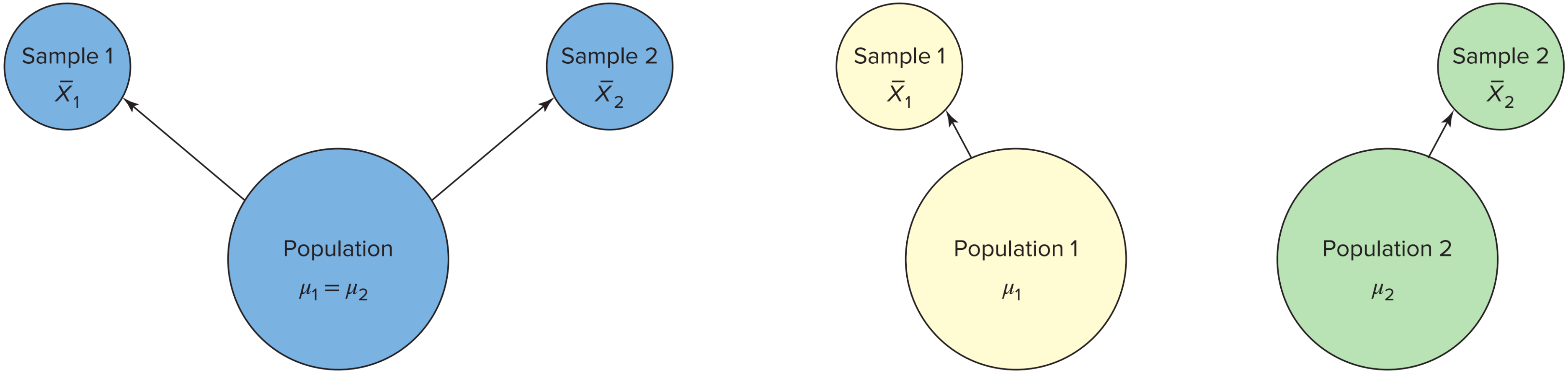
z Test Between Two Means

Formula for the z Test for Comparing Two Means from Independent Populations

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

$$\text{Test value} = \frac{(\text{observed value}) - (\text{expected value})}{\text{standard error}}$$

Situations in the Comparison of Means



(a) Difference is not significant. The means of the populations are the same.

Do not reject H_0: $\mu_1 = \mu_2$ since $\bar{X}_1 - \bar{X}_2$ is not significant.

(b) Difference is significant. The means of the populations are different.

Reject H_0: $\mu_1 = \mu_2$ since $\bar{X}_1 - \bar{X}_2$ is significant.

These tests can also be one-tailed, using the following hypotheses:

Right-tailed			Left-tailed		
H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 > \mu_2$	or	H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 > 0$	H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 < \mu_2$	or	H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 < 0$

z Test Between Two Means (p.435)

Summary

Let $X_1, \ldots, X_{n_X}$ and $Y_1, \ldots, Y_{n_Y}$ be *large* (e.g., $n_X > 30$ and $n_Y > 30$) samples from populations with means μ_X and μ_Y and standard deviations σ_X and σ_Y, respectively. Assume the samples are drawn independently of each other.

To test a null hypothesis of the form $H_0\colon \mu_X - \mu_Y \le \Delta_0$, $H_0\colon \mu_X - \mu_Y \ge \Delta_0$, or $H_0\colon \mu_X - \mu_Y = \Delta_0$:

- Compute the z-score: $z = \dfrac{(\overline{X} - \overline{Y}) - \Delta_0}{\sqrt{\sigma_X^2/n_X + \sigma_Y^2/n_Y}}$. If σ_X and σ_Y are unknown they may be approximated with s_X and s_Y, respectively.
- Compute the P-value. The P-value is an area under the normal curve, which depends on the alternate hypothesis as follows:

Alternate Hypothesis	*P*-value
$H_1\colon \mu_X - \mu_Y > \Delta_0$	Area to the right of z
$H_1\colon \mu_X - \mu_Y < \Delta_0$	Area to the left of z
$H_1\colon \mu_X - \mu_Y \ne \Delta_0$	Sum of the areas in the tails cut off by z and $-z$

Page 435

z Test Between Two Means

The P-value method for hypothesis testing between two means follows the same format as stated in previous sections:

- **Step 1** State the hypotheses and identify the claim.
- **Step 2** Compute the test value.
- **Step 3** Find the P-value.
- **Step 4** Make the decision.
- **Step 5** Summarize the results.

