EM384: Analytical Methods for Engineering Management

Lesson 29: Random Variables and Probability Distributions II

10 April 2023

Table of contents

Lesson Objectives

Lesson Objectives

- Describe the characteristics of four discrete probability distributions that can be incorporated into a simulation (General, Discrete Uniform, Bernoulli, Binomial, Poisson) using the PMF and CDF.
- Describe the characteristics of three continuous probability distributions that can be incorporated into a simulation (Uniform, Normal, Exponential) using the PDF and CDF
- Use Excel to simulate random variables from these distributions

Probability Distributions

Probability Distributions

- Refer to the Lesson 29 Reading for the discrete and continuous probability distributions we will use to generate random variables (RVs) as inputs to our models.
- For each distribution, you should be able to:
  - Know their sample space and how they are parameterized.
  - Recognize the shape of the PMF/PDF.
  - Recognize the shape of the CDF.
  - Be able to generate a RV from that distribution in Excel.
- Discrete general, discrete uniform, Bernoulli, Binomial, Poisson
- Continuous Uniform, Normal, Exponential

Probability Distributions

- A probability distribution is a function that describes the likelihood of observing different outcomes of a random event or possible values of a random variable.
- Probability distributions are used in simulation to generate random values that mimic real-world scenarios. Outcomes of simulations can also be represented with probability distributions.

Probability Distributions

Discrete probability distributions are used when the random variable takes only finite or countably infinite values, while continuous probability distributions are used when the random variable can take any real value within its sample space.

- Discrete probability distributions are defined by their Probability Mass Function, or PMF.
- Continuous probability distributions are defined by their Probability Distribution Functions, or PDF.

A discrete random variable can only assume a finite or countably infinite number of values, so its probability mass function $p(x)$, at every point x in the sample space is

$$p(x) = P(X = x)$$

A continuous random variable can assume an uncountably infinite number of values, so its probability density function $f(x)$ integrated over an interval $[a, b]$ is

$$\int_a^b f(x)dx = P(a \leq X \leq b)$$

Probability Distributions

A *cumulative distribution function (CDF)* $F(x)$ of a continuous random variable X defined on (a, b) gives the probability of getting z *or less* and is defined as

$$P(X \leq z) = F(z) = \int_a^z f(x)dx$$

A *cumulative distribution function (CDF)* $F(x)$ of a discrete random variable X defined on $[a, b]$ gives the probability of getting z *or less* and is defined as

$$P(X \leq z) = F(z) = \sum_{a \leq x \leq z} p(x)$$

Probability Distributions

Since a PMF and PDF assign probabilities (for discrete RVs) or a density (for continuous RVs) to all possible outcomes in their sample space S, we have the following by definition:

Discrete random variable X with PMF $p(x)$ defined on sample space S:

$$p(x) \geq 0 \ \forall \ x \in S$$

$$\sum_{x \in S} p(x) = 1$$

$$p(X = c) = p(c) \text{ for any value } c \in S$$

Continuous random variable X with PDF $f(x)$ defined on sample space S:

$$f(x) \geq 0 \ \forall \ x \in S$$

$$\int_{x \in S} f(x) dx = 1$$

$$p(X = c) = 0 \text{ for any value } c \in S$$

Probability Distributions

In EM384, we will use the following named distributions, with parameters indicated next to each one:

Discrete Distributions:

- General Discrete Distribution
- Uniform Discrete Distribution (a and b)
- Bernoulli Distribution (p)
- Binomial Distribution (n and p)
- Poisson Distribution (λ)

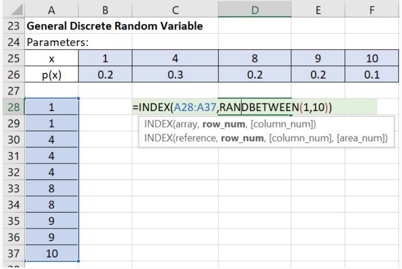
Continuous Distributions:

- Uniform Distribution (a and b)
- Normal Distribution (μ and σ)
- Exponential Distribution (λ)

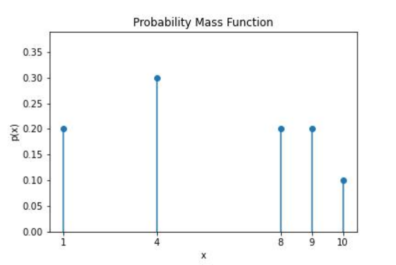
Discrete General Distribution

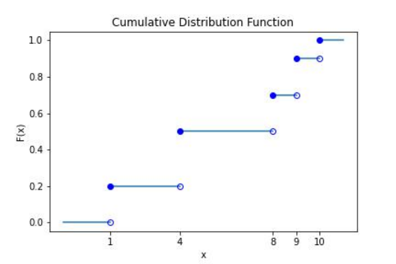
The discrete general distribution doesn't fit any of the other named distributions. It is fully described by the outcomes in the sample space and their associated probabilities.

