Random Binary Expansions

May 1, 2018

Approximating F

Question 1

The program is listed on page 6. I chose $N = 100\,000$ so that the cusps at the dyadic rationals would be easily visible.

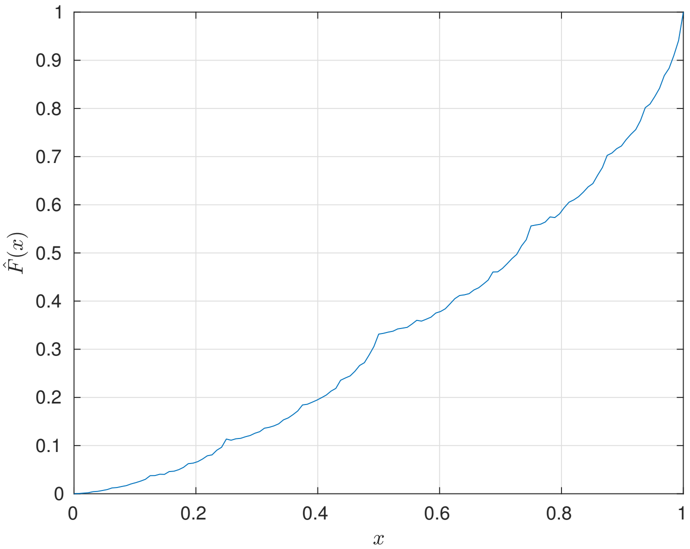

Figure 1: Graph of $\hat{F}(x)$ against x for $p = 2/3$ and $n = 30$.

Calculating F

Question 2

$F(x) = \mathbb{P}(X \leq x) = \sum_{k=1}^{n} x_k p^{|i<k|x_i=1|} q^{1+|i<k|x_i=0|}$. We may derive this by replacing the kth 1 in the binary expansion with 0, and then finding the probability that x begins with the digits up to this 0, and then summing over k from 1 up to n. The formula may alternatively be thought of as a polynomial in p where the degree of the polynomial is the number of 1's in the binary expansion minus 1 and the coefficient of p^i is q^{k-i} where k is the position of the ith 1 in the binary expansion.

Question 3

The program to produce this graph is listed on page 7.

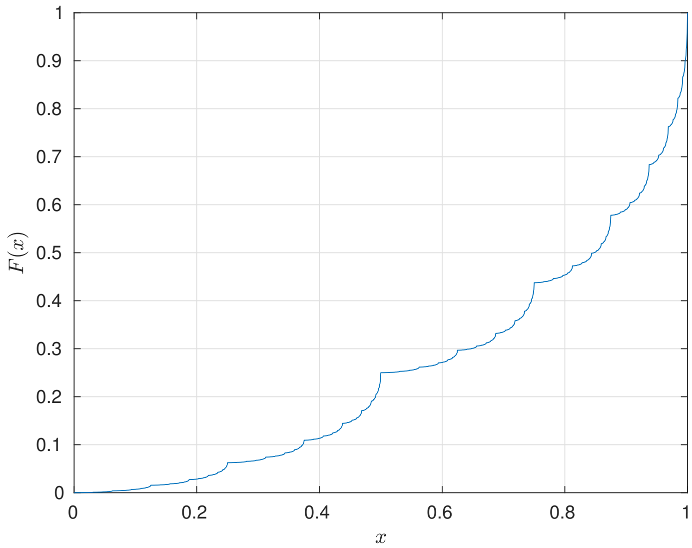

Figure 2: Graph of $F(x)$ against x for $p = 3/4$ and $n = 11$.

Compared to the graph in Question 1, the slope appears lower on the left and steeper on the right. This is because x is more likely to be greater as the probability of each digit being a 1 is greater. Moreover since the graph in Question 1 was empirical, it has more noise and the cusps are not as well defined. The complexity of the algorithm (to work out the value of F) in Question 1 is $O(Nn)$ and in Question 2 is $O(n)$. If we wanted to determine the complexity of producing the graph we would just need to multiply this by the number of samples chosen.

