

Answer :

The point estimate for the average height of active individuals is, mean, 171.1

The median of active individuals is Median 170.3



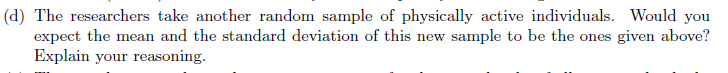
PE Mean = 171.1 and PE Median = 170.3



Answer :

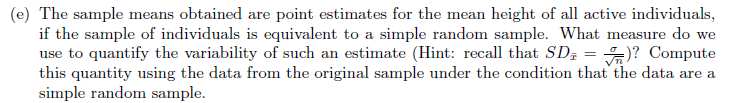
It depends on what is considered unusual. 180 cms. is above Q3 (75th percentile) of the point sample and 155 cms. is below Q1 (25th percentile). If 'unusual' means above the Q3 percentile or below the Q1 percentile, then both data fit the definition. If 'unusual' means 2 standard deviations from the mean, then both data are NOT unusual since both are within 1 standard deviation of the mean. But since this is only a point sample of the entire population, no conclusions should be drawn about the population data.

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.



Answer :

No, the mean and standard deviation would mostly be different. Since the random sample is from the population of active individuals, the mean and the standard deviation should be different , It will be close.

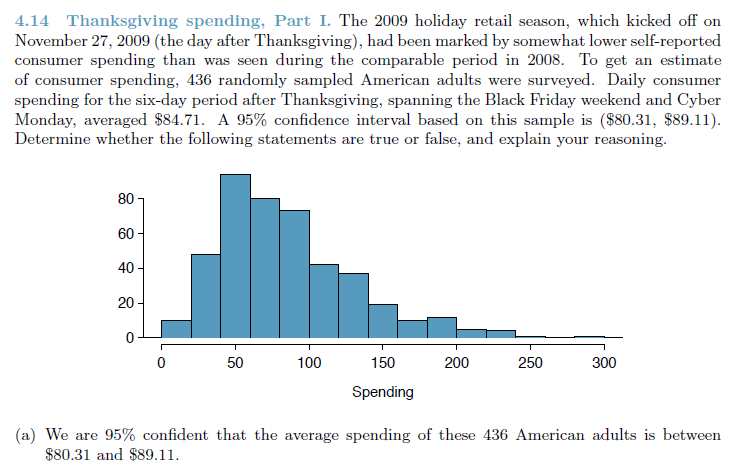


Answer :

SD(mean of sample X) = SD of sample X/sqrt of sample n

= 9.4 / SQRT(507)

= 0.42



Answer :

We can say that its for the sample and population in whole. In this case it would be FALSE to say that only 436 will be under the same confidence variable of 80.31 to 89.11.



Answer :

False - data is right skewed, but its very few outliers as the sample size would increase it would tend to form a normal distribution .



Answer :

FALSE; This mean was specific to the sample , a different sample may have different mean.



Answer :

Contradictory to A , its TRUE as we are using confidence variable for Population.



Answer :

TRUE, since the number of people that might fall in the range would be less than the 95% of the population.



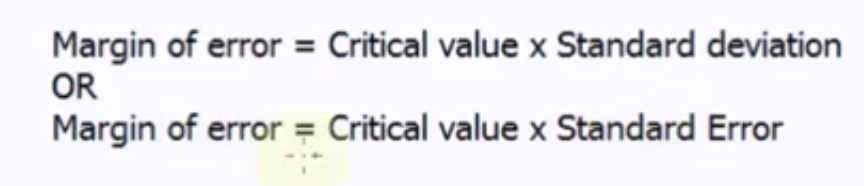
Answer :

A sample size of 3 times bigger is not enough, as the standard error equals to se = sd/sqrt(n)

. To make the confidence interval shrink to one third of what it is now, we would need a

