AP Statistics

Chapter 7 notes and solutions

Example:

population, parameter, sample, and statistic

The Gallup Poll asked a random sample of 515 US adults whether or not they believe in ghosts. Of the respondents, 160 said "Yes."

population: All US adults

parameter: p, the proportion of all US adults who believe in ghosts

sample: 515 people interviewed

statistic: $\hat{p}$ = 160/515 = 0.31, the proportion of the sample who say they believe in ghosts

Solutions: 7.1

P. 436, #1,3,5

1a. Pop: all people who signed the card

parameter: proportion of pop who actually quit smoking

sample: random sample of 1000 people who signed the cards

statistic: the proportion of the sample who actually quit smoking

$\hat{p}$ = 0.21

1b. pop: all the turkey meat

parameter: minimum temp in all of the turkey meat.

sample: four randomly chosen location in the turkey

statistic: minimum temp in the sample of four locations; sample min = 170 F σ

3. $\mu = 2.5003$ is a parameter and

$\bar{x} = 2.5009$ is a statistic

5. $\hat{p} = .48$ is a statistic and

$p = .52$ is a parameter

For #7, use small population {2,6,8,10,10,12}

a. list all 15 possible SRSs of size n = 2 from the population. Find the value of $\bar{x}$ for each example

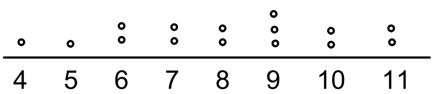
2 and 6, ($\bar{x} = 4$)
2 and 8, (5)
2 and 10 (6)
2 and 10 (6)
2 and 12 (7)

6 and 8. ($\bar{x} = 7$)
6 and 10, (8)
6 and 10, (8)
6 and 12, (9)

8 and 10, ($\bar{x} = 9$)
8 and 10, (9)
8 and 12, (10)

10 and 10, ($\bar{x} = 10$)
10 and 12, (11)
10 and 12, ($\bar{x} = 11$)

b. make a graph of the sampling distribution of $\bar{x}$. Describe what you see.



skewed left

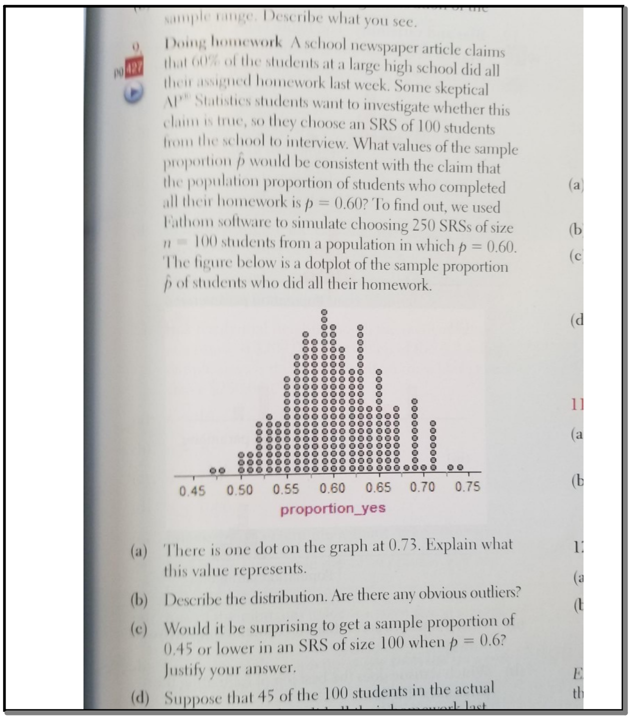
unimodal

mean is 8, which is mean of the pop.

the values of $\bar{x}$ vary from 4 to 11

9. Doing HW. A school newspaper article claims that 60% of the students at a large high school did all of their assigned hw last week. Some skeptical AP students want to investigate whether this claim is true, so they choose an SRS of 100 from the school to interview. What values of the sample proportion $\hat{p}$ would be consistent with the claim that the population proportion of the students who completed all their hw is p = 0.60? To find out, we used Fathom software to choosing 250 SRSs of size n = 100 students from a population in which $\hat{p}$ = 0.60. The figure below is a dotplot of the sample proportion p of students who did all of their hw.

a. Page 437... the one dot at 0.73... Explain what it means.



sample range. Describe what you see.

