Chapter 20 – Testing Hypotheses about Proportions

1. Hypotheses.

a) H_0 : The governor's "negatives" are 30%. ($p = 0.30$)
H_A : The governor's "negatives" are less than 30%. ($p < 0.30$)

b) H_0 : The proportion of heads is 50%. ($p = 0.50$)
H_A : The proportion of heads is not 50%. ($p \neq 0.50$)

c) H_0 : The proportion of people who quit smoking is 20%. ($p = 0.20$)
H_A : The proportion of people who quit smoking is greater than 20%. ($p > 0.20$)

2. More hypotheses.

a) H_0 : The proportion of high school graduates is 40%. ($p = 0.40$)
H_A : The proportion of high school graduates is not 40%. ($p \neq 0.40$)

b) H_0 : The proportion of cars needing transmission repair is 20%. ($p = 0.20$)
H_A : The proportion of cars needing transmission repair is less than 20%. ($p < 0.20$)

c) H_0 : The proportion of people who like the flavor is 60%. ($p = 0.60$)
H_A : The proportion of people who like the flavor is greater than 60%. ($p > 0.60$)

3. Negatives.

Statement d is the correct interpretation of a *P*-value.

4. Dice.

Statement d is the correct interpretation of a *P*-value.

5. Relief.

It is *not* reasonable to conclude that the new formula and the old one are equally effective. Furthermore, our inability to make that conclusion has nothing to do with the *P*-value. We can not prove the null hypothesis (that the new formula and the old formula are equally effective), but can only fail to find evidence that would cause us to reject it. All we can say about this *P*-value is that there is a 27% chance of seeing the observed effectiveness from natural sampling variation if the new formula and the old one are equally effective.

6. Cars.

It is reasonable to conclude that a greater proportion of high schoolers have cars. If the proportion were no higher than it was a decade ago, there is only a 1.7% chance of seeing such a high sample proportion just from natural sampling variability.

7. He cheats!

a) Two losses in a row aren't convincing. There is a 25% chance of losing twice in a row, and that is not unusual.

b) If the process is fair, three losses in a row can be expected to happen about 12.5% of the time. (0.5)(0.5)(0.5) = 0.125.

c) Three losses in a row is still not a convincing occurrence. We'd expect that to happen about once every eight times we tossed a coin three times.

d) Answers may vary. Maybe 5 times would be convincing. The chances of 5 losses in a row are only 1 in 32, which seems unusual.

8. Candy.

a) $P(\text{first three vanilla}) = \left(\frac{6}{12}\right)\left(\frac{5}{11}\right)\left(\frac{4}{10}\right) \approx 0.091$

b) It seems reasonable to think there really may have been six of each. We would expect to get three vanillas in a row about 9% of the time. That's unusual, but not *that* unusual.

c) If the fourth candy was also vanilla, we'd probably start to think that the mix of candies was not 6 vanilla and 6 peanut butter. The probability of 4 vanilla candies in a row is:

$$P(\text{first four vanilla}) = \left(\frac{6}{12}\right)\left(\frac{5}{11}\right)\left(\frac{4}{10}\right)\left(\frac{3}{9}\right) \approx 0.03$$

We would only expect to get four vanillas in a row about 3% of the time. That's unusual.

9. Cell phones.

1) Null and alternative hypotheses should involve p, not $\hat{p}$.

2) The question is about *failing* to meet the goal. H_A should be $p < 0.96$.

3) The student failed to check $nq = (200)(0.04) = 8$. Since $nq < 10$, the Success/Failure condition is violated. Similarly, the 10% Condition is not verified.

4) $SD(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.96)(0.04)}{200}} \approx 0.014$. The student used $\hat{p}$ and $\hat{q}$.

5) Value of z is incorrect. The correct value is $z = \frac{0.94 - 0.96}{0.014} \approx -1.43$.

6) P-value is incorrect. $P = P(z < -1.43) = 0.076$

7) For the P-value given, an incorrect conclusion is drawn. A P-value of 0.12 provides no evidence that the new system has failed to meet the goal. The correct conclusion for the corrected P-value is: Since the P-value of 0.076 is fairly low, there is weak evidence that the new system has failed to meet the goal.

10. Got milk?

1) Null and alternative hypotheses should involve p, not $\hat{p}$.

2) The question asks if there is evidence that the 90% figure is *not accurate*, so a two-sided alternative hypothesis should be used. H_A should be $p \neq 0.90$.

3) One of the conditions checked appears to be $n > 10$, which is not a condition for hypothesis tests. The Success/Failure Condition checks $np = (750)(0.90) = 675 > 10$ and $nq = (750)(0.10) = 75 > 10$. Also, the 10% condition is not verified.

4) $SD(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.90)(0.10)}{750}} \approx 0.011$. The student used rounded values of $\hat{p}$ and $\hat{q}$.

5) Value of z is incorrect. The correct value is $z = \dfrac{0.876 - 0.90}{0.011} \approx -2.18$.

