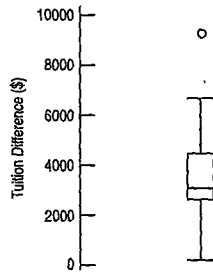


Based on these data, with 95% confidence, the average braking distance for these tires on dry pavement is between 133.6 and 145.2 feet.

- b) Not simple random samples, but most likely representative; stops most likely independent of each other; less than 10% of all possible wet stops; Normal probability plots are relatively straight. Based on these data, with 95% confidence, the average increase in distance for these tires on wet pavement is between 51.4 and 74.6 feet
25. a) Paired Data Assumption: Data are paired by college. Randomization Condition: This was a random sample of public colleges and universities. 10% Condition: these are fewer than 10% of all public colleges and universities.



Normal Population Assumption: U.C. Irvine seems to be an outlier; we might consider removing it.

- b) Having deleted the observation for U.C.-Irvine, whose difference of \$9300 was an outlier, we are 90% confident, based on the remaining data, that nonresidents pay, on average, between \$2615.31 and \$3918.02 more than residents. If we retain the outlier, the interval is (\$2759, \$4409).
- c) Assertion is reasonable; with or without the outlier, \$3500 is in the confidence interval.
26. Using a  $t$ -test for paired differences,  $t = -0.86$  and two-tailed  $P = 0.396$ . With a  $P$ -value so high, we fail to reject the null hypothesis of no mean difference. There is no evidence that sexual images in ads affects people's ability to remember the product being advertised.
27. a) 60% is 30 strikes;  $H_0: \mu = 30$  vs.  $H_A: \mu > 30$ .  $t = 6.07$ ,  $P$ -value =  $3.92 \times 10^{-6}$ . With a very small  $P$ -value, we reject  $H_0$ . There is very strong evidence that players can throw more than 60% strikes after training, based on this sample.
- b)  $H_0: \mu_D = 0$  vs.  $H_A: \mu_D > 0$ .  $t = 0.135$ ,  $P$ -value = 0.4472. With such a high  $P$ -value, we do not reject  $H_0$ . These data provide no evidence that the program has improved pitching in these Little League players.
28. If this group is representative of all students, we can be 95% confident that freshmen gain a mean of between 1.40 and 2.43 pounds during their first 12 weeks at college. That's strong evidence of a weight gain, but it's unlikely that it amounts to 15 pounds for the whole first year.

## PART VI REVIEW

1. a)  $H_0: \mu_{Jan} - \mu_{Jul} = 0$ ;  $H_A: \mu_{Jan} - \mu_{Jul} \neq 0$ .  $t = -1.94$ ,  $df = 43.68$ ,  $P$ -value = 0.0590. Since  $P < 0.10$ , reject the null.

These data show a significant difference in mean age to crawl between January and July babies.