In Excel, you can replicate the frequencies of the outcomes in a new table, and then pick uniformly from that table using INDEX and RANDBETWEEN.



	A	B	C	D	E	F
23	**General Discrete Random Variable**					
24	Parameters:					
25	x	1	4	8	9	10
26	p(x)	0.2	0.3	0.2	0.2	0.1
27						
28	1		=INDEX(A28:A37,RANDBETWEEN(1,10))			
29	1		INDEX(array, **row_num**, [column_num])			
30	4		INDEX(reference, **row_num**, [column_num], [area_num])			
31	4					
32	4					
33	8					
34	8					
35	9					
36	9					
37	10					

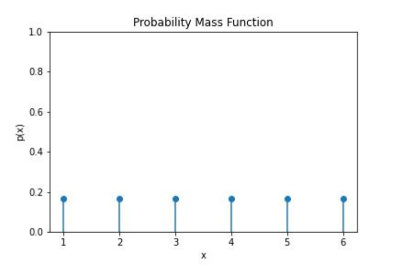


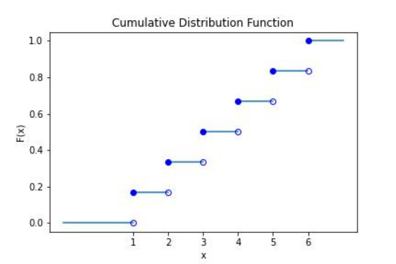


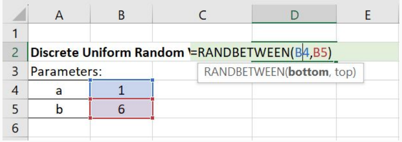
Discrete Uniform Distribution $X \sim \mathcal{U}\{a, b\}$

The discrete Uniform distribution is defined on an integer sample space $S = \{a, a+1, ..., b\}$ where the probability of all outcomes is equal.

In Excel, you can use the formula =RANDBETWEEN(a,b) to generate a discrete Uniform random variable $X \sim \mathcal{U}\{a, b\}$.



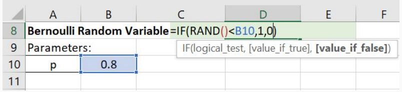




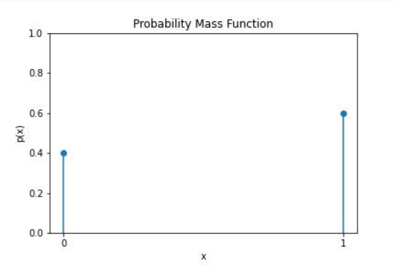
The Bernoulli distribution is defined on sample space $S = \{0, 1\}$ where the probability of outcome 1 is parameter p.

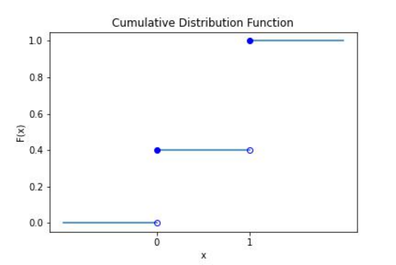
In Excel, you can use the formula =IF(RAND()<p,1,0) to generate a Bernoulli random variable $X \sim Bernoulli(p)$.

Note: if $X \sim Bernoulli(p)$ then $X \sim \mathcal{B}(1, p)$ as well.



	A	B	C	D	E	F
8	**Bernoulli Random Variable**	=IF(RAND()<B10,1,0)				
9	Parameters:		IF(logical_test, [value_if_true], **[value_if_false]**)			
10	p	0.8				
11						

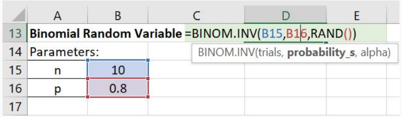




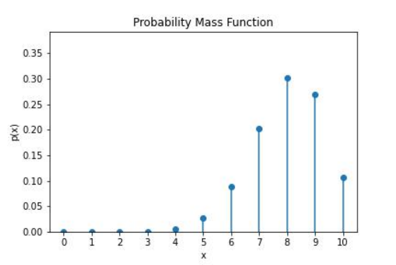
Binomial Distribution $X \sim \mathcal{B}(n, p)$

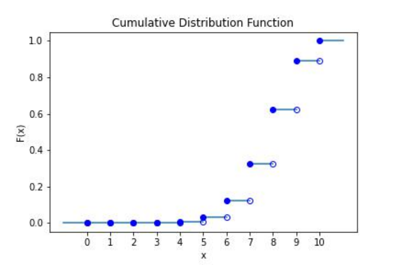
The Binomial distribution is defined on sample space $S = \{0, \ldots, n\}$ where each of n trials is Bernoulli with parameter p. Thus, a Binomial RV is the number of 1's we would get from n identically distributed Bernoulli RVs.

