Design of Engineering Experiments
Chapter 2 – Some Basic Statistical Concepts

- Describing sample data
  - Random samples
  - Sample mean, variance, standard deviation
  - Populations versus samples
  - Population mean, variance, standard deviation
  - Estimating parameters
- Simple **comparative** experiments
  - The hypothesis testing framework
  - The two-sample *t*-test
  - Checking assumptions, validity

Portland Cement Formulation (page 26)

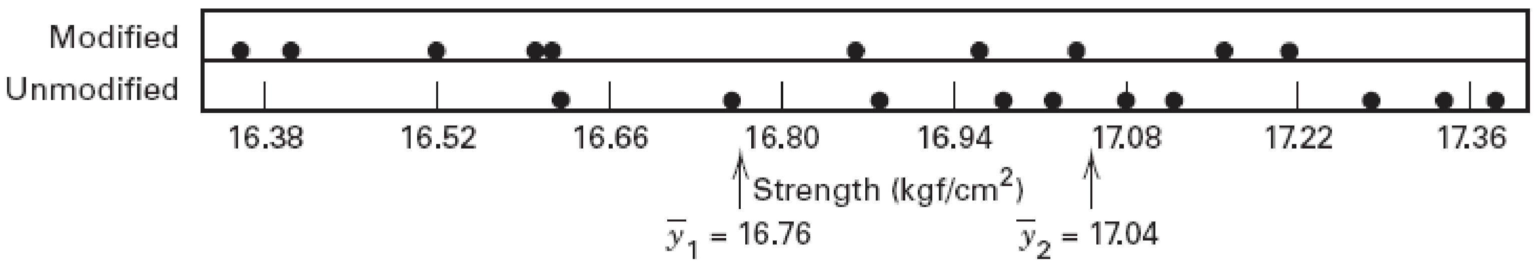
■ **TABLE 2.1**

Tension Bond Strength Data for the Portland Cement Formulation Experiment

j	Modified Mortar y_{1j}	Unmodified Mortar y_{2j}
1	16.85	16.62
2	16.40	16.75
3	17.21	17.37
4	16.35	17.12
5	16.52	16.98
6	17.04	16.87
7	16.96	17.34
8	17.15	17.02
9	16.59	17.08
10	16.57	17.27

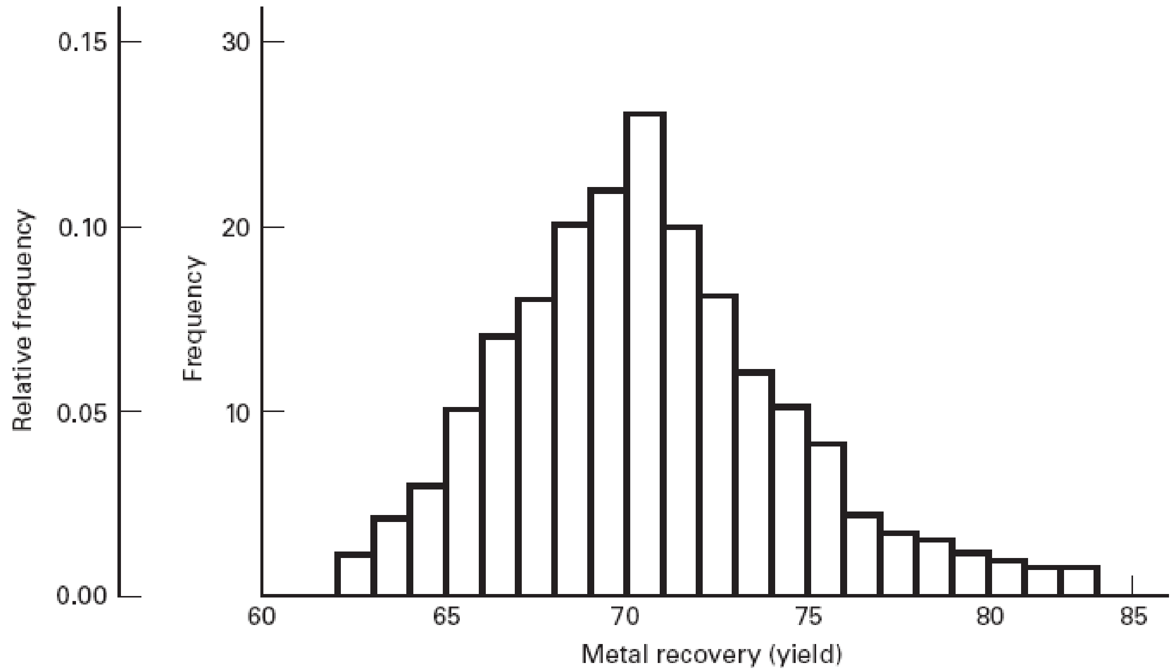
Graphical View of the Data

Dot Diagram, Fig. 2.1, pp. 26



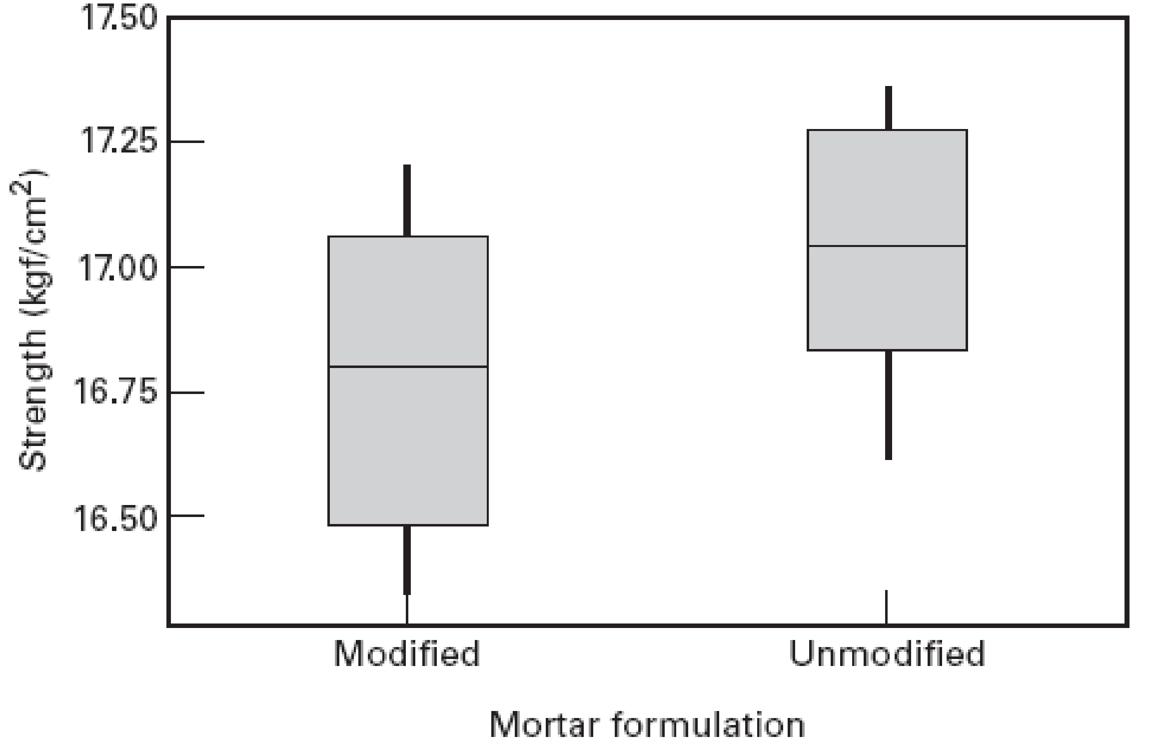
■ **FIGURE 2.1** **Dot diagram for the tension bond strength data in Table 2.1**

