Tutorial 4

1. Consider throwing n balls randomly into $b < n$ boxes. What is the probability, denoted by $P(k)$, that a given box has exactly $k \leq n$ balls in it? Can you guess at what value of k, $P(k)$ will be maximized?

2. Consider a shop is selling 100 blind boxes (盲盒) of toy figures. Among these 100 boxes, it is stated that 2 of them are special editions. Suppose you want to get one special edition and you plan to buy at most 5 boxes. That is, you will buy one by one. For example, after you buy one, you will immediately open the box to see if it is a special-edition figure. What is the probability that you can get one special edition within 5 purchases? Which distribution can you apply to approximate this probability?

3. A discrete random variable (RV) X has the following cumulative density function (CDF):

$$F(k) = \begin{cases} 0, & k < 1 \\ 0.2, & 1 \leq k < 2 \\ 0.4, & 2 \leq k < 3 \\ 0.6, & 3 \leq k < 4 \\ 0.8, & 4 \leq k < 5 \\ 1, & k \geq 5 \end{cases}$$

Determine the probability mass function (PMF) of X. Sketch the PMF.

4. Given that the PMF $P_X(r)$ of a discrete RV X has the form:

$$P_X(r) = \begin{cases} \alpha p^r, & r = 2, 3, \cdots \\ 0, & \text{otherwise} \end{cases}$$

Find all possible values for α and p.

5. Given that the CDF of a continuous RV X is:

$$F(x) = \begin{cases} 0, & x < 0 \\ x^4, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$$

Determine the probability that X has a value between 0.2 and 0.4.

6. Suppose X is a Poisson RV with PMF:

$$p(r) = P(X = r) = e^{-\lambda}\frac{\lambda^r}{r!}, \quad r = 0, 1, 2, \cdots$$

Find the PMF of $Y = 2X$.

7. Describe how you can utilize uniform RVs to generate Bernoulli RVs with $p = 0.5$. Consider using the MATLAB command `rand`.

Solution

1.

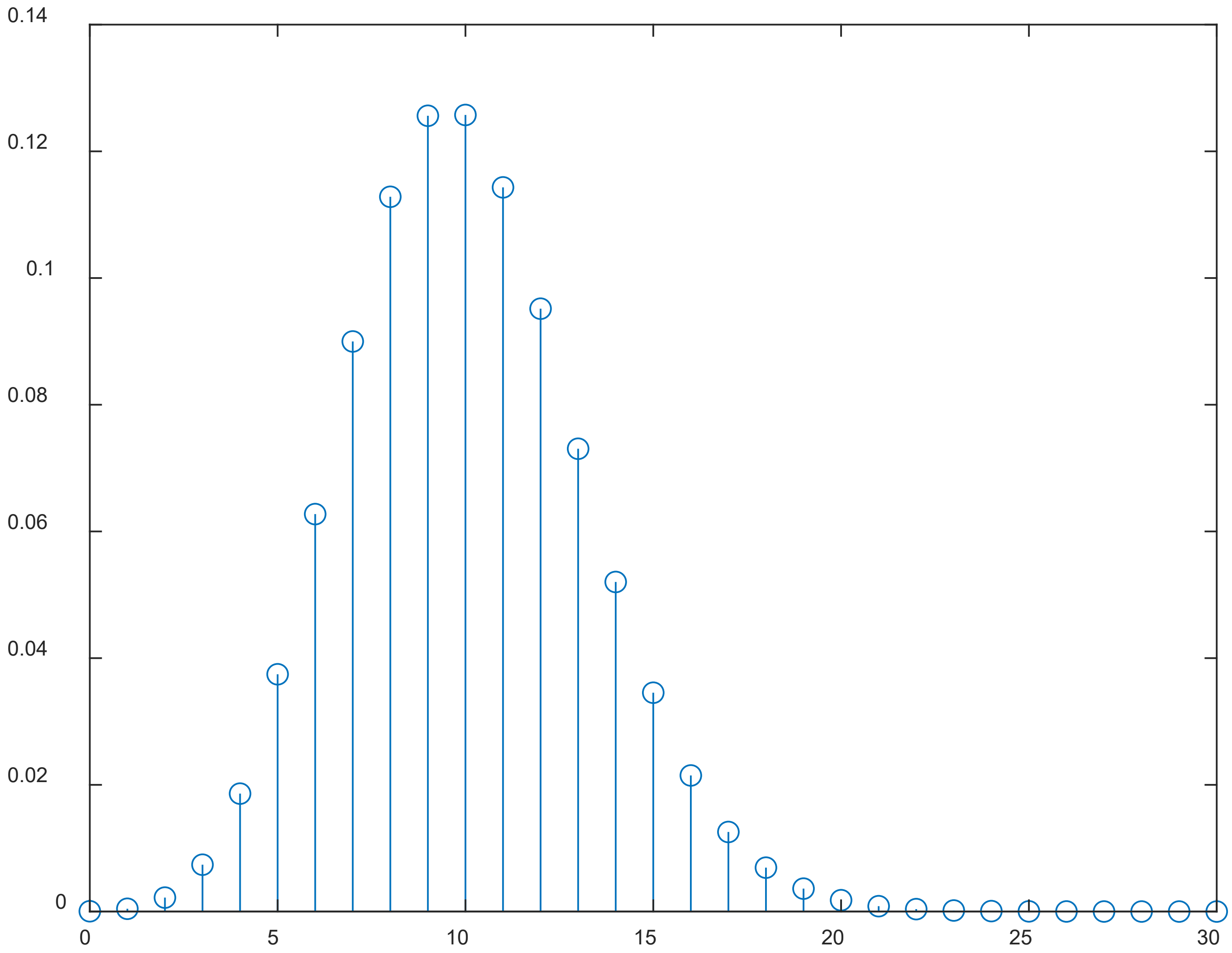
As there are b boxes, each ball has a chance of $1/b$ throwing into one of the boxes. Hence we may apply binomial distribution and employ $p = 1/b$ as the probability of success or the probability of throwing into a given box. Hence the probability is:

