Homework\_4

Priya Shaji

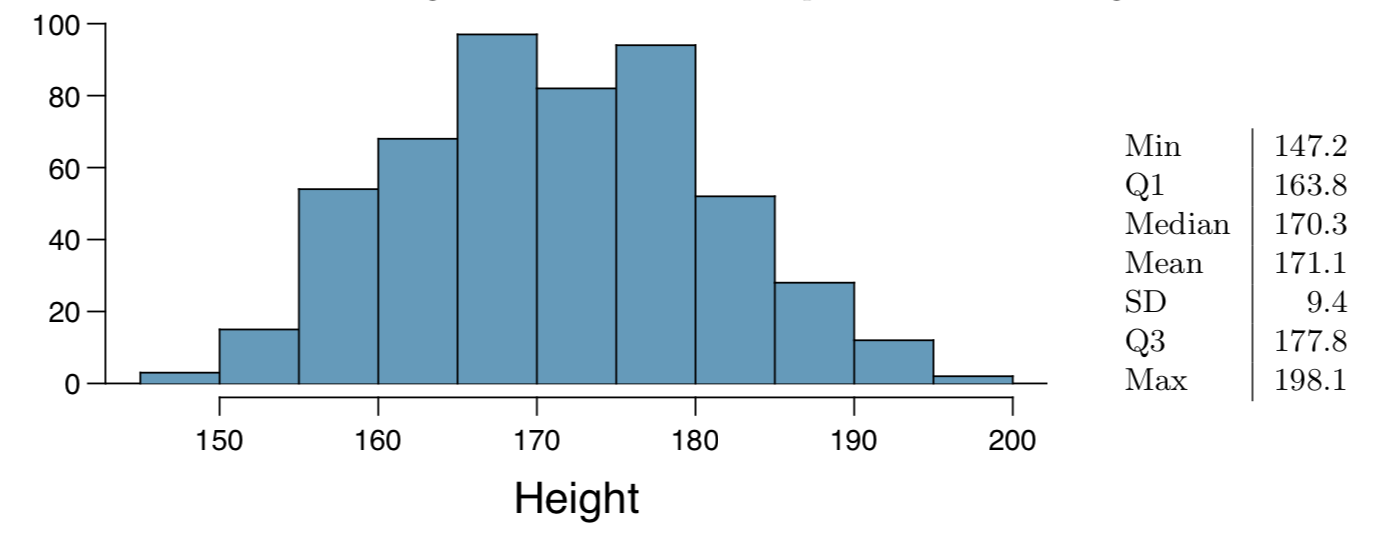
3/4/2019

## Foundations for Inference

### 4.4

Heights of adults. Researchers studying anthropometry collected body girth measurements and skeletal diameter measurements, as well as age, weight, height and gender, for 507 physically active individuals.

1. What is the point estimate for the average height of active individuals? What about the median? The histogram below shows the sample distribution of heights in centimeters



Answer a)

point estimate mean: 171.1 median: 170.3

1. What is the point estimate for the standard deviation of the heights of active individuals? What about the IQR?

Answer b)

Standard deviation: 9.4

# IQR: Q3-Q1  
177.8-163.8

## [1] 14

IQR: 14

1. Is a person who is 1m 80cm (180 cm) tall considered unusually tall? And is a person who is 1m 55cm (155cm) considered unusually short? Explain your reasoning.

Answer c)

180 cm is not so tall. 180cm is 1 PE SD above mean, and 155cm is 1.5 PE SD below mean. It is unusual for person to be this short.

1. The researchers take another random sample of physically active individuals. Would you expect the mean and the standard deviation of this new sample to be the ones given above? Explain your reasoning.

Answer d)

For another random sample the point estimates would be similar, but not same. It will be close.

1. The sample means obtained are point estimates for the mean height of all active individuals, if the sample of individuals is equivalent to a simple random sample. What measure do we use to quantify the variability of such an estimate (Hint: recall that SDx ̄ = pn )? Compute this quantity using the data from the original sample under the condition that the data are a simple random sample.