If you have a large sample, a histogram may be useful



■ **FIGURE 2.2** **Histogram for 200 observations on metal recovery (yield) from a smelting process**

Box Plots, Fig. 2.3, pp. 28

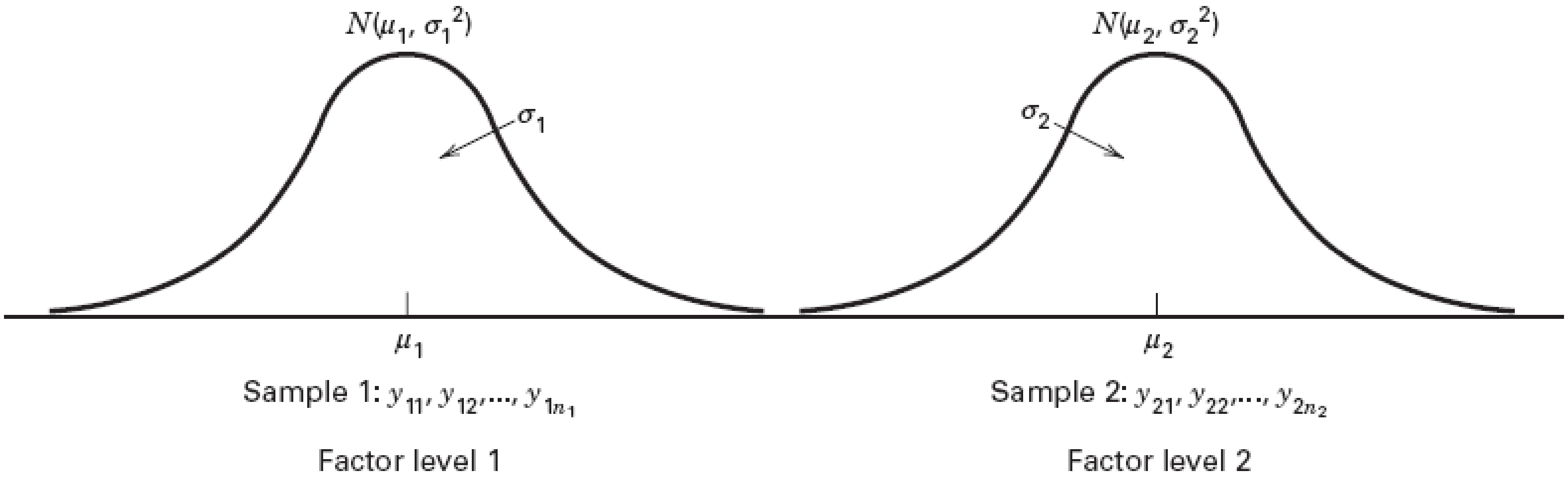


■ **FIGURE 2.3** **Box plots for the Portland cement tension bond strength experiment**

The Hypothesis Testing Framework

- **Statistical hypothesis testing** is a useful framework for many experimental situations
- Origins of the methodology date from the early 1900s
- We will use a procedure known as the **two-sample *t*-test**

The Hypothesis Testing Framework



■ **FIGURE 2.9** The sampling situation for the two-sample *t*-test

- Sampling from a **normal** distribution
- Statistical hypotheses:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Estimation of Parameters

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ estimates the population mean } \mu$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \text{ estimates the variance } \sigma^2$$

Summary Statistics (pg. 38)

Modified Mortar

"New recipe"

$$\bar{y}_1 = 16.76$$

$$S_1^2 = 0.100$$

$$S_1 = 0.316$$

$$n_1 = 10$$

Unmodified Mortar

"Original recipe"

$$\bar{y}_2 = 17.04$$

$$S_2^2 = 0.061$$

$$S_2 = 0.248$$

$$n_2 = 10$$

How the Two-Sample *t*-Test Works:

Use the sample means to draw inferences about the population means

$\bar{y}_1 - \bar{y}_2 = 16.76 - 17.04 = -0.28$

$$\frac{\text{Difference in sample means}}{\text{Standard deviation of the difference in sample means}}$$

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}, \quad \text{and } \sigma_{\bar{y}_1 - \bar{y}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}, \bar{y}_1 \text{ and } \bar{y}_2 \text{ independent}$$

This suggests a statistic:

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

If the variances were known we could use the normal distribution as the basis of a test

Z_0 has a N(0,1) distribution if the two population means are equal

If we knew the two variances how would we use Z_0 to test H_0?

Ho: μ1-μ2=0

H1: μ1-μ2≠0

Suppose that $\sigma_1 = \sigma_2 = 0.30$. Then we can calculate

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{-0.28}{\sqrt{\frac{0.3^2}{10} + \frac{0.3^2}{10}}} = \frac{-0.28}{0.1342} = -2.09$$

How "unusual" is the value $Z_0 = -2.09$ if the two population means are equal?

It turns out that 95% of the area under the standard normal curve (probability) falls between the values $Z_{0.025} = 1.96$ and $-Z_{0.025} = -1.96$.

So the value $Z_0 = -2.09$ is pretty unusual in that it would happen less that 5% of the time if the population means were equal

Standard Normal Table (see appendix)

I Cumulative Standard Normal Distribution[a]

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

z	0.00	0.01	0.02	0.03	0.04	z
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.0
0.1	0.53983	0.54379	0.54776	0.55172	0.55567	0.1
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.2
0.3	0.61791	0.62172	0.62551	0.62930	0.63307	0.3
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.4
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.5
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.6
0.7	0.75803	0.76115	0.76424	0.76730	0.77035	0.7
0.8	0.78814	0.79103	0.79389	0.79673	0.79954	0.8
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.9
1.0	0.84134	0.84375	0.84613	0.84849	0.85083	1.0
1.1	0.86433	0.86650	0.86864	0.87076	0.87285	1.1
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	1.2
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	1.3
1.4	0.91924	0.92073	0.92219	0.92364	0.92506	1.4
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	1.5
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	1.6
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	1.7
1.8	0.96407	0.96485	0.96562	0.96637	0.96711	1.8
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	1.9
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	2.0
2.1	0.98214	0.98257	0.98300	0.98341	0.93882	2.1
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	2.2
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	2.3
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	2.4
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	2.5
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	2.6
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	2.7
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	2.8
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	2.9
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	3.0
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	3.1
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	3.2
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	3.3
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	3.4
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	3.5
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	3.6
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	3.7
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	3.8
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	3.9

[a]Reproduced with permission from *Probability and Statistics in Engineering and Management Science*, 3rd edition, by W. W. Hines and D. C. Montgomery, Wiley, New York, 1990.