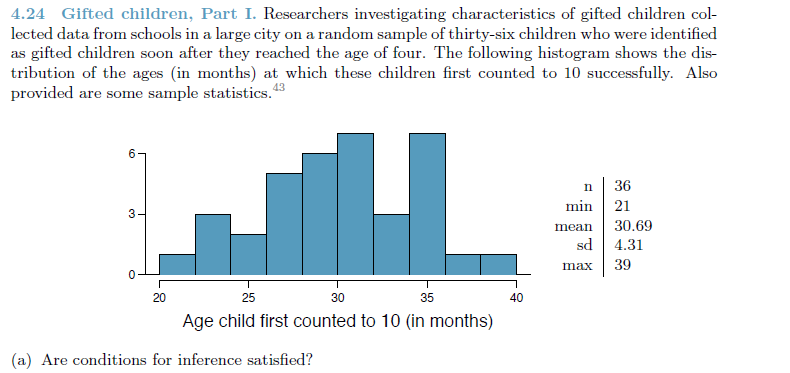
sample size of 9 times bigger.



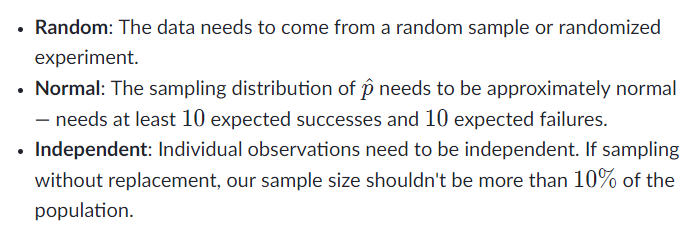


Answer :

(89.11- 80.31)/2=8.8/2 = 4.4 => Margin of error



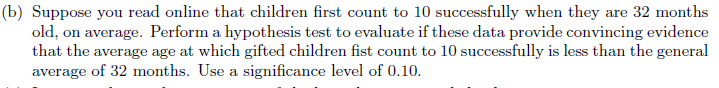
SELF NOTE:



Answer :

1. Independence; Yes As data is collected from Schools of the city , assuming all types of schools are considered.
2. sample size >= 30;
3. The distribution is bi-modal and not too much skewed.
4. Normal : The sampling distribution of  p^​ is approximately normal as long as the expected number of successes and failures are both at least 10.

​expected successes: np≥10  
expected failures: n(1−p)≥10​



Answer :

HO: The mean age is 32 months to count to 10.

HA: The mean age is less than 32 months to count to 10.

HO:μ=32  
HA:μ<32

Decision rule: Reject the null if the pvalue is less than 10%.

Calculate Z:

samp\_meanHW <- 30.69

samp\_sdMW <- 4.32

N <- 36

stdError <- samp\_sdMW / sqrt(N)

ztest <-(samp\_meanHW - 32) / stdError

p=pnorm(ztest)