In Excel, you can use the formula =BINOM.INV(n,p,RAND()) to generate a Binomial random variable X with parameters n and p where $X \sim \mathcal{B}(n, p)$.



	A	B	C	D	E
13	**Binomial Random Variable**	=BINOM.INV(B15,B16,RAND())			
14	Parameters:		BINOM.INV(trials, **probability_s**, alpha)		
15	n	10			
16	p	0.8			
17					

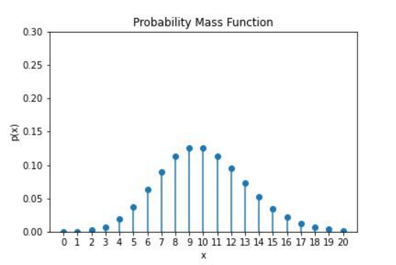


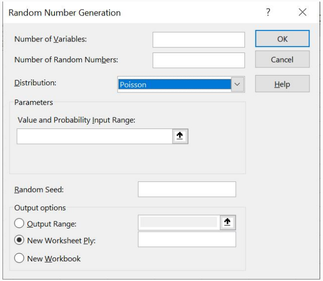


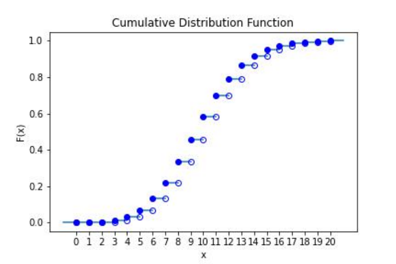
Poisson Distribution $X \sim Pois(\lambda)$

The Poisson distribution is defined on sample space $S = \{0, 1, ...\} = \mathbb{Z}^*$ where the outcome denotes the number of observations of a phenomenon that happens at a mean rate λ.

In Excel, you have to use the Data Analysis Tookpak to generate a Poisson random variable $X \sim Pois(\lambda)$.



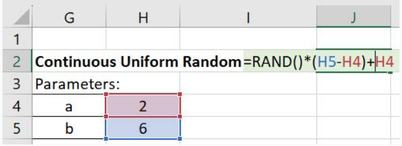

Random Number Generation
Number of Variables:
Number of Random Numbers:
Distribution: Poisson
OK
Cancel
Help
Parameters
Value and Probability Input Range:
Random Seed:
Output options
Output Range:
New Worksheet Ply:
New Workbook

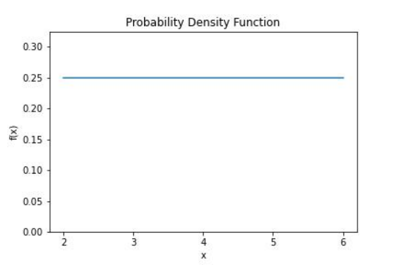


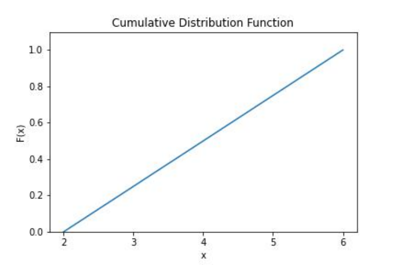
Uniform Distribution $X \sim \mathcal{U}[a, b]$

The Uniform distribution (not to be confused with discrete uniform distribution) is defined on sample space $S = [a, b]$ where all outcomes are equally likely.

In Excel, you can use the formula =RAND()*(b-a)+a to generate a Uniform random variable $X \sim \mathcal{U}[a, b]$.



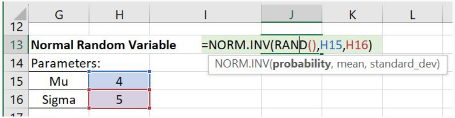


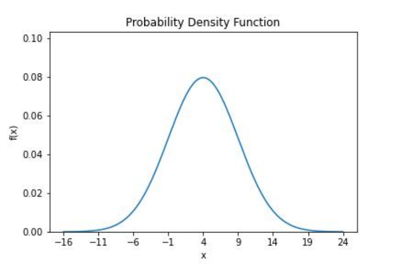


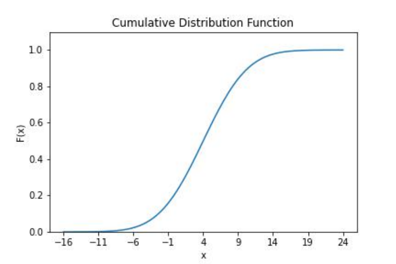
Normal Distribution $\quad X \sim \mathcal{N}(\mu, \sigma)$

The Normal distribution is defined on sample space $S = (-\infty, \infty) = \mathbb{R}$.

In Excel, you can use the formula =NORM.INV(RAND(),μ,σ) to generate a Normal random variable $X \sim \mathcal{N}(\mu, \sigma)$ where μ is the mean and σ is the standard deviation.