I Cumulative Standard Normal Distribution (*Continued*)

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

$Z_{0.025} = 1.96$

z	0.05	0.06	0.07	0.08	0.09	z
0.0	0.51994	0.52392	0.52790	0.53188	0.53586	0.0
0.1	0.55962	0.56356	0.56749	0.57142	0.57534	0.1
0.2	0.59871	0.60257	0.60642	0.61026	0.61409	0.2
0.3	0.63683	0.64058	0.64431	0.64803	0.65173	0.3
0.4	0.67364	0.67724	0.68082	0.68438	0.68793	0.4
0.5	0.70884	0.71226	0.71566	0.71904	0.72240	0.5
0.6	0.74215	0.74537	0.74857	0.75175	0.75490	0.6
0.7	0.77337	0.77637	0.77935	0.78230	0.78523	0.7
0.8	0.80234	0.80510	0.80785	0.81057	0.81327	0.8
0.9	0.82894	0.83147	0.83397	0.83646	0.83891	0.9
1.0	0.85314	0.85543	0.85769	0.85993	0.86214	1.0
1.1	0.87493	0.87697	0.87900	0.88100	0.88297	1.1
1.2	0.89435	0.89616	0.89796	0.89973	0.90147	1.2
1.3	0.91149	0.91308	0.91465	0.91621	0.91773	1.3
1.4	0.92647	0.92785	0.92922	0.93056	0.93189	1.4
1.5	0.93943	0.90462	0.94179	0.94295	0.94408	1.5
1.6	0.95053	0.95154	0.95254	0.95352	0.95448	1.6
1.7	0.95994	0.96080	0.96164	0.96246	0.96327	1.7
1.8	0.96784	0.96856	0.96926	0.96995	0.97062	1.8
1.9	0.97441	0.97500	0.97558	0.97615	0.97670	1.9
2.0	0.97982	0.98030	0.98077	0.98124	0.98169	2.0
2.1	0.98422	0.98461	0.98500	0.98537	0.98574	2.1
2.2	0.98778	0.98809	0.98840	0.98870	0.98899	2.2
2.3	0.99061	0.99086	0.99111	0.99134	0.99158	2.3
2.4	0.99286	0.99305	0.99324	0.99343	0.99361	2.4
2.5	0.99461	0.99477	0.99492	0.99506	0.99520	2.5
2.6	0.99598	0.99609	0.99621	0.99632	0.99643	2.6
2.7	0.99702	0.99711	0.99720	0.99728	0.99736	2.7
2.8	0.99781	0.99788	0.99795	0.99801	0.99807	2.8
2.9	0.99841	0.99846	0.99851	0.99856	0.99861	2.9
3.0	0.99886	0.99889	0.99893	0.99897	0.99900	3.0
3.1	0.99918	0.99921	0.99924	0.99926	0.99929	3.1
3.2	0.99942	0.99944	0.99946	0.99948	0.99950	3.2
3.3	0.99960	0.99961	0.99962	0.99964	0.99965	3.3
3.4	0.99972	0.99973	0.99974	0.99975	0.99976	3.4
3.5	0.99981	0.99981	0.99982	0.99983	0.99983	3.5
3.6	0.99987	0.99987	0.99988	0.99988	0.99989	3.6
3.7	0.99991	0.99992	0.99992	0.99992	0.99992	3.7
3.8	0.99994	0.99994	0.99995	0.99995	0.99995	3.8
3.9	0.99996	0.99996	0.99996	0.99997	0.99997	3.9

So if the variances were known we would conclude that we should reject the null hypothesis at the **5% level of significance**

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

and conclude that the alternative hypothesis is true.

This is called a fixed significance level test, because we compare the value of the test statistic to a **critical value** (1.96) that we selected in advance before running the experiment.

The standard normal distribution is the **reference distribution** for the test.

Another way to do this that is very popular is to use the *P*-value approach. The *P*-value can be thought of as the observed significance level.

For the Z-test it is easy to find the *P*-value.

Normal Table

I Cumulative Standard Normal Distribution (*Continued*)

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

z	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**	*z*
0.0	0.51994	0.52392	0.52790	0.53188	0.53586	0.0
0.1	0.55962	0.56356	0.56749	0.57142	0.57534	0.1
0.2	0.59871	0.60257	0.60642	0.61026	0.61409	0.2
0.3	0.63683	0.64058	0.64431	0.64803	0.65173	0.3
0.4	0.67364	0.67724	0.68082	0.68438	0.68793	0.4
0.5	0.70884	0.71226	0.71566	0.71904	0.72240	0.5
0.6	0.74215	0.74537	0.74857	0.75175	0.75490	0.6
0.7	0.77337	0.77637	0.77935	0.78230	0.78523	0.7
0.8	0.80234	0.80510	0.80785	0.81057	0.81327	0.8
0.9	0.82894	0.83147	0.83397	0.83646	0.83891	0.9
1.0	0.85314	0.85543	0.85769	0.85993	0.86214	1.0
1.1	0.87493	0.87697	0.87900	0.88100	0.88297	1.1
1.2	0.89435	0.89616	0.89796	0.89973	0.90147	1.2
1.3	0.91149	0.91308	0.91465	0.91621	0.91773	1.3
1.4	0.92647	0.92785	0.92922	0.93056	0.93189	1.4
1.5	0.93943	0.90462	0.94179	0.94295	0.94408	1.5
1.6	0.95053	0.95154	0.95254	0.95352	0.95448	1.6
1.7	0.95994	0.96080	0.96164	0.96246	0.96327	1.7
1.8	0.96784	0.96856	0.96926	0.96995	0.97062	1.8
1.9	0.97441	0.97500	0.97558	0.97615	0.97670	1.9
2.0	0.97982	0.98030	0.98077	0.98124	0.98169	2.0
2.1	0.98422	0.98461	0.98500	0.98537	0.98574	2.1
2.2	0.98778	0.98809	0.98840	0.98870	0.98899	2.2
2.3	0.99061	0.99086	0.99111	0.99134	0.99158	2.3
2.4	0.99286	0.99305	0.99324	0.99343	0.99361	2.4
2.5	0.99461	0.99477	0.99492	0.99506	0.99520	2.5
2.6	0.99598	0.99609	0.99621	0.99632	0.99643	2.6
2.7	0.99702	0.99711	0.99720	0.99728	0.99736	2.7
2.8	0.99781	0.99788	0.99795	0.99801	0.99807	2.8
2.9	0.99841	0.99846	0.99851	0.99856	0.99861	2.9
3.0	0.99886	0.99889	0.99893	0.99897	0.99900	3.0
3.1	0.99918	0.99921	0.99924	0.99926	0.99929	3.1
3.2	0.99942	0.99944	0.99946	0.99948	0.99950	3.2
3.3	0.99960	0.99961	0.99962	0.99964	0.99965	3.3
3.4	0.99972	0.99973	0.99974	0.99975	0.99976	3.4
3.5	0.99981	0.99981	0.99982	0.99983	0.99983	3.5
3.6	0.99987	0.99987	0.99988	0.99988	0.99989	3.6
3.7	0.99991	0.99992	0.99992	0.99992	0.99992	3.7
3.8	0.99994	0.99994	0.99995	0.99995	0.99995	3.8
3.9	0.99996	0.99996	0.99996	0.99997	0.99997	3.9

Find the probability above $Z_0 = -2.09$ from the table.

This is 1 – 0.98169 = 0.01832

The *P*-value is twice this probability, or 0.03662.

