Lab 5

Kristina Arevalo

Fri Oct 02 2020

Contents

Problem 1

Trust but verify. We trust that the rnorm() will generate random deviates in accordance with the definition of the normal distribution. For example, we learned in this lab, that a normal distribution with mean = 0, and sd =1 , should only produce values larger than 2.5 with a specific small probability, that is P(x>2.5) = 0.006209665. Verify this is approximately the case by randomly sampling 1 million numbers from this distribution, and calculate what proportion of numbers are larger than 2.5. (1 point)

```
sample <- rnorm(1000000, 0 ,1)

larger_than_2.5 <- sample[sample > 2.5]

length(larger_than_2.5)/1000000
```

```
## [1] 0.006196
```

confidence = 100

Problem 2

If performance on a standardized test was known to follow a normal distribution with mean 100 and standard deviation 10, and 10,000 people took the test, how many people would be expected to achieve a score higher than 3 standard deviations from the mean? (1 point)

```
10000*pnorm(130, 100 , 10, lower.tail = FALSE)
```

```
## [1] 13.49898
```

confidence = 100

Problem 3

You randomly sample 25 numbers from a normal distribution with mean = 10 and standard deviation = 20. You obtain a sample mean of 12. You want to know the probability that you could have received a sample mean of 12 or larger.

Create a sampling distribution of the mean for this scenario with at least 10,000 sample means (1 point). Then, calculate the proportion of sample means that are 12 or larger (1 point).

```
sample_dist <- replicate(10000,mean(rnorm(25,10,20)))

length(sample_dist[sample_dist >=12])/10000
```

```
## [1] 0.2964
```

confidence = 100

Problem 4

You randomly sample 100 numbers from a normal distribution with mean = 10 and standard deviation = 20. You obtain a sample mean of 12. You want to know the probability that you could have received a sample mean of 12 or larger. Create a sampling distribution of the mean for this scenario with at least 10,000 sample means. Then, calculate the proportion of sample means that are 12 or larger. Is the proportion different from question 3, why? (1 point).

```
sample_dist2 <- replicate(10000,mean(rnorm(100,10,20)))
length(sample_dist2[sample_dist2 >=12])/10000
```

```
## [1] 0.159
```

The proportion is different from question 3 because you have a larger n so you will have a lower proportion of having a mean that is farther away from 10. Having a larger sample size increases the power and will more accurately reflect the population mean.

confidence = 100 # Problem 5 You randomly sample 25 numbers from a normal distribution with mean = 10 and standard deviation = 20. You obtain a sample standard deviation of 15. You want to know the probability that you could have received a sample standard deviation of 15 or less. Create a sampling distribution of standard deviations for this scenario with at least 10,000 sample standard deviations. Then, calculate the proportion of sample standard deviations that are 15 or less. (1 point)

```
sample_dist3 <- replicate(10000,sd(rnorm(25,10,20)))
length(sample_dist3[sample_dist3 <=15])/10000
```

```
## [1] 0.0403
```

confidence = 100