Chapter 9

9.2 A small p-value indicates an unlikely sample result when the null hypothesis is assumed to be correct. The observation of sample data that is highly unlikely casts doubt on the correctness of the null hypothesis.

9.4 We can test the statement on the box by first assuming that $\mu = 1{,}000$, where μ is the mean or expected life(hours) of the light bulb. We have to randomly draw a sample of n packages (with $n \geq 30$, so that we have an approximately normal distribution of sample mean). Using the sample mean $\bar{x}$ from this sample, we can calculate the test statistic to test the correctness of the claim.

$$z = \frac{\bar{x} - \mu_0}{F/\sqrt{n}} = \frac{\bar{x} - 1000}{50/\sqrt{n}}$$ where $\bar{x}$ and n are obtained from the

sample. If this calculated test statistic is “large” then we can reject the claim that the expected life of the bulbs is 1,000 hours. If z is “small” then we would not reject this claim.

9.6 a. H_0: $\mu = 87.4$ vs. H_A: $\mu \neq 87.4$, where μ is historical mean number of applicants for a training program in sales at IBM.

b. H_0: $\mu = 0.05$ vs. H_A: $\mu \neq 0.05$, where μ is the inflation rate in the area.

c. H_0: $\mu \geq 3$ vs. H_A: $\mu < 3$, where μ is the mean sick days per employee per month for the entire workforce.

d. H_0: $B \leq 0.1$ vs. H_A: $B > 0.1$, where B is the overall defect rate for the machine.

e. H_0: $\mu \leq 0.018$ vs. H_A: $\mu > 0.018$, where μ is the population mean stretch of the new line at 15 percent of its working load.

f. H_0: $\mu = 378$ vs. H_A: $\mu \neq 378$, where μ is the population mean cost($) of piece work.

g. H_0: $\mu \leq 1000$ vs. H_A: $\mu > 1000$, where μ is the population mean breaking strength(lbs.) of the new composite per square inch.

h. H_0: $B \leq 0.07$ vs. H_A: $B > 0.07$, where B is the population proportion of unemployment in the next recession.

i. H_0: B ≥ 0.06 vs. H_A: B < 0.06, where B is the population proportion of overdue accounts receivable.

9.8 a. Two-tailed z test is needed. Thus, $z_{critical} = \pm z_{"/2} = \pm z_{0.01} = \pm 2.33$, by Appendix Table A.3, or from EXCEL 2.326 =NORMSINV(0.99)

b. One-tailed t test is needed because the population standard deviation is unknown. Thus, $t_{critical} = t_{("=0.05, df=15)} = 1.753$, by Appendix Table A.4, or from EXCEL 1.7531 =TINV(0.1,15)

c. Two-tailed z test is needed. Thus $z_{critical} = \pm z_{"/2} = \pm z_{0.025} = \pm 1.96$, by Appendix Table A.3, or from EXCEL 1.95996 =NORMSINV(0.975)

9.10 A Type I error occurs if the null hypothesis is rejected in favor of the alternative when in fact the statement in the null hypothesis is true. The probability of a Type I error is ". Thus, 1 - " is the probability of accepting the null hypothesis when it is correct. This does not equal the probability of making a Type II error, which is the probability of not rejecting the null hypothesis when it is false. To calculate the probability of a Type II error we would need to know the true value of the population parameter of interest. If we knew this value, however, there would be no need to do the test.

9.12 Peter Chevalier was asserting that the FDA is using too small of ". That is, the FDA set the probability of a Type I error too low compared to that used by other nations.

9.14 If the null hypothosis is not rejected (at an " Type I error level) then the value of the population parameter specified in the null hypothesis will be inside the (1-")100 percent confidence interval constructed with the same sample.

9.16 a. H_0: μ ≥ 10,000 vs. H_A: μ < 10,000, where μ is the average mileage of the company's cars.

b. Using the ASCII data file EX9-16.PRN, the mean and the standard deviation of the given sample data is 9296.556 miles and 2862.787 miles respectively. We need to assume that the population distribution is normal. Because the population standard deviation is unknown and the sample size is 9, we have to use the t distribution in calculating the p-value. The p-value is

$$P(\bar{x} \le 9297) = P(t \le \frac{9297 - 10000}{2863/\sqrt{9}}) = P(t \le -0.74) = 0.240$$

where the degrees of freedom of the t is 8.

