Statistics 572
Homework 3

Alejandro Robles

September 25th, 2018

1 Acceptance/Rejection Method

Using Acceptance/Rejection method, generate a random variable, X, having size of 100. The true density of x follows the distribution described below.

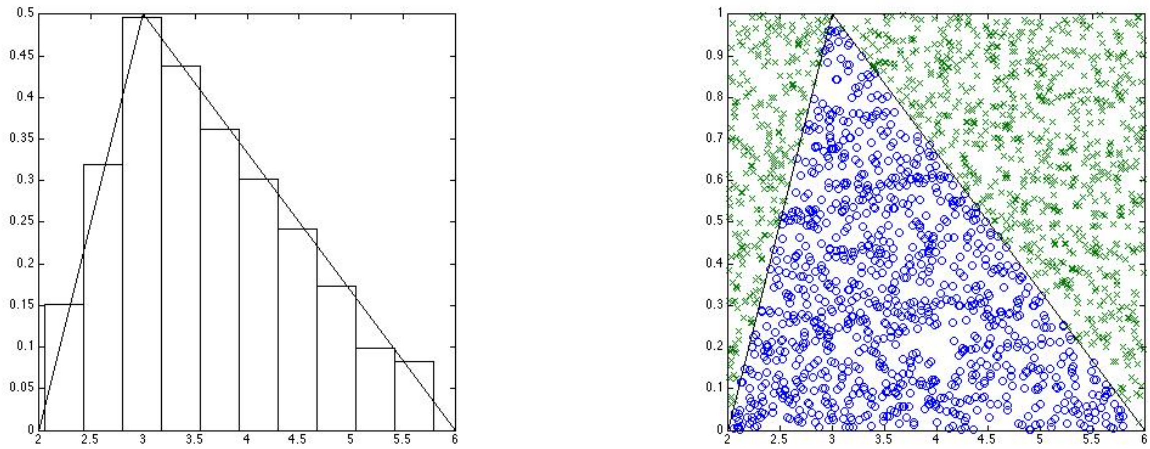
$$f(x) = \begin{cases} \frac{x-2}{2} & 2 \leq x \leq 3 \\ \frac{100-x}{100} & 3 < x \leq 6 \end{cases}$$

MATLAB Code:

```
cla
clf('reset');
c = 1;   % constant
n=1000;  % generate 100 rv's
% set up the arrays to store variates
x = zeros(1,n);   % random variates
xy = zeros(1,n); % corresponding y values
rej = zeros(1,n); % rejected variates
rejy = zeros(1,n); % corresponding y values
irv=1;
irej=1;
while irv <= n
   y = unifrnd(2,6);  % random number from g(y)
   u = rand(1);  % random number for comparison
   if  y <= 3 && u <= (y-2)/c;
      x(irv)=y;
      xy(irv) = u*c;
      irv=irv+1;
   elseif y > 3 && u <= (2 - y/3)/c;
      x(irv)=y;
      xy(irv) = u*c;
      irv=irv+1;
   else
            rej(irej)= y;
      rejy(irej) = (u*c); % really comparing u*c<=2*y
      irej = irej + 1;
   end
end
acc = irv/(irv+irej)
```

```
% Plotting
figure(1)
plot(x, xy, 'o', rej, rejy,'x');
axis square
hold on
plot([2,3], [0,1], 'k');
plot([3,6], [1,0], 'k');
hold off
figure(2)
[fr, x] = hist(x);
h=x(2)-x(1);
bar(x,fr/(n*h),1,'W')
axis square
hold on
plot([2,3], [0,0.5], 'k');
plot([3,6], [0.5,0], 'k');
hold off
```

Result:



(a) Histogram

(b) Acceptance/Rejection

Figure 1: Acceptance/Rejection

Discussion: Using Y $\sim$ Uniform(2,6), we generate our random sample. In Figure 1(a), we see that the empirical histogram nicely fits the theoretical probability distribution. In Figure 1(b), we see how this visually represents our accuracy of 0.4895.

2 Exercise 4.2

Write the MATLAB code to implement example 4.5. Generate 500 random variables from this distribution and construct a histogram (hist function) to verify your code.

