1. HYPOTHESES TESTING FOR PROPORTIONS

1. A clean standard requires that vehicle exhaust not exceed specified limits for various pollutants. Many states require that cars be tested annually to be sure they meet these standards. Suppose state regulators double check a random sample of cars that a suspect repair shop has certified as okay. They will revoke the shop's license if they find significant evidence that the shop is certifying vehicle that do not meet standards.

   The appropriate hypotheses are:

   A. **Null hypothesis:** The regulators decide that the shop is meeting standards.
      **Alternative hypothesis:** The regulators decide that the shop is not meeting standards.
   B. **Null hypothesis:** The shop’s emission standards are higher than for other shops.
      **Alternative hypothesis:** The shop’s emission standards are not higher than for other shops.
   C. **Null hypothesis:** The shop is meeting the emissions standards.
      **Alternative hypothesis:** The shop is not meeting the emission standards.
   D. Null hypothesis: The repair shop’s license is revoked.
      Alternative hypothesis: The repair shop’s license is not revoked.

QUESTIONS 2 - 4
The National center for Education Statistics monitors many aspects of elementary and secondary education nationwide. Their 1996 numbers are often used as a baseline to assess changes. In 1996 31% of students reported that their mothers had graduated from college. In 2000, responses from 8368 students found that this figure had grown to 32%. Is this evidence that there has been an increase in education level among mothers?

2. The appropriate hypotheses are:

   A. $H_0: \ p = 0.31; \ H_A: \ p > 0.31$
   B. $H_0: \ p>0.31; \ H_A: \ p<0.31$
   C. $H_0: \ p = 0.32; \ H_A: \ p \neq 0.32$
   D. $H_0: \ p<0.31; \ H_A: \ p>0.32$

3. Let all necessary conditions be satisfied. Find the test statistic (z), and the P-value.

   A. $z = 1.978$, P-value = 0.047
   B. $z = 1.978$, P-value = 0.024
   C. $z = 0.0051$, P-value = 1.978
   D. $z = 0.048$, P-value = 1.978

4. State your conclusion.

   A. With a P-value of 1.978, we fail to reject the null hypothesis.
   B. No conclusion can be drawn because the P-value is too big to be considered in this context.
   C. **With a P-value of 0.024, we reject the null hypothesis. There is evidence to suggest that the percentage of students whose mothers are college graduates has changed since 1996. In fact, the evidence suggests that the percentage has increased.**
   D. The confidence level was not given, so no conclusion can be arrived at.
5. National data in the 1960s showed that about 44% of the adult population had never smoked cigarettes. In 1995 a national health survey interviewed a random sample of 881 adults and found that 52% had never been smokers.

(a) Create a 95% confidence interval for the proportion of adults (in 1995) who had never been smokers.
(b) Does this provide evidence of a change in behavior among Americans? Using your confidence interval, test an appropriate hypothesis and state your conclusion.

ANSWERS
(a) (0.1531, 0.5789)  (b) Null Hypo: p = 0.44; Alter Hypo: p not equal to 0.44. Since 44% is not in the 95% CI, we will reject the Null Hypo. There is strong evidence that, in 1995, the percentage of adults who have never smoked was not 44%.

6. A company hopes to improve customer satisfaction, setting as a goal no more than 5% negative comments. A random survey of 350 customers found only 10 with complaints.

(a) Create a 95% confidence interval for the true level of dissatisfaction among customers.
(b) Does this provide evidence that the company has reached its goal? Using your confidence interval, test an appropriate hypothesis and state your conclusion.

ANSWERS
(a) (0.011, 0.046)  (b) Null Hypo: p = 0.05; Alter. Hypo: p < 0.05. Since 5% is not in the 95% CI, we will reject the Null Hypo. There is strong evidence that less that 5% of customers have complaints.

7. A company with a fleet of 150 cars found that the emissions systems of 7 out of 22 they tested failed to meet pollution control guidelines. Is this strong evidence that more than 20% of the fleet might be out of compliance? Test an appropriate hypothesis and state your conclusion. Be sure the appropriate assumptions and conditions are satisfied before you proceed.