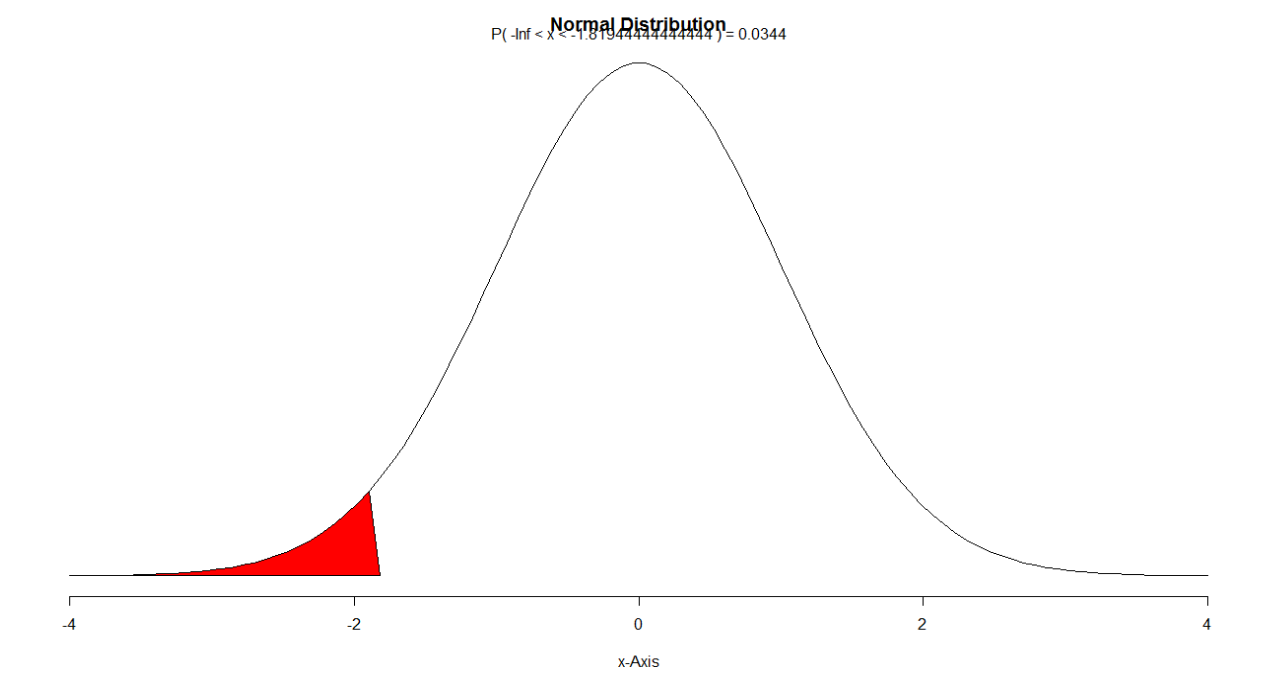
DATA606::normalPlot(bounds = c(-Inf, ztest))

lowerHW <- samp\_meanHW - Z \* samp\_sdMW / sqrt(N)

upperHW <- samp\_meanHW + Z \* samp\_sdMW / sqrt(N)  
# lowerHW <- samp\_meanHW - ztest \* samp\_sdMW / sqrt(N)

# upperHW <- samp\_meanHW +ztest \* samp\_sdMW / sqrt(N)

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





Answer :

if the null hypothesis is true, the probability of observing a sample mean of 30.69 is 0.34. This is greater than significance 0.10 which means we cannot reject the null hypothesis.



Answer :

SE = 4.31/SQRT(36) = 0.72

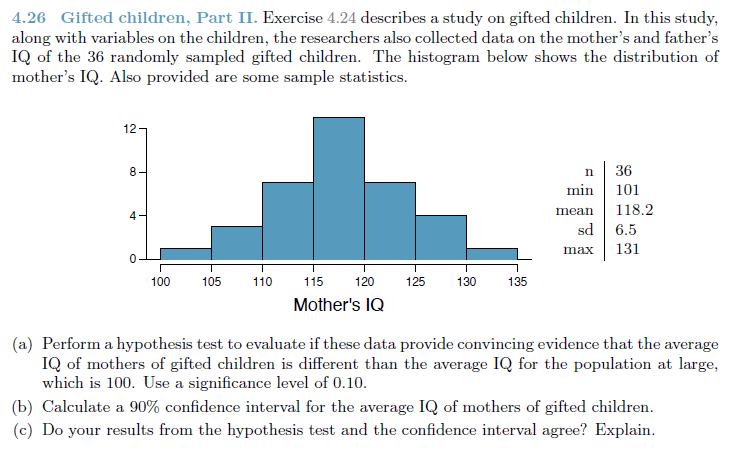
30.87 +/- (z-score for 90% confidence interval \* SE)

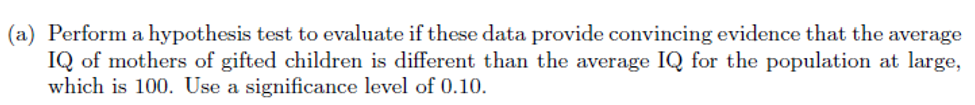
30.69 +/- (1.645 \* 0.72) = (29.51, 31.87)



Answer :

Yes our results agree if we see the range of confidence interval is less than 32. For 90% of the time.





