AP Statistics Test A – Inference for Proportions – Part V – Key*

1. B 2. D 3. E 4. E 5. D 6. D 7. B 8. B 9. E 10. E

11. President's approval rating: The smaller poll would have more variability and would thus be more likely to vary from the actual approval rating of 54%. We would expect the larger poll to be more consistent with the 54% rating. So, it is more likely that the smaller poll would report that the President's approval rating is below 50%.

12. Cereal: Two methods can be used to solve this problem:

Method 1:

Let B = weight of one box of cereal and T = weight of 12 boxes of cereal.
We are told that the contents of the boxes are approximately Normal, and we can assume that the content amounts are independent from box to box.

$E(\mathrm{T}) = E(\mathrm{B}_1 + \mathrm{B}_2 + \cdots + \mathrm{B}_{12}) = E(\mathrm{B}_1) + E(\mathrm{B}_2) + \cdots + E(\mathrm{B}_{12}) = 156$ ounces

Since the content amounts are independent,

$Var(\mathrm{T}) = Var(\mathrm{B}_1 + \mathrm{B}_2 + \cdots + \mathrm{B}_{12}) = Var(\mathrm{B}_1) + Var(\mathrm{B}_2) + \cdots + Var(\mathrm{B}_{12}) = 3$

$SD(\mathrm{T}) = \sqrt{Var(\mathrm{T})} = \sqrt{3} = 1.73$ ounces

We model T with $N(156, 1.73)$.

$z = \frac{160-156}{1.73} = 2.31$ and $P(\mathrm{T} > 160) = P(z > 2.31) = 0.0104$

There is a 1.04% chance that a case of 12 cereal boxes will weigh more than 160 ounces.

Method 2:

Using the Central Limit Theorem approach, let $\bar{y}$ = average content of boxes in the case. Since the contents are Normally distributed, $\bar{y}$ is modeled by $N\left(13, \frac{0.5}{\sqrt{12}}\right)$.

$$P\left(\bar{y} > \frac{160}{12}\right) = P(\bar{y} > 13.33) = P\left(z > \frac{13.33-13}{0.5/\sqrt{12}}\right) = P(z > 2.31) = 0.0104.$$

There is a 1.04% chance that a case of 12 cereal boxes will weigh more than 160 ounces.

13. Exercise: A random sample of 150 men found that 88 of the men exercise regularly, while a random sample of 200 women found that 130 of the women exercise regularly.

a. Conditions:

* Randomization Condition: We are told that we have random samples.
* 10% Condition: We have less than 10% of all men and less than 10% of all women.
* Independent samples condition: The two groups are clearly independent of each other.
* Success/Failure Condition: Of the men, 88 exercise regularly and 62 do not; of the women, 130 exercise regularly and 70 do not. The observed number of both successes and failures in both groups is at least 10.

With the conditions satisfied, the sampling distribution of the difference in proportions is approximately Normal with a mean of $p_M - p_W$, the true difference between the population proportions. We can find a two-proportion z-interval.

We know: $n_M = 150$, $\hat{p}_M = \frac{88}{150} = 0.587$, $n_W = 200$, $\hat{p}_W = \frac{130}{200} = 0.650$.

We estimate $SD(\hat{p}_M - \hat{p}_W)$ as

$$SE(\hat{p}_M - \hat{p}_W) = \sqrt{\frac{\hat{p}_M \hat{q}_M}{n_M} + \frac{\hat{p}_W \hat{q}_W}{n_W}} = \sqrt{\frac{(0.587)(0.413)}{150} + \frac{(0.65)(0.35)}{200}} = 0.0525$$

$$ME = z^* \times SE(\hat{p}_M - \hat{p}_W) = 1.96(0.0525) = 0.1029$$

The observed difference in sample proportions = $\hat{p}_M - \hat{p}_W$ = 0.587 – 0.650 = -0.063, so the 95% confidence interval is -0.063 ± 0.1029, or -16.6% to 4.0%.