9. pg 427 **Doing homework** A school newspaper article claims that 60% of the students at a large high school did all their assigned homework last week. Some skeptical AP® Statistics students want to investigate whether this claim is true, so they choose an SRS of 100 students from the school to interview. What values of the sample proportion $\hat{p}$ would be consistent with the claim that the population proportion of students who completed all their homework is $p = 0.60$? To find out, we used Fathom software to simulate choosing 250 SRSs of size $n = 100$ students from a population in which $p = 0.60$. The figure below is a dotplot of the sample proportion $\hat{p}$ of students who did all their homework.

(a) There is one dot on the graph at 0.73. Explain what this value represents.

(b) Describe the distribution. Are there any obvious outliers?

(c) Would it be surprising to get a sample proportion of 0.45 or lower in an SRS of size 100 when $p = 0.6$? Justify your answer.

(d) Suppose that 45 of the 100 students in the actual

7.2 Sample Proportions Continued

Sample problem worked out from Page 445 and 446.

Going to College: polling organization asks an SRS of 1500 first year college students how far away their home is. Suppose that 35% of all first year students attend college within 50 miles of home.

Problem: find the probability that the random sample of 1500 students will give a result within 2 percentage points of this true value.

Solution:

Step 1: distribution and values of interest

probability that $\hat{p}$ falls between 0.33 and 0.37

$P(0.33 < \hat{p} < 0.37)$

$n = 1500$ $p = 0.35$

$\mu_{\hat{p}} = p = 0.35$

Standard Deviation

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1 - p)}{n}} = \sqrt{\frac{.35(1 - .35)}{1500}} = 0.0123$$

Large Counts condition:

np and n(1 - p) have to be greater than 10

np = 1500(0.35) = 525

n(1 - p) = 1500(1 - 0.35) = 975

So, Normal approximation will be quite accurate

Perform Calculations using Table A.

Find $P(0.33 < \hat{p} < 0.37)$

$$\frac{X - \mu}{\sigma} = \frac{0.33 - 0.35}{.0123} = -1.63$$

$$\frac{X - \mu}{\sigma} = \frac{0.37 - 0.35}{.0123} = 1.63$$

z = -1.63 = .9484

x = 1.63 = .0516

And... .9484 - .0516 = 0.8968

So, about 90% of all SRS of 1500 will give a results within 2 percentage points of the truth about the population.

Solutions: 7.2

Page 447, #27, 29, 33

27. a. We would not be surprised to find 8 (32%) orange candies because values this small happened fairly often in the simulation. However, there were a few samples in which there were 5 (20%) or fewer orange candies. Thus, getting 5 orange candies would be surprising.

b. A sample of 50, because we expect to be closer to p = 0.45 in larger samples.

29. 45% orange candies. SRS = 25

observe the sample proportion $\hat{p}$ of orange candies.

a. $\mu_{\hat{p}} = p = 0.45$

the mean of the sampling distribution is equal to the parameter p.

b. $\sigma_{\hat{p}} = \sqrt{\frac{p(1 - p)}{n}} = \sqrt{\frac{.45(1 - .45)}{25}} = 0.0955$

10% condition is met because there are more than 250 candies in the large machine.

c. Yes, because np = 25(0.45) = 11.25

and n(1 - p) = 25(1 - .45) = 13.75 are both at least 10.

Large Counts condition: $np \geq 10$ and $n(1 - p) \geq 10$

d. If sample size is 100 rather than 25, how does this change the sampling distribution of $\hat{p}$?

a. $\mu_{\hat{p}} = p = 0.45$

the mean of the sampling distribution is equal to the parameter p.

b. $\sigma_{\hat{p}} = \sqrt{\frac{p(1 - p)}{n}} = \sqrt{\frac{.45(1 - .45)}{100}} = 0.0497$

33. 3000 unionized workers. 30% are hispanic. 15-member union exec. committee contains 3 hispanics. what would be the probability of 3 or fewer Hispanics if the exec. committee were chosen at random from all the workers?

Large counts condition

$np = 15(0.3) = 4.5...$ $4.5 < 10$ so the condition is NOT met.

On your own... with a little help.

Page 448, #35, 37, and 39

35. Drink the cereal milk?

1012 U.S. adults... what do you do with the milk after the cereal is eaten? Dairy industry claims that 70% drink it. Suppose this is true.

a. What is the mean of the sampling distribution of $\hat{p}$?