So we would reject the null hypothesis at any level of significance that is larger than or equal to 0.03662.

Typically 0.05 is used as the cutoff.

The *t* -Test

- The Z-test just described would work perfectly if we knew the two population variances.
- Since we usually don't know the true population variances, what would happen if we just plugged in the sample variances?
- The answer is that if the sample sizes were large enough (say both $n > 30$ or 40) the Z-test would work just fine. It is a good **large-sample test** for the difference in means.
- But many times that isn't possible.
- It turns out that if the sample size is small we can no longer use the N(0,1) distribution as the reference distribution for the test.

How the Two-Sample *t*-Test Works:

Use S_1^2 and S_2^2 to estimate σ_1^2 and σ_2^2

The previous ratio becomes $\dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

However, we have the case where $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Pool the individual sample variances:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

How the Two-Sample *t*-Test Works:

The test statistic is

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- Values of t_0 that are near zero are consistent with the null hypothesis
- Values of t_0 that are very different from zero are consistent with the alternative hypothesis
- t_0 is a "distance" measure-how far apart the averages are expressed in standard deviation units
- Notice the interpretation of t_0 as a **signal-to-noise** ratio

The Two-Sample (Pooled) *t*-Test

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} = \frac{9(0.100) + 9(0.061)}{10 + 10 - 2} = 0.081$$

$$S_p = 0.284$$

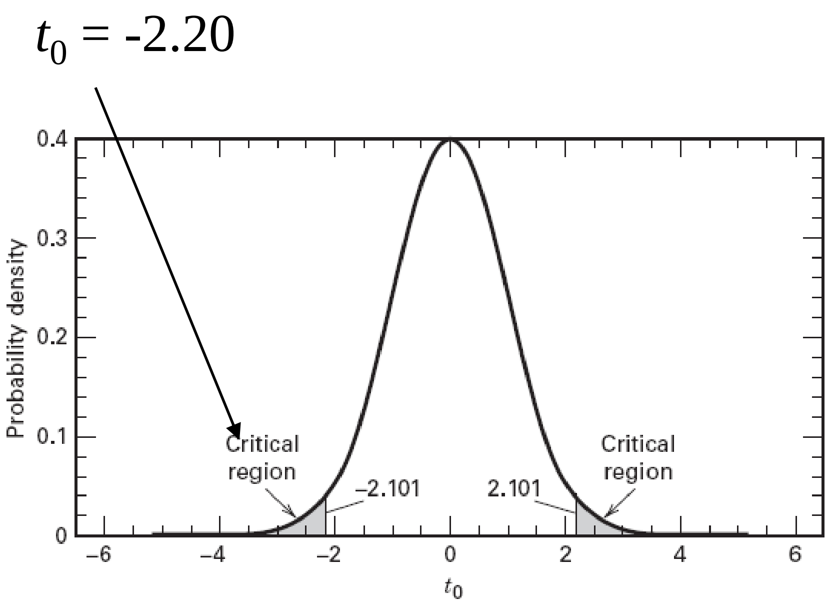
$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{16.76 - 17.04}{0.284\sqrt{\frac{1}{10} + \frac{1}{10}}} = -2.20$$

The two sample means are a little over two standard deviations apart

Is this a "large" difference?

The Two-Sample (Pooled) *t*-Test

- We need an **objective** basis for deciding how large the test statistic t_0 really is
- In 1908, W. S. Gosset derived the **reference distribution** for t_0 … called the *t* distribution
- Tables of the *t* distribution – see textbook appendix page 614



■ FIGURE 2.10 The *t* distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$

II Percentage Points of the t Distribution[a]

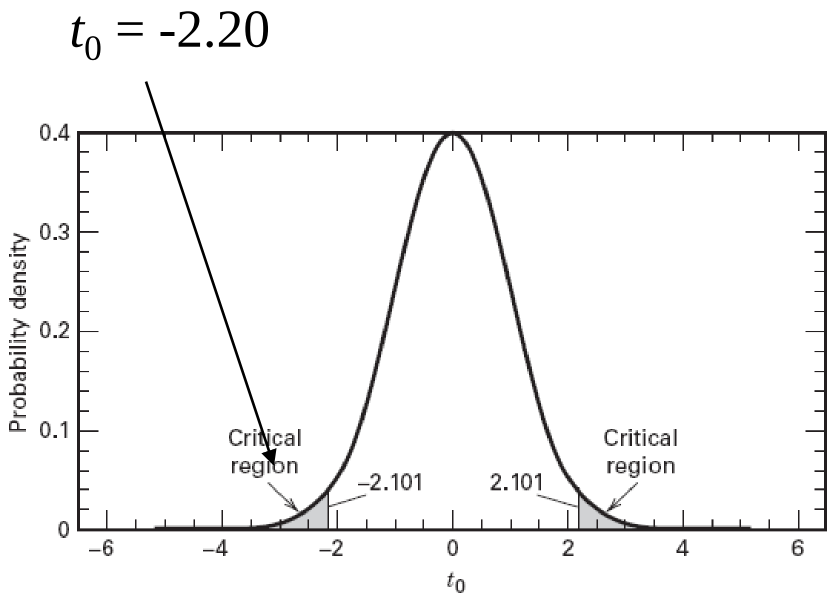
ν \ α	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0025	0.001	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	14.089	23.326	31.598
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	7.453	10.213	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	0.265	0.727	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	4.019	4.785	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	3.326	3.787	4.140
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.038	3.396	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	0.254	0.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	2.860	3.160	3.373
∞	0.253	0.674	1.282	1.645	1.960	2.326	2.576	2.807	3.090	3.291

ν = Degrees of freedom.

[a]Adapted with permission from *Biometrika Tables for Statisticians*, Vol. 1, 3rd edition, by E. S. Pearson and H. O. Hartley, Cambridge University Press, Cambridge, 1966.

The Two-Sample (Pooled) *t*-Test

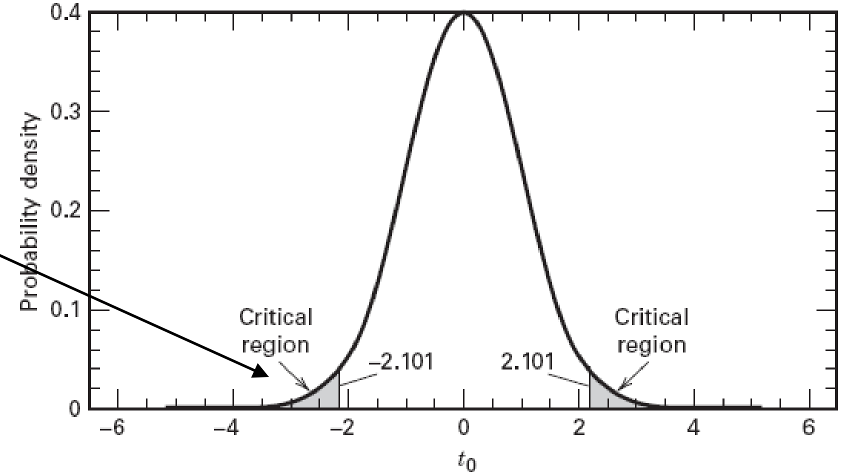
- A value of t_0 between –2.101 and 2.101 is consistent with equality of means
- It is possible for the means to be equal and t_0 to exceed either 2.101 or –2.101, leads to the conclusion that the means are different
- Could also use the ***P*-value** approach



■ FIGURE 2.10 The t distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$

The Two-Sample (Pooled) *t*-Test

$t_0 = -2.20$



■ **FIGURE 2.10** **The *t* distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$**

- The ***P*-value** is the area (probability) in the tails of the *t*-distribution beyond -2.20 + the probability beyond +2.20 (it's a two-sided test)
- The *P*-value is a measure of how unusual the value of the test statistic is given that the null hypothesis is true
- The *P*-value the risk of **wrongly rejecting** the null hypothesis of equal means (it measures rareness of the event)
- The exact *P*-value in our problem is $P = 0.042$ (found from a computer)

Approximating the *P*-value

Our *t*-table only gives probabilities greater than positive values of *t*. So take the absolute value of t_0 = -2.20 or $|t_0|$= 2.20.

