606 HW4 (Part 2)

Ahsanul Choudhury

October 10, 2016

Load Library

```r
library(IS606)
```

```
##
## Welcome to CUNY IS606 Statistics and Probability for Data Analytics
## This package is designed to support this course. The text book used
## is OpenIntro Statistics, 3rd Edition. You can read this by typing
## vignette('os3') or visit www.OpenIntro.org.
##
## The getLabs() function will return a list of the labs available.
##
## The demo(package='IS606') will list the demos that are available.
```

```
##
## Attaching package: 'IS606'
```

```
## The following object is masked from 'package:utils':
##
##     demo
```

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.

(a) Are conditions for inference satisfied?

-Yes, the sample of randomly selected although from a very specific geographic location and is over minimum trashold of 30, the disribution is bimodal and not strongly skwed.

(b) 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.

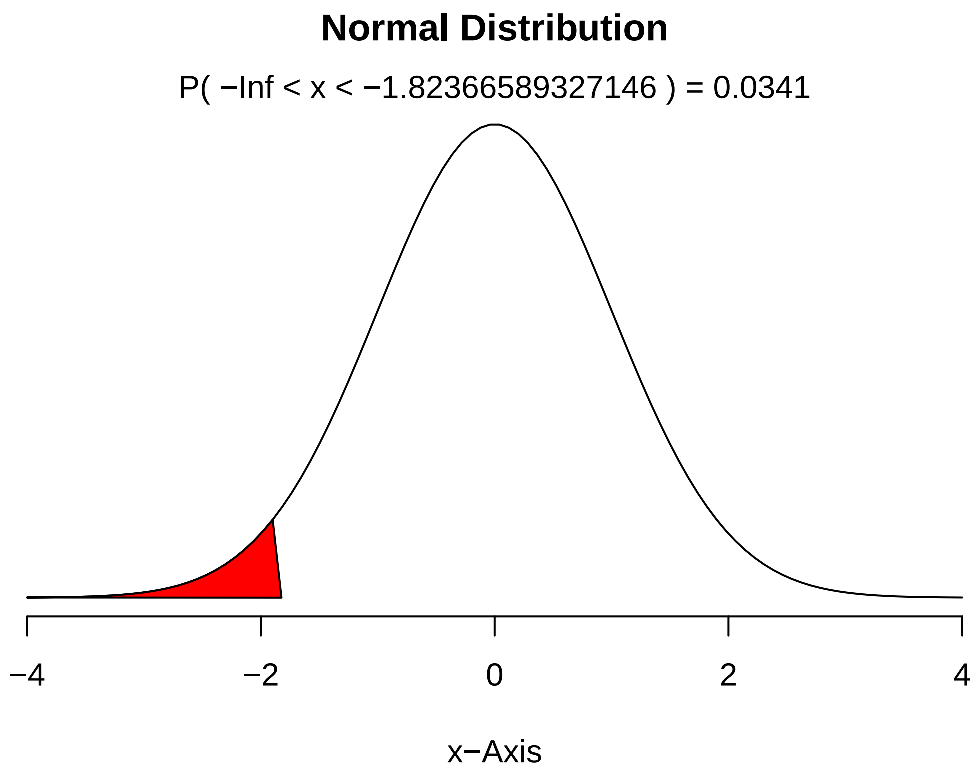
Test

```
a <- 0.10
xbar <- 30.69
sd <- 4.31
n <- 36
SE <- sd / sqrt(n)
zbar <- (xbar - 32) / SE
zbar
```

```
## [1] -1.823666
```

One sided hypothesis test

```
normalPlot(bounds = c(-Inf, zbar))
```



P-value

```
pval <- pnorm(zbar)
pval
```

```
## [1] 0.0341013
```

(c) Interpret the p-value in context of the hypothesis test and the data.

In our hypothesis test and the data the p-vlue of 0.0341013 is musch lower than 0.1, which tells us that the gifted children counting to 10 mean 30.69 months is not close to 32 months average and we reject the null hypothesis in favor of alternative.

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

90% confidence interval

```
zbar90 <- 1.645
intvl <- c(xbar - zbar90*SE, xbar + zbar90*SE)
intvl
```

```
## [1] 29.50834 31.87166
```

(e) Do your results from the hypothesis test and the confidence interval agree? Explain.

