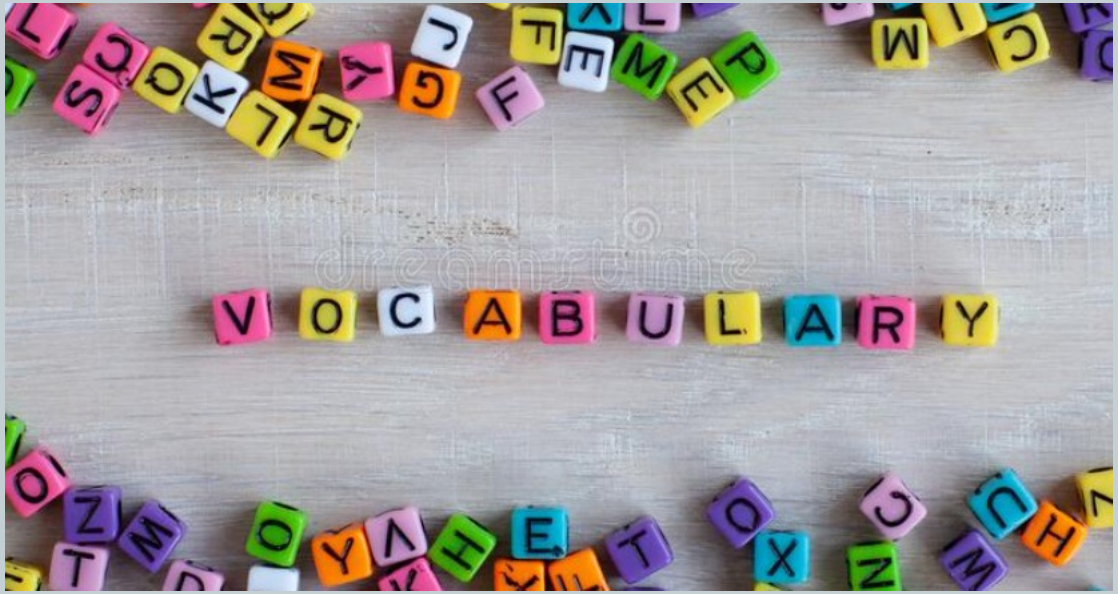
CHAPTER FOUR

Discrete Random Variables

X

Vocabulary

Listen for these words today to determine what they mean.



Experiment

Sample Space Ω

Probability Function

Random Variable

(two flavors: Discrete and Continuous)

Probability Mass Function

Cumulative Distribution Function

Experiment: Roll two dice

Sample Space Ω

$\Omega = \{(\omega_1, \omega_2) : \omega_1, \omega_2 \ \in \{1, 2, 3, 4, 5, 6\}\ \}$

$\Omega = \{(1,1), (1,2), (1,3), \dots, (2,1), \dots, (6,5), (6,6)\}$

We are only interested in the sum: $\{2, 3, \dots, 12\}$.

So, we are interested in the value of

the function $S: \Omega \to \mathbb{R}$ given by

$S(\omega_1, \omega_2) = \omega_1 + \omega_2 = k$ for $(\omega_1, \omega_2) \in \Omega$.

$k \in \{2, 3, \dots, 12\}$

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

Discrete Random Variable aka DRV

The function $S(\omega_1, \omega_2) = \omega_1 + \omega_2 = k$ for $(\omega_1, \omega_2) \in \Omega$
is a discrete random variable.

$S: \Omega \longrightarrow \mathbb{R}.$

That is to say S transforms
$\Omega = \{(1,1), (1,2), (1,3), \dots, (2,1), \dots, (6,5), (6,6)\}$ to

$$\widetilde{\Omega} = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Notice that a DRV is a function.
Do be confused with the terminology being used: 'variable' and 'function'

$\{S = k\}$ refers to the subset of ordered pairs $(\omega_1, \omega_2) \in \Omega$ for which $S(\omega_1, \omega_2) = \omega_1 + \omega_2 = k$

What are the **outcomes** in the event $\{S = 8\}$?

What are the **outcomes** in the event $\{S = 8\}$?