	A	
1	9876	
2	11346	
3	12321	
4	8564	
5	12590	
6	3210	
7	7487	
8	8886	
9	9389	
10	9296.555556	=AVERAGE(A1:A9)
11	2862.787283	=STDEV(A1:A9)
12	954.2624277	=A11/SQRT(9)
13	-0.73716037	=(A10-10000)/A12
14	0.241041095	=TDIST(-A13,8,1)

Because the p-value = 0.24 is relatively large, the null hypothesis cannot be rejected(i.e., the company's claim cannot be accepted as true) unless the level of " is set above 24 percent, which is hardly usual.

9.18 We need to test H_0: $\mu \le 80$ vs. H_A: $\mu > 80$, where μ is the average piece rate(units per worker per hour). Although the population standard deviation is unknown, we can use z distribution for testing because the sample size 47 is large. The p-value is

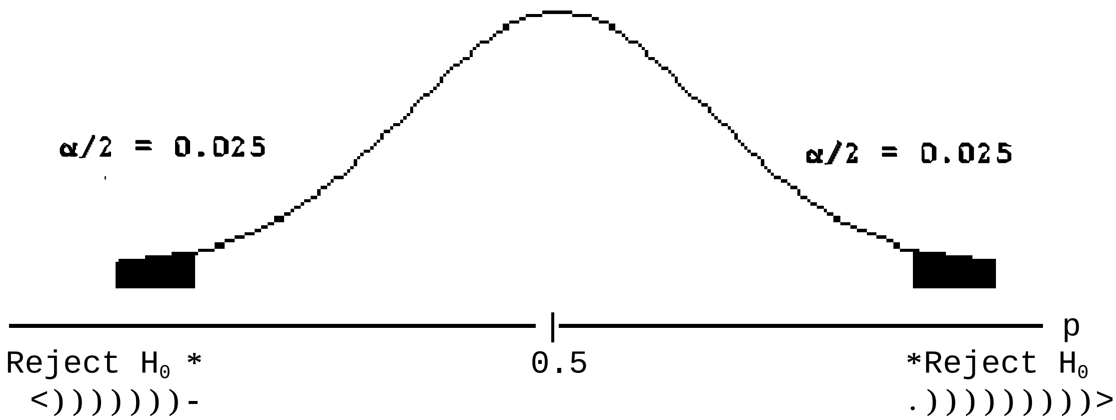
$$P(\bar{x} \ge 96) = P(z \ge \frac{96 - 80}{21/\sqrt{47}}) = P(z \ge 5.22) = 0.0000$$

This p-value is much smaller than the " level 0.05. Thus, the null hypothesis is rejected in favor of the alternative. We conclude that the new process is better than the old.

9.20 To test the hypothesis H_0: $\mu \leq 1.0$ vs. H_A: $\mu > 1.0$, where μ is the mean number of defects per car, we need the sample size and the sample standard deviation(or the population standard deviation) for the sample yielding a mean of 5.1. We also need the desired probability of a Type I error.

9.22 a. H_0: B = 0.5 vs. H_A: B $\neq$ 0.5, where B is the winners' chance of repeating their success.

b. Rejection region for the hypothesis test for the chances of repeating success



c. Assuming H_0 is correct, the standard error of the estimator(sample proportion), i.e., the standard deviation of the sample proportion, is calculated to be

$$F_p = \sqrt{\frac{B_0(1-B_0)}{n}} = \sqrt{\frac{(0.5)(0.5)}{106}} = 0.049$$ when B_0 is the

hypothesized value of the population proportion in the null hypothesis.

d. From the quoted statement, we know that the null hypothesis was not rejected. Assuming the test was done at the 0.05 " level, this implies that,

$$either \quad P(p \le b) = P\left(z \le \frac{b - 0.5}{\sqrt{\frac{(0.5)(0.5)}{106}}}\right)$$

$$or \quad P(p \ge b) = P\left(z \ge \frac{b - 0.5}{\sqrt{\frac{(0.5)(0.5)}{106}}}\right)$$

could have been as low as 0.025, where b is the value of the sample proportion). Thus, the calculated z could have been as high as 1.96 and as low as -1.96 in the limiting case. This implies that the value b of the point estimator p could have been between 0.405(40.5 percent) and 0.595(59.5 percent). With any estimated value for b in this range, the p-value cannot be smaller than 0.05 and thus the null hypothesis cannot be rejected.