Now with 18 degrees of freedom, find the values of *t* in the table that bracket this value.

These are 2.101 < $|t_0|$= 2.20 < 2.552. The right-tail probability for *t* = 2.101 is 0.025 and for *t* = 2.552 is 0.01. Double these probabilities because this is a two-sided test.

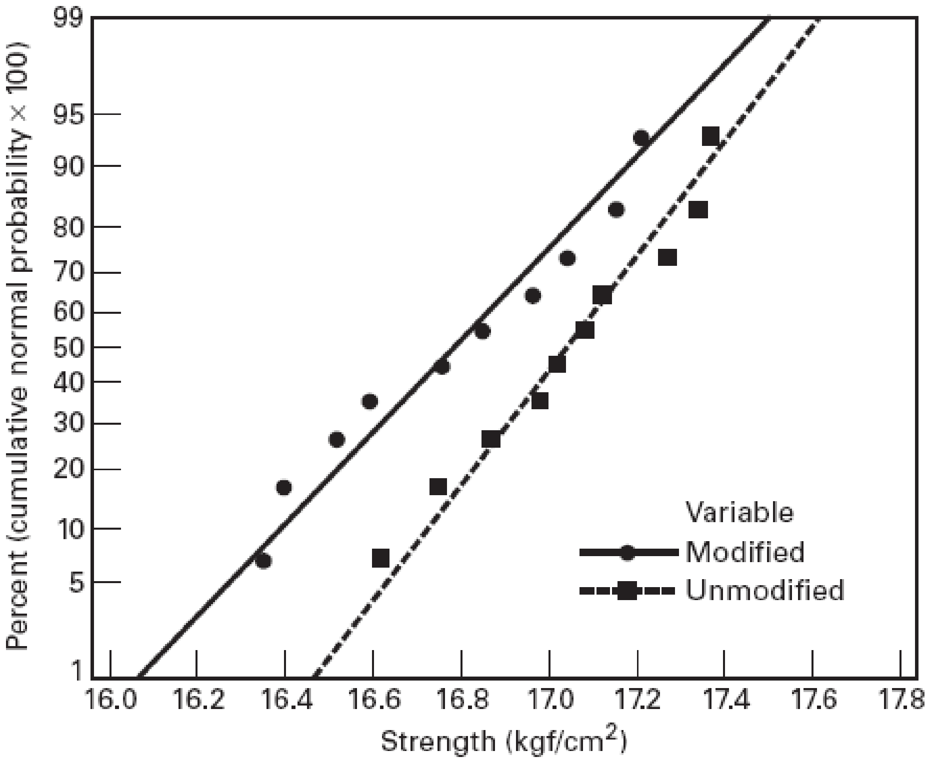
Therefore the *P*-value must lie between these two probabilities, or

$$0.05 < P\text{-value} < 0.02$$

These are upper and lower bounds on the P-value.

We know that the actual *P*-value is 0.042.

Checking Assumptions – The Normal Probability Plot



■ **FIGURE 2.11** Normal probability plots of tension bond strength in the Portland cement experiment

Importance of the *t*-Test

- Provides an **objective** framework for simple comparative experiments
- Could be used to test all relevant hypotheses in a two-level factorial design, because all of these hypotheses involve the mean response at one "side" of the cube versus the mean response at the opposite "side" of the cube

Confidence Intervals (See pg. 43)

- Hypothesis testing gives an objective statement concerning the difference in means, but it doesn't specify "how different" they are
- **General form** of a confidence interval

$$L \leq \theta \leq U \text{ where } P(L \leq \theta \leq U) = 1 - \alpha$$

- The 100(1- α)% **confidence interval** on the difference in two means:

$$\bar{y}_1 - \bar{y}_2 - t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)} \leq \mu_1 - \mu_2 \leq$$
$$\bar{y}_1 - \bar{y}_2 + t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)}$$

The actual 95 percent confidence interval estimate for the difference in mean tension bond strength for the formulations of Portland cement mortar is found by substituting in Equation 2.30 as follows:

$$16.76 - 17.04 - (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}} \leq \mu_1 - \mu_2$$
$$\leq 16.76 - 17.04 + (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}}$$
$$-0.28 - 0.27 \leq \mu_1 - \mu_2 \leq -0.28 + 0.27$$
$$-0.55 \leq \mu_1 - \mu_2 \leq -0.01$$

Thus, the 95 percent confidence interval estimate on the difference in means extends from -0.55 to -0.01 kgf/cm^2. Put another way, the confidence interval is $\mu_1 - \mu_2 = -0.28 \pm 0.27$ kgf/cm^2, or the difference in mean strengths is -0.28 kgf/cm^2, and the accuracy of this estimate is ± 0.27 kgf/cm^2. Note that because $\mu_1 - \mu_2 = 0$ is *not* included in this interval, the data do not support the hypothesis that $\mu_1 = \mu_2$ at the 5 percent level of significance (recall that the *P*-value for the two-sample *t*-test was 0.042, just slightly less than 0.05). It is likely that the mean strength of the unmodified formulation exceeds the mean strength of the modified formulation. Notice from Table 2.2 that both Minitab and JMP reported this confidence interval when the hypothesis testing procedure was conducted.

What if the Two Variances are Different?

EXAMPLE 2.1

Nerve preservation is important in surgery because accidental injury to the nerve can lead to post-surgical problems such as numbness, pain, or paralysis. Nerves are usually identified by their appearance and relationship to nearby structures or detected by local electrical stimulation (electromyography), but it is relatively easy to overlook them. An article in *Nature Biotechnology* ("Fluorescent Peptides Highlight Peripheral Nerves During Surgery in Mice," Vol. 29, 2011) describes the use of a fluorescently labeled peptide that binds to nerves to assist in identification. Table 2.3 shows the normalized fluorescence after two hours for nerve and muscle tissue for 12 mice (the data were read from a graph in the paper).

