Key

Chapter 9 Review

1) What is Inference?

   An educated guess about a parameter using data

2) What is inference based on?

   sampling distributions (histograms of samples)

3) When dealing with proportions, what symbol represents the unknown population proportion (claim)?

   p

4) When dealing with proportions, what symbol represents the sample proportion?

   $\hat{p}$

5) What is the [redacted] form for a confidence interval?

   estimate ± margin of error

   $$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

6) What 2 things can you do to decrease the margin of error in a confidence interval?

   1) decrease confidence level
   2) increase sample size

7) What does being 95% confident **mean**?

   In 95% of all samples, our interval will have our true proportion in it.

8) What are the 4 steps to a test of significance?

   1- Hypotheses
   2- Test Statistic
   3- P-value
   4- Conclusion

9) What do these tests compare?

   sample $(\hat{p})$ to claim (p)

10) What is the generic form for a test statistic?

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

11) What is a P-value? What is it really telling us about our sample?

The probability (chance) of getting our sample (or something more extreme). It tells us how likely our sample is, if our Ho is true.

12) What is alpha (α)? What do we use it for?

The significance level. We compare our p-value to α.

13) When do we reject our Ho? When do we fail to reject it?

when p-value < α. When p-value > α.

14) What does it mean if a sample is significant?

The pvalue < α, so we reject Ho.

α p-value

15) I have a significance level of 0.10. I get a sample that happens 8% of the time. Is it significant? What is my conclusion (reject Ho or fail to reject Ho)?

yes. p-value < α = 0.10.

α p-value

16) I have a significance level of 0.05. I get a sample that happens 8% of the time. Is it significant? What is my conclusion (reject Ho or fail to reject Ho)?

no. p-value > α = 0.05.

17) If we reject our Ho at α=0.05, will we reject it at α=0.01? Why or why not?

No! Example: $\alpha = 0.03$

18) If we reject our Ho at α=0.01, will we reject it at α=0.05? Why or why not?

yes! anything less than 0.01 is always less than 0.05

19) Tests of significance are really looking for evidence ...

<u>AGAINST</u> the Ho (claim).

20) Whenever we do a test of significance or a confidence interval, what are the 3 things we ***need*** to check for?

1) SRS
2) $pop \geq 10 \cdot n$
3) $n \geq 30$

---

21) I have a 98% confidence interval of certain data that is (0.65, 0.78).

a. Give a POSSIBLE 92% confidence interval for the same data

smaller: (0.68, 0.75)

b. Give a POSSIBLE [illegible] confidence interval for the sample data
99%

larger: (0.63, 0.80)

22) An entomologist samples a field for traces of a harmful insect by placing a square yard frame at random locations and carefully examining the ground within the frame for these traces. An SRS of 75 locations selected from a county's pastureland finds traces of the harmful insect in 13 locations.

a. Give a 95% confidence interval for the proportion of all possible locations that are infested, and interpret.

SRS ✓

$pq \geq 10 \cdot 75$ ✓

$n \geq 30$ ✓

$\hat{p} = \frac{13}{75}$ $\quad \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.08766, 0.259)$

We are 95% confident that the true percent of infested locations is between 8.76% and 25.9%.

b. We read a magazine that claims that 15% of all possible locations are contaminated by this harmful insect. Knowing this, find the minimum sample size needed to calculate a 92% confidence interval with a 6% margin of error.

$0.06 = 1.75\sqrt{\frac{(0.15)(0.85)}{n}}$ $\quad n = 109$

c. Using the claim in part (b), perform a test of significance to see if the true proportion of harmful insects has increased. Use the info in the beginning of the problem

$H_0: p = 0.15$

$H_a: p > 0.15$

$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = 0.566$

$P(z > 0.566) = 0.286$

We fail to reject Ho because p-value > $\alpha = 0.05$. We have sufficient evidence that the true ~~proportion~~ percent of infested locations is still 15%.