We are 95% confident that the proportion of women who exercise regularly is between 4.0% lower and 16.6% higher than the proportion of men who exercise regularly.

b. Since zero is contained in my confidence interval, I cannot say that a higher proportion of women than men exercise regularly. My confidence interval does not support my friend's claim.

14. Internet access:

Since $ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$, we have or $z^* = \frac{0.03}{\sqrt{\frac{(0.28)(0.72)}{1028}}} \approx 2.14$.

Our confidence level is approximately $P(-2.14 < z < 2.14) = 0.9676$, or 97%..

15. Sleep

Hypothesis: $\text{H}_0 : p = 0.50$ $\text{H}_\text{A} : p > 0.50$

Plan: Okay to use the Normal model because the trials are independent (random sample of U.S. adults), these 1003 U.S. adults are less than 10% of all U.S. adults, and $np_0 = (1003)(0.50) = 501.5 \geq 10$ and $nq_0 = (1003)(0.50) = 501.5 \geq 10$.

We will do a one-proportion z-test.

Mechanics: $SD(p_0) = \sqrt{\frac{p_0 q_0}{n}} = \sqrt{\frac{(0.50)(0.50)}{1003}} = 0.0158$; sample proportion: $\hat{p} = 0.55$

$$P(\hat{p} > 0.55) = P(z > \frac{0.55 - 0.50}{0.0158}) = P(z > 3.16) = 0.0008$$

With a *P*-value of 0.0008, I reject the null hypothesis. There is strong evidence that the proportion of U.S. adults who feel they get enough sleep is more than 50%.

20. a) (0.004, 0.209)
b) We are 95% confident, based on these data, that the proportion of heart disease patients who die within 4 years is between 0.4% and 20.9% higher for depressed patients than for nondepressed patients.
c) 95% of all random samples will produce intervals that contain the true value.

21. a) No; subjects weren't assigned to treatment groups. It's an observational study.
b) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 \neq 0$. $z = 3.56$, P-value $= 0.0004$. With a P-value this low, we reject H_0. There is a significant difference in the clinic's effectiveness. Younger mothers have a higher birth rate than older mothers. Note that the Success/Failure Condition is met based on the pooled estimate of p.
c) We are 95% confident, based on these data, that the proportion of successful live births at the clinic is between 10.0% and 27.8% higher for mothers under 38 than in those 38 and older. However, the Success/Failure Condition is not met for the older women, since # Successes < 10. We should be cautious in trusting this confidence interval.

22. a) No; subjects weren't assigned to treatment groups. It's an observational study.
b) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 > 0$. $z = 2.56$, P-value $= 0.005$ (with technology, $z = 2.71$, P-value $= 0.003$). With a P-value this low we reject H_0. This study shows a significantly higher rate of low-weight babies born to mothers exposed to high levels of air pollution during pregnancy.
c) We are 90% confident that the proportion of low-birth weight babies will be between 0.8% and 7.6% higher for mothers exposed to high levels of air pollution than those who were not.

23. a) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 > 0$. $z = 1.18$, P-value $= 0.118$. With P-value this high, we fail to reject H_0. These data do not show evidence of a decrease in the voter support for the candidate.
b) Type II

24. a) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 > 0$. $z = 0.78$, P-value $= 0.2166$. With a P-value this high, we fail to reject H_0. There is no evidence, based on this information, that men are more likely than women to make online purchases of books.
b) Type II c) (−0.04, 0.11)
d) With 95% confidence we estimate that the proportion of men who buy books online could be 4 percentage points lower than women, or up to 11 percentage points higher.