MATLAB Code:

```
prob_mass_func=[.15,.22,.33,.10,.20];
n=500; acc=zeros(1,n); rej=zeros(1,n);
c=1.65; irej=1; iacc=1;
while iacc<=n
    y=unidrnd(5);
    u=rand(1);
    if u <= prob_mass_func(y)/(c*(1/5));
        acc(iacc)=y;
        iacc=1+iacc;
    else
        rej(irej)=y;
        irej=1+irej;
    end
end
accept=hist(acc,5);
bar(accept);
title('Histogram')
xlabel('X');
ylabel('Frequency');
disp(accept/500);
```

Result:

```
>> exercise_4_2
    0.1520    0.2580    0.2680    0.1140    0.2080
```

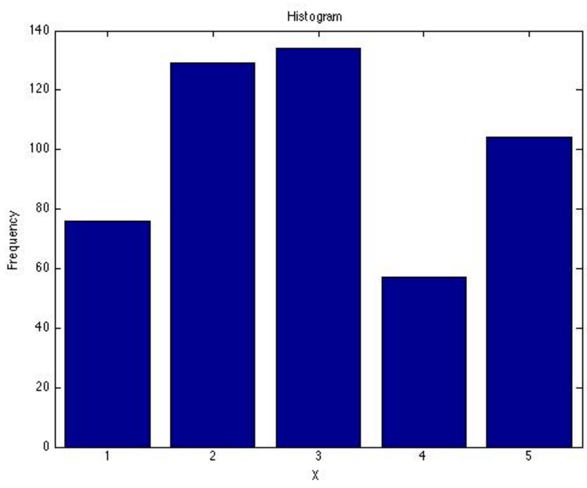


Figure 2: Histogram

Discussion: As we can see from the histogram (Figure 2) and from the proportion of occurrences for each element in the support, that we closely match the theoretical probability distribution function.

3 Exercise 4.4 (Inverse Transformation Method)

Write a MATLAB function that will return random numbers that are uniformly distributed over the interval (a, b).

MATLAB Code:

```
function uniform_rv = uniform_inverse(a, b, n)
    u = rand(1,n);
    uniform_rv = (b - a)*u + a;

% Calling Function
a = 5; b = 10; n = 1000;
uniform_rvs = uniform_inverse(a,b,n);
hist(uniform_rvs);
```

Result:

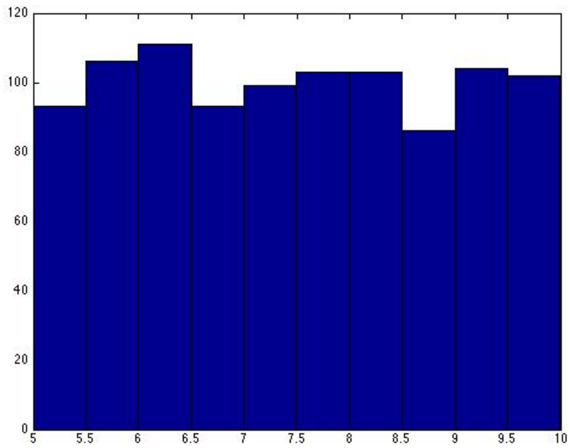


Figure 3: Histogram for Uniform(5,10), Inverse Transformation Method

Discussion: Uniform distribution on the interval [a, b] follows f(x) = 1/(b - a) for $a < x < b$

The CDF is F(t) = (t - a)/(b - a) for $a < t < b$.

Then $F^{-1}(u) = (b - a)u + a$. So if U $\sim$ U[0, 1], then $(b - a)U + a \sim U[a, b]$.

Using the Inverse Transform Method, we see from Figure 3 that this method approximately generates random samples from Uniform(5, 10).

4 Exercise 4.4 (Alternative Method)

Write a MATLAB function that will return random numbers that are uniformly distributed over the interval (a, b).

MATLAB Code:

```
function norm_rv = unif_alt(a, b, n)
    % Parameters
    x = a:0.01:b;
    norm_rv = zeros(1,n);
    % Approximate CDF values with user defined function
    F = zeros(1,length(x));
    for i = 1:length(x)
        F(i) = (x(i)-a)/(b-a);
    end
    % Alternative Approach using Uniform Distribution
    for i = 1:n
        u = rand;
        index = min(find(F >= u, 1 ));
        norm_rv(i) = x(index);
    end
    % Checking fit, with bin size = 30
    hist(norm_rv)

% Calling Function
unif_alt(5,10, 1000)
```

Result:

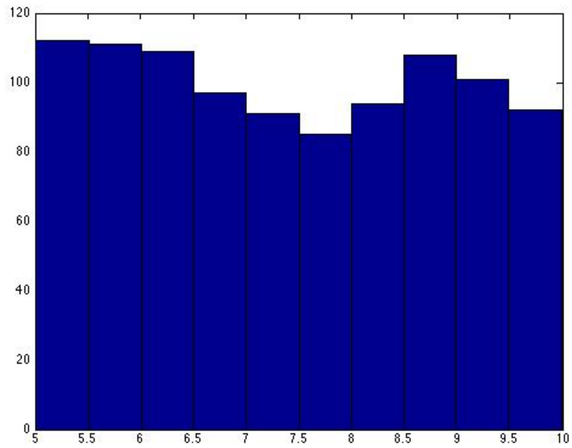


Figure 4: Histogram for Uniform(5,10), Alternative Method

Discussion: Comparing Figure 3 with Figure 4, we see similar histograms. It's hard to tell which one is more accurate but the Inverse Method looks slightly more accurate. Regardless both methods generate reasonable random samples from Uniform(5,10).