Example: Leisure Time

A study using two random samples of 35 people each found that the average amount of time those in the age group of 26–35 years spent per week on leisure activities was 39.6 hours, and those in the age group of 46–55 years spent 35.4 hours. Assume that the population standard deviation for those in the first age group found by previous studies is 6.3 hours, and the population standard deviation of those in the second group found by previous studies was 5.8 hours.

At $\alpha = 0.05$, can it be concluded that there is a significant difference in the average times each group spends on leisure activities?

Example: Leisure Time…

Step 1: State the hypotheses and identify the claim

H_o: $\mu_1 = \mu_2$ **H_1: $\mu_1 \neq \mu_2$ (claim)**

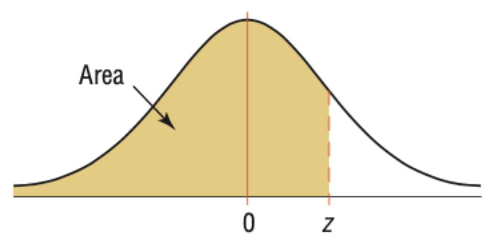
Step 2: Compute the test value

$$z = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(39.6 - 35.4) - 0}{\sqrt{\frac{6.3^2}{35} + \frac{5.8^2}{35}}} = \frac{4.2}{1.447} = 2.90$$

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

For z values greater than 3.49, use 0.9999.



Example: Leisure Time…

Step 3: Find the P-value

- The test value for a two-tailed test is 2.90, then the **P-value** obtained from the table is **0.0038**.
- This value is obtained by looking up the area for z = 2.90, which is **0.9981**.
- Then 0.9981 is subtracted from 1.0000 to get 0.0019. Finally, **this value is doubled** to get 0.0038 since the test is two-tailed.

Example: Leisure Time…

Step 4: Make the decision

- At $\alpha = 0.05$, the decision is to reject the null hypothesis, since P-value < α (that is, $0.0038 < 0.05$)

Step 5: Summarize the results

- There is enough evidence to support the claim that the means are not equal.

Example 6.9

The article "Effect of Welding Procedure on Flux Cored Steel Wire Deposits" (N. Ramini de Rissone, I. de S. Bott, et al., *Science and Technology of Welding and Joining*, 2003:113–122) compares properties of welds made using carbon dioxide as a shielding gas with those of welds made using a mixture of argon and carbon dioxide. One property studied was the diameter of inclusions, which are particles embedded in the weld. A sample of 544 inclusions in welds made using argon shielding averaged 0.37 μm in diameter, with a standard deviation of 0.25 μm. A sample of 581 inclusions in welds made using carbon dioxide shielding averaged 0.40 μm in diameter, with a standard deviation of 0.26 μm. (Standard deviations were estimated from a graph.) Can you conclude that the mean diameters of inclusions differ between the two shielding gases?

p.433

Example 6.9 - SOLUTION

$\bar{X} \quad = 0.37\ \mu\text{m}$ $\qquad$ $\bar{Y} \quad = 0.40\ \mu\text{m}$

$s_x = 0.25\ \mu\text{m}$ $\qquad$ $s_y = 0.26\ \mu\text{m}$

$n_x = 544$ $\qquad$ $n_x = 581$

<u>Hypotheses:</u>

$\mathrm{H_0}: \mu_\mathrm{x} - \mu_\mathrm{y} = 0$ $\qquad$ $\mathrm{H_1}: \mu_\mathrm{x} - \mu_\mathrm{y} \neq 0$ (claim)

Example 6.9 – SOLUTION…

$$z = \frac{(\bar{X} - \bar{Y}) - (\mu_x - \mu_y)}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$$