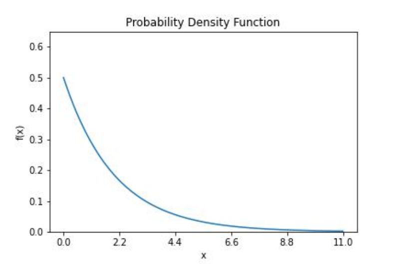


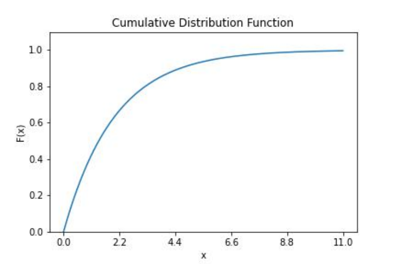
Exponential Distribution $X \sim Exp(\lambda)$

The Exponential distribution is defined on sample space $S = [0, \infty)$.

In Excel, you can use the formula =$-1/\lambda$*LN(RAND()) to generate an exponential random variable X with rate parameter λ (where the mean is $\mu = 1/\lambda$), where $X \sim Exp(\lambda)$.

	G	H	I	J	K
8	**Exponential Random Variable**		=-H11*LN(RAND())		
9	Parameters:				
10	Lambda	0.5			
11	Mu	2			

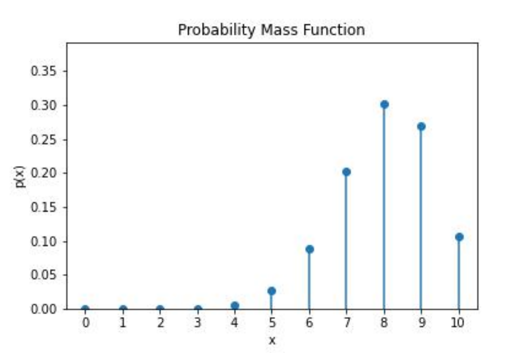


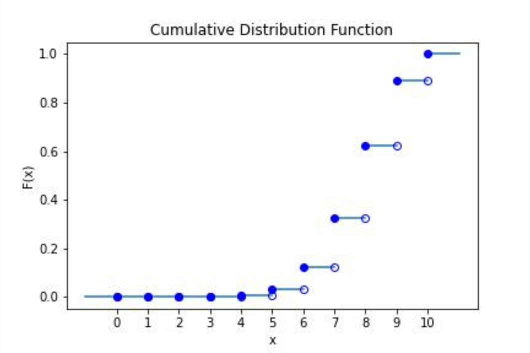


Check on Learning

Check on Learning

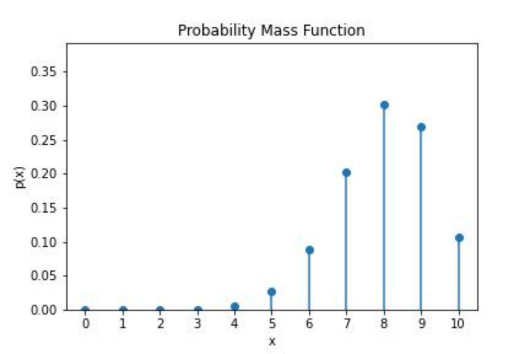
- What distribution does the following graph belong to (parameterized by n and p)?
- From the graph, what is its sample space?

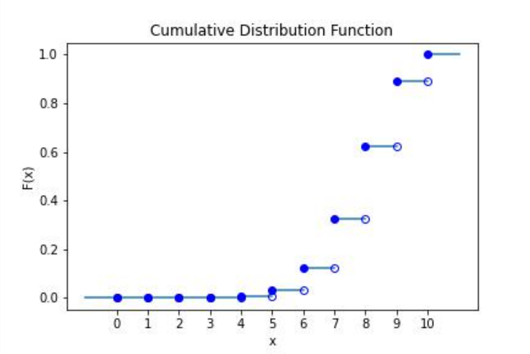




Check on Learning

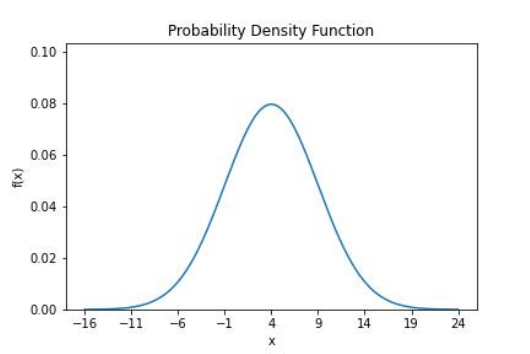
- What distribution does the following graph belong to? What distribution does the follow graph belong to (parameterized by n and p)? Binomial Distribution
- From the graph, what is its sample space? $\{0, 1, 2, \ldots, 10\}$