Properties of F

Question 4

The plots suggest that F is continuous everywhere, and this is indeed the case. Given a sequence $x_1, x_2, \ldots$ we have $\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots) = 0$ as this would be an infinite product of p's and q's where $0 < p < 1$ and $0 < q < 1$, which would tend to 0. Since x has only two finite binary expansions, it follows that $\mathbb{P}(X = x) = 0$, and therefore F is continuous everywhere since

$$\lim_{y \uparrow x}(F(x) - F(y)) = \lim_{y \uparrow x} \mathbb{P}(y < X \leq x) = \mathbb{P}(X = x) = 0 \implies F(x) = \lim_{y \uparrow x} F(y)$$

and

$$\lim_{y \downarrow x}(F(y) - F(x)) = \lim_{y \downarrow x} \mathbb{P}[x < X \leq y] = \mathbb{P}(X \in \varnothing) = 0 \implies F(x) = \lim_{y \downarrow x} F(y)$$

Question 5

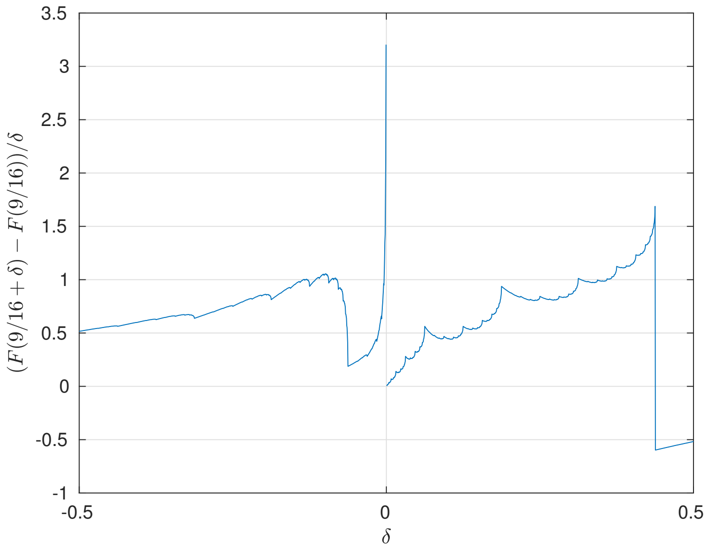

Figure 3: Graph of $(F(c+\delta)-F(c))/\delta$ against δ for $p = 3/4$, $n = 11$ and $c = 9/16$.

The plot suggests F is not left-differentiable at c since the derivative approaches ∞ but is right-differentiable at c with a right-derivative of 0. I chose δ to be of the form $-1/2 + i/2^n$ for $i = 0, \ldots, 2^n$, so that both δ and $c + \delta$ have finite binary expansions.

Question 6

If $p = 1/2$ then F is differentiable everywhere with derivative 1. Indeed in this case we have $F(x) = x$. If $p > 1/2$ then F is not left-differentiable as the left-derivative tends to infinity but it is right-differentiable with right-derivative 0. If $p < 1/2$ then it is not right-differentiable as the right-derivative tends to infinity but it is left-differentiable with left-derivative 0. This is clearly supported by the plots where we have cusps at the dyadic rationals with a steep slope at one side and a flat slope on the other.

Proof: From the formula in Question 2 we find that for n sufficently large, the right difference quotient $\frac{F(c+\delta)-F(c)}{\delta}$ with $\delta = 1/2^n$ is equal to $\frac{p^k q^{n-k}}{1/2^n} = (p/q)^k(2q)^n$ where k is the number of 1's in the binary expansion of c. This tends to 0 if $2q < 1$ i.e. if $p > 1/2$, equals 1 if $p = q = 1/2$ and tends to infinity if $2q > 1$ i.e. if $p < 1/2$. By carrying and summing the resulting geometric series we find that the analogous formula for the left difference quotient with $\delta = -1/2^n$ is $(p/q)^{l-k+1}(2p)^n$ where k and l are constants, and then a very similar argument to the above applies. For general δ it suffices to note that F

is an increasing function, since then if $1/2^{n+1} < \delta < 1/2^n$ we have

$$\left|\frac{f(c+\delta)-f(c)}{\delta}\right| \leq \frac{f(c+1/2^n)-f(c)}{1/2^{n+1}} = 2\left(\frac{f(c+1/2^n)-f(c)}{1/2^n}\right) \to 0$$

with an analogous argument for $-1/2^n < \delta < -1/2^{n+1}$. $\square$

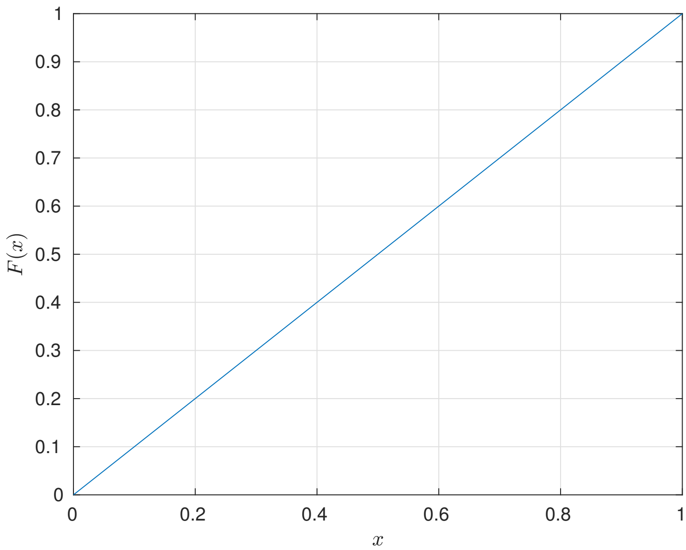

Figure 4: Graph of $F(x)$ against x for $p = 1/2$.