$$P(k) = \binom{n}{k} \left(\frac{1}{b}\right)^k \left(1 - \frac{1}{b}\right)^{n-k}$$

Since there are n balls and b boxes, the most probable case might be when each box gets n/b balls. Analogously, when flipping 50 fair coins, we expect that the most probable case corresponds to 25 heads and 25 tails.

Hence, $P(k)$ will reach its maximum value when k is the integer nearest to n/b.

$n = 1000$, $b = 100 \Rightarrow k = 10$ and $p = 0.01$



2.
There are 5 chances to get a special-edition figure. You may get in the first try. If not, you need to have a second try, and so on. Hence the probability is:

2/100+
98/100*2/99+
98/100*97/99*2/98+
98/100*97/99*96/98*2/97+
98/100*97/99*96/98*95/97*2/96 = 0.0980

We may use geometric distribution to perform approximation. Applying the CDF of geometric RV:

$$F(r) = P(X \leq r) = 1 - (1-p)^r$$

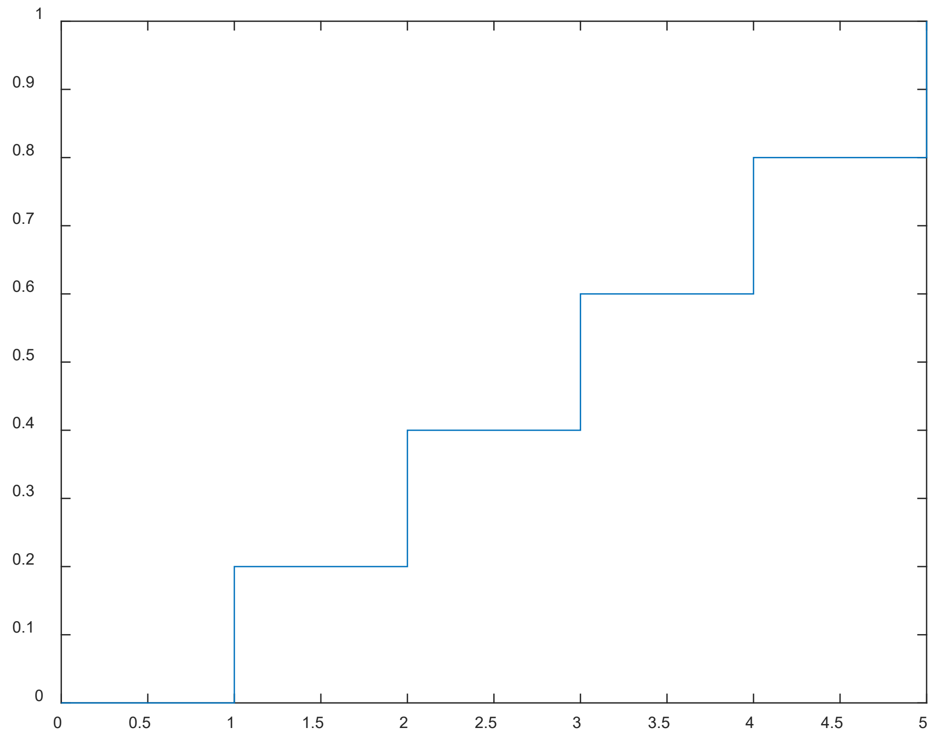
Now we have $r = 5$ and $p = 0.02$, $1 - (1-p)^r = 0.0961$.