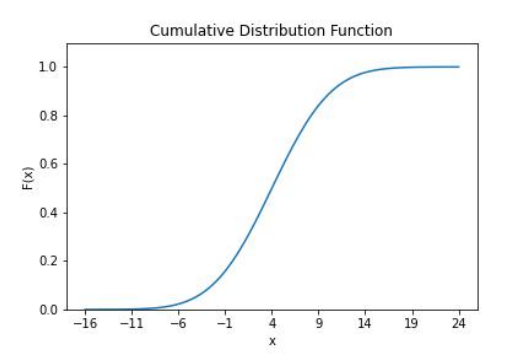




Check on Learning

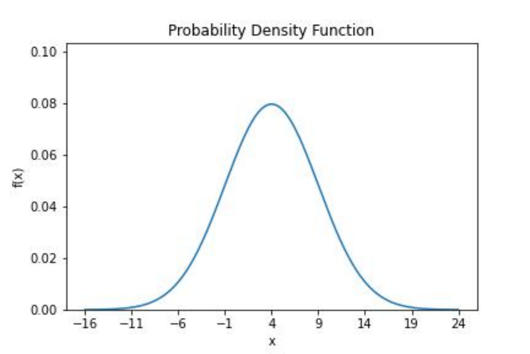
- What distribution does the following graph belong to?
- From the graph, what is its sample space?

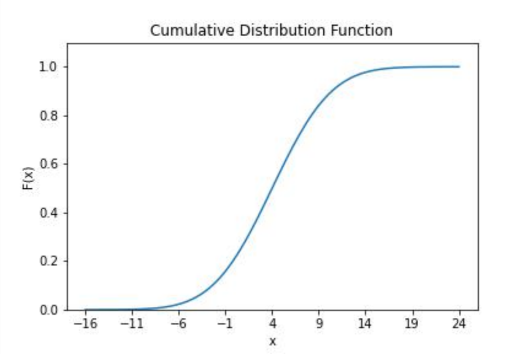




Check on Learning

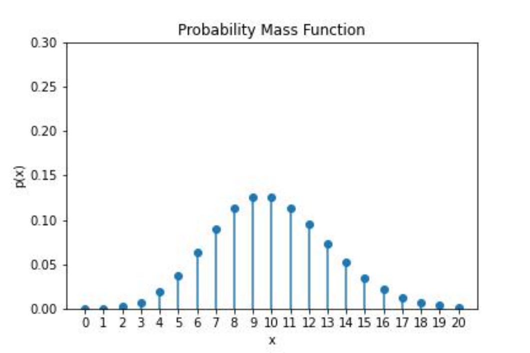
- What distribution does the following graph belong to? Normal Distribution with mean $\mu = 3$
- From the graph, what is its sample space? $(-\infty, \infty)$

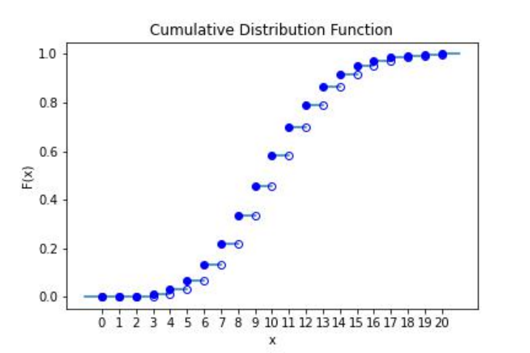




Check on Learning

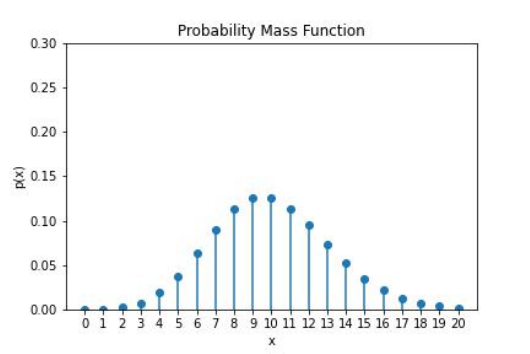
- What distribution does the following graph belong to (parameterized by λ)?
- From the graph, what is its sample space?