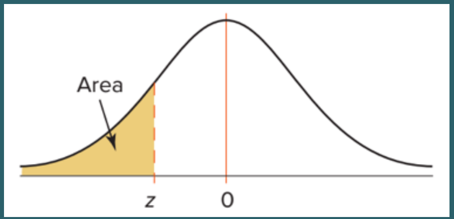
$$z = \frac{(0.37 - 0.40) - (0)}{\sqrt{\frac{0.25^2}{544} + \frac{0.26^2}{581}}}$$

$$\mathbf{z = -1.97}$$

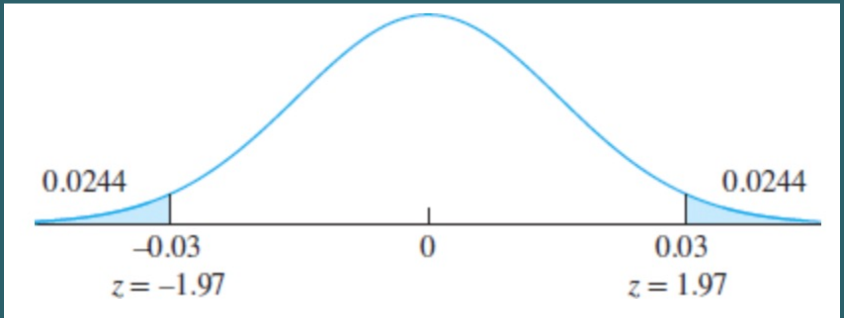
Example 6.9 – SOLUTION…

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559



Two-tailed test: The P-value = 0.0244*2 = 0.0488



Example 6.9 – SOLUTION…

- A follower of the 5% rule (that is $\alpha = 0.05$) would reject the null hypothesis.
- It is certainly reasonable to be skeptical about the truth of H_0.

- It can be concluded that the mean diameters of inclusions differ between the two shielding gases.

z Test Between Two Means…

- Sometimes, we are interested in testing a specific difference in means other than zero
- The methods described in this section can be used to test a hypothesis that two population means differ by a specified constant.

Example 6.10

Refer to Example 6.9. Can you conclude that the mean diameter for carbon dioxide welds (μ_Y) exceeds that for argon welds (μ_X) by more than 0.015 μm?

Solution

The null and alternate hypotheses are

$$H_0: \mu_X - \mu_Y \geq -0.015 \quad \text{versus} \quad H_1: \mu_X - \mu_Y < -0.015$$

We observe $\overline{X} = 0.37$, $\overline{Y} = 0.40$, $s_X = 0.25$, $s_Y = 0.26$, $n_X = 544$, and $n_Y = 581$. Under H_0, we take $\mu_X - \mu_Y = -0.015$. The null distribution of $\overline{X} - \overline{Y}$ is given by expression (6.2) to be

$$\overline{X} - \overline{Y} \sim N(-0.015,\ 0.01521^2)$$

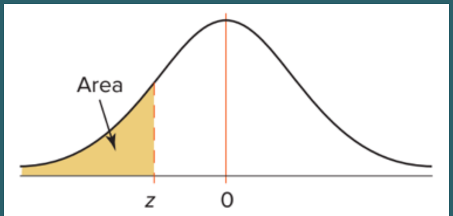
We observe $\overline{X} - \overline{Y} = 0.37 - 0.40 = -0.03$. The z-score is

$$\boxed{z = \frac{-0.03 - (-0.015)}{0.01521} = -0.99}$$

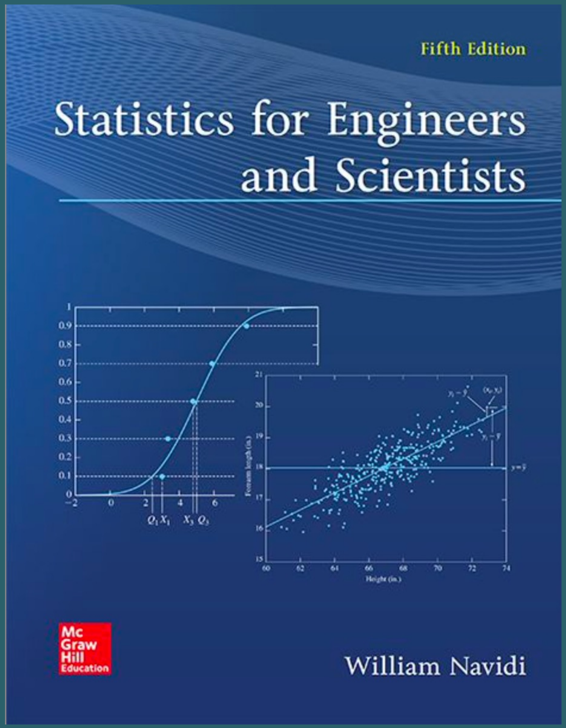
This is a one-tailed test. The P-value is 0.1611. We cannot conclude that the mean diameter of inclusions from carbon dioxide welds exceeds that of argon welds by more than 0.015 μm.