23) Find the minimum sample size needed if we want to calculate a 90% confidence interval with a margin of error of 4%.

$0.04 = 1.65\sqrt{\frac{(0.5)(0.5)}{n}}$ $\quad n = 426$

24) Of the 500 respondent households in a SRS, 43% had less than 2 cars at home. There was a census done a few years ago that said that 46% had less than 2 cars at home. Test whether the census is still accurate for the proportion of households with less than 2 cars at home or if it has gone down since the census was done.

SRS ✓

$n \geq 30$ ✓

pop ≥ 5000 ✓

$H_0: p = 0.46$

$H_a: p < 0.46$

$\hat{p} = 0.43$ $\alpha = 0.05$

$n = 500$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = -1.346$$

$P(z < -1.346) = 0.089$

We fail to reject Ho b/c p-value $> \alpha = 0.05$.

We have sufficient evid. that the true proportion of households w/ less than 2 cars is still 0.46.

25) Suppose that 84% of a sample of 125 nurses working the ER in city hospitals express positive job satisfaction, ~~while only 72% of a sample of 150 nurses working the general floors express similar fulfillment.~~ Establish and interpret a 90% confidence interval estimate for the ~~difference.~~ ER nurses.

$\hat{p} = 0.84 = \frac{105}{125}$

$n = 125$

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.78606, 0.89394)$$

SRS ✓

$n \geq 30$ ✓

pop ≥ 1250 ✓

We are 90% confident that the true percent of ER nurses expressing positive job satisfaction is between 78.61% and 89.39%.

26) You want to estimate the proportion of students at your college or university who are employed for 10 or more hours per week while classes are in session. You plan to present your results by way of a 95% confidence interval. Using the guessed value $\hat{p} = 0.35$, find the sample size required if the interval is to have an approximate margin of error of 0.05.

$z^* = 1.96$

$\hat{p} = 0.35$

$m = 0.05$

$$0.05 = 1.96\sqrt{\frac{(0.35)(0.65)}{n}}$$

$n = 350$

27) A recent study claims that by May, 65% of statistics students will have "senioritis." You believe that this proportion is actually lower (one reason being that not all statistics students are seniors). You take a SRS in May of 93 stat students (between all three CB High Schools) and find that 42 of them claim to have "senioritis." Test your claim at a significance level of 0.05.

SRS ✓

$n \geq 30$ ✓

$pop \geq 930$ ✓

$H_0: p = 0.65$

$H_a: p \leq 0.65$

$$z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = -4.011$$

$P(z < -4.011) = 3.023 \times 10^{-5}$

$\alpha = 0.05$

- We reject Ho b/c p-value $< \alpha = 0.05$.
- We have sufficient evidence that the true percent of stat students who will have senioritis by May is less than 65%..

28) A company is marketing its new toy for children ages 3-8, however they are interested in the proportion of 4 to 5-year-old children like the toy. They take a SRS of 53 4 to 5-year-old children and perform a series of tests to determine whether the child likes the toy or not. They determine that of the children in their sample, 28 like their toy. Using a 96% level of confidence, estimate the percent of children that like the toy.

SRS ✓

$n \geq 30$ ✓

$pop \geq 530$ ✓

$\hat{p} = \frac{28}{53}$

96% conf.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0.38748, 0.66913)$$

We are 96% conf. that the true proportion of 4-5 year olds that like the toy is btw 0.38748 and 0.66913.

29) We have a confidence interval is (0.03834, 0.12166). This interval is created from a sample of 150 people.

a. What is the sample proportion? 0.08

b. What is the margin of error? 0.04166

c. With what level of confidence was this interval created?

$$0.04166 = z^* \sqrt{\frac{(0.08)(0.92)}{150}}$$

$1.88 = z^*$

94% confidence

MULTIPLE CHOICE

A recent Gallup Poll asked "Do you consider the amount of federal income tax you have to pay as too high, about right, or too low?" 69% of the sample answered "Too high." Gallup says that for results based on the sample of national adults ($n = 1{,}055$) surveyed April 6-7, 1999, the margin of sampling error is ± 3 percentage points. **The next two questions** concern this poll.