25. a) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 \neq 0$. $z = -0.39$, P-value $= 0.6951$. With a P-value this high, we fail to reject H_0. There is no evidence of racial differences in the likelihood of multiple births, based on these data.
b) Type II

26. a) $H_0: p_1 - p_2 = 0; H_A: p_1 - p_2 > 0$. $z = 0.24$, P-value $= 0.4068$. With a P-value this high, we fail to reject H_0. These data do not suggest that mammograms are effective in reducing breast cancer deaths.
b) Type II

27. a) We are 95% confident, that between 67.0% and 83.0% of patients with joint pain will find medication A effective.
b) We are 95% confident, that between 51.9% and 70.3% of patients with joint pain will find medication B effective.
c) Yes, they overlap. This might indicate no difference in the effectiveness of the medications. (Not a proper test.)
d) We are 95% confident that the proportion of patients with joint pain who will find medication A effective is between 1.7% and 26.1% higher than the proportion who will find medication B effective.
e) No. There is a difference in the effectiveness of the medications.
f) To estimate the variability in the difference of proportions, we must add variances. The two one-sample intervals do not. The two-sample method is the correct approach.

28. a) We are 95% confident, based on the data, that between 47.5% and 56.5% of male voters will vote for the candidate.
b) We are 95% confident, based on the data, that between 40.7% and 49.3% of female voters will vote for the candidate.
c) Yes, they overlap. There appears to be no discernible gender gap, but this is not the proper approach.
d) We are 95% confident, that the proportion of men who will vote for the candidate is between 0.8% and 13.2% higher than the proportion of female voters who will vote for him.
e) No. There is a gender gap.
f) To estimate the variability in the difference of proportions, we must add variances. The two one-sample intervals do not. The two-sample method is the correct approach.

29. The conditions are satisfied to test $H_0: p_{young} = p_{old}$ against $H_A: p_{young} > p_{old}$. The one-sided P-value is 0.0619, so we may reject the null hypothesis. Although the evidence is not strong, *Time* may be justified in saying that younger men are more comfortable discussing personal problems.

30. The conditions are met to test the hypothesis that public and private colleges have the same retention rates. The two-sided P-value is 0.1996, too high to reject the null. We don't have evidence of an overall difference.

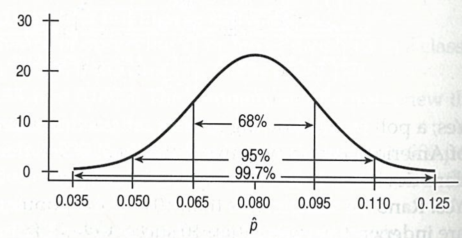
31. Yes. With a low P-value of 0.003, reject the null hypothesis of no difference. There's evidence of an increase in the proportion of parents checking the Web sites visited by their teens.

32. No. Based on a high P-value of 0.11, fail to reject the hypothesis that there's no age-based difference in computer gaming among teenage boys.

PART V REVIEW

1. H_0: There is no difference in cancer rates, $p_1 - p_2 = 0$. H_A: The cancer rate in those who use the herb is higher, $p_1 - p_2 > 0$.

2. a) Yes, $0.08 \times 325 = 26$, so we expect more than 10 successes and more than 10 failures.
b) $\mu = 0.08, \sigma = 0.015$
c)



d) There is about a 68% chance of observing between 6.5% and 9.5% of colorblind males in a class of this size and a 95% chance of having between 5% and 11% colorblind males. Almost all classes of this size will have a percentage between 3.5% and 12.5% colorblind males.

3. a) 10.29
b) Not really. The z-score is −1.11. Not any evidence to suggest that the proportion for Monday is low.
c) Yes. The z-score is 2.26 with a P-value of 0.024 (two-sided).
d) Some births are scheduled for the convenience of the doctor and/or the mother.

4. a) No. There is always natural sampling variation.
b) 47.6% and 53.8% (using 1.96 SDs).
c) We can't use a Normal model; we expect only 4 "successes".
d) Less. Proportions farther from 0.5 have smaller standard deviations.