Cumulative Standard Normal Distribution

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148



One-tailed test: The P-value = 0.1611



6-7 Small-Sample Tests for the Difference Between Two Means

Introduction

- In many situations the conditions for using *z* test cannot be met—in particular, the population standard deviations are not known or when samples are small.
- In these cases, a *t* test can be used to test the difference between means when the two samples are independent and when the samples are taken from two normally or approximately normally distributed populations.
- Samples are independent samples when they are not related.

t test for for the Difference Between Two Means

Formula for the t Test for Testing the Difference Between Two Means, Independent Samples

Variances are assumed to be unequal:

$$\boxed{t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}}$$

where the degrees of freedom are equal to the smaller of $n_1 - 1$ or $n_2 - 1$.

Example: Work Absences

A study was done to see if there is a difference between the number of sick days men take and the number of sick days women take.

A random sample of 9 men found that the mean of the number of sick days taken was 5.5. The standard deviation of the sample was 1.23.

A random sample of 7 women found that the mean was 4.3 days and a standard deviation of 1.19 days.

At **$\alpha = 0.05$**, can it be concluded that there is a difference in the means?

Example: Work Absences…

1. Hypothesis and claim

$H_0\colon \mu_1 = \mu_2$ and $H_1\colon \mu_1 \neq \mu_2$ (claim)

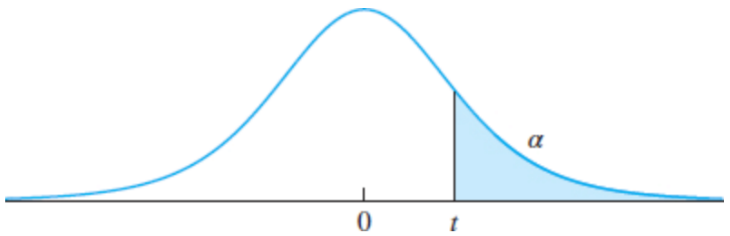
2. The test value

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{(5.5 - 4.3) - 0}{\sqrt{\dfrac{1.23^2}{9} + \dfrac{1.19^2}{7}}} = 1.972$$

Example: Work Absences…

3. The P-value

TABLE A.3 Upper percentage points for the Student's *t* distribution



v	α = 0.40	0.25	0.10	0.05	0.025	0.01	0.005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355

Two-tailed test: The P-value is between 2*0.025 = 0.05 and 2*0.05 = 0.10 (calculator gives 0.0961)

Example: Work Absences…

4. Decision

The P-value > α = 0.05, thus we do not reject H_0

5. Conclusion

There is not enough evidence to support the claim that the mean are different.

Calculation of Degrees of Freedom

- In many statistical software packages, a different method is used to compute the degrees of freedom for this t test.
- They are determined by the formula

$$\text{d.f.} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1 - 1) + (s_2^2/n_2)^2/(n_2 - 1)}$$

- This method is also used in your textbook, but it is **NOT REQUIRED** in this course.

Example 6.12

Good website design can make Web navigation easier. The article "The Implications of Visualization Ability and Structure Preview Design for Web Information Search Tasks" (H. Zhang and G. Salvendy, International Journal of Human-Computer Interaction, 2001:75–95) presents a comparison of item recognition between two designs.

- A sample of 10 users using a conventional Web design averaged 32.3 items identified, with a standard deviation of 8.56.
- A sample of 10 users using a new structured Web design averaged 44.1 items identified, with a standard deviation of 10.09.

Can we conclude that the mean number of items identified is greater with the new structured design?