We would like to test the hypothesis that the mean normalized fluorescence after two hours is greater for nerve tissue then for muscle tissue. That is, if μ_8 is the mean normalized fluorescence for nerve tissue and is the mean normalized fluorescence for muscle tissue, we want to test

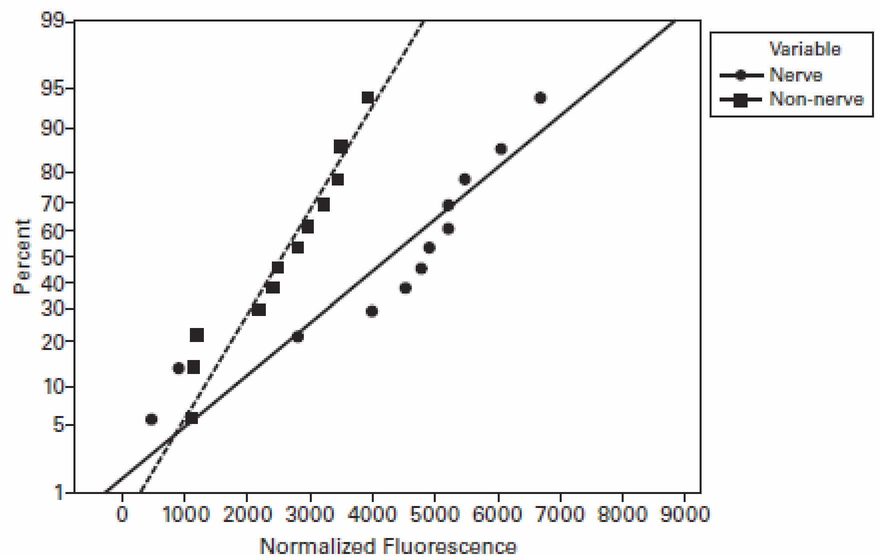
$$H_0{:}\mu_1 = \mu_2$$
$$H_1{:}\mu_1 > \mu_2$$

TABLE 2.3
Normalized Fluorescence After Two Hours

Observation	Nerve	Muscle
1	6625	3900
2	6000	3500
3	5450	3450
4	5200	3200
5	5175	2980
6	4900	2800
7	4750	2500
8	4500	2400
9	3985	2200
10	900	1200
11	450	1150
12	2800	1130

The descriptive statistics output from Minitab is shown below:

```
Variable    N   Mean  StDev  Minimum  Median  Maximum
Nerve      12   4228   1918      450    4825     6625
Non-nerve  12   2534    961     1130    2650     3900
```



■ **FIGURE 2.14** **Normalized Fluorescence Data from Table 2.3**

If we are testing

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

and cannot reasonably assume that the variances σ_1^2 and σ_2^2 are equal, then the two-sample t-test must be modified slightly. The test statistic becomes

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \tag{2.31}$$

This statistic is not distributed exactly as t. However, the distribution of t_0 is well approximated by t if we use

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2/n_1)^2}{n_1 - 1} + \frac{(S_2^2/n_2)^2}{n_2 - 1}} \tag{2.32}$$

as the number of degrees of freedom. A strong indication of unequal variances on a normal probability plot would be a situation calling for this version of the t-test. You should be able to develop an equation for finding that confidence interval on the difference in mean for the unequal variances case easily.

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{4228 - 2534}{\sqrt{\frac{(1918)^2}{12} + \frac{(961)^2}{12}}} = 2.7354$$

The number of degrees of freedom are calculated from Equation 2.32:

$$v = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{(S_1^2 / n_1)^2}{n_1 - 1} + \frac{(S_2^2 / n_2)^2}{n_2 - 1}} = \frac{\left(\frac{(1918)^2}{12} + \frac{(961)^2}{12}\right)^2}{\frac{[(1918)^2 / 12]^2}{11} + \frac{[(961)^2 / 12]^2}{11}} = 16.1955$$

If we are going to find a P-value from a table of the t-distribution, we should round the degrees of freedom down to 16. Most computer programs interpolate to determine the P-value. The Minitab output for the two-sample t-test is shown below. Since the P-value reported is small (0.015), we would reject the null hypothesis and conclude that the mean normalized fluorescence for nerve tissue is greater than the mean normalized fluorescence for muscle tissue.

```
Difference = mu (Nerve) - mu (Non-nerve)
Estimate for difference:  1694
95% lower bound for difference:  613
T-Test of difference = 0 (vs >): T-Value = 2.74  P-Value = 0.007  DF = 16
```

Other Chapter Topics

- Hypothesis testing when the variances are known
- One sample inference (t and Z tests)
- Hypothesis tests on variances (F tests)

2-4.3 Confidence Intervals

Although hypothesis testing is a useful procedure, it sometimes does not tell the entire story. It is often preferable to provide an interval within which the value of the parameter or parameters in question would be expected to lie. These interval statements are called **confidence intervals**. In many engineering and industrial experiments, the experimenter already knows that the means μ_1 and μ_2 differ; consequently, hypothesis testing on $\mu_1 = \mu_2$ is of little interest. The experimenter would usually be more interested in a confidence interval on the difference in means $\mu_1 - \mu_2$.

To define a confidence interval, suppose that θ is an unknown parameter. To obtain an interval estimate of θ, we need to find two statistics L and U such that the probability statement

$$P(L \leqslant \theta \leqslant U) = 1 - \alpha \tag{2-27}$$

is true. The interval

$$L \leqslant \theta \leqslant U \tag{2-28}$$

is called a **100(1 − α) percent confidence interval** for the parameter θ. The interpretation of this interval is that if, in repeated random samplings, a large number of such intervals are constructed, $100(1 - \alpha)$ percent of them will contain the true value of θ. The statistics L and U are called the **lower** and **upper confidence limits**, respectively, and $1 - \alpha$ is called the **confidence coefficient**. If $\alpha = 0.05$, Equation 2-28 is called a 95 percent confidence interval for θ. Note that confidence intervals have a frequency interpretation; that is, we do not know if the statement is true for this specific sample, but we do know that the *method* used to produce the confidence interval yields correct statements $100(1 - \alpha)$ percent of the time.

Suppose that we wish to find a $100(1 - \alpha)$ percent confidence interval on the true difference in means $\mu_1 - \mu_2$ for the portland cement problem. The interval can be derived in the following way. The statistic

$$\frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$$

is distributed as $t_{n_1+n_2-2}$. Thus,

$$P\left(-t_{\alpha/2,n_1+n_2-2} \leqslant \frac{\bar{y}_1 - \bar{y}_2 - (\mu_1 - \mu_2)}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \leqslant t_{\alpha/2,n_1+n_2-2}\right) = 1 - \alpha$$

or

$$P\left(\bar{y}_1 - \bar{y}_2 - t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leqslant \mu_1 - \mu_2 \leqslant \bar{y}_1 - \bar{y}_2 + t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}\right) = 1 - \alpha \quad (2\text{-}29)$$

Comparing Equations 2-29 and 2-27, we see that

$$\bar{y}_1 - \bar{y}_2 - t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leqslant \mu_1 - \mu_2 \leqslant \bar{y}_1 - \bar{y}_2 + t_{\alpha/2,n_1+n_2-2}S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \quad (2\text{-}30)$$

is a $100(1 - \alpha)$ percent confidence interval for $\mu_1 - \mu_2$.