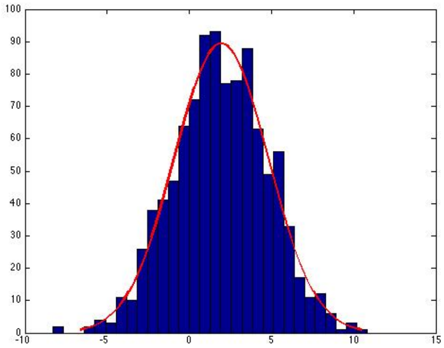


Figure 5: Histogram of sampled random numbers following N(2,3)

5 Exercise 4.5

Write a MATLAB function that will return random numbers from the normal distribution with mean μ and variance σ^2 . The user should be able to set values for the mean and variance as input arguments.

MATLAB Code:

```
function norm_rv = norm_alt(mu, var, n)
    % Parameters
    sigma = sqrt(var);
    end_point = 5*sigma;
    start_point = -end_point;
    x = start_point:0.01:end_point;
    norm_rv = zeros(1,n);
    % Approximate CDF values with user defined function
    F = zeros(1,length(x));
    for i = 1:length(x)
        F(i) = integral(@(b) normpdf(b,mu, sigma) , -inf, x(i));
    end
    % Alternative Approach using Uniform Distribution
    for i = 1:n
        u = rand;
        index = min(find(F >= u, 1 ));
        norm_rv(i) = x(index);
    end
    % Checking fit, with bin size = 30
    histfit(norm_rv, 30, 'normal')
% Calling function in another script
mu = 2; variance = 9; n = 1000;
normal_rvs = norm_alt(mu, variance, n);
fprintf('the mean of the sample is %0.5f\n',  mean(normal_rvs));
fprintf('the variance of the sample is %0.5f\n', var(normal_rvs));
```

Result:

```
the mean of the sample is 1.93115
the variance of the sample is 8.10390
```

Discussion: As we can see from Figure 5, the sampled random numbers do approximately follow a Normal distribution. Calculating the sample mean and variance, we also see that it approximately has the parameters we requested for.

6 Exercise 4.12

Generate 1000 binomial random variables for n = 5 and p = 0.3, 0.5, 0.8. In each case, determine the observed relative frequencies and the corresponding theoretical probabilities. How is the agreement between them? Note: Use Inverse CDF Method.

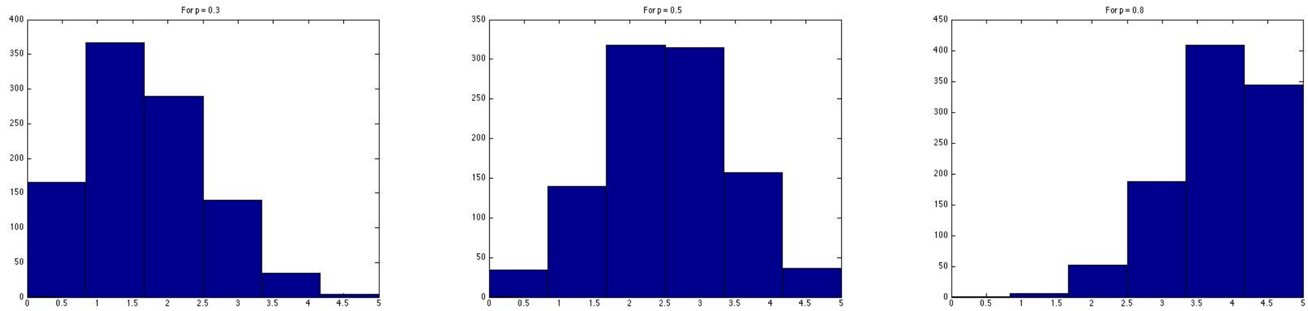
MATLAB Code:

```
size = 1000;
coefficients = [1, 5, 10 ,10, 5, 1];
n = 5;
k = [0.3, 0.5, 0.8];

for p = k;
    fprintf('For p = %0.1f, \n', p);
    cdf = zeros(1,n+1);
    theoretical_probs = zeros(1,n+1);
    binomial_rvs = zeros(1,size);
    cdf_value = 0;

    for i = 1:n+1;
        theoretical_probs(i) = coefficients(i)*(p^(i-1))*((1-p)^(5 - i + 1));
        cdf_value = cdf_value + theoretical_probs(i);
        cdf(i) = cdf_value;
    end

    % Disclaimer:
    % Because the index of the cdf array corresponds to the support of a Binomial(n=5) minus 1
    % Then we don't need to do if-else statements
    % Instead we just find the minimum index where cdf is larger than u and subtract by 1.
    for i = 1:size;
        u = rand;
        binomial_rvs(i) = find(cdf >= u,1) - 1;
    end;

    binomial_rvs = sort(binomial_rvs);
    [occurences,xi]=hist(binomial_rvs,unique(binomial_rvs));
    frequency = occurences/size;
    for i = xi;
        fprintf('Relative Frequency      : P(X = %d) = %0.4f\n', i, frequency(i+1));
        fprintf('Theoretical Probability : P(X = %d) = %0.4f\n', i, theoretical_probs(i+1));
        disp('--------------------------------------------------------')
    end
    figure;
    hist(binomial_rvs, 6)
    title(sprintf('For p = %0.1f', p))

end
```