Solution

Let $\overline{X} = 44.1$ be the sample mean for the structured Web design. Then $s_X = 10.09$ and $n_X = 10$. Let $\overline{Y} = 32.3$ be the sample mean for the conventional Web design. Then $s_Y = 8.56$ and $n_Y = 10$. Let μ_X and μ_Y denote the population mean measurements made by the structured and conventional methods, respectively. The null and alternate hypotheses are

$$\boxed{H_0: \mu_X - \mu_Y \leq 0 \quad \text{versus} \quad H_1: \mu_X - \mu_Y > 0 \text{ (c)}}$$

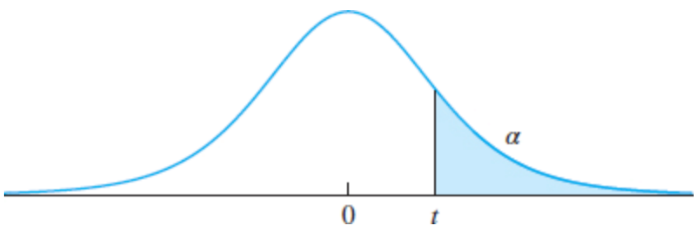
The test statistic is

$$t = \frac{(\overline{X} - \overline{Y}) - 0}{\sqrt{s_X^2/n_X + s_Y^2/n_Y}}$$

Substituting values for $\overline{X}, \overline{Y}, s_X, s_Y, n_X$, and n_Y, we compute the value of the test statistic to be $\boxed{t = 2.820.}$ Under H_0, this statistic has an approximate Student's t distribution, with the number of degrees of freedom given by

$$\boxed{\nu = \frac{\left(\frac{10.09^2}{10} + \frac{8.56^2}{10}\right)^2}{\frac{(10.09^2/10)^2}{9} + \frac{(8.56^2/10)^2}{9}} = 17.53 \approx 17}$$

TABLE A.3 Upper percentage points for the Student's *t* distribution



	α						
v	**0.40**	**0.25**	**0.10**	**0.05**	**0.025**	**0.01**	**0.005**
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861

One-tailed test: 0.005 <P-value < 0.01 (calculator = 0.005899)

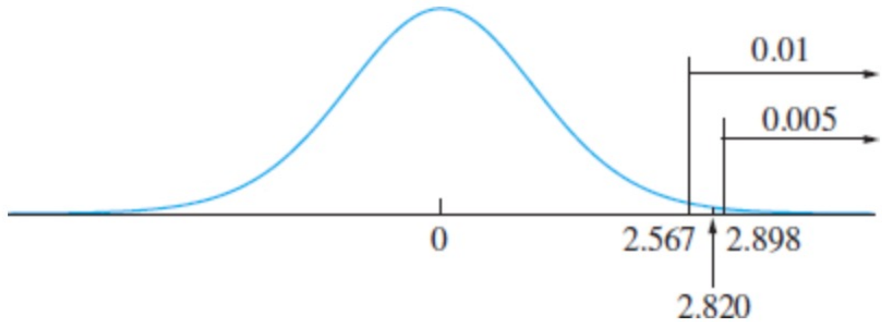
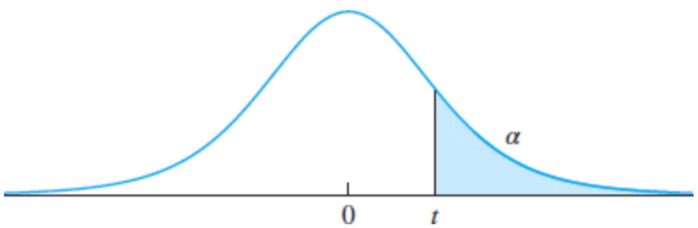


FIGURE 6.14 Solution to Example 6.12. The P-value is the area in the right-hand tail, which is between 0.005 and 0.01.

Conclusion: There is strong evidence that the mean number of items identified is greater for the new design.

TABLE A.3 Upper percentage points for the Student's *t* distribution



Simplified method

v	0.40	0.25	0.10	0.05	0.025	0.01	0.005
	α						
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861

One-tailed test: 0.01 <P-value < 0.025 (calculator = 0.010024)