Ans: $\{S = 8\} = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

$\{S = 8\}$ refers to the subset of ordered pairs $(\omega_1, \omega_2) \in \Omega$ for which

$$S(\omega_1, \omega_2) = \omega_1 + \omega_2 = 8$$

Now we can talk about the probability of the event $\{S = k\}$,

or $P(S = k)$, for any k.

	⚀	⚁	⚂	⚃	⚄	⚅
⚀	2	3	4	5	6	7
⚁	3	4	5	6	7	8
⚂	4	5	6	7	8	9
⚃	5	6	7	8	9	10
⚄	6	7	8	9	10	11
⚅	7	8	9	10	11	12

$$P(S = 1) = P(\emptyset) = 0$$

$$P(S = 2) = P\big((1,1)\big) = \frac{1}{36}$$

$$P(S = 3) = P(\{(1,2),(2,1)\}) = \frac{2}{36}$$

$$P(S = 8) = P(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = \frac{5}{36}$$

$$P(S = 13) = P(\emptyset) = 0$$

Now we can in essence ignore Ω.

It is sufficient to merely list the values of S and the corresponding probabilities.

This information is contained in the Probability Mass Function, p of S.

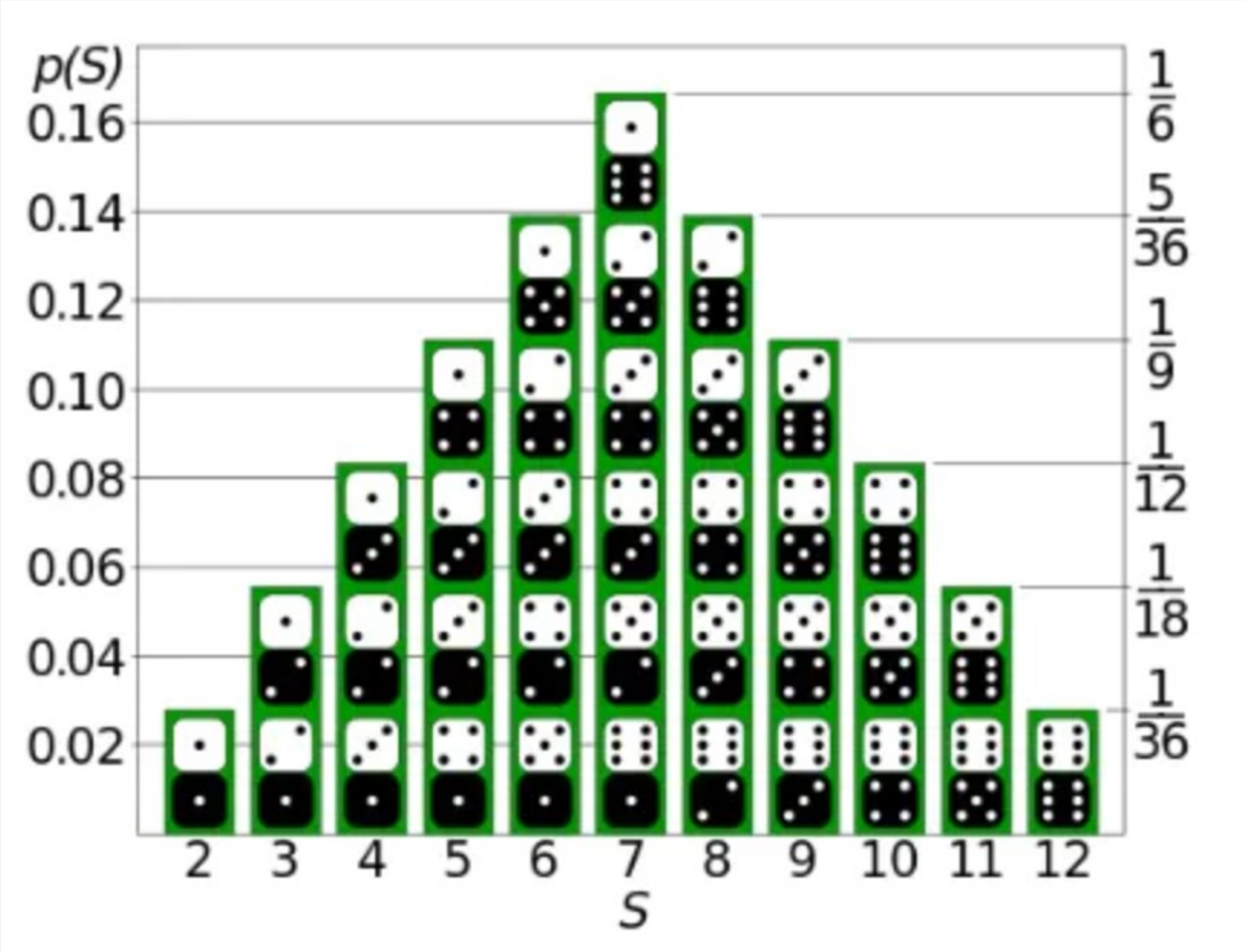
$p\colon \mathbb{R} \to [0,1]$ such that $p(a) = P(S = a)$.