$\mu_{\hat{p}} = p =$

b. Find the standard deviation. Check to see if the 10% condition is met.

*Is 1012 less than 10% of all US population? if so, the condition has been met.

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

c. Is the sampling distribution of p approximately normal?

If the large counts condition has been met, then the distribution is approx. normal.

find np =

find n(1 - p) =

both have to be at least 10 to meet the condition.

d. Of the poll respondents, 67% say they drink the milk. Find the probability of obtaining a sample of 1012 adults in which 67% or fewer say they drink the milk.

Find $P(\hat{p}<0.67)$

$$\frac{X - \mu}{\sigma} = \frac{0.67 - 0.70}{.0144} = -2.08$$

From Table A: $P(Z < -2.08) = 0.0188$

so, there is a .0188 probability of obtaining a sample in which 67% or fewer drink the milk.

37.

39.

A tire manufacturer designed a new tread pattern for its all-weather tires. Repeated tests were conducted on cars of approximately the same weight traveling at 60 miles per hour. The tests showed that the new tread pattern enables the cars to stop completely in an average distance of 125 feet with a standard deviation of 6.5 feet and that the stopping distances are approximately normally distributed.

(a) What is the 70th percentile of the distribution of stopping distances?

A tire manufacturer designed a new tread pattern for its all-weather tires. Repeated tests were conducted on cars of approximately the same weight traveling at 60 miles per hour. The tests showed that the new tread pattern enables the cars to stop completely in an average distance of 125 feet with a standard deviation of 6.5 feet and that the stopping distances are approximately normally distributed.

(a) What is the 70th percentile of the distribution of stopping distances?

$$Z = \frac{x-\mu}{\sigma}$$

AREA Z

$.70 = .525$

$$.525 = \frac{(X-125)}{6.5}$$

$$6.5 \cdot .525 = \frac{(X-125)}{6.5} \cdot 6.5$$

$$3.4125 = X - 125$$

$+125 \qquad +125$

$$128.4 = X$$

* from standard normal table

or... TI-83: 2nd VARS 3:invNorm(

then enter the percentile to

get the Z-score

How to Find Inverse Normal on the TI-83 with the InvNorm Command

The InvNorm function (Inverse Normal Probability Distribution Function) on the TI-83 gives you an x-value if you input the area (probability region) to the left of the x-value. The area must be between 0 and 1. You must also input the mean and standard deviation.

Step 1: Press 2nd then VARS to access the DISTR menu.

Step 2: Arrow down to 3:invNorm(and press ENTER.

Step 3: Type the area, mean and standard deviation in the following format:

invNorm (probability,mean,standard deviation).

invNorm(.70, 125,6.5) = 128.408

(b) What is the probability that at least 2 cars out of 5 randomly selected cars in the study will stop in a distance that is greater than the distance calculated in part (a) ?

The probability that the stopping distance is GREATER than 128.4 is

1 - 0.70 = 0.30

Let Y = the number out of 5 cars that stop in a distance greater than 128.4 feet. Y is a BINOMIAL RANDOM VARIABLE.

n = 5 and p = 0.30 k = 2

BINOMIAL PROBABILITY

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

MATH - PRB - 3:nCr

$P(Y \geq 2) = 1 - P(Y < 2) =$

$$=1- \left[\binom{5}{0}(.3)^0(.7)^5 + \binom{5}{1}(.3)^1(.7)^4\right]$$

.16807 .36015

$P(Y \geq 2) = 1 - .52822 = .47478$

47% that at least 2 of the 5 cars selected will stop, taking more than 128.5 feet to do so.

Binomial CDF (better way)

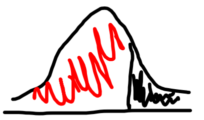
"cumulative Distribution function)

2nd VARS - A: binomcdf(

(Trials, Prob, upper bound X)

(5, .3, 1)

X = 0.52822



and... 1 - .052822 = .47178

So, P(Y≥2) = .47178

What is the probability that a randomly selected sample of 5 cars in the study will have a mean stopping distance of at least 130 feet?

Let $\bar{x}$ be the mean of the stopping distances of 5 randomly selected cars.

remember... $\mu_{\bar{x}} = \mu$

remember... $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$

So, x is normal with a mean of 125 ft and a standard deviation of $6.5/\sqrt{5}$

To find P(X ≥ 130), use normalCDF function

normalcdf(lower, upper, υ, σ)

2nd VARS - normalcdf(130,1000,125,6.5/sqrt5)

So, P(x > 130) = 0.0427

or, there is a .043 probability that these 5 cars (on average) will take 130 feet or more to stop.

