HW 9

Chapter 9,Page 339, Question 11

Priya Shaji

10/20/2019

Write a computer program to simulate 10,000 Bernoulli trials with probability .3 for success on each trial. Have the program compute the 95 percent confidence interval for the probability of success based on the proportion of successes. Repeat the experiment 100 times and see how many times the true value of .3 is included within the confidence limits.

Answer

Given:

1. Simulate 10,000 Bernoulli trials:

```
central_limit <- runif(10000)
# Display first few values of the trials
head(central_limit)
```

```
## [1] 0.88369227 0.55247345 0.77807274 0.27479928 0.92067348 0.04389163
```

2. Let's calculate the interval conditions according to the given requirements:

```
# Minimum limit

min(central_limit)
```

```
## [1] 9.920984e-05
```

```
# Maximum Limit

max(central_limit)
```

```
## [1] 0.9996779
```

```
# Length of simulations

length(central_limit)
```

```
## [1] 10000
```

3. Computer program to simulate 10,000 Bernoulli trials with probability 0.3

```
success = 0 # to calculate success on each trial
for(trial in central_limit){
  if(trial <= 0.3){ # simulate 10,000 Bernoulli trials with probability 0.3
    success = success + 1
  }
}
success/length(central_limit)
```

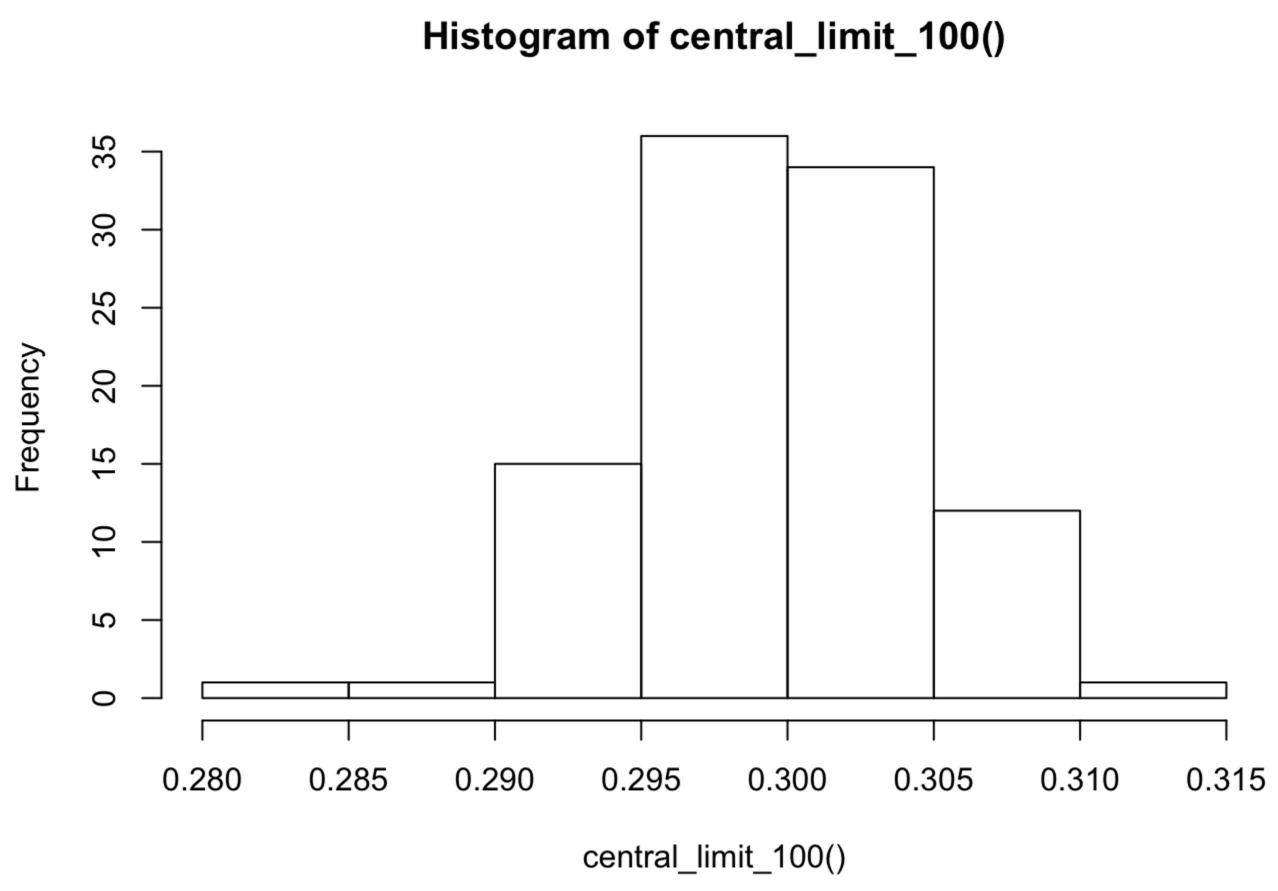
```
## [1] 0.2966
```

4. Repeat the experiment 100 times and see how many times the true value of 0.3 is included within the confidence limits.

```
central_limit_100 <- function(){
  vec = vector()
  for(i in 1:100){
    central_limit <- runif(10000)
    success = 0
    for(trial in central_limit){
      if(trial <= 0.3){
      success = success + 1
      }
    }
    vec[i] = success/length(central_limit)
  }
return(vec)
}
```

5. Let's plot histogram to analyze CLT demonstrated:

```
hist(central_limit_100())
```



The CLT demonstrated here: Data was taken from a uniform distribution, and also, proportion of successes are normally distributed about the expected value of 0.3.

We will now use the central_limit_100 function to test against the 95% confidence intervals $p \pm 1.96 * \sqrt{pq/n}$:

where: p = 0.3, q = 0.7, n = 10000

```
p = 0.3
q = 0.7
n = 10000
trials <- central_limit_100()
# calculate lower bound
lower_bound = p - 1.96*(p*q/n)^0.5
lower_bound
```

```
## [1] 0.2910182
```

```
# calculate upper bound
upper_bound = p + 1.96*(p*q/n)^0.5
upper_bound
```

```
## [1] 0.3089818
```

6. As a final step let's see how many times the true value of 0.3 is included within the confidence limits.

```
ans <- trials <= upper_bound & trials >= lower_bound
table(ans)
```

```
## ans
## FALSE  TRUE
##     4    96
```

Therefore out of 100, 'TRUE Value' is expected to be within the 95% confidence interval.