Answer e)

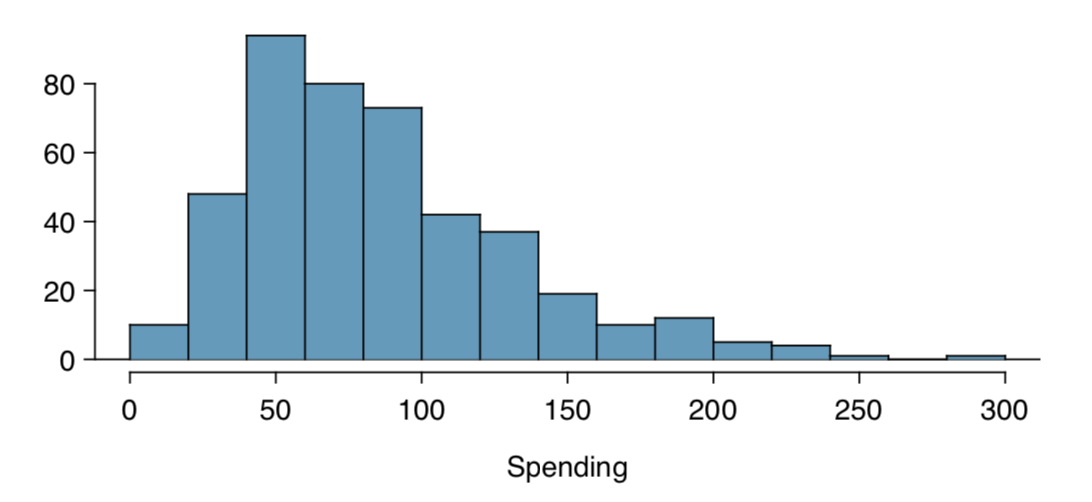
Compute Standard Error for sample mean using below formula

9.4/sqrt(507)

## [1] 0.4174687

### 4.14

Thanksgiving spending, Part I. The 2009 holiday retail season, which kicked o↵ on November 27, 2009 (the day after Thanksgiving), had been marked by somewhat lower self-reported consumer spending than was seen during the comparable period in 2008. To get an estimate of consumer spending, 436 randomly sampled American adults were surveyed. Daily consumer spending for the six-day period after Thanksgiving, spanning the Black Friday weekend and Cyber Monday, averaged $84.71. A 95% confidence interval based on this sample is ($80.31, $89.



1. We are 95% confident that the average spending of these 436 American adults is between $80.31 and $89.11.

Answer a)

False- The point estimate is always in the confidence interval.

1. This confidence interval is not valid since the distribution of spending in the sample is right skewed.

Answer b)

False - data is right skewed but not stronlgy skewed.

1. 95% of random samples have a sample mean between $80.31 and $89.11.

Answer c)

False

1. We are 95% confident that the average spending of all American adults is between $80.31 and $89.11.

Answer d)

True

1. A 90% confidence interval would be narrower than the 95% confidence interval since we don’t need to be as sure about our estimate.

Answer e)

True

1. In order to decrease the margin of error of a 95% confidence interval to a third of what it is now, we would need to use a sample 3 times larger.

Answer f)

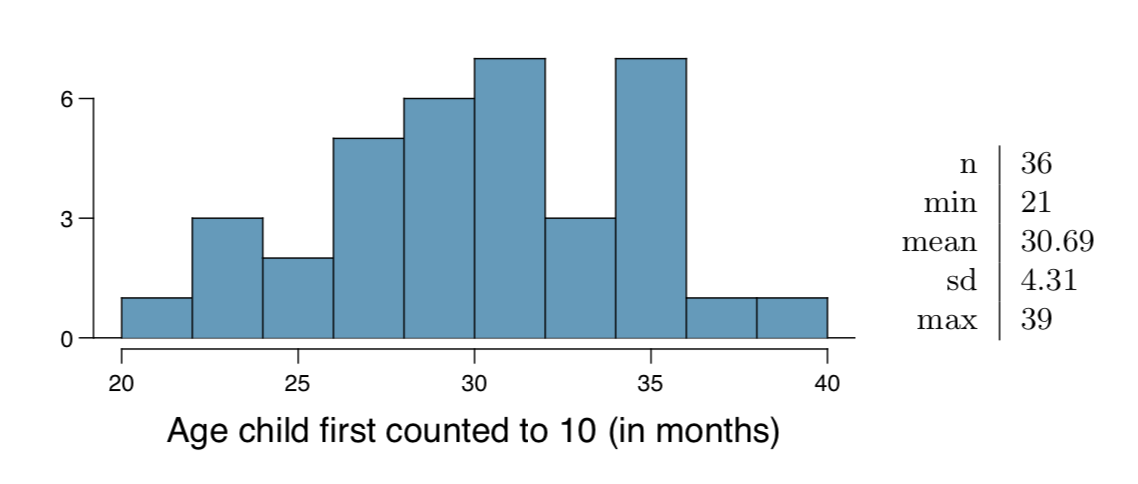
False- To decrease the MOE to onethird of current value we need 9 times current sample.