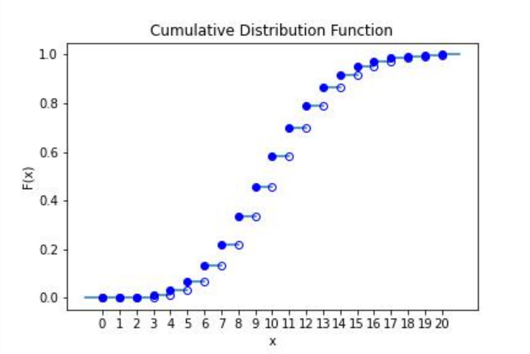




Check on Learning

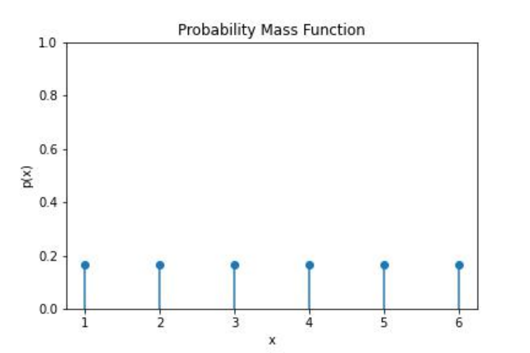
- What distribution does the following graph belong to? (parameterized by λ)? Poisson Distribution
- From the graph, what is its sample space? $\{0, 1, 2, 3, \ldots\}$

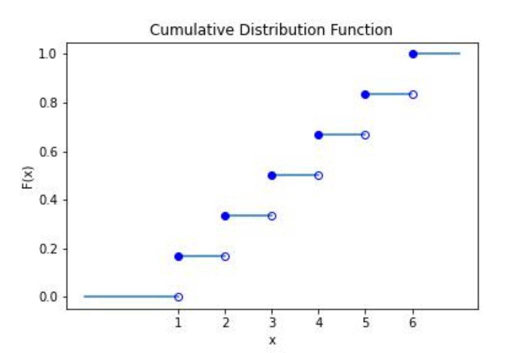




Check on Learning

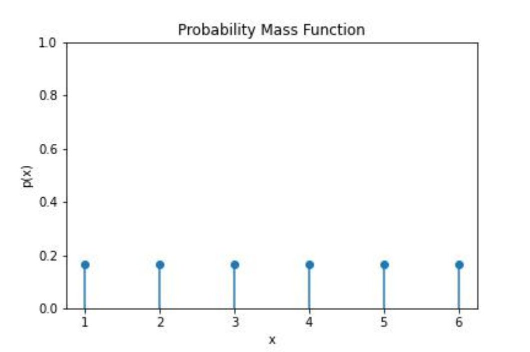
- What distribution does the following graph belong to?
- From the graph, what is its sample space?

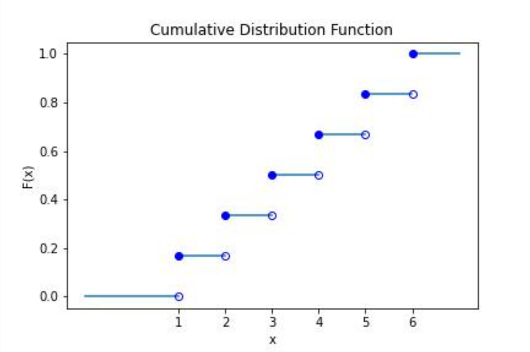




Check on Learning

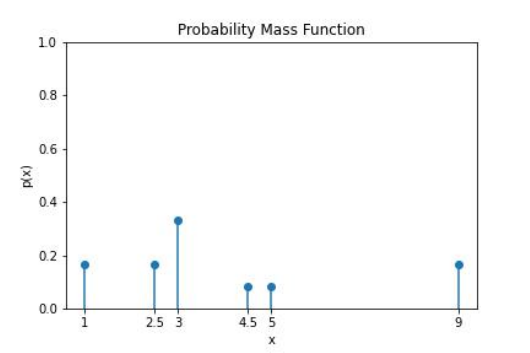
- What distribution does the following graph belong to? Discrete Uniform Distribution
- From the graph, what is its sample space? $\{1, 2, 3, 4, 5, 6\}$

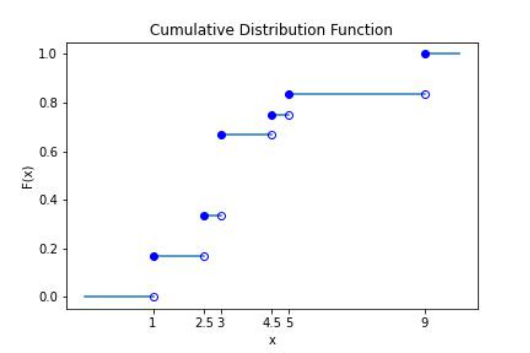




Check on Learning

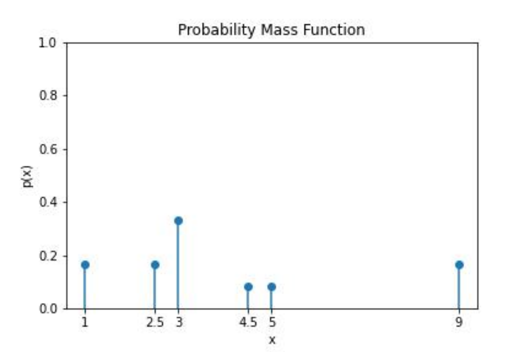
- What distribution does the following graph belong to?
- From the graph, what is its sample space?