5. a) $H_0: p_1 = 0.40; H_A: p_1 < 0.40$
b) Random sample; less than 10% of all California gas stations, $0.4(27) = 10.8, 0.6(27) = 16.2$. Assumptions and conditions are met.

c) $z = -1.49$, P-value $= 0.0677$

d) With a P-value this high, we fail to reject H_0. These data do not provide evidence that the proportion of leaking gas tanks is less than 40% (or that the new program is effective in decreasing the proportion).

e) Yes, Type II.

f) Increase α, increase the sample size.

g) Increasing α—increases power, lowers chance of Type II error, but increases chance of Type I error.
Increasing sample size—increases power, costs more time and money.

6. a) An experiment.

b) H_0: There is no difference, $p_1 - p_2 = 0$.
H_A: Patients with carbolic acid are more likely to live, $p_1 - p_2 > 0$.
$z = 2.91$, P-value $= 0.0018$; with a P-value so low, we reject H_0. These data show that carbolic acid is effective in increasing the chances of surviving an amputation.

c) We are not told whether the patients were randomized to the treatments. We are not told whether the experiment was double-blinded or even blinded at all. This could have biased the results toward a more favorable outcome.

7. a) The researcher believes that the true proportion of "A's" is within 10% of the estimated 54%, namely, between 44% and 64%.

b) Small sample c) No, 63% is contained in the interval.

8. 0.647; $\bar{y} \approx N(3.5, 0.538)$. Rolls are independent.

9. a) Pew believes that the true proportion is within 3% of the 33% from the sample; that is, between 30% and 36%.

b) Larger, since it's a smaller sample.

c) We are 95% confident that the proportion of active traders who rely on the Internet for investment information is between 38.7% and 51.3%, based on this sample.

d) Larger, since it's a smaller sample.

10. a) 42.7% to 51.1%

b) Since the interval extends above 50%, it is possible that the death penalty does have majority support.

c) About 3382 people

11. a) Bimodal!

b) μ, the population mean. Sample size does not matter.

c) $\sigma/\sqrt{n}$; sample size does matter.

d) It becomes closer to a Normal model and narrower as the sample size increases.

12. a) Individuals were probably independent of one another, they are less than 10% of all African-American women, and there were clearly at least 10 successes and 10 failures. Is the sample a random sample from the population of African-American women?

b) 39.5% to 44.5%

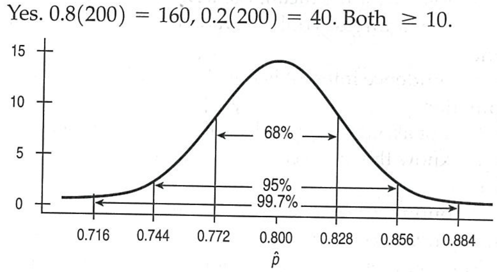
c) We are 95% confident that between 39.5% and 44.5% of African-American women have a vitamin D deficiency.

d) 95% of all such random samples will produce intervals that contain the true proportion.

13. a) $\mu = 0.80$, $\sigma = 0.028$

b) Yes. $0.8(200) = 160$, $0.2(200) = 40$. Both ≥ 10.

c)



d) 0.039

14. a) (−4.5%, 6.5%)

b) No. The interval for the difference includes 0.

15. H_0: There is no difference, $p_1 - p_2 = 0$. H_A: Early births have increased, $p_1 - p_2 < 0$. $z = -0.729$, P-value $= 0.2329$. Because the P-value is so high, we do not reject H_0. These data do not show an increase in the incidence of early birth of twins.

16. a) We are 95% confident that the increase in the proportion of women who may develop side effects from magnesium sulfide is between 18.1% and 20.8%.

b) H_0: There is no difference, $p_1 - p_2 = 0$. H_A: Treatment prevents eclampsia, $p_1 - p_2 < 0$. $z = -4.84$, P-value $= 6.4 \times 10^{-7}$. Because the P-value is so low, we reject H_0. This study shows evidence that treatment with magnesium sulfide is effective in preventing eclampsia.