1. The margin of error is 4.4.

True

### 4.24

Gifted children, Part I. Researchers investigating characteristics of gifted children collected data from schools in a large city on a random sample of thirty-six children who were identified as gifted children soon after they reached the age of four. The following histogram shows the distribution of the ages (in months) at which these children first counted to 10 successfully. Also provided are some sample statistics



1. Are conditions for inference satisfied?

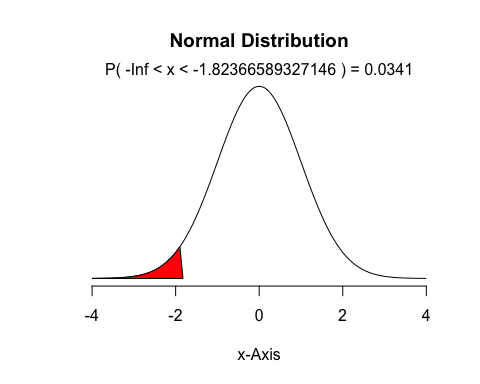
Answer a)

Yes.There is Random and large sample and no skeweredness.

1. Suppose you read online that children first count to 10 successfully when they are 32 months old, on average. Perform a hypothesis test to evaluate if these data provide convincing evidence that the average age at which gifted children fist count to 10 successfully is less than the general average of 32 months. Use a significance level of 0.10.

Answer b)

n = 36  
mn = 30.69  
sd = 4.31  
se <- sd / sqrt(n)  
z = (mn - 32) / se  
  
p=pnorm(z)  
normalPlot(bounds = c(-Inf, z))



1. Interpret the p-value in context of the hypothesis test and the data.

Answer c)

p

## [1] 0.0341013

When compared with significance level of .1, we can reject null hypothesis.

1. Calculate a 90% confidence interval for the average age at which gifted children first count to 10 successfully.

Answer d)

z90= 1.645  
lower = mn-z90 \*se  
upper = mn+z90 \*se  
c(lower,upper)

## [1] 29.50834 31.87166

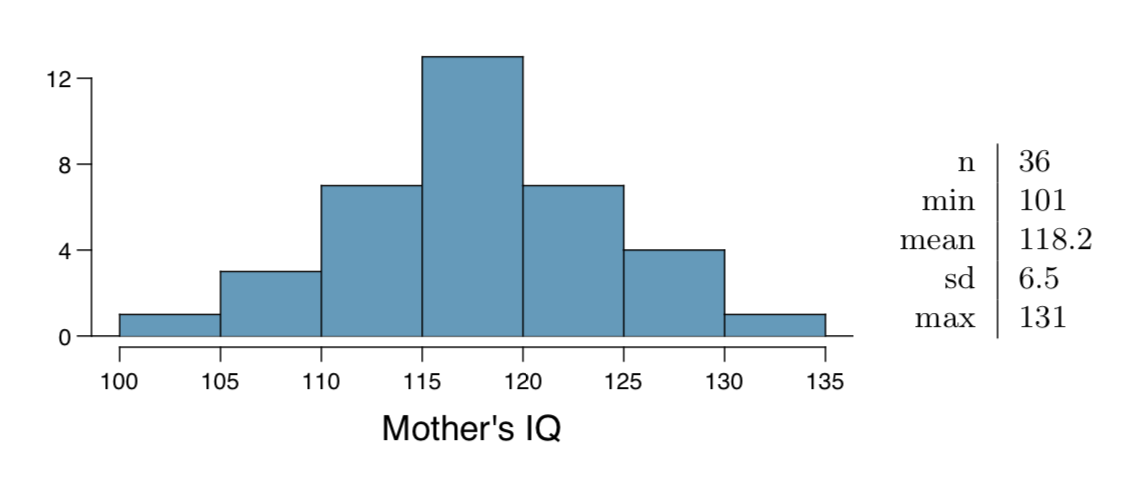
1. Do your results from the hypothesis test and the confidence interval agree? Explain.

Answer e)

Yes our results agree if we see the range of confidence interval is less than 32.

### 4.26

Gifted children, Part II. Exercise 4.24 describes a study on gifted children. In this study, along with variables on the children, the researchers also collected data on the mother’s and father’s IQ of the 36 randomly sampled gifted children. The histogram below shows the distribution of mother’s IQ. Also provided are some sample statistics.



1. Perform a hypothesis test to evaluate if these data provide convincing evidence that the average IQ of mothers of gifted children is different than the average IQ for the population at large, which is 100. Use a significance level of 0.10.