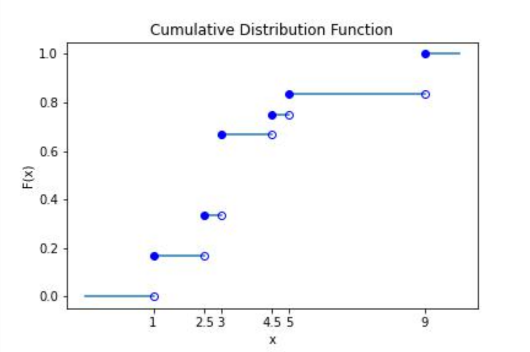




Check on Learning

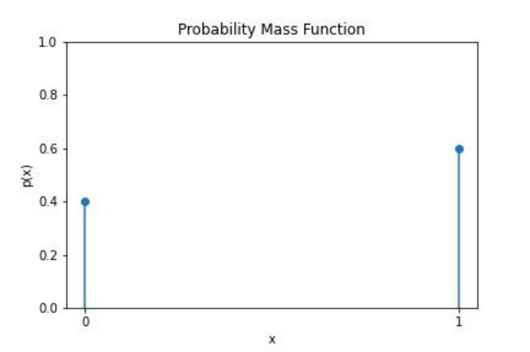
- What distribution does the following graph belong to? Discrete General Distribution
- From the graph, what is its sample space? $\{1, 2.5, 3, 4.5, 5, 9\}$

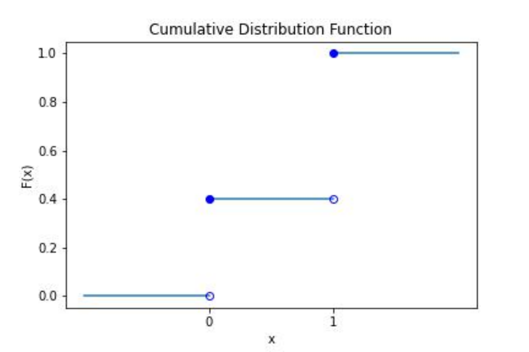




Check on Learning

- What distribution does the following graph belong to?
- From the graph, what is its sample space?

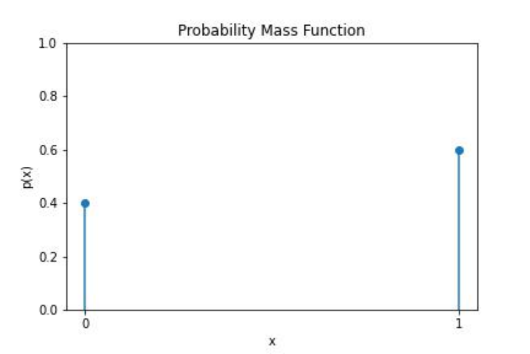


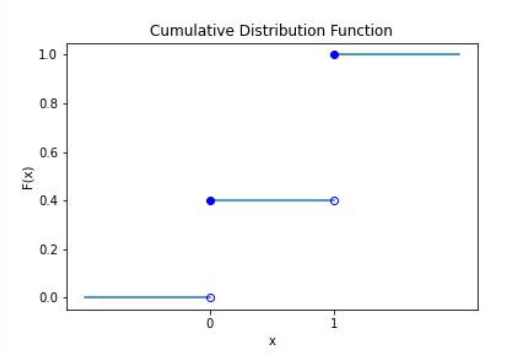


Check on Learning

- What distribution does the following graph belong to? Bernoulli Distribution, with parameter $p = 0.6$

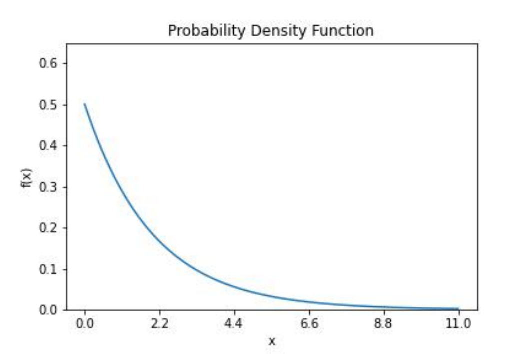
- From the graph, what is its sample space? $\{0, 1\}$

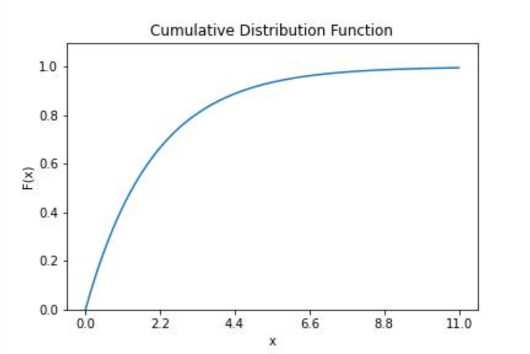




Check on Learning

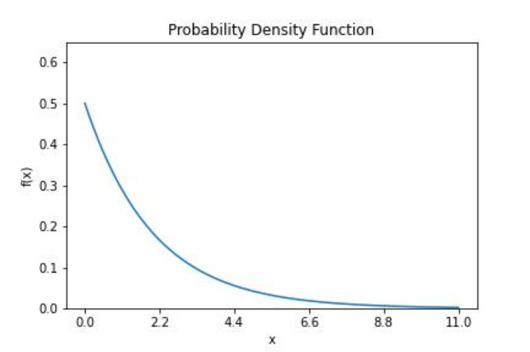
- What distribution does the following graph belong to?
- From the graph, what is its sample space?