6) The P-value calculated is in the wrong direction. To test the given hypothesis, the lower-tail probability should have been calculated. The correct, two-tailed P-values is $P = 2P(z < -2.18) = 0.029$.

7) The P-value is misinterpreted. Since the P-value is so low, there is moderately strong evidence that the proportion of adults who drink milk is different than the claimed 90%. In fact, our sample suggests that the proportion may be lower. There is only a 2.9% chance of observing a $\hat{p}$ as far from 0.90 as this simply from natural sampling variation.

11. Dowsing.

a) H_0 : The percentage of successful wells drilled by the dowser is 30%. ($p = 0.30$)
H_A : The percentage of successful wells drilled by the dowser is greater than 30%. ($p > 0.30$)

b) **Independence assumption:** There is no reason to think that finding water in one well will affect the probability that water is found in another, unless the wells are close enough to be fed by the same underground water source.
Randomization condition: This sample is not random, so hopefully the customers you check with are representative of all of the dowser's customers.
10% condition: The 80 customers sampled may be considered less than 10% of all possible customers.
Success/Failure condition: $np = (80)(0.30) = 24$ and $nq = (80)(0.70) = 56$ are both greater than 10, so the sample is large enough.

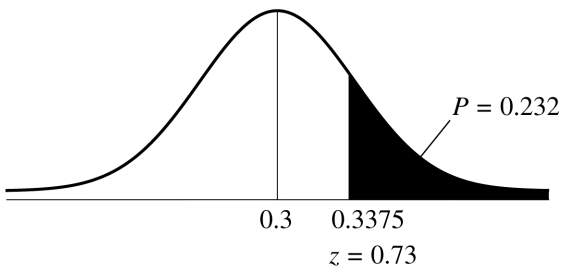
c) The sample of customers may not be representative of all customers, so we will proceed cautiously. A Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.30$ and $\sigma(\hat{p}) = \sqrt{\dfrac{pq}{n}} = \sqrt{\dfrac{(0.30)(0.70)}{80}} \approx 0.0512$.

We can perform a one-proportion z-test. The observed proportion of successful wells is $\hat{p} = \dfrac{27}{80} = 0.3375$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\dfrac{pq}{n}}}$$

$$z = \frac{0.3375 - 0.30}{\sqrt{\dfrac{(0.30)(0.70)}{80}}}$$

$$z \approx 0.73$$



d) If his dowsing has the same success rate as standard drilling methods, there is more than a 23% chance of seeing results as good as those of the dowser, or better, by natural sampling variation.

e) With a P-value of 0.232, we fail to reject the null hypothesis. There is no evidence to suggest that the dowser has a success rate any higher than 30%.

12. Abnormalities.

a) H_0 : The percentage of children with genetic abnormalities is 5%.($p = 0.05$)
H_A : The percentage of with genetic abnormalities is greater than 5%. ($p > 0.05$)

b) **Independence assumption:** There is no reason to think that one child having genetic abnormalities would affect the probability that other children have them.
Randomization condition: This sample may not be random, but genetic abnormalities are plausibly independent. The sample is probably representative of all children, with regards to genetic abnormalities.
10% condition: The sample of 384 children is less than 10% of all children.
Success/Failure condition: $np = (384)(0.05) = 19.2$ and $nq = (384)(0.95) = 364.8$ are both greater than 10, so the sample is large enough.

c) The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.05$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.05)(0.95)}{384}} \approx 0.0111$.

We can perform a one-proportion z-test. The observed proportion of children with genetic abnormalities is $\hat{p} = \frac{46}{384} \approx 0.1198$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.1198 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{384}}}$$

$$z \approx 6.28$$

The value of z is approximately 6.28, meaning that the observed proportion of children with genetic abnormalities is over 6 standard deviations above the hypothesized proportion. The P-value associated with this z score is 2×10^{-10}, essentially 0.

d) If 5% of children have genetic abnormalities, the chance of observing 46 children with genetic abnormalities in a random sample of 384 children is essentially 0.

e) With a P-value of this low, we reject the null hypothesis. There is strong evidence that more than 5% of children have genetic abnormalities.

f) We don't know that environmental chemicals cause genetic abnormalities. We merely have evidence that suggests that a greater percentage of children are diagnosed with genetic abnormalities now, compared to the 1980s.