Answer a)

mn = 118.2  
n=36  
sd=6.5  
se=sd/sqrt(n)  
  
z = (mn-100)/se  
pnorm(z)

## [1] 1

p

## [1] 0.0341013

Reject the null hypothesis, conclusion the avg IQ of mother is diff that population at large.

1. Calculate a 90% confidence interval for the average IQ of mothers of gifted children.

Answer b)

lower = mn-z90 \*se  
upper = mn+z90 \*se  
c(lower,upper)

## [1] 116.4179 119.9821

1. Do your results from the hypothesis test and the confidence interval agree? Explain.

Answer c)

If the null hypothesis is true, the probability of observing a sample mean of 118.2 is only 0.003. This is a lot less than the significance value of 0.10. Thus we can reject the null hypothesis and say that for this population (mother’s of gifted children), their IQ is greater than the general population.

### 4.34

CLT. Define the term “sampling distribution” of the mean, and describe how the shape, center, and spread of the sampling distribution of the mean change as sample size increases.

Answer:

A sampling distribution shows the distribution of n samples from a population.As sample size increases, the shape approaches normal distribution, center is more pronounced and the spread becomes lean with more smaples near the mean.

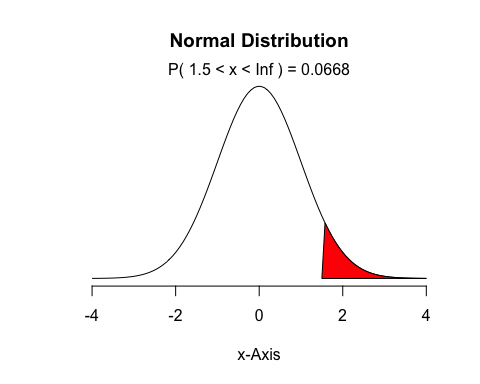
### 4.40

CFLBs. A manufacturer of compact fluorescent light bulbs advertises that the distribution of the lifespans of these light bulbs is nearly normal with a mean of 9,000 hours and a standard deviation of 1,000 hours.

1. What is the probability that a randomly chosen light bulb lasts more than 10,500 hours?

Answer a)

mn=9000  
sd= 1000  
z = (10500 - mn) / sd  
p = 1 - pnorm(z)  
normalPlot(bounds = c(z, Inf))



1. Describe the distribution of the mean lifespan of 15 light bulbs.

Answer b)

SE = sd/sqrt(15)  
SE

## [1] 258.1989

Distribution is normal.

1. What is the probability that the mean lifespan of 15 randomly chosen light bulbs is more than 10,500 hours?

Answer c)

z= 10500-9000/se  
z

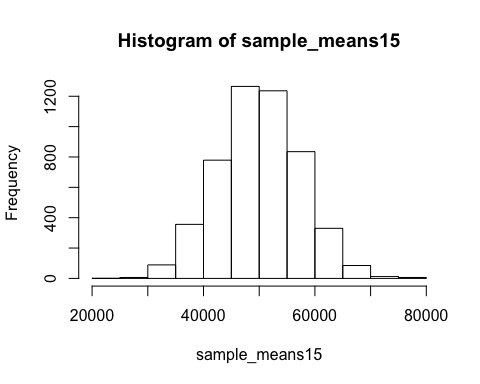
## [1] 2192.308

z score is high. It is highly unlikely that mean lifespan will be greater that 10500.

1. Sketch the two distributions (population and sampling) on the same scale.

Answer d)

p <- 100000  
  
sample\_means15 <- rep(NA, 5000)  
  
for(i in 1:5000){  
 samp <- sample(p, 15)  
 sample\_means15[i] <- mean(samp)  
 }  
  
hist(sample\_means15)



1. Could you estimate the probabilities from parts (a) and (c) if the lifespans of light bulbs had a skewed distribution?

Answer e)

No the estimate require a dist with little skewness.

### 4.48

Same observation, different sample size. Suppose you conduct a hypothesis test based on a sample where the sample size is n = 50, and arrive at a p-value of 0.08. You then refer back to your notes and discover that you made a careless mistake, the sample size should have been n = 500. Will your p-value increase, decrease, or stay the same? Explain.

Answer)

As the sample size increases the spread becomes narrower and the sd deviation becomes smaller. A smaller SD results in a larger z-score which decreases the p-value. That is, if the null hypothesis is true, as you increase the sample size, you stregthen the case that the null hypothesis is true (p-value decreases).