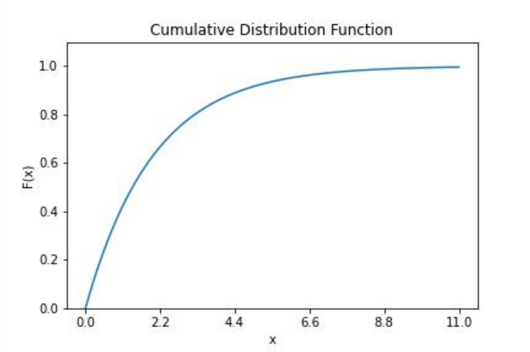




Check on Learning

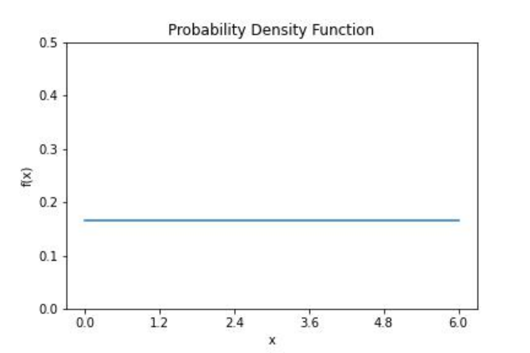
- What distribution does the following graph belong to? Exponential Distribution
- From the graph, what is its sample space? $[0, \infty)$

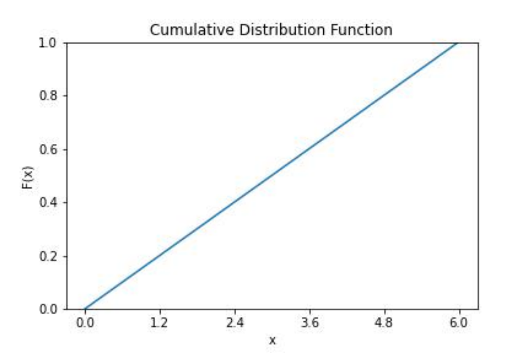




Check on Learning

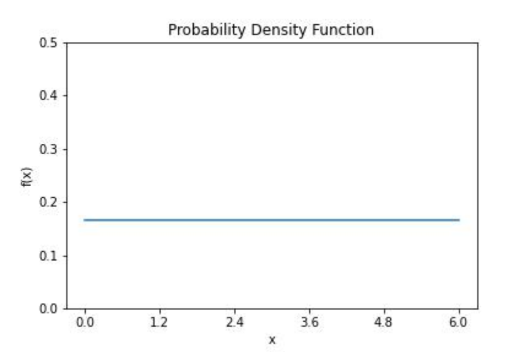
- What distribution does the following graph belong to?
- From the graph, what is its sample space?

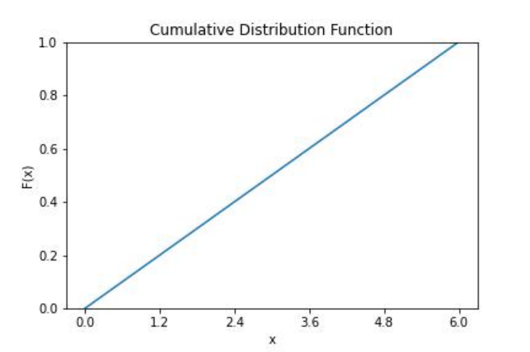




Check on Learning

- What distribution does the following graph belong to? Continuous Uniform
- From the graph, what is its sample space? [0, 6]





Practical Exercise

PE1

We can create random variables in Excel using either formulas (Excel will regenerate them when the sheet is changed), or using the Data Analysis Toolpak.

- Open the Lesson 29 PE - RVs from Probability Distributions (Shell)

Conclusion

Next Class

Homework:

- Re-read PDF on probability distributions.
- Read Chap 24.5 - 24.6 (in PDF)

Next Lesson:

- Understand Monte Carlo simulation.
- Create a Monte Carlo simulation in Excel.
- Interpret the results of a Monte Carlo Simulation using expected value, standard deviation, and a histogram.
- Find the probability of a result using the results of a Monte Carlo simulation