13. Absentees.

a) H_0 : The percentage of students in 2000 with perfect attendance the previous month is 34% ($p = 0.34$)
H_A : The percentage of students in 2000 with perfect attendance the previous month is different from 34% ($p \neq 0.34$)

b) **Independence assumption:** It is reasonable to think that the students' attendance records are independent of one another.
Randomization condition: Although not specifically stated, we can assume that the National Center for Educational Statistics used random sampling.
10% condition: The 8302 students are less than 10% of all students.
Success/Failure condition: $np = (8302)(0.34) = 2822.68$ and $nq = (8302)(0.66) = 5479.32$ are both greater than 10, so the sample is large enough.

c) Since the conditions for inference are met, a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.34$ and

$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.34)(0.66)}{8302}} \approx 0.0052$$

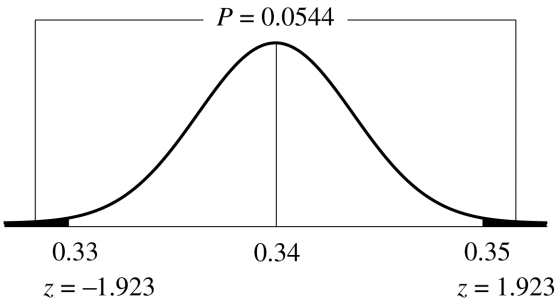
We can perform a two-tailed one-proportion z-test. The observed proportion of perfect attendees is $\hat{p} = 0.33$.

d) With a P-value of 0.0544, we reject the null hypothesis. There is some evidence to suggest that the percentage of students with perfect attendance in the previous month has changed in 2000.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.33 - 0.34}{\sqrt{\frac{(0.34)(0.66)}{8302}}}$$

$$z \approx -1.923$$



e) This result is not meaningful. A difference this small, although statistically significant, is of little practical significance.

14. Educated mothers.

a) H_0 : The percentage of students in 2000 whose mothers had graduated college is 31% ($p = 0.31$)
H_A : The percentage of students in 2000 whose mothers had graduated college is different than 31% ($p \neq 0.31$)

b) **Independence assumption:** It is reasonable to think that the students' responses are independent of one another.
Randomization condition: Although not specifically stated, we can assume that the National Center for Educational Statistics used random sampling.
10% condition: The 8368 students are less than 10% of all students.
Success/Failure condition: $np = (8368)(0.31) = 2594.08$ and $nq = (8368)(0.69) = 5773.92$ are both greater than 10, so the sample is large enough.

c) Since the conditions for inference are met, a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.31$ and

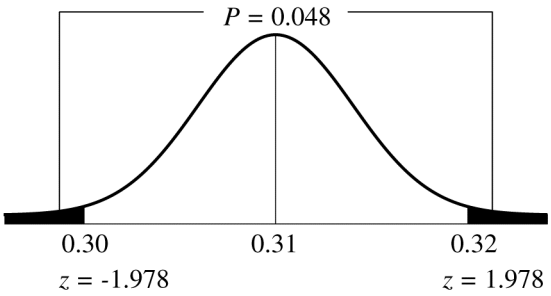
$$\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.31)(0.69)}{8368}} \approx 0.0051$$

We can perform a one-proportion two-tailed z-test.
The observed proportion of students whose mothers are college graduates is $\hat{p} = 0.32$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.32 - 0.31}{\sqrt{\frac{(0.31)(0.69)}{8368}}}$$

$$z \approx 1.978$$



d) With a P-value of 0.048, 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.

e) This result is not meaningful. A difference this small, although statistically significant, is of little practical significance.

15. Contributions, please, part II.

a) H_0 : The contribution rate is 5% ($p = 0.05$)
H_A : The contribution rate is less than 5% ($p < 0.05$)

b) **Independence assumption:** There is no reason to believe that one randomly selected potential donor's decision will affect another's decision.
Randomization condition: The sample was 100,000 randomly selected potential donors.
10% condition: We will assume that the entire mailing list has over 1,000,000 names.
Success/Failure condition: $np = 5000$ and $nq = 95{,}000$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.05$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.05)(0.95)}{100{,}000}} \approx 0.0007$.

We can perform a one-proportion z-test. The observed contribution rate is

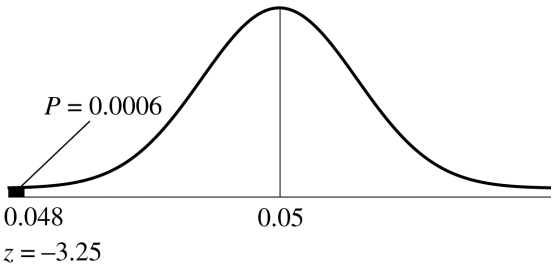
$$\hat{p} = \frac{4{,}781}{100{,}000} = 0.04781.$$

c) Since the P-value = 0.0006 is low, we reject the null hypothesis. There is strong evidence that contribution rate for all potential donors is lower than 5%.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.048 - 0.05}{\sqrt{\frac{(0.05)(0.95)}{100{,}000}}}$$

$$z \approx -3.25$$



16. Take the offer, part II.

a) H_0 : The success rate is 2% ($p = 0.02$)
H_A : The success rate is something other than 2% ($p \neq 0.02$)

b) **Independence assumption:** There is no reason to believe that one randomly selected cardholder's decision will affect another's decision.
Randomization condition: The sample was 50,000 randomly selected cardholders.
10% condition: We will assume that the number of cardholders is more than 500,000.
Success/Failure condition: $np = 1000$ and $nq = 49{,}000$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.02$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.02)(0.98)}{50{,}000}} \approx 0.0006$.