9.24 We need to test the following hypothesis:

H_0: B $\le \frac{1}{3}$ vs. H_A: B $> \frac{1}{3}$, where B is the proportion of consumers who prefer bottle A.

The sample proportion was observed to be 0.36 for a sample of 100 consumers. Thus, the p-value is

$$P(p \ge 0.36) = P\left(z \ge \frac{0.36 - 1/3}{\sqrt{\frac{(1/3)(2/3)}{100}}}\right) = P(z \ge 0.57) = 0.2843$$

This approximate p-value can be obtained directly from EXCEL by “=1-NORMDIST(0.36,1/3,SQRT((1/3)*(2/3)/100),1)” = 0.2858. Because the p-value is greater than the " level 0.1, H_0 is rejected. (Or equivalently, the test statistic z = 0.65 is not beyond the critical value $z_{0.1}$ = 1.28 and thus H_0 is rejected.) We cannot conclude that more than one third of the customers prefer bottle A.

9.26 We test H_0: $\mu = \mu_o$ vs. H_A: $\mu \ne \mu_0$ where μ is the composition of vapors of a batch of alloys and μ_0 is the (mean) norm.
Type I error is to conclude that the composition of vapors deviate from the norm when in fact it does not.
Type II error is to conclude that the composition of vapors does not deviate from the norm when it does.

9.28 a. Using the ASCII data file EX9-28.PRN, the sample mean and the sample standard deviation are calculated to be

6.5 and 2.828 respectively. Using this sample information, we are to test H_0: μ ≤ 6 vs. H_A: μ > 6, where μ is the mean number of defects per car. Because the population standard deviation is unknown and the sample is small, we use the t distribution for testing, assuming the population distribution is normal.

b. The p-value is

$$P(\bar{x} \geq 6.5) = P(t \geq \frac{6.5-6}{2.828/\sqrt{16}}) = P(t \geq 0.707) = 0.245$$

where the degrees of freedom are 15. Using EXCEL:

	A	B	
1	3		
2	6		
3	8		
4	9		
5	5		
6	7		
7	10		
8	4		
9	5		
10	12		
11	8		
12	4	6.5	=AVERAGE(A1:A16)
13	6	2.828427125	=STDEV(A1:A16)
14	9	0.707106781	=B13/SQRT(16)
15	1	0.707106791	=(B12-6)/B14
16	7	0.245170239	=TDIST(B15,15,1)

Because the p-value is greater than the 0.05 " level, the null hypothesis is not rejected. (Or equivalently, the calculated t = 0.707 is not beyond the critical $t_{("=0.05,\ df=15)}$ = 1.753 and thus the null hypothesis is not rejected.) There is not sufficient evidence to conclude that the mean is more than six defects per car.

9.30 We need to test H_0: B = 0.5 vs. H_A: B ≠ 0.5, where B is the proportion of companies that showed a stock price fall after moving. The sample proportion is 0.567(=17/30) for a sample of 30. Thus, the probability of getting a sample proportion 0.567 or higher, assuming H_0 is true, will be

$$P(p \geq 0.567) = P(z \geq \frac{0.567-0.5}{\sqrt{\frac{(0.5)(0.5)}{30}}}) = P(z \geq 0.73) = 0.2327$$

This approximate probability is obtained directly from EXCEL by “=1-NORMDIST(.567,.5,SQRT((.5)*(.5)/30),1)” = 0.2315. Thus, the two-tail p-value is 0.46(=2(0.23)). Because this p-value is greater than a typical " level of 0.05, H_0 is not rejected. (Or equivalently, the test statistic z = 0.73 is not beyond the critical $z_{0.025}$ = 1.96 so H_0 is not rejected.) The test result does not contradict the *WSJ*'s hypothesis.

9.32 Using the ASCII file EX9-32.PRN, the sample mean and the sample standard deviation are calculated to be 40 and 26.248 respectively. We are to test H_0: μ = 42 vs. H_A: μ ≠ 42, where μ is the mean passenger miles flown. Using the t distribution, the probability of getting a sample mean 40 or less, assuming H_0 is true, will be

$$P(\overline{X} \leq 40) = P(t \leq \frac{40-42}{26.248/\sqrt{10}}) = P(t \leq -0.24) = 0.408$$

where the degrees of freedom of the t is 9. Thus, for a two-tail test the p-value is about 0.816. From EXCEL:

	A	B	
1	76.9		
2	76		
3	59		
4	51.5		
5	41.2		
6	35.6	40	=AVERAGE(A1:A10)
7	34.2	26.24830153	=STDEV(A1:A10)
8	11.1	8.300441755	=B7/SQRT(10)
9	10	-0.240951031	=(B6-42)/B8
10	4.5	0.814992532	=TDIST(-B9,9,2)

Thus, because the two-tail p-value of 0.815 is greater than a typical 0.05 " level, H_0 is not rejected. We cannot reject the regulators' claim.