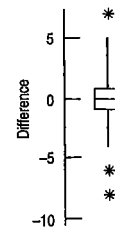
- b)  $H_0: \mu_{Apr} - \mu_{Oct} = 0$ ;  $H_A: \mu_{Apr} - \mu_{Oct} \neq 0$ .  $t = -0.92$ ;  $df = 59.40$ ;  $P$ -value = 0.3610. Since  $P > 0.10$ , do not reject the null; these data do not show a significant difference between April and October with regard to the mean age at which crawling begins.
- c) These results are not consistent with the claim.
2.  $H_0: \mu_D = 0$ ;  $H_A: \mu_D > 0$ .  $t = 1.36$ ;  $df = 20$ ;  $P$ -value = 0.0949. Because the  $P$ -value is high, we do not reject  $H_0$ . These data do not show that floral scent improved the average maze completion time between scented and unscented.
3.  $H_0: p = 0.26$ ;  $H_A: p \neq 0.26$ .  $z = 0.946$ ;  $P$ -value = 0.3443. Because the  $P$ -value is high, we do not reject  $H_0$ . These data do not show that the Denver-area rate is different from the national rate in the proportion of businesses with women owners.
4. a) We are 95% confident the average savings in Canada for prescription drugs is between \$77.57 and \$174.43.  
b) We are 95% confident that the average savings in Canada for prescription drugs is between 40.1% and 64.2%.  
c) Using percents makes the histogram more unimodal and symmetric.  
d) Probably would change. The pharmacy may have listed only the 12 drugs with the "best" savings.
5. Based on these data, we are 95% confident that the mean difference in aluminum oxide content is between  $-3.37$  and  $1.65$ . Since the interval contains 0, the means in aluminum oxide content of the pottery made at the two sites could reasonably be the same.
6. We are 95% confident that the proportion of streams in the Adirondacks with shale substrates is between 32.8% and 47.4%.
7. a)  $H_0: p_{ALS} - p_{Other} = 0$ ;  $H_A: p_{ALS} - p_{Other} > 0$ .  $z = 2.52$ ;  $P$ -value = 0.0058. With such a low  $P$ -value, we reject  $H_0$ . This is strong evidence that there is a higher proportion of varsity athletes among ALS patients than those with other disorders.  
b) Observational retrospective study. To make the inference, one must assume the patients studied are representative.
8. Paired samples; boxplot shows no strong skewness or outliers. One might wonder how the individuals in the study were selected. We are 95% confident that average percentage of 15-year-old males who have been drunk is between 4.5% and 11.4% more than 15-year-old females for these countries. We cannot infer that these percentages are true for other countries.
9.  $H_0: \mu = 7.41$ ;  $H_A: \mu \neq 7.41$ .  $t = 2.18$ ;  $df = 111$ ;  $P$ -value = 0.0313. With such a low  $P$ -value, we reject  $H_0$ . Assuming that Missouri babies fairly represent the United States, these data suggest that American babies are different from Australian babies in birth weight; it appears American babies are heavier, on average.
10. a) 88.6%    b) 82.2%  
c) The petition would be certified when there are not enough valid signatures.  
d) A correct petition is not certified.  
e)  $H_0: p = 0.822$ ;  $H_A: p > 0.822$ .  $z = 7.48$ ;  $P$ -value =  $3.64 \times 10^{-14}$ . With such a low  $P$ -value, we reject  $H_0$ . This sample provides sufficient evidence for certification of the petition.  
f) Increase sample size.
11. a) If there is no difference in the average fish sizes, the chance of seeing an observed difference this large just by natural sampling variation is less than 0.1%.  
b) If cost justified, feed them a natural diet.    c) Type I
12. We have two independent samples, but we don't know how these vehicles were chosen. Even if we consider them representative, the samples are small, and the data for SUVs are skewed to the right. These data are not appropriate for inferences.
13. a) Assuming the conditions are met, from these data we are 95% confident that patients with cardiac disease average between 3.39 and 5.01 years older than those without cardiac disease.

- b) Older patients are at greater risk from a variety of other health issues, and perhaps more depressed.
14. a) We are 95% confident that the proportion of smokers is between 3.67% and 12.93% more in patients with heart disease than in those without heart disease, based on this information.
- b) Yes. The interval lies completely above 0.
- c) Could be a confounding variable. The smokers are at even greater risk for heart attack.
15. a) Stratified sample survey.
- b) We are 95% confident that the proportion of boys who play computer games is between 7.0 and 17.0 percentage points higher than among girls.
- c) Yes. The entire interval lies above 0.
16. a) Based on the data, we are 95% confident that the difference in recruiting is between -4.2% and 8.2%.
- b) No. The interval contains 0 and negative values—the new strategy may actually lower acceptance rates.
17. Based on the data, we are 95% confident that the mean difference in words misunderstood is between -3.76 and 3.10. Because 0 is in the confidence interval, we would conclude that the two tapes could be equivalent.
18.  $H_0: p_{VT} - p_{NH} = 0$ ;  $H_A: p_{VT} - p_{NH} \neq 0$ .  $z = -0.59$ ; P-value = 0.5563. With such a high P-value, we do not reject  $H_0$ . These data show no evidence of a difference in the rates of cesarean deliveries between Vermont and New Hampshire.
19. a)
- 

The countries that appear to be outliers are Spain, Italy, and Portugal. They are all Mediterranean countries.