We can perform a one-proportion z-test. The observed success rate is $\hat{p} = \frac{1{,}184}{50{,}000} = 0.02368$.

c) Since the P-value is less than 0.0001, we reject the null hypothesis. There is strong evidence that success rate for all cardholders is not 2%. In fact, this sample suggests that the success rate is higher than 2%.

17. Law School.

a) H_0 : The law school acceptance rate for *LSATisfaction* is 63% ($p = 0.63$)
H_A : The law school acceptance rate for *LSATisfaction* is greater than 63% ($p > 0.63$)

b) **Randomization condition:** These 240 students may be considered representative of the population of law school applicants.
10% condition: There are certainly more than 2,400 law school applicants.
Success/Failure condition: $np = 151.2$ and $nq = 88.8$ are both greater than 10, so the sample is large enough.

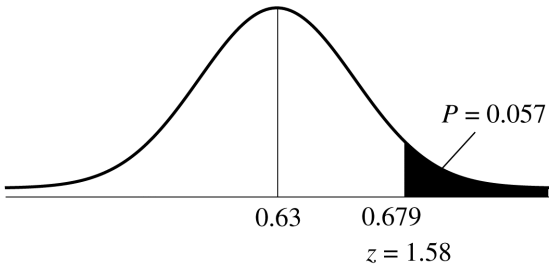
The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.63$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.63)(0.37)}{240}} \approx 0.0312$.

We can perform a one-proportion z-test. The observed success rate is $\hat{p} = \frac{163}{240} = 0.6792$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.679 - 0.63}{\sqrt{\frac{(0.63)(0.37)}{240}}}$$

$$z \approx 1.58$$



c) Since the *P*-value = 0.057 is fairly low, we reject the null hypothesis. There is weak evidence that the law school acceptance rate is higher for *LSATisfaction* applicants. Candidates should decide whether they can afford the time and expense.

18. Med School.

a) H_0 : The med school acceptance rate for Striving College is 46% ($p = 0.46$)
H_A : The law school acceptance rate for Striving College is less than 46% ($p < 0.46$)

b) **Randomization condition:** Assume that these 180 students are representative of all applicants from this college.
10% condition: 180 students represent less than 10% of all applicants.
Success/Failure condition: $np = 82.8$ and $nq = 97.2$ are both greater than 10, so the sample is large enough.

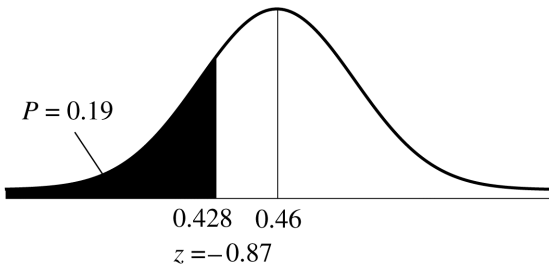
The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.46$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.46)(0.54)}{180}} \approx 0.0531$.

We can perform a one-proportion *z*-test. The observed success rate is $\hat{p} = \frac{77}{180} = 0.428$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.428 - 0.46}{\sqrt{\frac{(0.46)(0.54)}{180}}}$$

$$z \approx -0.87$$



c) Since the *P*-value = 0.19 is high, we fail to reject the null hypothesis. There is no evidence that the med school acceptance rate at Striving College is significantly lower than 46%. This could simply be year-to-year variation, as the president suggests.

19. Pollution.

H_0 : The percentage of cars with faulty emissions is 20%. ($p = 0.20$)
H_A : The percentage of cars with faulty emissions is greater than 20%. ($p > 0.20$)

Two conditions are not satisfied. 22 is greater than 10% of the population of 150 cars, and $np = (22)(0.20) = 4.4$, which is not greater than 10. It's not advisable to proceed with a test.

20. Scratch and dent.

H_0 : The percentage of damaged machines is 2%, and the warehouse is meeting the company goal. ($p = 0.02$)
H_A : The percentage of damaged machines is greater than 2%, and the warehouse is failing to meet the company goal. ($p > 0.02$)

An important condition is not satisfied. $np = (60)(0.02) = 1.2$, which is not greater than 10. The Normal model is not appropriate for modeling the sampling distribution.

21. Twins.

H_0 : The percentage of twin births to teenage girls is 3%. ($p = 0.03$)
H_A : The percentage of twin births to teenage girls differs from 3%. ($p \neq 0.03$)

Independence assumption: One mother having twins will not affect another. Observations are plausibly independent.
Randomization condition: This sample may not be random, but it is reasonable to think that this hospital has a representative sample of teen mothers, with regards to twin births.
10% condition: The sample of 469 teenage mothers is less than 10% of all such mothers.
Success/Failure condition: $np = (469)(0.03) = 14.07$ and $nq = (469)(0.97) = 454.93$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.03$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.03)(0.97)}{469}} \approx 0.0079$.