9.34 a. We need to test H_0: μ ≤ 0.19 vs. H_A: μ > 0.19, where μ is Folger's average market share.

b. Using the ASCII data file EX9-34.PRN, the sample mean and the sample standard deviation for the given sample data are calculated to be 0.201 and 0.031 respectively. Using the t distribution, the p-value will be

$$P(\overline{x} \geq 0.201) = P\left(t \geq \frac{0.201-0.19}{0.031/\sqrt{10}}\right) = P(t \geq 1.12) = 0.146$$

where the degrees of freedom of the t is 9. From EXCEL:

	A	B	
1	0.2		
2	0.19		
3	0.17		
4	0.21		
5	0.15		
6	0.2	0.201	=AVERAGE(A1:A10
7	0.25	0.030713732	=STDEV(A1:A10)
8	0.24	0.009712535	=B7/SQRT(10)
9	0.22	1.132557068	=(B6-.19)/B8
10	0.18	0.143335222	=TDIST(B9,9,1)

Because this p-value is greater than a typical 0.05 " level, H_0 is not rejected. We cannot conclude that the market share exceeds 19 percent this year.

9.36 We test the hypotheses H_0: B ≤ 0.5 vs. H_A: B > 0.5, where B is the proportion of GE 60 watt light bulbs lasting longer than 1000 hours. Based on the sample proportion 0.639(=23/36) for a sample of 36, the p-value will be

$$P(p \geq 0.639) = P\left(z \geq \frac{0.639-0.50}{\sqrt{\frac{(0.50)(0.50)}{36}}}\right) = P(z \geq 1.67) = 0.0475$$

This approximate probability is obtained directly from EXCEL by “=1-NORMDIST(23/36,.5,SQRT((.5)*(.5)/36),1)” = 0.4779. Because this p-value is less than the 0.05 " level, H_0 is rejected. (Or equivalently, the test statistic z = 1.68 is beyond the critical value $z_{0.05}$ = 1.645 and thus H_0 is rejected.) We conclude that more than 50 percent of all GE 60 watt bulbs will last longer than 1000 hours.

9.38 This additional information will raise the calculated z statistic but again this is meaningless because sampling is not random.

9.40 a. If some consumers refuse to answer, the sample of those consumers who participate cannot be interpreted as random.

b. We need to test H_0: B = 0.36 vs. H_A: B $\neq$ 0.36, where B is the proportion of American consumers who would not respond to questions. From a simple random sample ($n \geq 30$) of the phone logs we can obtain an estimate of B. This estimate can then be used to calculate the z test statistic. If this calculated value is "large" we reject H_0 in favor of H_A.

9.42 a. The hypotheses are H_0: $\mu \geq$ 500,000 vs. H_A: μ < 500,000 where μ is the mean dollar amount of commission. Using the sample mean of $450,000 and the sample standard deviation of $50,000, the p-value is

$$P(\overline{x} \leq 450000) = P\left(t \leq \frac{450000 - 500000}{50000/\sqrt{25}}\right) = P(t \leq -5) = 0.0000$$

where the degrees of freedom are 24. Because this p-value is much smaller than a typical 0.05 " level 0.05, H_0 is rejected. We can conclude that the mean commission is smaller than the claimed $500,000.

b. The competitor's sample cannot be viewed as random because it included only those brokers who applied for the fictitious position. Lack of randomness may cause a lower sample mean, thus a false conclusion in part a.

9.44 The ASCII data file EX9-44.PRN shows the sample mean and sample standard deviation to be 11.93 and 1.241 percent. We are to test H_0: μ = 12 vs. H_A: $\mu \neq$ 12, where μ is the mean percentage of health care cost. Using the t distribution, the test statistic is

$$t = \frac{11.93 - 12}{1.241/\sqrt{10}} = -0.178$$

Because this calculated t statistic -0.178 is not beyond the critical value $-t_{("=0.025,\ df=9)}$ = -2.262

at the 0.05 " level, the null hypothesis is not rejected. Based on these this sample data the claim that 12 percent of payroll is health care cost cannot be rejected.