The actual 95 percent confidence interval estimate for the difference in mean tension bond strength for the formulations of portland cement mortar is found by substituting in Equation 2-30 as follows:

$$16.76 - 17.92 - (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}} \leqslant \mu_1 - \mu_2 \leqslant 16.76 - 17.92 + (2.101)0.284\sqrt{\tfrac{1}{10} + \tfrac{1}{10}}$$

$$-1.16 - 0.27 \leqslant \mu_1 - \mu_2 \leqslant -1.16 + 0.27$$

$$-1.43 \leqslant \mu_1 - \mu_2 \leqslant -0.89$$

Thus, the 95 percent confidence interval estimate on the difference in means extends from -1.43 kgf/cm^2 to -0.89 kgf/cm^2. Put another way, the confidence interval is $\mu_1 - \mu_2 = -1.16$ kgf/cm^2 $\pm$ 0.27 kgf/cm^2, or the difference in mean strengths is -1.16 kgf/cm^2, and the accuracy of this estimate is ± 0.27 kgf/cm^2. Note that because $\mu_1 - \mu_2 = 0$ is *not* included in this interval, the data do not support the hypothesis that $\mu_1 = \mu_2$ at the 5 percent level of significance. It is likely that the mean strength of the unmodified formulation exceeds the mean strength of the modified formulation. Notice from Table 2-2 that Minitab also reported this confidence interval when the hypothesis testing procedure was conducted.

2-4.4 The Case Where $\sigma_1^2 \neq \sigma_2^2$

If we are testing

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

and cannot reasonably assume that the variances σ_1^2 and σ_2^2 are equal, then the two-sample t-test must be modified slightly. The test statistic becomes

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \tag{2-31}$$

This statistic is not distributed exactly as t. However, the distribution of t_0 is well-approximated by t if we use

$$v = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}} \tag{2-32}$$

as the degrees of freedom. A strong indication of unequal variances on a normal probability plot would be a situation calling for this version of the t-test. You should be able to develop an equation for finding that confidence interval on the difference in mean for the unequal variances case easily.

2-4.5 The Case Where σ_1^2 and σ_2^2 Are Known

If the variances of both populations are *known*, then the hypotheses

$$H_0: \mu_1 = \mu_2$$
$$H_1: \mu_1 \neq \mu_2$$

may be tested using the statistic

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} \tag{2-33}$$

If both populations are normal, or if the sample sizes are large enough so that the central limit theorem applies, the distribution of Z_0 is $N(0, 1)$ if the null hypothesis is true. Thus, the critical region would be found using the normal distribution rather than the t. Specifically, we would reject H_0 if $|Z_0| > Z_{\alpha/2}$, where $Z_{\alpha/2}$ is the upper $\alpha/2$ percentage point of the standard normal distribution.

Unlike the t-test of the previous sections, the test on means with known variances does not require the assumption of sampling from normal populations. One can use the central limit theorem to justify an approximate normal distribution for the difference in sample means $\bar{y}_1 - \bar{y}_2$.

The $100(1 - \alpha)$ percent confidence interval on $\mu_1 - \mu_2$ where the variances are known is

$$\bar{y}_1 - \bar{y}_2 - Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{y}_1 - \bar{y}_2 + Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad (2\text{-}34)$$

As noted previously, the confidence interval is often a useful supplement to the hypothesis testing procedure.

2-4.6 Comparing a Single Mean to a Specified Value

Some experiments involve comparing only one population mean μ to a specified value, say, μ_0. The hypotheses are

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

If the population is normal with known variance, or if the population is nonnormal but the sample size is large enough so that the central limit theorem applies, then the hypothesis may be tested using a direct application of the normal distribution. The test statistic is

$$Z_0 = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}} \qquad (2\text{-}35)$$

If $H_0: \mu = \mu_0$ is true, then the distribution of Z_0 is $N(0, 1)$. Therefore, the decision rule for $H_0: \mu = \mu_0$ is to reject the null hypothesis if $|Z_0| > Z_{\alpha/2}$. The value of the mean μ_0 specified in the null hypothesis is usually determined in one of three ways. It may result from past evidence, knowledge, or experimentation. It may be the result of some theory or model describing the situation under study. Finally, it may be the result of contractual specifications.

The $100(1 - \alpha)$ percent confidence interval on the true population mean is

$$\bar{y} - Z_{\alpha/2}\sigma/\sqrt{n} \leq \mu \leq \bar{y} + Z_{\alpha/2}\sigma/\sqrt{n} \qquad (2\text{-}36)$$

EXAMPLE 2-1 ....

A vendor submits lots of fabric to a textile manufacturer. The manufacturer wants to know if the lot average breaking strength exceeds 200 psi. If so, she wants to accept the lot. Past experience indicates that a reasonable value for the variance of breaking strength is $100(\text{psi})^2$. The hypotheses to be tested are

$$H_0: \mu = 200$$
$$H_1: \mu > 200$$

Note that this is a one-sided alternative hypothesis. Thus, we would accept the lot only if the null hypothesis $H_0: \mu = 200$ could be rejected (i.e., if $Z_0 > Z_\alpha$).

Four specimens are randomly selected, and the average breaking strength observed is $\bar{y} = 214$ psi. The value of the test statistic is

$$Z_0 = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}} = \frac{214 - 200}{10/\sqrt{4}} = 2.80$$

If a type I error of $\alpha = 0.05$ is specified, we find $Z_\alpha = Z_{0.05} = 1.645$ from Appendix Table I. Thus H_0 is rejected, and we conclude that the lot average breaking strength exceeds 200 psi.

If the variance of the population is unknown, we must make the additional assumption that the population is normally distributed, although moderate departures from normality will not seriously affect the results.

To test $H_0: \mu = \mu_0$ in the variance unknown case, the sample variance S^2 is used to estimate σ^2. Replacing σ with S in Equation 2-35, we have the test statistic

$$t_0 = \frac{\bar{y} - \mu_0}{S/\sqrt{n}} \tag{2-37}$$

The null hypothesis $H_0: \mu = \mu_0$ would be rejected if $|t_0| > t_{\alpha/2,n-1}$, where $t_{\alpha/2,n-1}$ denotes the upper $\alpha/2$ percentage point of the t distribution with $n - 1$ degrees of freedom. The $100(1 - \alpha)$ percent confidence interval in this case is

$$\bar{y} - t_{\alpha/2,n-1}S/\sqrt{n} \leq \mu \leq \bar{y} + t_{\alpha/2,n-1}S/\sqrt{n} \tag{2-38}$$

2-4.7 Summary

Tables 2-3 and 2-4 summarize the test procedures discussed above for sample means. Critical regions are shown for both two-sided and one-sided alternative hypotheses.