NormalCDF means normal cumulative distribution function. We use it to find the probability that a variable will fall into a certain interval... in this case, over 130 feet.

7.3 Study help

Movies

The number of movies viewed in the last year by HS students has an average of 19.3 with a standard deviation of 15.8. Suppose we take a SRS of 100 HS students and calculate the mean number of movies viewed by the members of the sample.

a. What is the mean of the sampling distribution of $\overline{x}$?

$\mu_x = \mu = 19.3$ movies.... Because $\overline{x}$ is an unbiased estimator of μ

b. What is the standard deviation of the sampling distribution of x? Check that the 10% condition is met.

$\sigma_x = \frac{\sigma}{\sqrt{n}} = \frac{15.8}{\sqrt{100}} = 1.58$ movies

The 10% condition is met because there are more than 10(100) = 1000 HS Students

Google, "Online statbook sampling distributions applet"

7.3-1: Sample Means

Page 461, #49, 51, 53

49. $\mu_{\bar{x}} = \mu = 225$ seconds.

Because the sample size (10) is less than 10% of the population of songs on David's ipod,

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{10}} = 18.974$ seconds

7.1 to 7.3-1 Student Questions

I´m confused on the formulas for the probability, i know like the standard deviation and mean for the most part but the other things have me very confused.

A refresher of how to use the binomial probability formula would be awesome.

I think a rough overview of everything then let people ask questions or ask you to specify. I personally can say I am not at all confident in my abilities during this chapter.

I don't think I have questions, but you could talk about section 7.3, the #49 problem :)

I don't really have any question.

Nothing I need, I am all good sir

The 7.2 Sample Proportions Supplemental homework for problems 37 and 39. As well as the 7.2 Sample Means homework for all of the three problems.

We should talk more about the Formulas to find the mean and standard devotion of the sampling distributions.

I think some extra in-class questions, or posting some of the college board practice quizzes would be good.

I have no doubts so far

nothin

7.3 answers

49. $\mu_{\bar{x}} = \mu = 225$ seconds

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{10}} = 18.974$ seconds

* sample size of 10 is less than 10% of population of songs on David's Ipod.

NEW HOMEWORK!

Page 466, #1, 4, 6

51. $30 = \frac{60}{\sqrt{n}}$

Solving for n... n = 4 songs.

53. a. Normal with $\mu_{\bar{x}}$ = 188 mg/dl

Because the sample size (100) is less than 10% of all men age 20 to 34,

$\sigma_{\bar{x}} = \frac{41}{\sqrt{100}}$ = 4.1 mg/dl

b. there is a 0.5357 probability that $\bar{x}$ estimates μ within +/- 3 mg/dl

c. There is a .9790 probability that $\bar{x}$ estimates μ within +/- 3 mg/dl

Welcome! Monday, February 15th, 2021

Try this one on for size!

"Has anyone ever noticed that things in our world haven't gone so well ever since the Planter's Baby Nut showed up?"

Buy Me Some Peanuts and Sample Means

1. At the P. Nutty Peanut Company, dry roasted, shelled peanuts are placed in jars by a machine. The distribution of weights in the bottles is approximately Normal, with a mean of 16.1 ounces and a standard deviation of 0.15 ounces.

(a) Without doing any calculations, explain which outcome is more likely, randomly selecting a single jar and finding the contents to weigh less than 16 ounces or randomly selecting 10 jars and finding the average contents to weigh less than 16 ounces.

(b) Find the probability of each event described above. Since the distribution is normal you can use "normalcdf" on your calculator.

Single jar weighing 16 oz or less:

A. Since averages are less variable than individual measurements, I would expect the sample mean of 10 jars to be closer, on average, to the true mean of 16.1 ounces. Thus, it is more likely that a single jar would weigh less than 16 ounces than the average of 10 jars to be less than 16 ounces.

Buy Me Some Peanuts and Sample Means

1. At the P. Nutty Peanut Company, dry roasted, shelled peanuts are placed in jars by a machine. The distribution of weights in the bottles is approximately Normal, with a mean of 16.1 ounces and a standard deviation of 0.15 ounces.

(b) Find the probability of each event described above. Since the distribution is normal you can use "normalcdf" on your calculator.