In fact, this approximation will become even more accurate when the number of boxes is larger, say, 1000:

20/1000 +
980/1000*20/999 +
980/1000*979/999*20/998 +
980/1000*979/999*978/998*20/997 +
980/1000*979/999*978/998*977/997*20/996=0.0963

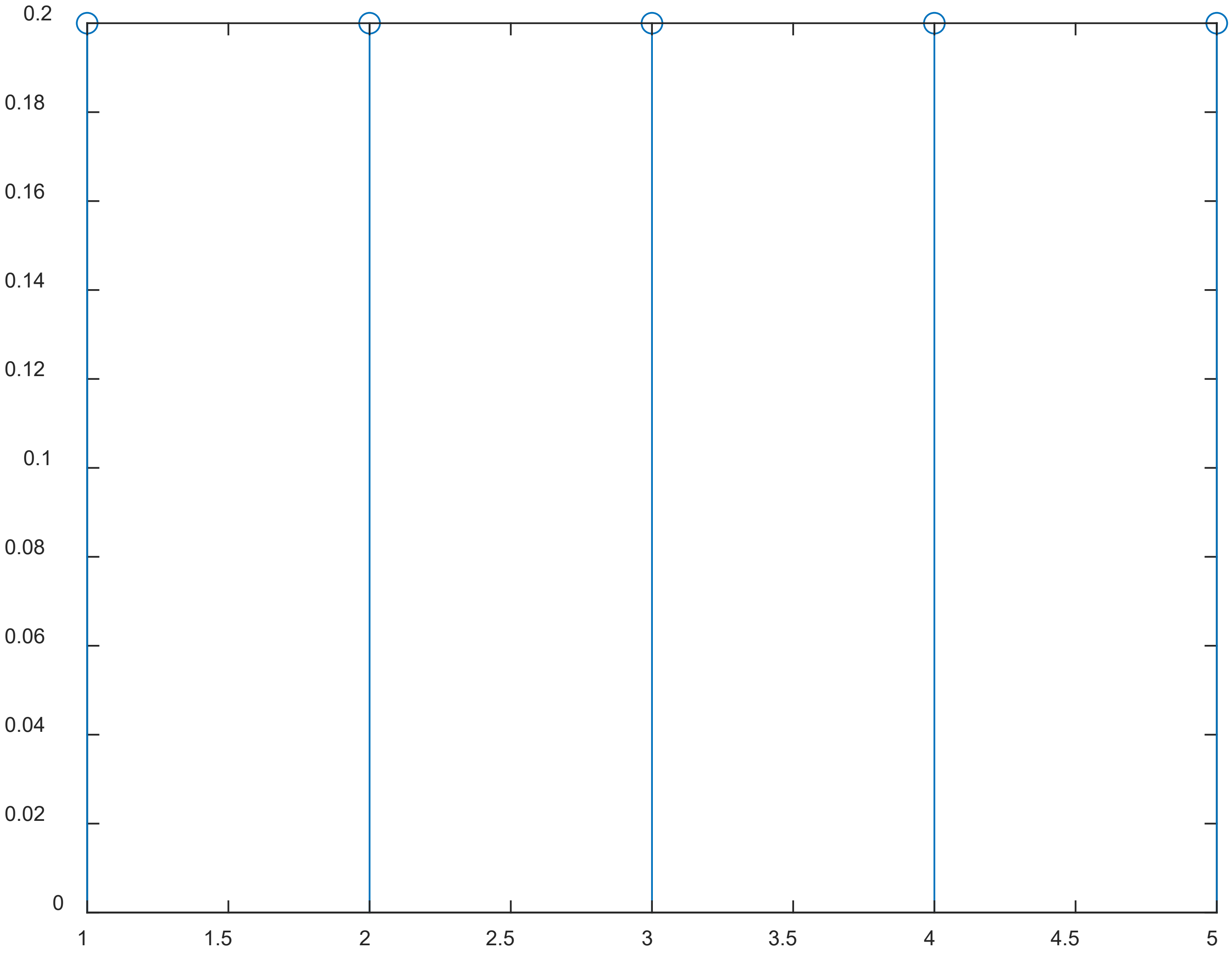
3.