Alternatively, and equivalently, we could do this test with a two-tail p-value, which is obtained from EXCEL as follows:

	A	B	
	------	---------------	
1	14.1		
2	10.2		
3	12.1		
4	10.3		
5	11.8		
6	11.6	11.93	=AVERAGE(A1:A10)
7	12.1	1.241012132	=STDEV(A1:A10)
8	13.7	0.392442494	=B7/SQRT(10)
9	11.5	-0.178370082	=(B6-12)/B8
10	11.9	0.862382796	=TDIST(-B9,9,2)"

Because this 0.86 two-tail p-value is larger than the 0.05 Type I error level, we do not reject the null hypothesis.

9.46 The test is H_0: B = 0.5 vs. H_A: B $\neq$ 0.5, where B is the proportion of law firms having marketing specialists. For a sample proportion 0.48 and sample of 400, the p-value is

$$2P(p \leq 0.48) = 2P\left(z \leq \frac{0.48-0.50}{\sqrt{\frac{(0.50)(0.50)}{400}}}\right) = 2P(z \leq -0.8)$$

$$= 2(0.2119) = 0.4238$$

which can be obtained directly from EXCEL by entering "=2*NORMDIST(.48,.5,SQRT((.5)*(.5)/400),1)" = 0.4237. Because this two-tail p-value is greater than the 0.05 " level, H_0 is not rejected. (Or equivalently, the test statistic z = -0.8 is not beyond the critical value $-z_{0.025}$ = -1.96 so H_0 is not rejected.) We conclude that the observed 48 percent using marketing specialists is not significantly different from the claimed 50 percent.

9.48 The test is H_0: B = 0.5 vs. H_A: B $\neq$ 0.5, where B is the proportion of decisions in favor of plaintiffs. Based on the sample proportion 0.778(=14/18) for a sample of 18, the calculated z statistic will be

$$z = \frac{0.778 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{18}}} = 2.36$$

Because the calculated z = 2.36 is beyond the critical $z_{0.025}$ = 1.96, H_0 is rejected at the 0.05 " level. We conclude that there is a significant difference between Han's sample and the national norm. Equivalently, the EXCEL command “=2*NORMDIST(.778,.5,SQRT((.5)*(.5)/18),1)” yields a two-tail p-value of 0.01833, which is small, and H_0 is rejected.

9.50 The observation that five (39 percent) of the 13 died implies nothing about the hypothesis that drinking prior to riding increases the risk of falling and dying. To test this hypothesis we need to know the proportion of those who drank and fell and the proportion of those who drank alcohol and did not fall.

9.52 The smallest standard deviation is 7.83 miles per hours as determined in EXCEL in the following way:

	a	
1	1.788546342	=TINV(0.08,48)
2	14	=(67-65)*SQRT(49)
3	7.827585828	=A2/A1

9.54 The probability of 177 out of 236 newsletters making predictions that underperform the S&P by chance (Β =0.5), is 7.99361E-15 =1-NORMDIST(177/236,0.5,SQRT((0.5)*(0.5)/236),1) which is approximately zero. Thus, these newsletters are doing worse than what could be accomplished on average with the flip of a coin.

9.56 If 14 of 55 restaurants increased prices on fries and Cokes an average of 6.1 cents and the other 41 restaurants did not change prices at all then the sample mean increase was 1.553 cents. Although a standard deviation was not given, if price increases ranged from 0 to 14 cents, as implied by the article, then an approximate standard deviation is about 2 cents, assuming we can get 2 standard deviations in between the highest and lowest values. (This guess at the standard deviation is likely low because using the freguency data of x = 6.1, with frequency 14, and x = 0, with frequency 41, yields a standard deviation of 2.68 cents, and this calculation ignores variation around the 6.1 cents.) The calculated test statistic is now $(1.553-0)/2/\sqrt{55} = 0.1047$, which is relatively small. The *WSJ* does not provide sufficient evidence to reject the null hypothesis that the mean price increase for fries and Cokes for all McDonald's restaurants is zero. (Some EXCEL calculations follow.)

x	f	fx	(x-mean)^2	f(x-mean)^2
6.1	14	85.4	20.6776893	289.4876496
0	41	0	2.41096198	98.84944132
	mean=	1.553		7.191427609 =s^2
				2.681683727 =s

0.104683013 =(1.5527-0)/2/SQRT(55)