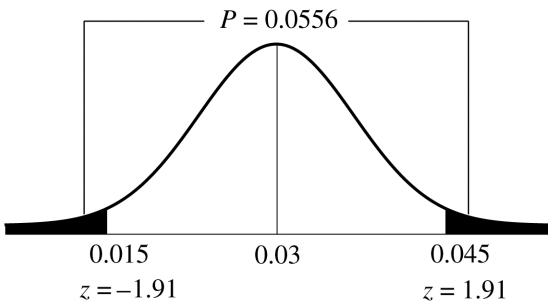
We can perform a one-proportion z-test. The observed proportion of twin births to teenage mothers is $\hat{p} = \frac{7}{469} \approx 0.015$.

Since the P-value = 0.0556 is fairly low, we reject the null hypothesis. There is some evidence that the proportion of twin births for teenage mothers at this large city hospital is lower than the proportion of twin births for all mothers.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.015 - 0.03}{\sqrt{\frac{(0.03)(0.97)}{469}}}$$

$$z \approx -1.91$$



22. Football 2006.

H_0 : The percentage of home team wins is 50%. ($p = 0.50$)
H_A : The percentage of home team wins is greater than 50%. ($p > 0.50$)

Independence assumption: Results of one game should not affect others.
Randomization condition: This season should be representative of other seasons, with regards to home team wins.
10% condition: 240 games represent less than 10% of all games, in all seasons.
Success/Failure condition: $np = (240)(0.50) = 120$ and $nq = (240)(0.50) = 120$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.50$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.5)(0.5)}{240}} \approx 0.0323$.

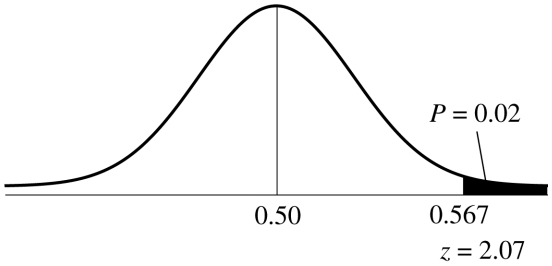
We can perform a one-proportion z-test. The observed proportion of home team wins is $\hat{p} = \frac{136}{240} = 0.567$.

Since the *P*-value = 0.02 is low, we reject the null hypothesis. There is strong evidence that the proportion of home teams wins is greater than 50%. This provides evidence of a home team advantage.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.567 - 0.50}{\sqrt{\frac{(0.50)(0.50)}{240}}}$$

$$z \approx 2.07$$



23. Webzine.

H_0 : The percentage of readers interested in an online edition is 25%. ($p = 0.25$)
H_A : The percentage of readers interested in an online edition is greater than 25%. ($p > 0.25$)

Independence assumption: Interest of one reader should not affect interest of other readers.
Randomization condition: The magazine conducted an SRS of 500 current readers.
10% condition: 500 readers are less than 10% of all potential subscribers.
Success/Failure condition: $np = (500)(0.25) = 125$ and $nq = (500)(0.75) = 375$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.25$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.25)(0.75)}{500}} \approx 0.0194$.

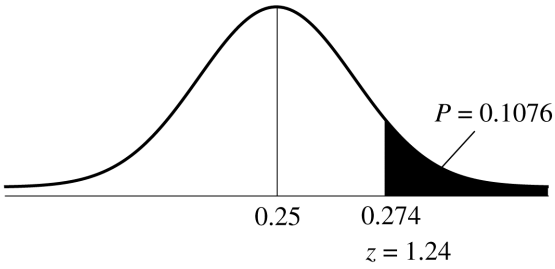
We can perform a one-proportion *z*-test. The observed proportion of interested readers is $\hat{p} = \frac{137}{500} = 0.274$.

Since the *P*-value = 0.1076 is high, we fail to reject the null hypothesis. There is little evidence to suggest that the proportion of interested readers is greater than 25%. The magazine should not publish the online edition.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.274 - 0.25}{\sqrt{\frac{(0.25)(0.75)}{500}}}$$

$$z \approx 1.24$$



24. Seeds.

H_0 : The germination rate of the green bean seeds is 92%. ($p = 0.92$)
H_A : The germination rate of the green bean seeds is less than 92%. ($p < 0.92$)

Independence assumption: Seeds in a single packet may not germinate independently. They have been treated identically with regards to moisture exposure, temperature, etc. They may have higher or lower germination rates than seeds in general.
Randomization condition: The cluster sample of one bag of seeds was not random.
10% condition: 200 seeds is less than 10% of all seeds.
Success/Failure condition: $np = (200)(0.92) = 184$ and $nq = (200)(0.08) = 16$ are both greater than 10, so the sample is large enough.