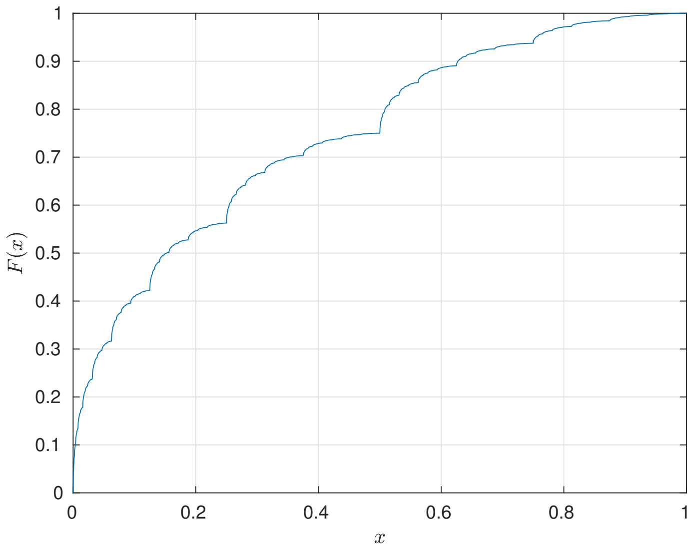

Figure 5: Graph of $F(x)$ against x for $p = 1/4$.

Program `fhat.m` for Question 1.

```
function [output] = fhat(x,p,n,N)
%FHAT A Monte Carlo simulation to approximation the cumulative
%distribution function F.
U = zeros(1,n);
b = zeros(1,n);
X = zeros(1,N);
f = zeros(1,N);
for i=1:n
    b(i)=1/2^i;
end
for i=1:N
    for j=1:n
        U(j)=binornd(1,p);
    end
    X(i)=dot(U,b);
    if X(i) <= x
        f(i) = 1;
    else
        f(i) = 0;
    end
end
output = sum(f)/N;
end
```

Program `plot1.m` to produce Figure 1 for Question 1.

```
m=2^7 + 1;
p=2/3;
n=30;
N=100000;
x = linspace(0,1,m);
y = zeros(1,m);
for i = 1:m
    y(i) = fhat(x(i),p,n,N);
end
plot(x,y)
grid on
ylabel('$$\hat{F}(x)$$','Interpreter','LaTeX')
xlabel('$$x$$','Interpreter','LaTeX')
```

Program `f.m` for Question 3.

```
function [output] = f(x,p,n)
if x==1
    output=1;
    return
end
q=zeros(1,n+1);
r=zeros(1,n);
q(n+1)=x*2^n;
for i=1:n
    q(n-i+1)=floor(q(n-i+2)/2);
    r(n-i+1)=mod(q(n-i+2),2); % determining the x_i's
end
m=zeros(1,n+1);
for i=1:n
    if r(i)==1
        m(i+1)=m(i)+1;
    else
        m(i+1)=m(i);
    end
end
s=0;
for i=1:n
    s=s+r(i)*p^m(i)*(1-p)^(i-m(i));
end
output=s;
end
```

Program `plot2.m` to produce Figure 2 for Question 3.

```
n=11;
p=1/4;
x = zeros(1,2^n+1);
y = zeros(1,2^n+1);
for i = 1:2^n+1
    x(i) = (i-1)/2^n;
    y(i) = f(x(i),p,n);
end
plot(x,y)
grid on
ylabel('$$F(x)$$','Interpreter','LaTeX')
xlabel('$$x$$','Interpreter','LaTeX')
```

Program `plot03.m` to produce Figure 3 for Question 5.

```
c=9/16;
n=11;
p=3/4;
d = zeros(1,2^n+1);
y = zeros(1,2^n+1);
for i = 1:2^n+1
    d(i) = (i-1)/2^n - 1/2;
    y(i) = (f(c+d(i),p,n)-f(c,p,n))/d(i);
end
plot(d,y)
grid on
ylabel('$$(F(c + \delta)-F(c))/\delta$$','Interpreter','LaTeX')
xlabel('$$\delta$$','Interpreter','LaTeX')
```