Single jar weighing 16 oz or less:

normalcdf(lower limit, upper limit, mean, standard deviation)

2nd - VARS - 2:normalcdf(

0,16,16.1,.15) = .2525

(b) Find the probability of each event described above. Since the distribution is normal you can use "normalcdf" on your calculator.

Single jar weighing 16 oz or less:

normcdf(0, 16, 16.1, 0.15) = 0.2525.

10 jars weighing 16 oz or less:

mean: 16.1..... s.d. = s.d./sqrt (n)

normcdf(0, 16, 16.1, $0.15/\sqrt{10}$) = 0.0175.

This answer agrees with the answer to part (a) because this probability is much smaller than 0.2525.

Welcome! Tuesday, February 16th, 2021

Almost finished!

0:00:00

Suppose that the number of texts sent during a typical day by a randomly selected high school student follows a right skewed distribution with a mean of 15 and a standard deviation of 35. Assuming that students at your school are typical texters, how likely is it that a random sample of 50 students will have sent more than a total of 1000 texts in the last 24 hours?

Mean: 15

Standard Deviation: $35/\sqrt{50}$

Independent: In a large school, it is reasonable to assume there are at least 500 students, so the sample size of 50 would be less than 10% of the population.

Normal: Since n is large (50 > 30), it is reasonable to consider the distribution is approximately Normal.

normalcdf (20, 5000, 15, $\sqrt{35}/50$) = 0.1562.

Conclude: There is about a 16% chance that a random sample of 50 high school students will send more than 1000 texts in a day.

Page 466, #1,4,6

1. Bad Eggs

Population: set of all eggs shipped in one day.

Sample: 200 eggs examined

parameter: proportion p = 0.03 of all eggs shipped that day that have salmanella.

statistic is proportion $\hat{p}$ = 9/200 = 0.045 of eggs in the sample that had salmanella

4. Do you jog?

a. $\mu_{\hat{p}} = p = 0.15$

b. $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \frac{.15(.85)}{1540}$

= 0.0091

c. yes, becasue np = 1540(0.15) = 231

and n(1 - p) = 1540(0.85) = 1309 are both at least 10.

d. We want to find $P(0.13 < \hat{p} < .17)$

$z = \frac{.13 - .15}{0.0091} = -2.20$

$z = \frac{.17 - .15}{0.0091} = 2.20$

So, desired probability is $P(-2.20 \leq Z \leq 2.20)$

= .9720

There is a .9720 probability of obtaining a sample in which between 13% and 17% are joggers.

6. a. X = WAIS score for randomly selected individual follows an N(100, 15) distribution and we want to find $P(X \geq 105)$

$z = \frac{105 - 100}{15} = 0.33$

$P(Z \geq 0.33) = 0.3707$

b. $\mu_{\bar{x}} = \mu = 100$

Because the sample size 60 is less than 10% of all adults,

$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{60}} = 1.9365$

c. $\bar{x}$ follows an N(100, 1.9365) distribution and we want to find $P(\bar{x} \geq 105)$

$z = \frac{105 - 100}{1.9365} = 2.58$

$P(Z \geq 2.58) = 0.0049$

There is a .0049 probability fo selecting a sample of 60 adults whose mean WAIS score is at least 105.

The Candy Machine

1. Suppose a VERY large candy machine has 15% orange candies. Imagine taking an SRS of 25 candies from the machine

and observing the sample proportion ̂? of orange candies.

(a) What is the mean of the sampling distribution of ̂? Why?

$\mu_{\hat{p}} = p = 0.15$

(b) Check to see if the 10% condition is met.

10% condition is met because there is likely more than 250 candies in the machine.

(c) Find the standard deviation of the sampling distribution of ̂

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.15(.85)}{25}} = 0.0714$$

(d) Is the sampling distribution of ̂ approximately Normal? Check to see if the Normal condition is met.

The sampling distribution is not approx normal because np = 25(.15) = 3.75 is less than 10. n(1 - p) = 25(.85) = 21.25

*both numbers have to be at least 10 for this to be approximately normal.

(e) If the sample size were 75 rather than 25, how would this change the sampling distribution of ̂? How would this impact the Normal condition?

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.15(.85)}{75}} = 0.0412$$

np = 75(.15) = 11.25

n(1 - p) = 75(.85) = 63.75

both are least 10