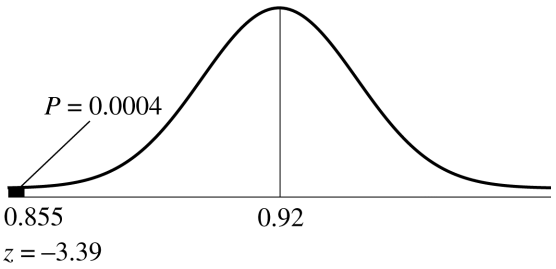
The conditions have *not* been satisfied. We will assume that the seeds in the bag are representative of all seeds, and cautiously use a Normal model to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.92$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.92)(0.08)}{200}} \approx 0.0192$.

We can perform a one-proportion z-test. The observed proportion of germinated seeds is $\hat{p} = \frac{171}{200} = 0.85$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.855 - 0.92}{\sqrt{\frac{(0.92)(0.08)}{200}}}$$

$$z \approx -3.39$$



Since the P-value = 0.0004 is very low, we reject the null hypothesis. There is strong evidence that the germination rate of the seeds in less than 92%. We should use extreme caution in generalizing these results to all seeds, but the manager should be safe, and avoid selling faulty seeds. The seeds should be thrown out.

25. Women executives.

H_0 : The proportion of female executives is similar to the overall proportion of female employees at the company. ($p = 0.40$)
H_A : The proportion of female executives is lower than the overall proportion of female employees at the company. ($p < 0.40$)

Independence assumption: It is reasonable to think that executives at this company were chosen independently.
Randomization condition: The executives were not chosen randomly, but it is reasonable to think of these executives as representative of all potential executives over many years.
10% condition: 43 executives are less than 10% of all possible executives at the company.
Success/Failure condition: $np = (43)(0.40) = 17.2$ and $nq = (43)(0.60) = 25.8$ are both greater than 10, so the sample is large enough.

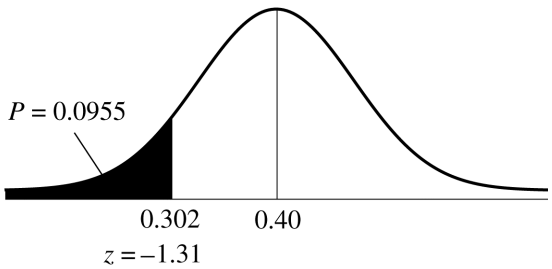
The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.40$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.40)(0.60)}{43}} \approx 0.0747$.

We can perform a one-proportion z-test. The observed proportion is $\hat{p} = \frac{13}{43} \approx 0.302$.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.302 - 0.40}{\sqrt{\frac{(0.40)(0.60)}{43}}}$$

$$z \approx -1.31$$



Since the P-value = 0.0955 is high, we fail to reject the null hypothesis. There is little evidence to suggest proportion of female executives is any different from the overall proportion of 40% female employees at the company.

26. Jury.

H_0 : The proportion of Hispanics called for jury duty is similar to the proportion of Hispanics in the county, 19%. ($p = 0.19$)
H_A : The proportion of Hispanics called for jury duty is less than the proportion of Hispanics in the county, 19%. ($p < 0.19$)

Independence assumption /Randomization condition: Assume that potential jurors were called randomly from all of the residents in the county. This is really what we are testing. If we reject the null hypothesis, we will have evidence that jurors are not called randomly.
10% condition: 72 people are less than 10% of all potential jurors in the county.
Success/Failure condition: $np = (72)(0.19) = 13.68$ and $nq = (72)(0.81) = 58.32$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.19$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.19)(0.81)}{72}} \approx 0.0462$.

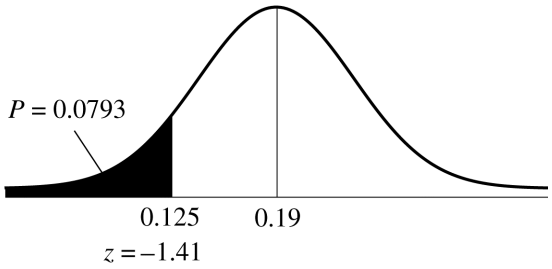
We can perform a one-proportion z-test. The observed proportion of Hispanics called for jury duty is $\hat{p} = \frac{9}{72} \approx 0.125$.

Since the P-value = 0.0793 is somewhat high, we fail to reject the null hypothesis. We are not convinced that Hispanics are underrepresented in the jury selection system. However, this P-value isn't extremely high. There is some evidence that the selection process may be biased. We should examine some other groups called for jury duty and take a closer look.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.125 - 0.19}{\sqrt{\frac{(0.19)(0.81)}{72}}}$$

$$z \approx -1.41$$



27. Dropouts.

H_0 : The proportion of dropouts at this high school is similar to 10.3%, the proportion of dropouts nationally. ($p = 0.103$)
H_A : The proportion of dropouts at this high school is greater than 10.3%, the proportion of dropouts nationally. ($p > 0.103$)

Independence assumption /Randomization condition: Assume that the students at this high school are representative of all students nationally. This is really what we are testing. The dropout rate at this high school has traditionally been close to the national rate. If we reject the null hypothesis, we will have evidence that the dropout rate at this high school is no longer close to the national rate.
10% condition: 1782 students are less than 10% of all students nationally.
Success/Failure condition: $np = (1782)(0.103) = 183.546$ and $nq = (1782)(0.897) = 1598.454$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, $\mu_{\hat{p}} = p = 0.103$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.103)(0.897)}{1782}} \approx 0.0072$.

