Lab Practical 7: Inverse transform method

Quang Trinh

04.18.18

## Task 1: Simulating from a logistic distribution

Use the inverse transform method to construct a simulator for the distribution with cdf

where and . True mean and variance are and respectively.

### The problem

Formulate the inverse transform method algorithm to simulate from a logistic distribution. Recall you will need to find the inverse of the cumulative distribution function. As a reminder, the general inverse transform method algorithm is shown at the top of the file “lab6\_ivt.pdf” on the course Blackboard site under the folder for this lab. Simulate 1000 variates from the logistic distribution using your algorithm.

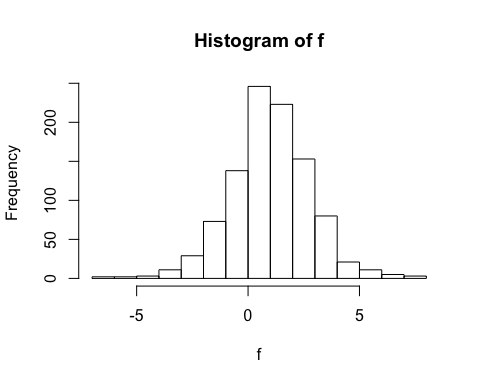
# Generate U(0,1)  
x = runif(1000,0,1)  
# Inverse transformation  
f = 1-log((1/x) - 1) # Inverse func: a - b\*log(1/x - 1)  
  
mean(f)

## [1] 0.9981201

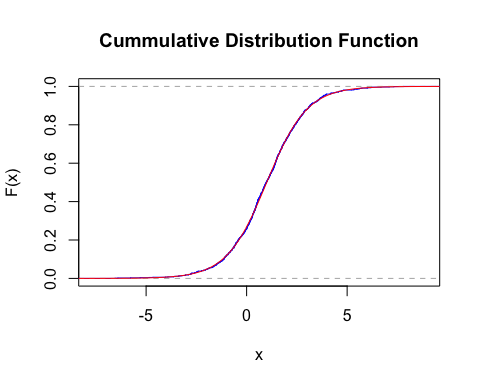
var(f)

## [1] 3.301781

hist(f)



# Plotting true CDF and empirical CDF  
plot(ecdf(f), col = 'blue', xlab='x', ylab='F(x)', main='Cummulative Distribution Function')  
trueX = seq(-10,10,0.05)  
trueCDF = 1 / (1 + exp(- (trueX -1)))  
lines(trueX,trueCDF,col = 'red')



### Report the following

* Present the inverse transform method formula for

*The inverse function is a - b \* log((1/x) - 1)*

* Present the mean and variance of the 1000 variates you simulate. Compare these empirical values to the truth. *The empirical mean is 0.96 and the empirical variance is 3.04. The true mean is 1 and the true variance is pi2\*b2/3 = 3.29. Thus, empirical values are very close to truth values.*
* Present a histogram of your 1000 variates.
* Plot the empirical cumulative distribution function of your 1000 variates. Overlay the true cdf on the plot and compare. In R, ecdf function will plot the empirical cdf. For the true cdf, compute the cdf over the sequence seq(-10, 10, 0.05). Recall the lines function to draw the true cdf on the empirical cdf plot. Make sure to present the true cdf in a difference color than the empirical cdf using the col option in the lines function. *Label the axes on the plot.*

*The true cdf line is colored in red, and the empirical cdf line is colored in blue. As we can see, the 2 lines matches up each other very closely. As the result, we can conclude that our experiment gives out results as expected. That is, if we have to run infinite number of x values, the empirical cdf line should be the same with the true cdf line.*

## Task 2: Simulating circles

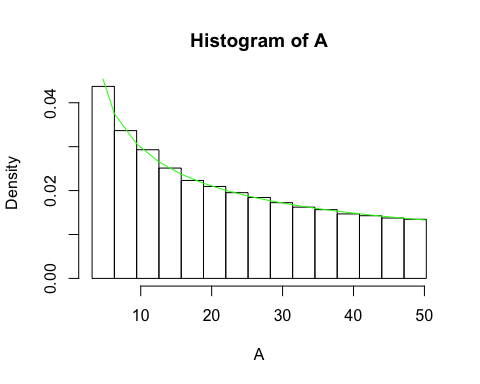
Dobrow 6.60: Let . Let denote the area of the circle of radius . We will simulate and then in turn obtain variates for the area .

### The problem

Simulate the area of the circle in two ways.

1. Simulate 100,000 variates for and then compute the area of the circle .
2. Use the inverse transform method to simulate directly from . The details are in the latter part of the pdf file “lab6\_IVT.pdf” on the course Blackboard site under the folder for this lab practical.

# Generate U(0,1)  
r = runif(100000,0,1)  
# Inverse transformation for   
A = pi\*(9\*r^2+6\*r+1)  
# True Values  
trueR = seq(pi,16\*pi,by=pi)  
truePDF = 1/(6\*sqrt(pi\*trueR))  
hist(A,breaks=trueR, freq=FALSE)  
lines(trueR,truePDF,col='green')



# Empirical values  
mean(A)

## [1] 21.93017

var(A)

## [1] 188.686

### Report the following

* Present the mean and standard deviation of the area . Compare to the true values derived in the “lab6\_IVT.pdf” file on the course Blackboard site. *The empirical mean is 22.04, and the empirical variance is 190.10. The true values of mean and variance are 21.99 and 189.496 respectively. As we can see, the empirical values are close (since 100000 variates) to the truth values.*
* Present a histogram of the variates from the inverse transform method simulation (problem 2 above). Use the sequence seq(pi, 16\*pi, by=pi) as the breaks in the histogram. Remember to set freq = FALSE as a histogram option for a density plot.
* Overlay the true pdf of on the histogram. The true pdf is found in the “lab6\_IVT.pdf” file on the course Blackboard site. Use the same sequence seq(pi, 16\*pi, by=pi) over which to compute the pdf. The lines function will overlay this true pdf on the histogram plot.

*As we can see from the histogram, the true pdf (green line) is nicely lay out on top of each column of the histogram. Hence, the histogram (which reflexes the empirical pdf) demonstrates the expected pdf (the truth).*