17. a) H_0: There is no difference, $p_1 - p_2 \geq 0$. H_A: Treatment prevents deaths from eclampsia, $p_1 - p_2 < 0$.

b) Samples are random and independent; less than 10% of all pregnancies (or eclampsia cases); more than 10 successes and failures in each group.

c) 0.8008

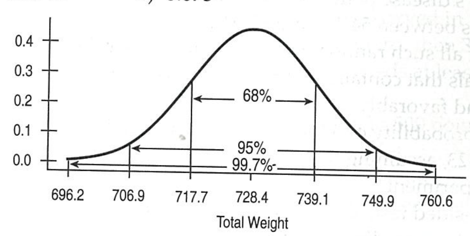
d) There is insufficient evidence to conclude that magnesium sulfide is effective in preventing eclampsia deaths.

e) Type II f) Increase the sample size, increase α.

g) Increasing sample size: decreases variation in the sampling distribution, is costly. Increasing α: Increases likelihood of rejecting H_0, increases chance of Type I error.

18. a) 33.7% b) 0.073

c)



19. a) It is not clear what the pollster asked. Otherwise they did fine.

b) Stratified sampling. c) 4%

d) 95% e) Smaller sample size.

f) Wording and order of questions (response bias).

20. About 1800

21. a) H_0: There is no difference, $p = 0.143$. H_A: The fatal accident rate is lower in girls, $p < 0.143$. $z = -1.67$, P-value $= 0.0479$. Because the P-value is low, we reject H_0. These data give some evidence that the fatal accident rate is lower for girls than for teens in general.

b) If the proportion is really 14.3%, we will see the observed proportion (11.3%) or lower 4.8% of the time by sampling variation.

22. a) H_0: There is no difference, $p_1 - p_2 = 0$. H_A: There is a difference, $p_1 - p_2 \neq 0$.

b) There is a difference in the proportion of students with perfect pitch; more Asians are likely to have it.

c) If there is no difference, the observed 25% difference will be seen by sampling variation less than about 1 time in 10,000.

d) The data do not "prove" anything about genetic differences causing differences in perfect pitch—merely that Asian students are more likely to have it.

23. a) One would expect many small fish, with a few large ones.

b) We don't know the exact distribution, but we know it's not Normal.

c) Probably not. With a skewed distribution, a sample size of five is not a large enough sample to say the sampling model for the mean is approximately Normal.

d) 0.961

24. a) Random sample; less than 10% of all high-school students; many more than 10 successes and 10 failures. 90% confidence interval for proportion who cheat: 72.9% to 75.1%.
b) Based on this information, we are 90% confident that the proportion of all high-school students who have cheated at least once on a test is between 72.9% and 75.1%.
c) 90% of all such random samples will produce confidence intervals that contain the true proportion of students who cheat.
d) Wider. More confidence means a larger margin of error.

25. a) Yes. $0.8(60) = 48, 0.2(60) = 12$. Both are ≥ 10.
b) 0.834
c) Higher. Bigger sample means smaller standard deviation for $\hat{p}$.
d) Answers will vary. For $n = 500$, the probability is 0.997.

26. a) H_0: Progress is on track, $p = 0.20$. H_A: Progress is off track, $p > 0.20$.
b) Random samples; less than 10% of all high-school students; many more than 10 failures and successes.
c) 0.0008
d) If the proportion were 20%, the probability of seeing a value as high as (or higher than) 23% in a random sample this large is about 0.0008.
e) By 2006, the rate of cigarette smoking in high-school students was higher than the target of 20%.
f) Type I

27. a) 54.4 to 62.5%
b) Based on this study, with 95% confidence the proportion of Crohn's disease patients who will respond favorable to infliximab is between 54.4% and 62.5%.
c) 95% of all such random samples will produce confidence intervals that contain the true proportion of patients who respond favorably.

28. No. The probability of 30% or more is very small: 6.1×10^{-5}.

29. At least 423, assuming that p is near 50%.

30. a) An experiment
b) A one-sided test, since they are interested only in a decrease in percentage needing repairs.
c) Deciding the additive reduces the number of repairs needed when there really is no difference.
d) Deciding the additive makes no difference when it really does reduce the number of repairs needed.
e) Type II
f) Given that the two groups received roughly the same use and care, yes. They can't necessarily claim it will work for all cars, only cars similar to their fleet.