1. The poll was carried out by telephone, so people without phones are always excluded from the sample. Any errors in the final result due to excluding people without phones
(a) are included in the announced margin of error
(b) are in addition to the announced margin of error
(c) can be ignored, because these people are not part of the population
(d) can be ignored, because this is a nonsampling error

2. If Gallup had used an SRS of size $n = 1055$ and obtained the sample proportion $\hat{p} = 0.69$, you can calculate that the margin of error for 95% confidence would be
(a) ± 0.02 percentage points
(b) ± 0.04 percentage points
(c) ± 1.4 percentage points
(d) ± 2.8 percentage points
(e) ± 3.0 percentage points

The student newspaper at a college asks an SRS of 250 undergraduates, "Do you favor eliminating the carnival from the term-end celebration?" In all 150 of the 250 are in favor. **The next five questions** concern this sample survey.

3. The _______ you want to estimate is the proportion p of all undergraduates who favor eliminating the carnival. That _______ should read
(a) bias (b) confidence level (c) mean
(d) parameter (e) statistic

4. To estimate p, you will use the proportion $\hat{p} = 150/250$ of your sample who favored eliminating the carnival. The number $\hat{p}$ is a
(a) bias (b) confidence level (c) mean
(d) parameter (e) statistic

5. A 95% confidence interval for the population proportion p is
(a) 150 ± 0.03 (b) 0.6 ± 0.03 (c) 150 ± 0.06
(d) 0.6 ± 0.06 (e) 1.67 ± 0.03

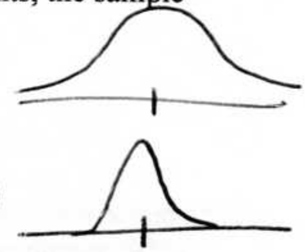
6. A 90% confidence interval based on this same sample would have
(a) the same center and a larger margin of error
(b) the same center and a smaller margin of error
(c) a larger margin of error and probably a different center
(d) a smaller margin of error and probably a different center
(e) the same center, but the margin of error changes randomly

8. You want to estimate the proportion of undergraduates at a college who favor eliminating evening exams. You will choose an SRS. If you enlarge your SRS from 250 to 1000 students, the sample proportion $\hat{p}$

(a) will have the same mean and the same standard deviation.
(b) will have smaller bias and the standard deviation will be 1/4 as large.
(c) will have smaller bias and the standard deviation will be 1/2 as large.
(d) will have the same mean and the standard deviation will be 1/4 as large.
(e) will have the same mean and the standard deviation will be 1/2 as large.

$n=250$

$n=1000$



9. The phrase "95% confidence" in a Gallup Poll press release means that
(a) our results are true for 95% of the population of all adults.
(b) 95% of the population falls within the margin of error we announce.
(c) the probability is 0.95 that a randomly chosen adult falls in the margin of error we announce.
(d) we got these results using a method that gives correct answers in 95% of all samples.

10. A recent Gallup Poll interviewed a random sample of 1523 adults. Of these, 868 bought a lottery ticket in the past year. A 95% confidence interval for the proportion of all adults who bought a lottery ticket in the past year is (assume Gallup used an SRS)
(a) 0.57 ± 0.00016
(b) 0.57 ± 0.00032
(c) 0.57 ± 0.013
(d) 0.57 ± 0.025
(e) 0.57 ± 0.03

$\hat{p} = \frac{868}{1523} = 0.57$ $\quad \hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$0.57 \pm$