Graphically, the CDF is:



It can be easily deduced that

$$P_X(k) = \begin{cases} 0.2, & k = 1, 2, 3, 4, 5 \\ 0, & \text{otherwise} \end{cases}$$


0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
1
1.5
2
2.5
3
3.5
4
4.5
5

4.

The sum of all PMFs should be equal to 1:

$$\sum_{r=2}^{\infty} p(r) = \sum_{r=2}^{\infty} \alpha p^r = \alpha p^2 [1 + p + p^2 + \cdots] = 1$$

First, the geometric sum must converge and hence $|p| < 1$. Together with the fact that the PMF must be nonnegative, we have $0 < p < 1$.

When the geometric sum converges, we have:

$$\alpha p^2 [1 + p + p^2 + \cdots] = \alpha p^2 \frac{1}{1-p} = \frac{\alpha p^2}{1-p} = 1 \Rightarrow \alpha = \frac{1-p}{p^2}$$

Hence their possible values are:

$$0 < p < 1, \quad \alpha = \frac{1-p}{p^2}$$

5.
Applying (2.10), the PDF is obtained as:

$$p(x) = \begin{cases} 4x^3, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Hence

$$P(0.2 \le X \le 0.4) = \int_{0.2}^{0.4} p(x)dx = x^4\Big|_{0.2}^{0.4} = 0.024$$

Alternatively, the probability can also obtained directly by using CDF:

$$P(0.2 \le X \le 0.4) = F(0.4) - F(0.2) = 0.024$$

6.
Since $Y = 2X$, we know that the admissible values of Y are 0, 2, 4, ….

The PMF of Y can be determined as:

$$P(Y/2 = r) = e^{-\lambda}\frac{\lambda^r}{r!}, \quad r = 0, 1, 2, \cdots$$

$$\Rightarrow P(Y = 2r) = e^{-\lambda}\frac{\lambda^r}{r!}, \quad r = 0, 1, 2, \cdots$$

$$\Rightarrow P(Y = k) = e^{-\lambda}\frac{\lambda^{k/2}}{(k/2)!}, \quad k = 0, 2, 4, \cdots$$

7.

The `rand` command produces a random number uniformly distributed between 0 and 1, while the Bernoulli RV is discrete and has 2 possible values, 0 with probability $1-p$, and 1 with probability p. Here, $p=0.5$ and we may simply assign RV=0 when the uniform number is between 0 and 0.5, and assign RV=1 when the uniform number is between 0.5 and 1, to produce a Bernoulli RV. For example,

```
>> rand
ans = 0.8147
```

We assign this as "1"

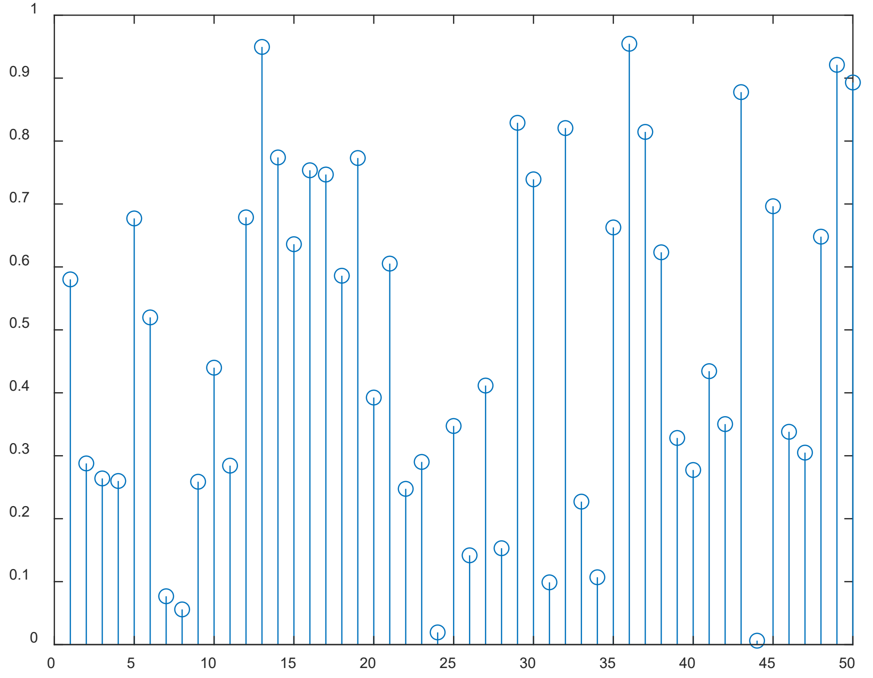
```
>> rand
ans = 0.9058
```

We assign this as "1"