-Yes they agree, we wanted to know if gifted childern learn to count before 32 months and the range falls under that age.

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.

(a) 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.

```
xbar <- 118.2
n <- 36
sd <- 6.5
SE <- sd / sqrt(n)
zbar <- (xbar - 100) / SE
pval <- 2 * (1-(pnorm(zbar)))
pval
```

```
## [1] 0
```

The p-vlue of is much lower than the significance level of 0.1, therfore we would reject the null hypothesis in favor of the alternative. In other words the IQ of gifted children's mother are higher than the general population.

(b) Calculate a 90% confidence interval for the average IQ of mothers of gifted children.

-90% confidence interval is 116.4179167, 119.9820833

(c) Do your results from the hypothesis test and the confidence interval agree? Explain

They do, the p-vlue is very small which makes us reject the null hypothesis and the confidence intervals are also way off the mean which is make us reject the null hypothesis at 90% confidence.

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.

-"sampling distribution" of the mean is the distribution of the mean taken for same size sample from a given population. As the sample size increases the the sample becomes more normally distributed, the mean get closer to the center and the spread becomes narrower.

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

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

The probability that a randomly chosen light bulb lasts more than 10,500 hours is 0.0668072

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

The standard error of 258.1988897 indicates the distrbution will be fairly normal.

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

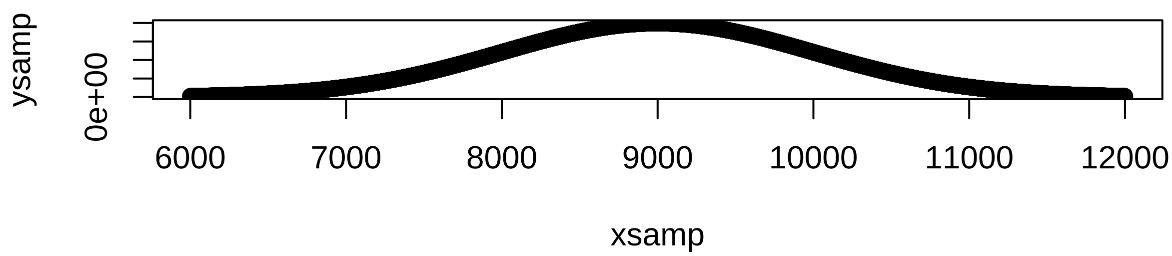
```
1-pnorm(10500,mean=9000,sd=258.1989)
```

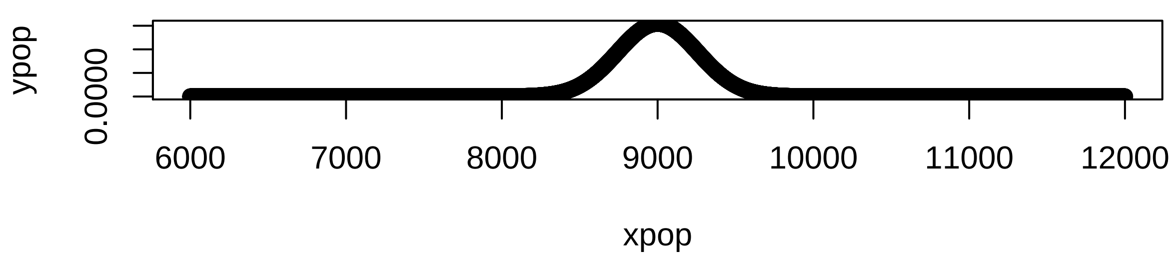
```
## [1] 3.133456e-09
```

The probability that the mean lifespan of 15 randomly chosen light bulbs is more than 10,500 hours is 3.133456e-09 which is very small.

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

```
par(mfrow = c(2, 1))
xsamp <- 6000:12000
xpop <- 6000:12000
ysamp <- dnorm(xsamp, mean = 9000, sd = 1000)
ypop <- dnorm(xpop, mean = 9000, sd = 258.1989)
plot(xsamp, ysamp)
plot(xpop, ypop)
```





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

-Z value assumes the distribution is normal

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.

-Standard error gets smaller as sample size gets bigger. Standard error is used to calculate z score and p-value. As we discover our mistake with sample size and increase it the z score will increase and p-value will decrease which can alter our decision on null hypothesis.