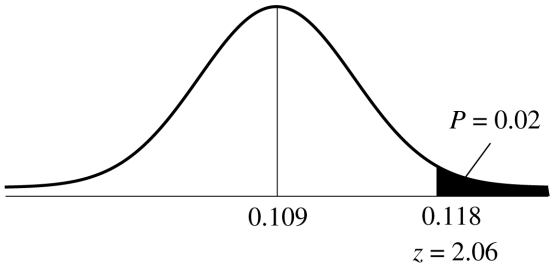
We can perform a one-proportion z-test. The observed proportion of dropouts is $\hat{p} = \frac{210}{1782} \approx 0.117845$.

Since the P-value = 0.02 is low, we reject the null hypothesis. There is evidence that the dropout rate at this high school is higher than 10.3%.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.117845 - 0.109}{\sqrt{\frac{(0.103)(0.897)}{1782}}}$$

$$z \approx 2.06$$



28. Acid rain.

H_0 : The proportion of trees with acid rain damage in Hopkins Forest is 15%, the proportion of trees with acid rain damage in the Northeast. ($p = 0.15$)
H_A : The proportion of trees with acid rain damage in Hopkins Forest is greater than 15%, the proportion of trees with acid rain damage in the Northeast. ($p > 0.15$)

Independence assumption /Randomization condition: Assume that the trees in Hopkins Forest are representative of all trees in the Northeast. This is really what we are testing. If we reject the null hypothesis, we will have evidence that the proportion of trees with acid rain damage is greater in Hopkins Forest than the proportion in the Northeast.
10% condition: 100 trees are less than 10% of all trees.
Success/Failure condition: $np = (100)(0.15) = 15$ and $nq = (100)(0.85) = 85$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.109$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.15)(0.85)}{100}} \approx 0.0357$.

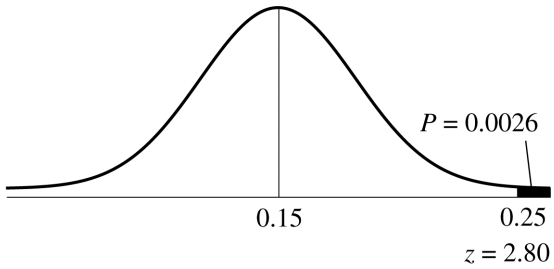
We can perform a one-proportion z-test. The observed proportion of damaged trees is $\hat{p} = \frac{25}{100} = 0.25$.

Since the P-value = 0.0026 is low, we reject the null hypothesis. There is strong evidence that the trees in Hopkins forest have a greater proportion of acid rain damage than the 15% reported for the Northeast.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.25 - 0.150}{\sqrt{\frac{(0.150)(0.850)}{100}}}$$

$$z \approx 2.80$$



29. Lost luggage.

H_0 : The proportion of lost luggage returned the next day is 90%. ($p = 0.90$)
H_A : The proportion of lost luggage returned the next day is lower than 90%. ($p < 0.90$)

Independence assumption: It is reasonable to think that the people surveyed were independent with regards to their luggage woes.
Randomization condition: Although not stated, we will hope that the survey was conducted randomly, or at least that these air travelers are representative of all air travelers for that airline.
10% condition: 122 air travelers are less than 10% of all air travelers on the airline.
Success/Failure condition: $np = (122)(0.90) = 109.8$ and $nq = (122)(0.10) = 12.2$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.90$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.90)(0.10)}{122}} \approx 0.0272$.

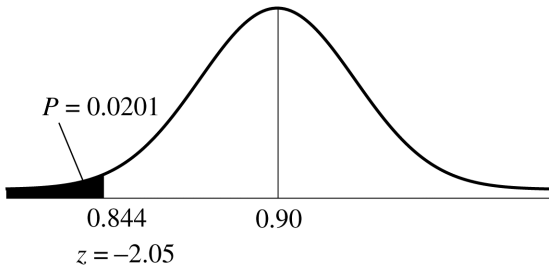
We can perform a one-proportion z-test. The observed proportion of dropouts is $\hat{p} = \frac{103}{122} \approx 0.844$.