ANSWER

Null Hypo: p = 0.20; Alter. Hypo: p > 0.20. Two conditions are not satisfied (verify!!!!!!!) We cannot proceed with a hypothesis test.

8. In a rural area only about 30% of the wells that are drilled find adequate water at a depth of 100 feet or less. A local man claims to be able to find water by “dowsing” – using a forked stick to indicate where the well should be drilled. You check with 80 of his customers and find that 27 have wells less than 100 feet deep. What do you conclude about this claim? (We consider a P – value of around 5% to represent strong evidence)

(a) Write appropriate hypotheses.
(b) Check the necessary assumptions.
(c) Perform the mechanics of the test. What is the P – value?
(d) Explain carefully what the P – value means in this context.
(e) What’s your conclusion?

FOR SOLUTION TO PROBLEM 8, SEE CLASS NOTES.

9. In the 1980s it was generally believed that congenital abnormalities affected about 5% of the nation’s children. Some people believe that the increase in the number of chemicals in the environment has led to an increase in the incidence of abnormalities. A recent study examined 384 children and found that 46 of them showed signs of an
abnormality. Is this strong evidence that the risk has increased? (We consider a P-value of around 5% to represent strong evidence)

(a) Write appropriate hypotheses.
(b) Check the necessary assumptions.
(c) Perform the mechanics of the test. What is the P-value?
(d) Explain carefully what the P-value means in this context.
(e) What's your conclusion?
(f) Do environmental chemicals cause congenital abnormalities?

FOR SOLUTION TO PROBLEM 9 SEE CLASS NOTES.

QUESTIONS 10 – 12

The National Center for Education Statistics monitors many aspects of elementary and secondary education nationwide. Their 1996 numbers are often used as a baseline to assess changes. In 1996, 34% of students had not been absent from school even once during the previous month. In the 2000 survey, responses from 8302 students showed that this figure had slipped to 33%. Officials would of course be concerned if student attendance were declining. Do these figures give evidence of a decrease in student attendance?

10. Write appropriate hypotheses.

A. \( H_0 : p = 0.33; \) \( H_A : p < 0.33 \)
B. \( H_0 : p > 0.34; \) \( H_A : p < 0.34 \)
C. \( H_0 : p = 0.34; \) \( H_A : p < 0.34 \)
D. \( H_0 : p < 0.33; \) \( H_A : p = 0.34 \)
E. \( H_0 : p = 0.34; \) \( H_A : p = 0.33 \)

11. Perform the test by finding the Test statistics and the P-value.

A. \( Z_0 = 0.0052; \text{P-value} = 0.027 \)
B. \( Z_0 = -1.923; \text{P-value} = 0.027 \)
C. \( Z_0 = 0.048; \text{P-value} = 0.0052 \)
D. \( Z_0 = 0.01; \text{P-value} = 0.048 \)
E. \( Z_0 = 1.923; \text{P-value} = 0.027 \)

12. State your conclusion.

A. With a \( P \)-value of 0.027, we fail to reject the null hypothesis. There is no evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
B. With a \( P \)-value of 0.0052, we reject the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
C. With a \( P \)-value of 0.048, essentially 0.05, equal to the significance level, we retain the null hypothesis. There is no evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
D. With a \( P \)-value of 0.027, we reject the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
E. With a \( P \)-value of 0.027, we retain the null hypothesis. There is evidence to suggest that the percentage of students with perfect attendance in the previous month has decreased in 2000.
In 2004, ACT Inc. reported that 74% of 1644 randomly selected college freshmen returned to college the next year. You are interested in estimating the national freshman-to-sophomore retention rate.

13. Construct a 98% confidence interval (CI) and interpret your interval.

A. CI = (0.715, 0.765); Between 71.5% and 76.5% of all college students return to college after freshman year.
B. CI = (0.765, 0.715); The proportion of all college students who return to college after freshman year is between 0.715 and 0.765.
C. CI = (0.011, 0.02); We are 98% confident that between 1.1% and 2% of all college students return to college after their freshman year.
D. CI = (0.74, 0.98); We are 98% confident that between 74% and 98% of all college students return to college after their freshman year.
E. CI = (0.715, 0.765); We are 98% confident that between 71.5% and 76.5% of all college students return to college after their freshman year.