Answer :

n = 36

mean = 118.2

s = 6.5

cl <- 10 # For alfa = 90%

alpha <- 1-(cl/100)

cp <- 1-(alpha/2)

z90 <- qnorm(cp)

n= 36

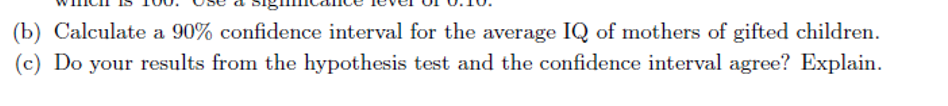
se=sd/sqrt(n)

z10 <- (mean - 100) / se

pnorm(z10)

## [1] 1

# Reject the null hypothesis, The data favor the claim that the mothers of gifted children have higher mean IQ than mothers in general.



Answer :

# Caculate Z score value

cl <- 90 # For alfa = 90%  
se=sd/sqrt(n)

alpha <- 1-(cl/100)

cp <- 1-(alpha/2)

z90 <- qnorm(cp)

z90

# Calcualte confidence Interval

mn = 118.2

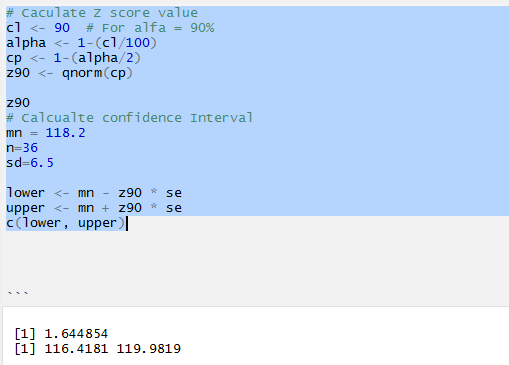
n=36

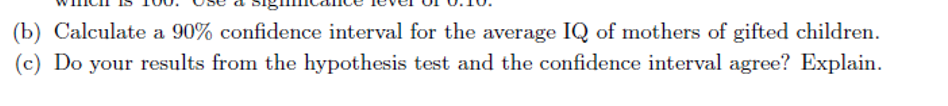
sd=6.5

lower <- mn - z90 \* se

upper <- mn + z90 \* se

c(lower, upper)





Answer :

The 90% CI (116.4,119.98) does not contain the proposed mean (100) from the null hypothesis. This is another way that we can see that we should reject the null hypothesis at the 10% level of significance.



Answer :

Sampling Distribution of the Mean is the plot of all the Mean of all the sample of size N from the given distribution.

Let’s say below is my given sample .

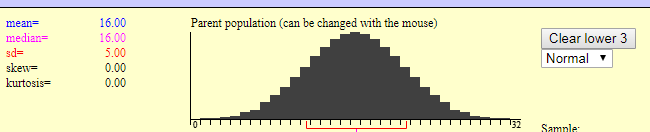
Sampling distribution of the mean is when we take random, independent samples of a constant sample size n. The graph showing the distribution of the values of the mean from all the samples taken is called the sampling distribution of the mean. The sampling distribution of the mean obeys the Central Limit Theorem in that it has a normal distribution (given sample size >= 30, and not overly skewed) and would tend towards the mean (sread becomes narrower) as sample size increases.