Result:

```
>> exercise_4_12
For p = 0.3,
Relative Frequency      : P(X = 0) = 0.1650
Theoretical Probability : P(X = 0) = 0.1681
--------------------------------------------------------
Relative Frequency      : P(X = 1) = 0.3670
Theoretical Probability : P(X = 1) = 0.3601
--------------------------------------------------------
Relative Frequency      : P(X = 2) = 0.2890
Theoretical Probability : P(X = 2) = 0.3087
--------------------------------------------------------
Relative Frequency      : P(X = 3) = 0.1400
Theoretical Probability : P(X = 3) = 0.1323
--------------------------------------------------------
Relative Frequency      : P(X = 4) = 0.0350
Theoretical Probability : P(X = 4) = 0.0283
--------------------------------------------------------
Relative Frequency      : P(X = 5) = 0.0040
Theoretical Probability : P(X = 5) = 0.0024
--------------------------------------------------------
For p = 0.5,
Relative Frequency      : P(X = 0) = 0.0340
Theoretical Probability : P(X = 0) = 0.0312
--------------------------------------------------------
Relative Frequency      : P(X = 1) = 0.1400
Theoretical Probability : P(X = 1) = 0.1562
--------------------------------------------------------
Relative Frequency      : P(X = 2) = 0.3180
Theoretical Probability : P(X = 2) = 0.3125
--------------------------------------------------------
Relative Frequency      : P(X = 3) = 0.3150
Theoretical Probability : P(X = 3) = 0.3125
--------------------------------------------------------
Relative Frequency      : P(X = 4) = 0.1570
Theoretical Probability : P(X = 4) = 0.1562
--------------------------------------------------------
Relative Frequency      : P(X = 5) = 0.0360
Theoretical Probability : P(X = 5) = 0.0312
--------------------------------------------------------
For p = 0.8,
Relative Frequency      : P(X = 0) = 0.0010
Theoretical Probability : P(X = 0) = 0.0003
--------------------------------------------------------
Relative Frequency      : P(X = 1) = 0.0060
Theoretical Probability : P(X = 1) = 0.0064
--------------------------------------------------------
Relative Frequency      : P(X = 2) = 0.0520
Theoretical Probability : P(X = 2) = 0.0512
--------------------------------------------------------
Relative Frequency      : P(X = 3) = 0.1880
Theoretical Probability : P(X = 3) = 0.2048
--------------------------------------------------------
Relative Frequency      : P(X = 4) = 0.4090
Theoretical Probability : P(X = 4) = 0.4096
```

```
--------------------------------------------------------
Relative Frequency       : P(X = 5) = 0.3440
Theoretical Probability : P(X = 5) = 0.3277
--------------------------------------------------------
```



(a) P = 0.3 (b) P = 0.5 (c) P = 0.8

Figure 6: Random Sample from Binomial with n = 5

Discussion:

Comparing the relative frequency versus the theoretical probability, we see that for the most part they are approximately equal. Figure 6 also visually reassures us that we generated random samples from the distributions we requested. It makes sense that when the probability of success is lower that we see x = 5 less likely and vice-versa. It also makes sense that if the probability of success and failure are equal then we should expect a symmetrical distribution.