A. If we were to select repeated samples like this, we would expect about 98% of the confidence intervals we created to contain the true proportion of all college students who return to college after their freshman year.
B. 98% of the confidence intervals contain the true proportion of all college students who return to college after their freshman year.
C. If we were to take several samples of the same size, we would expect just 2% of the confidence intervals to contain the true proportion of all college students who return to college after their freshman year.
D. We are 98% confident that all college students will return to college after their freshman year.
E. A, B, and D above.

15. Write the appropriate hypotheses.

A. H₀: The new signs provide the same visibility as the old signs.
   H₁: The new signs provide greater visibility than the old signs.
B. H₀: The new signs provide greater visibility than the old signs.
   H₀: The new signs provide the same visibility as the old signs.
C. H₀: The new signs provide the same visibility as the old signs.
   H₁: The new signs provided lesser visibility than the old signs.
D. H₀: The new signs provide lesser visibility than the old signs.
   H₀: The new signs provide the same visibility as the old signs.
E. H₀: The new signs provide the same visibility as the old signs.
   H₁: The new signs do not provide the same visibility as the old signs.
16. In this context, what would a Type I error be?

A. Type I error happens when the engineers decide that the new signs are not more visible when they actually are more visible.
B. **Type I error happens when the engineers decide that the new signs are more visible when they are not more visible.**
C. Type I error happens when the engineers decide that the new signs provide the same visibility as the old signs when in fact, the new signs provide lesser visibility than the old signs.
D. Type I error happens when the engineers decide that the old signs provide the same visibility as the new signs when they provide more visibility than the new signs.
E. Type I error happens when the highway safety engineers fail to persuade people to drive through a test course with several of the new and old style signs and rate which kind shows up the best.

17. In this context, what would a Type II error be?

A. Type II error happens when the highway safety engineers fail to persuade people to drive through a test course with several of the new and old style signs and rate which kind shows up the best.
B. Type II error happens when the engineers decide that the old signs provide the same visibility as the new signs when they provide more visibility than the new signs.
C. Type II error happens when the engineers decide that the new signs provide the same visibility as the old signs when in fact, the new signs provide lesser visibility than the old signs.
D. Type II error happens when the engineers decide that the new signs are more visible when they are not more visible.
E. **Type II error happens when the engineers decide that the new signs are not more visible when they actually are more visible.**

18. In this context, describe what is mean by the power of the test.

A. The power of the test is the probability that the engineers detect a sign that is falsely more visible.
B. The power of the test is the probability that the engineers detect that the new signs have the same visibility as the old signs.
C. **The power of the test is the probability that the engineers detect a sign that is truly more visible.**
D. The power of the test is the probability that the engineers reject a false null hypothesis.
E. The power of the test is the probability that the engineers retain a false null hypothesis.

19. Is this a one – tailed (left – tailed or right – tailed) or a two – tailed test? Why?

A. The test is left – tailed, because we are only interested in whether or not the signs are less visible. If the new design is more visible, we do not care how much more visible it is.
B. The test is two – tailed, because we are interested in whether or not the signs are more visible. If the new design is less visible, we do not care how much less visible it is.
C. The test is left – tailed, because we are only interested in whether or not the signs are more visible. If the new design is less visible, we do not care how much less visible it is.
D. **The test is right – tailed, because we are only interested in whether or not the signs are more visible. If the new design is less visible, we do not care how much less visible it is.**
E. The test is two – tailed, because the P – value is unknown. If we know the P – value, we would compare with the significance level = 0.05 and make the right decision as to whether the test is left – tailed or right – tailed.
20. If the hypothesis is tested at the 1% level of significance instead of 5%, how will this affect the power of the test?

A. **When the level of significance is dropped from 5% to 1%, power decreases. The null hypothesis is harder to reject, since more evidence is required.**

B. When the level of significance is dropped from 5% to 1%, power increases. The null hypothesis is harder to reject, since more evidence is required.