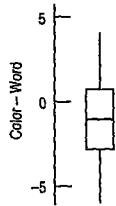
- b)  $H_0: \mu_D = 0$ ;  $H_A: \mu_D > 0$ .  $t = 5.56$ ;  $df = 10$ ; P-value = 0.0001. With such a low P-value, we reject  $H_0$ . These data show that European men are more likely than women to read newspapers.
20.  $H_0: \mu = \$10$ ;  $H_A: \mu > \$10$ .  $t = 0.66$ ;  $df = 13$ . P-value = 0.26. With such a high P-value, we do not reject  $H_0$ . These data do not provide evidence that he is likely to overspend his budget of \$10 per day.
21. We are 95% confident that the proportion of American adults who would agree with the statement is between 57.0% and 63.0%.
22.  $H_0: \mu_{Cert} - \mu_{UC} = 0$ ;  $H_A: \mu_{Cert} - \mu_{UC} > 0$ .  $t = 1.57$ ;  $df = 86$ ; P-value = 0.0598. The P-value is low enough to reject  $H_0$ . These data provide evidence that students of certified teachers achieve higher mean reading scores than students of uncertified teachers.
23. Data are matched pairs (before and after for the same rooms); less than 10% of all rooms in a large hotel; uncertain how these rooms were selected (are they representative?). Histogram shows that differences are roughly unimodal and symmetric, with no outliers. A 95% confidence interval for the difference, before - after, is (0.58, 2.65) counts. Since the entire interval is above 0, these data suggest that the new air-conditioning system was effective in reducing average bacteria counts.
24. There is a significant difference in Math performance; students of certified teachers seem to do better. There is a significant difference in language performance at  $\alpha = 0.05$  between students of certified teachers and those of uncertified teachers.
25. a) We are 95% confident that between 19.77% and 38.66% of children with bipolar symptoms will be helped with medication and psychotherapy, based on this study.
- b) 221 children
26. a)  $H_0: \mu_D = 0$ ;  $H_A: \mu_D \neq 0$ .  $t = 0.381$ ;  $df = 19$ ; P-value = 0.7078. With a high P-value, we do not reject  $H_0$ . There is little evidence that the tests differ in mean difficulty.
- b)  $H_0: \mu_D = 0$ ;  $H_A: \mu_D \neq 0$ . The boxplot of the differences shows three outliers, but they are symmetric. A paired  $t$ -test with

them included shows  $t = -0.2531$  with 19 df and a P-value of 0.8029, showing no evidence of a difference in mean scores. With the three outliers removed, the  $t$ -test shows  $t = -1.1793$  with 16 df and a P-value of 0.2555—still no evidence of a difference in means. We conclude the testing environment does not affect mean score.