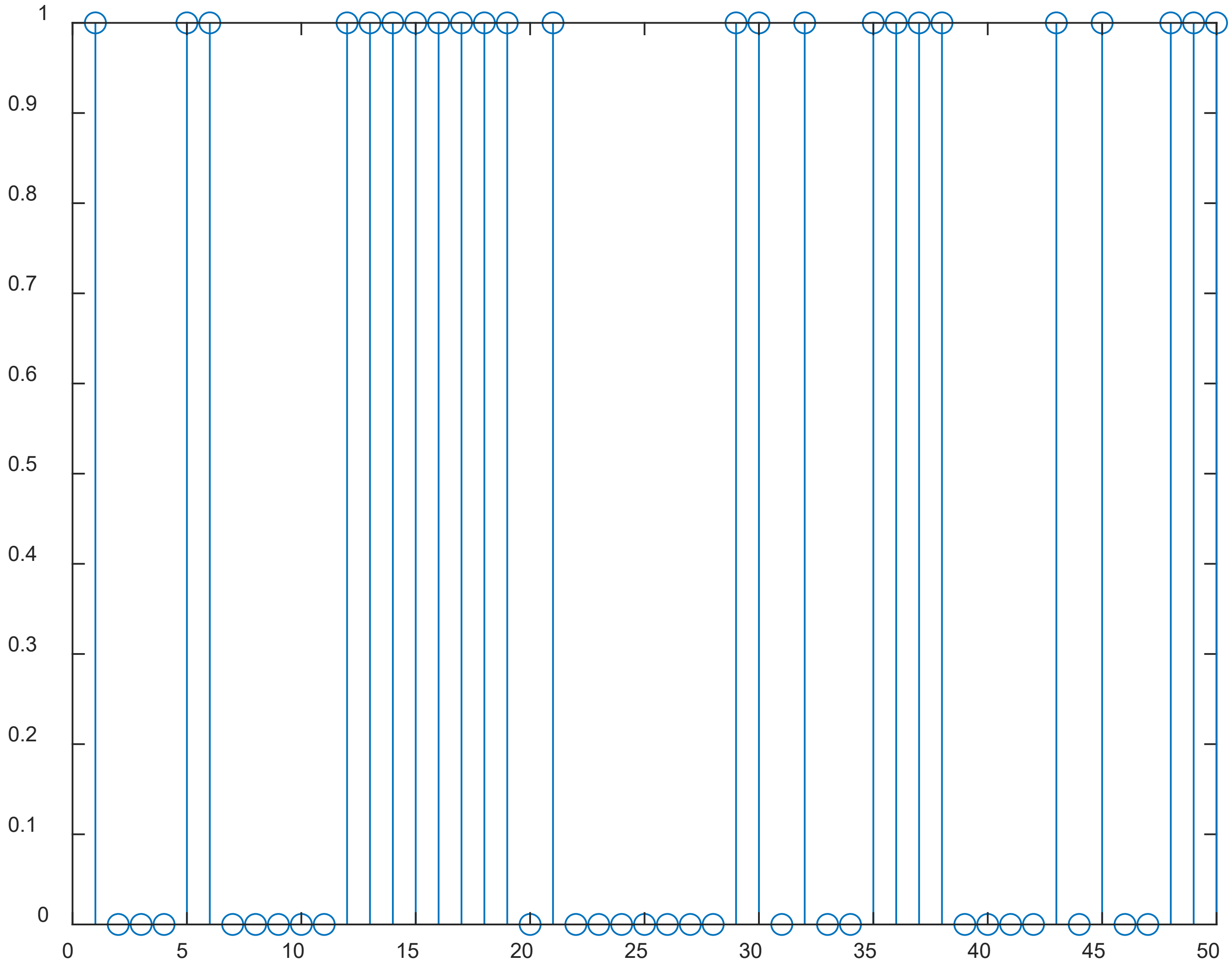
```
>> rand
ans = 0.1270
```

We assign this as "0"

```
N=50;
u=rand(1,N);
for i=1:N
if u(i)<0.5
    b(i)=0;
else
    b(i)=1;
end
end
```



```
stem(1:N, b)
```



`hist(b)`