C. When the level of significance is dropped from 5% to 1%, power stays the same. This is because power of a test is affected by the P-value rather than the significance level.

D. When the level of significance is dropped from 5% to 1%, power = 0.04. This is because the power of a test is affected by the significance level.

E. When the level of significance is dropped from 5% to 1%, the power = 0.04*(P-value).

21. The engineers hoped to base their decision on the reactions of 50 drivers, but time and budget constraints may force them to cut back to 20. How would this affect the power of the test? Explain.

A. If a sample of size 20 is used instead of 50, power will increase. A smaller sample size has more variability, lowering the ability of the test to detect falsehoods.

B. **If a sample of size 20 is used instead of 50, power will decrease. A smaller sample size has more variability, lowering the ability of the test to detect falsehoods.**

C. If a sample of size 20 is used instead of 50, the power will stay the same. The sample size n = 20 is too small and the 10% condition is not satisfied.

D. If a sample of size 20 is used instead of 50, the power will stay the same. The sample size n = 20 is too small and the Near Normal Condition fails to be satisfied.

E. If a sample of size 20 is used instead of 50, the power will stay the same. The sample size n = 20 is too small and the Success – Failure fails to be satisfied.

**QUESTIONS 22 – 25**

A company with a large fleet of cars hopes to keep gasoline costs down and sets a goal of attaining a fleet average of at least 26 miles per gallon. To see if the goal is being met, they check the gasoline usage for 50 company trips chosen at random, finding a mean of 25.36 mpg and a standard deviation of 4.21 mpg. Is this strong evidence they have failed to attain their fuel economy goal?

22. Write appropriate hypotheses.

A. \( H_0: \mu = 26; \ H_A: \mu > 26 \)

B. \( H_0: \mu > 26; \ H_A: \mu = 26 \)

C. \( H_0: \mu < 26; \ H_A: \mu = 26 \)

D. **\( H_0: \mu = 26; \ H_A: \mu < 26 \)**

E. \( H_0: \mu = 26; \ H_A: \mu \neq 26 \)

23. Are the necessary assumptions to perform inference satisfied?

A. No, the 10% condition is not satisfied.

B. No, the randomization condition is not satisfied.

C. No, the nearly normal condition is not satisfied.

D. No, the success failure condition is not satisfied.

E. **Yes, all the necessary assumptions and conditions are satisfied or can be assumed to be satisfied.**
24. Find the P-value (round to three decimal places) and explain what the P-value means in this context.

A. P-value = 0.144; The P-value is the probability of obtaining a sample mean of 25.36 mpg or more if the mean mileage of the fleet is 26 mpg.
B. P-value = 1.07494; The P-value is the probability of obtaining a sample mean of 25.36 mpg or less if the mean mileage of the fleet is 26 mpg.
C. P-value = -1.07494; The P-value is the probability of obtaining a sample mean of 25.36 mpg or less if the mean mileage of the fleet is 26 mpg.
D. P-value = 0.59538; The P-value is the probability that the mean mileage of the fleet is greater than or equal to the proposed mean of 26 mpg.
E. P-value = 0.144; The P-value is the probability of obtaining a sample mean of 25.36 mpg or less if the mean mileage of the fleet is 26 mpg.

25. State an appropriate conclusion.

A. Since the P-value = 0.144, we reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.
B. Since the P-value = 0.144, we fail to reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.
C. Since the P-value = -1.07494, a negatively high value, we fail to reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.
D. Since the P-value = 1.07494, a positively high value, we fail to reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.
E. Since the P-value = 0.59538, a relatively small value, we reject the null hypothesis. There is strong evidence to suggest that the mean mileage of cars in the fleet is more than 26 mpg.

26. Several factors are involved in the creation of a confidence interval. Among them are the sample size, the level of confidence, and the margin of error. Which of the following statements is incorrect?

A. For a given sample size, reducing the margin of error will mean lower confidence.
B. For a given sample size, higher confidence means a smaller margin of error.
C. For a certain confidence level, you can get a smaller margin of error by selecting a bigger sample.
D. For a fixed margin of error, smaller samples will mean lower confidence.
E. For a given confidence level, a sample 9 times as large will make a margin of error on third as big.