31. a) Random sample (?); certainly less than 10% of all preemies and normal babies; more than 10 failures and successes in each group. 1.7% to 16.3% greater for normal-birth weight children.
b) Since 0 is not in the interval, there is evidence that preemies have a lower high school graduation rate than children of normal birth weight.
c) Type I, since we rejected the null hypothesis.

32. a) We are 95% confident that between 11.6% and 16.4% of Texas children wear helmets when biking, roller skating, or skateboarding, based on these data.
b) The data might not be a random sample.
c) About 408, using the previous 14% as $\hat{p}$.

33. a) H_0: The computer is undamaged. H_A: The computer is damaged.
b) 20% of good PCs will be classified as damaged (bad), while all damaged PCs will be detected (good).
c) 3 or more. d) 20%
e) By switching to two or more as the rejection criterion, 7% of the good PCs will be misclassified, but only 10% of the bad ones will, increasing the power from 20% to 90%.

34. a) Increase b) Decrease

35. The null hypothesis is that Bush's disapproval proportion is 66%—the Nixon benchmark. The one-tailed test has a z-value of -2.00, so the P-value is 0.0228. It looks like Bush's May 2007 ratings were better than the Nixon benchmark low.

36. The null hypothesis is that the percentage of students who attain a GPA of at least 3.5 remained 20% in 2000. The sample proportion of 25% is more than four standard deviations above the hypothesized rate, strong evidence the results are not due to chance. This may be an indication of grade inflation.

37. a) The company is interested only in confirming that the athlete is well known.
b) Type I: the company concludes that the athlete is well known, but that's not true. It offers an endorsement contract to someone who lacks name recognition. Type II: the company overlooks a well-known athlete, missing the opportunity to sign a potentially effective spokesperson.
c) Type I would be more likely, Type II less likely.

38. a) Although 27% of the people polled could identify her, her name recognition rate in the whole population could be less than the required 25%.
b) Type II. c) Higher.

39. I am 95% confident that the proportion of U.S. adults who favor nuclear energy is between 7 and 19 percentage points higher than the proportion who would accept a nuclear plant near their area.

40. We're 95% confident that between 46% and 60% of anorexia patients will drop out of treatment programs. However, this wasn't a random sample of all patients; they were assigned a treatment rather than choosing one on their own, and they may have had different experiences if they were not part of an experiment.

CHAPTER 23

1. a) 1.74 b) 2.37 c) 0.0524 d) 0.0889

2. a) 2.36 b) 2.62 c) 0.9829 d) 0.0381

3. Shape becomes closer to Normal; center does not change; spread becomes narrower.

4. The critical value becomes smaller, approaching 1.96.

5. a) The confidence interval is for the population mean, not the individual cows in the study.
b) The confidence interval is not for individual cows.
c) We *know* the average gain in this study was 56 pounds!
d) The average weight gain of all cows does not vary. It's what we're trying to estimate.
e) No. There is not a 95% chance for another sample to have an average weight gain between 45 and 67 pounds. There is a 95% chance that another sample will have its average weight gain within two standard errors of the true mean.

6. a) Nine out of 10 intervals will contain the true mean salary; different samples will produce different intervals.
b) This is correct.
c) The interval is for the population mean, not individual teachers.
d) The interval is for the mean, not individual teachers.
e) The interval addresses only Nevada teachers, not the entire country.

7. a) No. A confidence interval is not about individuals in the population.
b) No. It's not about individuals in the sample, either.
c) No. We know the mean cost for students in the sample was \$1196.
d) No. A confidence interval is not about other sample means.
e) Yes. A confidence interval estimates a population parameter.

8. a) No. The confidence interval is not about the years in the sample.
b) No. The confidence interval does not predict what will happen in any one year.