27. a) From this histogram, about 115 loaves or more. (Not Normal.) This assumes the last 100 days are typical.
- b) Large sample size; CLT says  $\bar{y}$  will be approximately Normal.
- c) From the data, we are 95% confident that on average the bakery will sell between 101.2 and 104.8 loaves of bread a day.
- d) 25
- e) Yes, 100 loaves per day is too low—the entire confidence interval is above that.
28. a)
- 
- b) *Versicolor* generally has longer petals than *virginica*.
- c) (10.14, 14.46)
- d) We are 95% confident, based on the information given, that the average petal length for *versicolor* irises is between 10.14 and 14.46 mm longer than that of *virginica* irises.
- e) Yes. The entire interval is above 0.
29. a)  $H_0: p_{High} - p_{Low} = 0$ ;  $H_A: p_{High} - p_{Low} \neq 0$ .  $z = -3.57$ ; P-value = 0.0004. Because the P-value is so low, we reject  $H_0$ . These data suggest the IRS risk is different in the two groups; it appears people who consume dairy products often have a lower risk, on average.
- b) Doesn't indicate causality; this is not an experiment.
30. a)  $H_0: p = 0.16$ ;  $H_A: p > 0.16$ .  $z = 4.42$ ; P-value =  $5 \times 10^{-6}$ . Because the P-value is so low, we reject  $H_0$ . These data suggest that the proportion of tickets given to blacks on this section of the New Jersey Turnpike is unusually high.
- b) Doesn't prove it; there may be other factors.
- c) Answers will vary. Possibly what proportion of drivers on this highway is black?
31. Based on these data, we are 95% confident that seeded clouds will produce an average of between -4.76 and 559.56 more acre-feet of rain than unseeded clouds. Since the interval contains negative values, it may be that seeding is unproductive.
32. a)  $H_0: \mu = 35.4$ ;  $H_A: \mu < 35.4$
- b) There is one high value, but it is technically not an outlier. Otherwise the histogram appears roughly unimodal and symmetric (only 6 values).
- c)  $t = 0.726$ ; P-value = 0.7497. With such a large P-value, we do not reject  $H_0$ . The data do not show that the average weight is less than claimed.
- d)  $t = -0.53$ ; P-value = 0.3107. Without the high outlier, the data still do not show that the average weight is significantly less than claimed.
- e) We lack evidence to question the stated weight.
33. a) Randomizing order of the tasks helps avoid bias and memory effects. Randomizing the cards helps avoid bias as well.

- b)  $H_0: \mu_D = 0$ ;  $H_A: \mu_D \neq 0$



Boxplot of the differences looks symmetric with no outliers.

$t = -1.70$ ;  $P\text{-value} = 0.0999$ ; do not reject  $H_0$ , because  $P > 0.05$ . The data do not provide evidence that the color or written word dominates.

34. a) \$373.50  
 b) They are 95% confident that the average loss in a home burglary is between \$1644 and \$2391, based on their sample.  
 c) 95% of all random samples will produce confidence intervals that contain the true mean loss.
35. a) Different samples give different means; this is a fairly small sample. The difference may be due to natural sampling variation.  
 b)  $H_0: \mu = 100$ ;  $H_A: \mu < 100$   
 c) Batteries selected are a SRS (representative); fewer than 10% of the company's batteries; lifetimes are approximately Normal.  
 d)  $t = -1.0$ ;  $P\text{-value} = 0.1666$ ; do not reject  $H_0$ . This sample does not show that the average life of the batteries is significantly less than 100 hours.  
 e) Type II.
36. a) Based on these data, we are 90% confident that the average hamster litter will have between 7.11 and 8.33 babies.  
 b) Larger—to be more confident, we need a wider interval.  
 c) About 27 (based on  $t_{24}^*$ ).