QUESTIONS 27 - 28

27. A nutrition laboratory tests 50 “reduced sodium” hot dogs, finding that the mean sodium content is 321 mg, with a standard deviation of 36 mg. To construct a 95% confidence interval for the mean sodium content of this brand of hot dog, what assumptions would you make in this inference?

A. It is assumed that the hot dog weights are not multimodal and that the distribution of the population of hot dog weights does not contain any outliers.
B. It is assumed that the hot dog weights are random and that the distribution of the population of hot dog weights is not biased.
C. It is assumed that the hot dog weights are independent and that the distribution of the population of hot dog weights is normal.
D. It is assumed that np = 50·0.95 = 47.5 > 10 and that the distribution of the population of hot dog weights does not contain any outliers.
E. A and D above.
28. Which of the following condition(s) is (are) satisfied for inference?

I. Randomization condition
II. 10% condition
III. Nearly normal condition
IV. Success and Failure condition

A. I, II, and IV only
B. II and III only
C. I, III, and IV only
D. I, II, and III only
E. I, II, III, and IV

29. Production managers on an assembly line must monitor the output to be sure that the level of defective products remains small. They periodically inspect a random sample of the items produced. If they find a significant increase in the proportion of items that must be rejected, they will halt the assembly process until the problem can be identified and repaired.
(a) In this context, what is a Type I error?
(b) In this context, what is a Type II error?
(c) Which type of error would the factory owner consider more serious?
(d) Which type of error might customers consider more serious?

[Solution to problem 29 is posted on the class website]

30. Consider again the task of the quality control inspectors in Exercise 29.
(a) In this context, what is meant by the power of the test the inspectors conduct?
(b) They are currently testing 5 items each hour. Someone has proposed that they test 10 each hour instead. What are the advantages and disadvantages of such a change?
(c) Their test currently uses a 5% level of significance. What are the advantages and disadvantages of changing to an alpha level of 1%?
(d) Suppose that, as a day passes, one of the machines on the assembly line produces more and more items that are defective. How will this affect the power of the test?

[Solution to problem 30 is posted on the class website]

34. Doritos. Some students checked 6 bags of Doritos marked with a net weight of 28.3 grams. They carefully weighed the contents of each bag, recording the following weights (in grams): 29.2, 28.5, 28.7, 28.9, 29.1, 29.5.
   a) Do these data satisfy the assumptions for inference? Explain.
   b) Find the mean and standard deviation of the observed weights.
   c) Create a 95% confidence interval for the mean weight of such bags of chips.
Normal temperature. The researcher described in Exercise 9 also measured the body temperatures of that randomly selected group of adults. The data he collected are summarized below. We wish to estimate the average (or “normal”) temperature among the adult population.

<table>
<thead>
<tr>
<th>Summary</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count</td>
<td>52</td>
</tr>
<tr>
<td>Mean</td>
<td>98.285</td>
</tr>
<tr>
<td>Median</td>
<td>98.200</td>
</tr>
<tr>
<td>MidRange</td>
<td>98.600</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.6924</td>
</tr>
<tr>
<td>Range</td>
<td>2.800</td>
</tr>
<tr>
<td>IntQRRange</td>
<td>1.050</td>
</tr>
</tbody>
</table>

- a) Are the necessary conditions for a t-interval satisfied? Explain.
- b) Find a 98% confidence interval for mean body temperature.
- c) Explain the meaning of that interval.
- d) Explain what “98% confidence” means in this context.
- e) 98.6°F is commonly assumed to be “normal.” Do these data suggest otherwise? Explain.

### Parking

Hoping to lure more shoppers downtown, a city builds a new public parking garage in the central business district. The city plans to pay for the structure through parking fees. During a two-month period (44 weekdays), daily fees collected averaged $126, with a standard deviation of $15.