Since the P-value = 0.0201 is low, we reject the null hypothesis. There is evidence that the proportion of lost luggage returned the next day is lower than the 90% claimed by the airline.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.844 - 0.90}{\sqrt{\frac{(0.90)(0.10)}{122}}}$$

$$z \approx -2.05$$



30. TV ads.

H_0 : The proportion of respondents who recognize the name is 40%.($p = 0.40$)
H_A : The proportion of respondents who recognize the name is more than 40%. ($p > 0.40$)

Independence assumption: There is no reason to believe that the responses of randomly selected people would influence others.
Randomization condition: The pollster contacted the 420 adults randomly.
10% condition: A sample of 420 adults is less than 10% of all adults.
Success/Failure condition: $np = (420)(0.40) = 168$ and $nq = (420)(0.60) = 252$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.40$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.40)(0.60)}{420}} \approx 0.0239$.

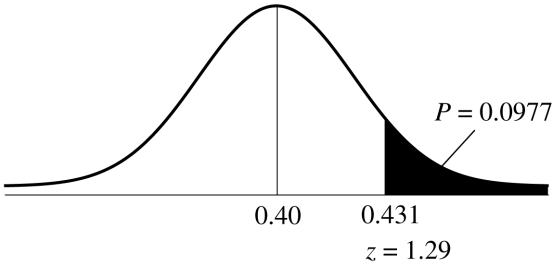
We can perform a one-proportion z-test. The observed proportion of dropouts is $\hat{p} = \frac{181}{420} \approx 0.431$.

Since the *P*-value = 0.0977 is fairly high, we fail to reject the null hypothesis. There is little evidence that more than 40% of the public recognizes the product.
Don't run commercials during the Super Bowl!

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.431 - 0.40}{\sqrt{\frac{(0.40)(0.60)}{420}}}$$

$$z \approx 1.29$$



31. John Wayne.

a) H_0 : The death rate from cancer for people working on the film was similar to that predicted by cancer experts, 30 out of 220.
H_A : The death rate from cancer for people working on the film was higher than the rate predicted by cancer experts.

The conditions for inference are not met, since this is not a random sample. We will assume that the cancer rates for people working on the film are similar to those predicted by the cancer experts, and a Normal model can be used to model the sampling distribution of the rate, with $\mu_{\hat{p}} = p = 30/220$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\left(\frac{30}{220}\right)\left(\frac{190}{220}\right)}{220}} \approx 0.0231$.

We can perform a one-proportion *z*-test. The observed cancer rate is $\hat{p} = \frac{46}{220} \approx 0.209$.

$$z = \frac{\hat{p} - p_0}{\sigma(\hat{p})}$$

$$z = \frac{\frac{46}{220} - \frac{30}{220}}{\sqrt{\frac{\left(\frac{30}{220}\right)\left(\frac{190}{220}\right)}{220}}}$$

$$z = 3.14$$

Since the *P*-value = 0.0008 is very low, we reject the null hypothesis. There is strong evidence that the cancer rate is higher than expected among the workers on the film.

b) This does not prove that exposure to radiation may increase the risk of cancer. This group of people may be atypical for reasons that have nothing to do with the radiation.

32. AP Stats.

H_0 : These students achieve scores of 3 or higher at a similar rate to the nation. ($p = 0.60$)
H_A : These students achieve these scores at a different rate than the nation. ($p \neq 0.60$)

Independence assumption: There is no reason to believe that students' scores would influence others.
Randomization condition: The teacher considers this class typical of other classes.
10% condition: A sample of 54 students is less than 10% of all students.
Success/Failure condition: $np = (54)(0.60) = 32.4$ and $nq = (54)(0.40) = 21.6$ are both greater than 10, so the sample is large enough.

The conditions have been satisfied, so a Normal model can be used to model the sampling distribution of the proportion, with $\mu_{\hat{p}} = p = 0.60$ and $\sigma(\hat{p}) = \sqrt{\frac{pq}{n}} = \sqrt{\frac{(0.60)(0.40)}{54}} \approx 0.0667$.

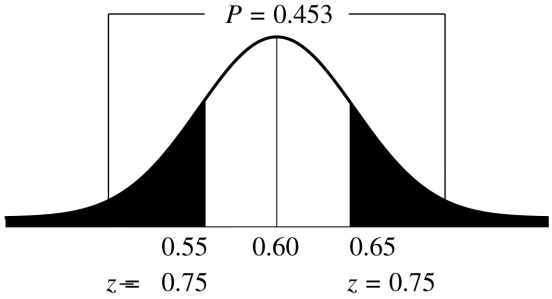
We can perform a one-proportion *z*-test. The observed pass rate is $\hat{p} = 0.65$.

Since the *P*-value = 0.453 is high, we fail to reject the null hypothesis. There is little evidence that the rate at which these students score 3 or higher on the AP Stats exam is any higher than the national rate.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{pq}{n}}}$$

$$z = \frac{0.65 - 0.60}{\sqrt{\frac{(0.60)(0.60)}{54}}}$$

$$z = 0.75$$



The teacher has no cause to brag. Her students did have a higher rate of scores of 3 or higher, but not so high that the results could not be attributed to sampling variability.