As the sample size increases -

1. shape - closer to a normal distribution (normal curve)

2. center - becomes taller (increase frequency of values that is close to the true population mean)

3. spread - becomes narrower

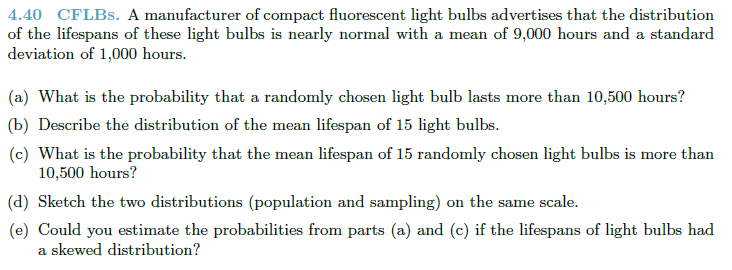


|  |  |  |  |
| --- | --- | --- | --- |
| Mean | SD | N | 10000 repetition. |
| 16.1 | 3.57 | 2 |  |
| 1.58 | 1.58 | 10 |  |
|  |  |  |  |

Here we can clearly say that if SD Decrease the curve graph, when sample size is more.

The plot would thinner compared to datapoints where SD is more.

<http://onlinestatbook.com/stat_sim/sampling_dist/index.html>



**(a) What is the probability that a randomly chosen light bulb lasts more than 10,500 hours?**

Answer :

z = (10,500 - 9,000)/1,000 = 1.5

1 - pnorm(1.5) = 0.07

**(b) Describe the distribution of the mean lifespan of 15 light bulbs.**

Answer :

Since the population distribution is nearly normal N(9000, 1000) is nearly normal, the SE for the sample is 1000/SQRT(15) = 258.

**(c) What is the probability that the mean lifespan of 15 randomly chosen light bulbs is more than 10,500 hours?**

Answer :

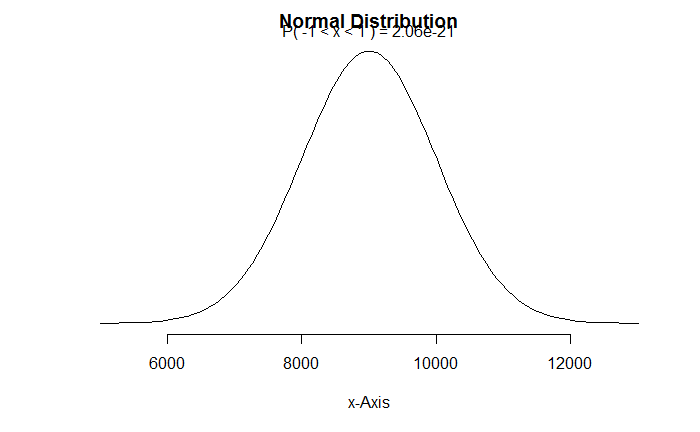
z= (10,500 - 9,000)/258 = 5.81

1 - pnorm(5.81) = 0 (no chance)

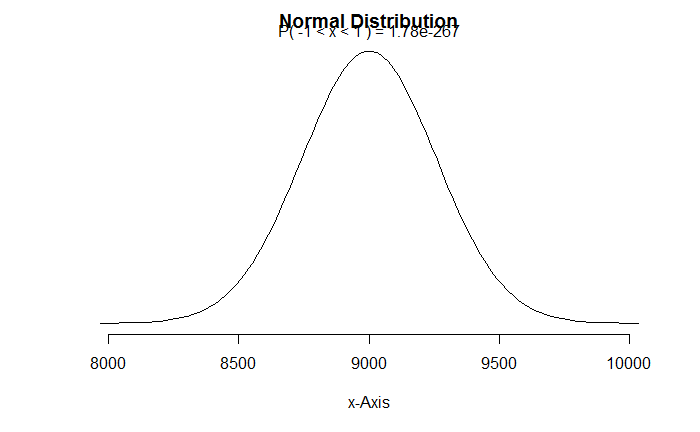
**(d) Sketch the two distributions (population and sampling) on the same scale.**

Answer :

normalPlot(mean = 9000, sd = 1000)



normalPlot(mean = 9000, sd = 258)



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

Answer :

No. The z values were based on the assumption of a normal distribution.

Inference for other estimators

#### 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).