- a) What assumptions must you make in order to use these statistics for inference?
- b) Write a 90% confidence interval for the mean daily income this parking garage will generate.
- c) Explain in context what this confidence interval means.
- d) Explain what “90% confidence” means in this context.
- e) The consultant who advised the city on this project predicted that parking revenues would average $130 per day. Based on your confidence interval, do you think the consultant was correct? Why?

Since the interval is completely below the $130 predicted by the consultant, there is evidence that the average daily parking revenue is lower than $130.
21. **Departures.** What are the chances your flight will leave on time? The U.S. Bureau of Transportation Statistics of the Department of Transportation publishes information about airline performance. Here are a histogram and summary statistics for the percentage of flights departing on time each month from 1995 thru 2006.

![Histogram of On Time Departure%](image)

- \( n = 144 \)
- \( \bar{y} = 81.1838 \)
- \( s = 4.47094 \)

There is no evidence of a trend over time. (The correlation of On Time Departure% with time is \( r = -0.016 \).)

a) Check the assumptions and conditions for inference.
b) Find a 90% confidence interval for the true percentage of flights that depart on time.
c) Interpret this interval for a traveler planning to fly.

22. **Late arrivals.** Will your flight get you to your destination on time? The U.S. Bureau of Transportation Statistics reported the percentage of flights that were late each month from 1995 through 2006. Here's a histogram, along with some summary statistics:

![Histogram of Late Arrival%](image)

- \( n = 144 \)
- \( \bar{y} = 20.0737 \)
- \( s = 4.08837 \)

We can consider these data to be a representative sample of all months. There is no evidence of a time trend (\( r = -0.07 \)).

a) Check the assumptions and conditions for inference about the mean.
b) Find a 99% confidence interval for the true percentage of flights that arrive late.
c) Interpret this interval for a traveler planning to fly.
7. Meal plan. After surveying students at Dartmouth College, a campus organization calculated that a 95% confidence interval for the mean cost of food for one term (of three in the Dartmouth trimester calendar) is ($1102, $1290). Now the organization is trying to write its report and is considering the following interpretations. 

Which of the interpretations is correct? 

a) 95% of all students pay between $1102 and $1290 for food. **Not correct**  
b) 95% of the sampled students paid between $1102 and $1290. **Not correct**  
c) We’re 95% sure that students in this sample averaged between $1102 and $1290 for food. **Not correct**  
d) 95% of all samples of students will have average food costs between $1102 and $1290. **Not correct**  
e) We’re 95% sure that the average amount all students pay is between $1102 and $1290. **Correct**

8. Rain. Based on meteorological data for the past century, a local TV weather forecaster estimates that the region’s average winter snowfall is 23", with a margin of error of ±2 inches. Assuming he used a 95% confidence interval, how should viewers interpret this news? 

Which of the interpretations is correct? 

a) During 95 of the last 100 winters, the region got between 21" and 25" of snow. **Not correct**  
b) There’s a 95% chance the region will get between 21" and 25" of snow this winter. **Not correct**  
c) There will be between 21" and 25" of snow on the ground for 95% of the winter days. **Not correct**  
d) Residents can be 95% sure that the area’s average snowfall is between 21" and 25". **Correct**  
e) Residents can be 95% confident that the average snowfall during the last century was between 21" and 25" per winter. **Not correct**

9. Pulse rates. A medical researcher measured the pulse rates (beats per minute) of a sample of randomly selected adults and found the following Student’s t-based confidence interval:

With 95.00% Confidence, 
70.887864 < μ(Pulse) < 74.697011

\[
\frac{(74.6 - 70.9)}{2} = 1.8 \text{ beats/min}
\]

a) Explain carefully what the software output means. 

b) What’s the margin of error for this interval? 

c) If the researcher had calculated a 99% confidence interval, would the margin of error be larger or smaller? Explain.