Table 2-3 Tests on Means with Variance Known

Hypothesis	Test Statistic	Criteria for Rejection
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$		$\|Z_0\| > Z_{\alpha/2}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$Z_0 = \frac{\bar{y} - \mu_0}{\sigma/\sqrt{n}}$	$Z_0 < -Z_\alpha$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$Z_0 > Z_\alpha$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$		$\|Z_0\| > Z_{\alpha/2}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z_0 < -Z_\alpha$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$		$Z_0 > Z_\alpha$

Table 2-4 Tests on Means of Normal Distributions, Variance Unknown

Hypothesis	Test Statistic	Criteria for Rejection
$H_0: \mu = \mu_0$ $H_1: \mu \neq \mu_0$		$\lvert t_0 \rvert > t_{\alpha/2,n-1}$
$H_0: \mu = \mu_0$ $H_1: \mu < \mu_0$	$t_0 = \dfrac{\bar{y} - \mu_0}{S/\sqrt{n}}$	$t_0 < -t_{\alpha,n-1}$
$H_0: \mu = \mu_0$ $H_1: \mu > \mu_0$		$t_0 > t_{\alpha,n-1}$
	if $\sigma_1^2 = \sigma_2^2$	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 \neq \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{S_p\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ $v = n_1 + n_2 - 2$	$\lvert t_0 \rvert > t_{\alpha/2,v}$
	if $\sigma_1^2 \neq \sigma_2^2$	
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 < \mu_2$	$t_0 = \dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$	$t_0 < -t_{\alpha,v}$
$H_0: \mu_1 = \mu_2$ $H_1: \mu_1 > \mu_2$	$v = \dfrac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{(S_1^2/n_1)^2}{n_1 - 1} + \dfrac{(S_2^2/n_2)^2}{n_2 - 1}}$	$t_0 > t_{\alpha,v}$

2-6 INFERENCES ABOUT THE VARIANCES OF NORMAL DISTRIBUTIONS

In many experiments, we are interested in possible differences in the mean response for two treatments. However, in some experiments it is the comparison of variability in the data that is important. In the food and beverage industry, for example, it is important that the variability of filling equipment be small so that all packages have close to the nominal net weight or volume of content. In chemical laboratories, we may wish to compare the variability of two analytical methods. We now briefly examine tests of hypotheses and confidence intervals for variances of normal distributions. Unlike the tests on means, the procedures for tests on variances are rather sensitive to the normality assumption. A good discussion of the normality assumption is in Appendix 2A of Davies (1956).

Suppose we wish to test the hypothesis that the variance of a normal population equals a constant, for example, σ_0^2. Stated formally, we wish to test

$$\begin{aligned} H_0&: \sigma^2 = \sigma_0^2 \\ H_1&: \sigma^2 \neq \sigma_0^2 \end{aligned} \tag{2-44}$$

The test statistic for Equation 2-44 is

$$\chi_0^2 = \frac{SS}{\sigma_0^2} = \frac{(n-1)S^2}{\sigma_0^2} \tag{2-45}$$

where $SS = \Sigma_{i=1}^{n}(y_i - y)^2$ is the corrected sum of squares of the sample observations. The appropriate reference distribution for χ_0^2 is the chi-square distribution with $n - 1$ degrees of freedom. The null hypothesis is rejected if $\chi_0^2 > \chi_{\alpha/2,n-1}^2$ or if $\chi_0^2 < \chi_{1-(\alpha/2),n-1}^2$, where $\chi_{\alpha/2,n-1}^2$ and $\chi_{1-(\alpha/2),n-1}^2$ are the upper $\alpha/2$ and lower $1 - (\alpha/2)$ percentage points of the chi-square distribution with $n - 1$ degrees of freedom, respectively. Table 2-7 (on the facing page) gives the critical regions for the one-sided alternative hypotheses. The $100(1 - \alpha)$ percent confidence interval on σ^2 is

$$\frac{(n-1)S^2}{\chi_{\alpha/2,n-1}^2} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi_{1-(\alpha/2),n-1}^2} \tag{2-46}$$

Now consider testing the equality of the variances of two normal populations. If independent random samples of size n_1 and n_2 are taken from populations 1 and 2, respectively, the test statistic for

$$\begin{aligned} H_0&: \sigma_1^2 = \sigma_2^2 \\ H_1&: \sigma_1^2 \neq \sigma_2^2 \end{aligned} \tag{2-47}$$

is the ratio of the sample variances

$$F_0 = \frac{S_1^2}{S_2^2} \tag{2-48}$$

The appropriate reference distribution for F_0 is the F distribution with $n_1 - 1$ numerator degrees of freedom and $n_2 - 1$ denominator degrees of freedom. The null hypothesis would be rejected if $F_0 > F_{\alpha/2,n_1-1,n_2-1}$ or if $F_0 < F_{1-(\alpha/2),n_1-1,n_2-1}$, where $F_{\alpha/2,n_1-1,n_2-1}$ and $F_{1-(\alpha/2),n_1-1,n_2-1}$ denote the upper $\alpha/2$ and lower $1 - (\alpha/2)$ percentage points of the F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom. Table IV of the Appendix gives only upper-tail percentage poins of F; however, the upper- and lower-tail points are related by

$$F_{1-\alpha,v_1,v_2} = \frac{1}{F_{\alpha,v_2,v_1}} \tag{2-49}$$

Table 2-7 Tests on Variances of Normal Distributions

Hypothesis	Test Statistic	Criteria for Rejection
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 \neq \sigma_0^2$		$\chi_0^2 > \chi^2_{\alpha/2,n-1}$ or $\chi_0^2 < \chi^2_{1-\alpha/2,n-1}$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 < \sigma_0^2$	$\chi_0^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi_0^2 < \chi^2_{1-\alpha,n-1}$
$H_0: \sigma^2 = \sigma_0^2$ $H_1: \sigma^2 > \sigma_0^2$		$\chi_0^2 > \chi^2_{\alpha,n-1}$
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 \neq \sigma_2^2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha/2,n_1-1,n_2-1}$ or $F_0 < F_{1-\alpha/2,n_1-1,n_2-1}$
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 < \sigma_2^2$	$F_0 = \dfrac{S_2^2}{S_1^2}$	$F_0 > F_{\alpha,n_2-1,n_1-1}$
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$	$F_0 = \dfrac{S_1^2}{S_2^2}$	$F_0 > F_{\alpha,n_1-1,n_2-1}$

Test procedures for more than two variances are discussed in Chapter 3, Section 3-4.3. We will also discuss the use of the variance or standard deviation as a response variable in more general experimental settings.

EXAMPLE 2-2

A chemical engineer is investigating the inherent variability of two types of test equipment that can be used to monitor the output of a production process. He suspects that the old equipment, type 1, has a larger variance than the new one. Thus, he wishes to test the hypothesis

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 > \sigma_2^2$$

Two random samples of $n_1 = 12$ and $n_2 = 10$ observations are taken, and the sample variances are $S_1^2 = 14.5$ and $S_2^2 = 10.8$. The test statistic is

$$F_0 = \frac{S_1^2}{S_2^2} = \frac{14.5}{10.8} = 1.34$$

From Appendix Table IV we find that $F_{0.05,11,9} = 3.10$, so the null hypothesis cannot be rejected. That is, we have found insufficient statistical evidence to conclude that the variance of the old equipment is greater than the variance of the new equipment.

The $100(1 - \alpha)$ confidence interval for the ratio of the population variances σ_1^2/σ_2^2 is

$$\frac{S_1^2}{S_2^2} F_{1-\alpha/2,n_2-1,n_1-1} \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{S_1^2}{S_2^2} F_{\alpha/2,n_2-1,n_1-1} \qquad (2\text{-}50)$$

To illustrate the use of Equation 2-50, the 95 percent confidence interval for the ratio of variances σ_1^2/σ_2^2 in Example 2-2 is, using $F_{0.025,9,11} = 3.59$ and $F_{0.975,9,11} = 1/F_{0.025,11,9} = 1/3.92 = 0.255$,

$$\frac{14.5}{10.8}(0.255) \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant \frac{14.5}{10.8}(3.59)$$

$$0.34 \leqslant \frac{\sigma_1^2}{\sigma_2^2} \leqslant 4.81$$