16. If the value of the test statistic $z = 2.5$,
(a) conclude that the null hypothesis and the alternative hypothesis are the same
(b) we reject the null hypothesis at the 5% significance level
(c) we fail to reject the null hypothesis at the 5% significance level
(d) we reject the alternative hypothesis at the 5% significance level
(e) should use a different null hypothesis

P-value:

$P(z > 2.5) =$

normcdf(2.5, E99, 0, 1)

$= 0.00621$

17. The null hypothesis is
(a) another name for the alternative hypothesis
(b) true with 95% probability
(c) usually a statement of "no effect" or "no difference"
(d) determined by looking at the data
(e) statistically significant

18. If a significance test gives a P-value of 0.50,
(a) the margin of error is 0.50
(b) the null hypothesis is very likely to be true
(c) we do not have good evidence against the null hypothesis
(d) we do have good evidence against the null hypothesis
(e) the effect of interest is practically significant

20. If a significance test gives P-value 0.005,
(a) the margin of error is 0.005.
(b) the null hypothesis is very likely to be true.
(c) we do not have good evidence against the null hypothesis.
(d) we do have good evidence against the null hypothesis.
(e) the effect of interest is practically significant.

21. The report of a sample survey of 1,014 adults says, "With 95% confidence, between 9% and 15% of all Americans expect to spend more money on gifts this year than last year." The phrase "95% confidence" means
(a) 95% of all Americans will spend between 9% and 15% of what they spent last year.
(b) 9% to 15% of all Americans will spend 95% of what they spent last year.
(c) there is a 95% chance that the percent who expect to spend more is between 9% and 15%.
(d) the method used to get the interval from 9% to 15%, when used over and over, produces intervals which include the true population percentage 95% of the time.
(e) we can be 95% confident that the method used to get the interval always gives the right answer.

22. A sample survey finds that 30% of a sample of 874 Ohio adults said good health was the thing they were most thankful for. If that sample were an SRS from the population of all Ohio adults, what would be the 99% confidence interval for the percent of all Ohio adults who feel that way?
(a) 25% to 35% (b) 26% to 34% (c) 27% to 33%
(d) 28% to 32% (e) 29% to 31%

23. If the 874 people in the previous question had called a 900 number to give their opinions, how would this affect your answer?
(a) Not at all, because the width of the confidence interval depends only on the sample size, and not on the population size.
(b) Not at all, because the width of the confidence interval depends only on the sample size, and not on how the sample was obtained.
(c) It would be wider because voluntary response polls have a bigger margin of error than SRSs.
(d) It would be narrower because voluntary response polls are less variable than SRSs.
(e) A confidence interval makes no sense for a voluntary response sample.

24. The name for the pattern of values that a statistic takes when we sample repeatedly from the same population is
(a) the bias of the statistic.
(b) the sampling distribution of the statistic.
(c) the scale of measurement of the statistic.
(d) the variability of the statistic.
(e) the sampling error.

28. A CBS News/New York Times opinion poll asked 1,190 adults whether they would prefer balancing the Federal budget over cutting taxes; 702 of those asked said "Yes." Take the sample to be an SRS from the population of all adults. Which of these is a correct 95% confidence interval for the proportion of all adults who prefer balancing the budget over cutting taxes?

(a) 0.59 ± 0.0004
(b) 0.59 ± 0.014
(c) 0.59 ± 0.0186
(d) 0.59 ± 0.0285
(e) 0.59 ± 0.037

29. Suppose that in fact 62% of all adults favor balancing the budget over cutting taxes. The number 62% is

(a) a bias.
(b) a margin of error.
(c) a statistic.
(d) a parameter.
(e) a coefficient of variation.

30. Suppose that in fact 62% of all adults favor balancing the budget over cutting taxes. If you take a large number of SRSs of size 1,190, the sample proportions who favor balancing the budget will vary. Some will be lower than 62% and some will be higher, but the average sample result will be very close to 62%. This fact is called

(a) small bias.
(b) small margin of error.
(c) high variability.
(d) large bias.
(e) low variability.