10. Crawling. Data collected by child development scientists produced this confidence interval for the average age (in weeks) at which babies begin to crawl:

\[
\text{t-Interval for } \mu \quad 29.202 < \mu(\text{age}) < 31.844
\]

95.00% Confidence:

\[
\frac{(31.8 - 29.2)}{2} = 2.6 \text{ weeks}
\]

a) Explain carefully what the software output means. 

b) What is the margin of error for this interval? 

c) If the researcher had calculated a 90% confidence interval, would the margin of error be larger or smaller? Explain.
Solutions to Type I, Type II Errors Problems

II. Homeowners.

a) The null hypothesis is that the level of home ownership does not rise. The alternative hypothesis is that it rises.

b) In this context, a Type I error is when the city concludes that home ownership is on the rise, but in fact, the tax breaks don't help.

c) In this context, a Type II error is when the city abandons the tax breaks, thinking they don't help, when in fact they were helping.

d) A Type I error causes the city to forego tax revenue, while a Type II error withdraws help from those who might have otherwise been able to buy a house.

e) The power of the test is the city's ability to detect an actual increase in home ownership.

III. Alzheimer's

a) The null hypothesis is that a person is healthy. The alternative is that they have Alzheimer's disease. There is no parameter of interest here.

b) A Type I error is a false positive. It has been decided that the person has Alzheimer's disease when they don't.

c) A Type II error is a false negative. It has been decided that the person is healthy, when they actually have Alzheimer's disease.

d) A Type I error would require more testing, resulting in time and money lost. A Type II error would mean that the person did not receive the treatment they needed. A Type II error is much worse.

e) The power of this test is the ability of the test to detect patients with Alzheimer's disease. In this case, the power can be computed as $1 - P(\text{Type II error}) = 1 - 0.08 = 0.92$.

IV. Testing cars.

$H_0$: The shop is meeting the emissions standards.

$H_A$: The shop is not meeting the emissions standards.

a) Type I error is when the regulators decide that the shop is not meeting standards when they actually are meeting the standards.

b) Type II error is when the regulators certify the shop when they are not meeting the standards.

c) Type I would be more serious to the shop owners. They would lose their certification, even though they are meeting the standards.

d) Type II would be more serious to environmentalists. Shops are allowed to operate, even though they are allowing polluting cars to operate.
Quality control.

14 H₀: The assembly process is working fine.
H₁: The assembly process is producing defective items.

a) Type I error is when the production managers decide that there has been an increase in the number of defective items and stop the assembly line, when the assembly process is working fine.

b) Type II error is when the production managers decide that the assembly process is working fine, but defective items are being produced.

c) The factory owner would probably consider Type II error to be more serious, depending of the costs of shutting the line down. Generally, because of warranty costs and lost customer loyalty, defects that are caught in the factory are much cheaper to fix than defects found after items are sold.

d) Customers would consider Type II error to be more serious, since customers don't want to buy defective items.

Cars again.

15 a) The power of the test is the probability of detecting that the shop is not meeting standards when they are not.

b) The power of the test will be greater when 40 cars are tested. A larger sample size increases the power of the test.

c) The power of the test will be greater when the level of significance is 10%. There is a greater chance that the null hypothesis will be rejected.

d) The power of the test will be greater when the shop is out of compliance “a lot”. Larger problems are easier to detect.

Production.

16 a) The power of the test is the probability that the assembly process is stopped when defective items are being produced.

b) An advantage of testing more items is an increase in the power of the test to detect a problem. The disadvantages of testing more items are the additional cost and time spent testing.

c) An advantage of lowering the alpha level is that the probability of stopping the assembly process when everything is working fine (committing a Type I error) is decreased. A disadvantage is that the power of the test to detect defective items is also decreased.

d) The power of the test will increase as a day passes. Bigger problems are easier to detect.
Equal opportunity?

$H_0$: The company is not discriminating against minorities.
$H_A$: The company is discriminating against minorities.

a) This is a one-tailed test. They wouldn’t sue if “too many” minorities were hired.

b) Type I error would be deciding that the company is discriminating against minorities when they are not discriminating.

c) Type II error would be deciding that the company is not discriminating against minorities when they actually are discriminating.

d) The power of the test is the probability that discrimination is detected when it is actually occurring.

e) The power of the test will increase when the level of significance is increased from 0.01 to 0.05.

f) The power of the test is lower when the lawsuit is based on 37 employees instead of 87. Lower sample size leads to less power.