## CHAPTER 26

- Chi-square test of independence. We have one sample and two variables. We want to see if the variable *Account Type* is independent of the variable *Trade Type*.
  - Other test. *Account Size* is quantitative, not counts.
  - Chi-square test of homogeneity. We want to see if the distribution of one variable, *Courses*, is the same for two groups (resident and nonresident students).
- Chi-square goodness-of-fit test. We want to see if the distribution of defects is uniform over the work days.
  - Other test. *Cholesterol Level* is quantitative, not counts.
  - Chi-square test of independence. We have data on two variables, *Political Learning* and *Major*, for one group of students.
- 10
  - Goodness-of-fit
  - $H_0$ : The die is fair (all faces have  $p = 1/6$ ).  
 $H_A$ : The die is not fair.
  - Count data; rolls are random and independent; expected frequencies are all bigger than 5.
  - 5
  - $\chi^2 = 5.600$ ,  $P\text{-value} = 0.3471$
  - Because the  $P\text{-value}$  is high, do not reject  $H_0$ . The data show no evidence that the die is unfair.
- Yellow, red: 21.2; orange, blue, green: 10.6; brown; 31.8
  - Goodness-of-fit
  - $H_0$ : The distribution is as specified by the company.  
 $H_A$ : The distribution is not as specified.
  - Count data; bag may not be a random sample, but most likely representative; expected counts are all bigger than 5.
  - 5
  - $\chi^2 = 9.315$ ,  $P\text{-value} = 0.0972$
  - Because the  $P\text{-value}$  is high, do not reject  $H_0$ . These data do not provide evidence that the distribution is other than specified.
- Weights are quantitative, not counts.
  - Count the number of each kind of nut, assuming the company's percentages are based on counts rather than weights.
- Data are averages, not counts.
- $H_0$ : The police force represents the population (29.2% white, 28.2% black, etc.).  $H_A$ : The police force is not representative of the population.  $\chi^2 = 16516.88$ ,  $df = 4$ ,  $P\text{-value} = 0.0000$ . Because the  $P\text{-value}$  is so low, we reject  $H_0$ . These data show that the police force is not representative of the population. In particular, there are too many white officers in relationship to their membership in the community.
- $H_0$ : Murders among women have the same distribution of weapons as all murders (63.4% guns, etc.).  $H_A$ : Murders among women have a different distribution of weapons than all murders.  $\chi^2 = 389.54$ ,  $df = 3$ ,  $P\text{-value} < 0.0001$ . Because the  $P\text{-value}$  is so low, we reject  $H_0$ . Women's murders do not have the same distribution of weapons as all murders nationwide. Women are much less likely to be killed by other weapons and more likely to be killed by personal attack.
- $\chi^2 = 5.671$ ,  $df = 3$ ,  $P\text{-value} = 0.1288$ . With a  $P\text{-value}$  this high, we fail to reject  $H_0$ . Yes, these data are consistent with those predicted by genetic theory.
  - $\chi^2 = 11.342$ ,  $df = 3$ ,  $P\text{-value} = 0.0100$ . Because of the low  $P\text{-value}$ , we reject  $H_0$ . These data provide evidence that the distribution is not as specified by genetic theory.
  - With small samples, many more data sets will be consistent with the null hypothesis. With larger samples, small discrepancies will show evidence against the null hypothesis.
- $H_0$ : Digits are all equally likely (all occur with frequency  $1/10$ ).  
 $H_A$ : Digits are not all equally likely.  $\chi^2 = 5.509$ ,  $df = 9$ ,  $P\text{-value} = 0.7879$ . Because the  $P\text{-value}$  is large, we do not reject  $H_0$ . These data provide no evidence that the digits in  $\pi$  are not all equally likely.
- $96/16 = 6$
  - Goodness of Fit
  - $H_0$ : The number of large hurricanes remains constant over decades.  
 $H_A$ : The number of large hurricanes has changed.
  - 15
  - $P\text{-value} = 0.63$
  - The very high  $P\text{-value}$  means these data offer no evidence that the numbers of large hurricanes has changed.
  - The final period is only 6 years rather than 10 and already 7 large hurricanes have been observed. Perhaps this decade will have an unusually large number of such hurricanes.
- Goodness of Fit
  - $655/49 = 13.367$
  - $H_0$ : All numbers are equally likely;  
 $H_A$ : Some numbers are more likely than others.
  - 48
  - 0.93
  - The very high  $P\text{-value}$  means these data offer no evidence that some numbers are more likely to come up than others.
- Independence
  - $H_0$ : Breastfeeding success is independent of having an epidural.  
 $H_A$ : There's an association between breastfeeding success and having an epidural.
- Homogeneity
  - $H_0$ : The same proportion of articles used statistics in the three time periods surveyed.  $H_A$ : The proportion of articles using statistics changed over time.
- 1
  - 159.34
  - Breastfeeding behavior should be independent for these babies. They are fewer than 10% of all babies; we assume they are representative. We have counts, and all the expected counts are at least 5.
- 2
  - 21.85
  - These are counted data. One article shouldn't affect another (except perhaps for the rare article based on a previous one in an earlier year cohort included in this study). We can regard the selected years as representative of those nearby, and the authors (judging by their title) seem to want to regard these articles as representative of those appearing in other similar-quality medical journals, so they're fewer than 10% of all articles. All expected counts are at least 5.