In this two-dice summing example, we would have:

$$p(2) = \frac{1}{36}\ , p(3) = \frac{2}{36}\ , p(4) = \frac{3}{36}\ , p(5) = \frac{4}{36}\ , \ \ldots , p(12) = \frac{1}{36}\ .$$

Notice that $p(2) + p(3) + \cdots + p(12) = 1$.

AN ALTERNATIVE LOOK AT THE PMF FOR THIS DICE EXAMPLE.

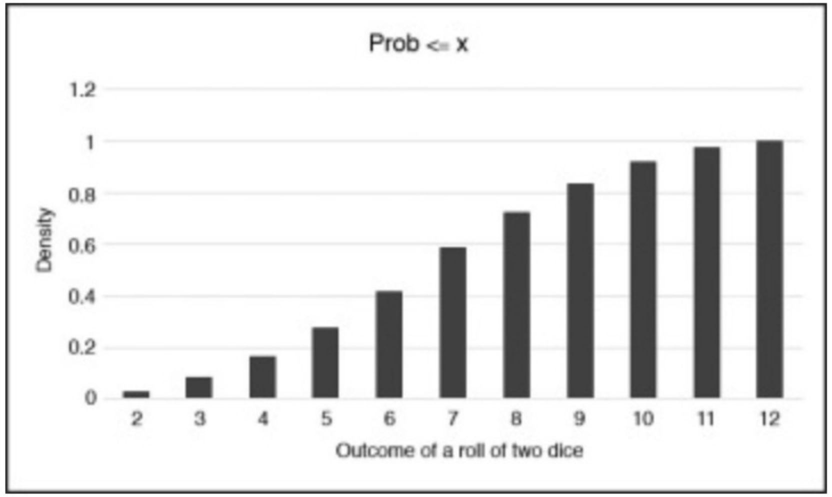


This brings us to the Cumulative Distribution Function (or CDF) of S.

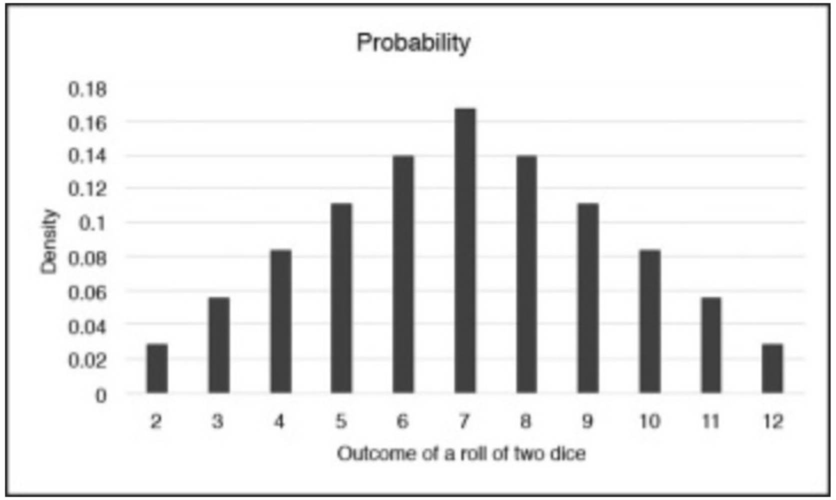
$F\colon \mathbb{R} \to [0,1]$, defined by $F(a) = P(S \leq a)$.

The probability distribution of S is determined by the CDF or the PMF.

CDF



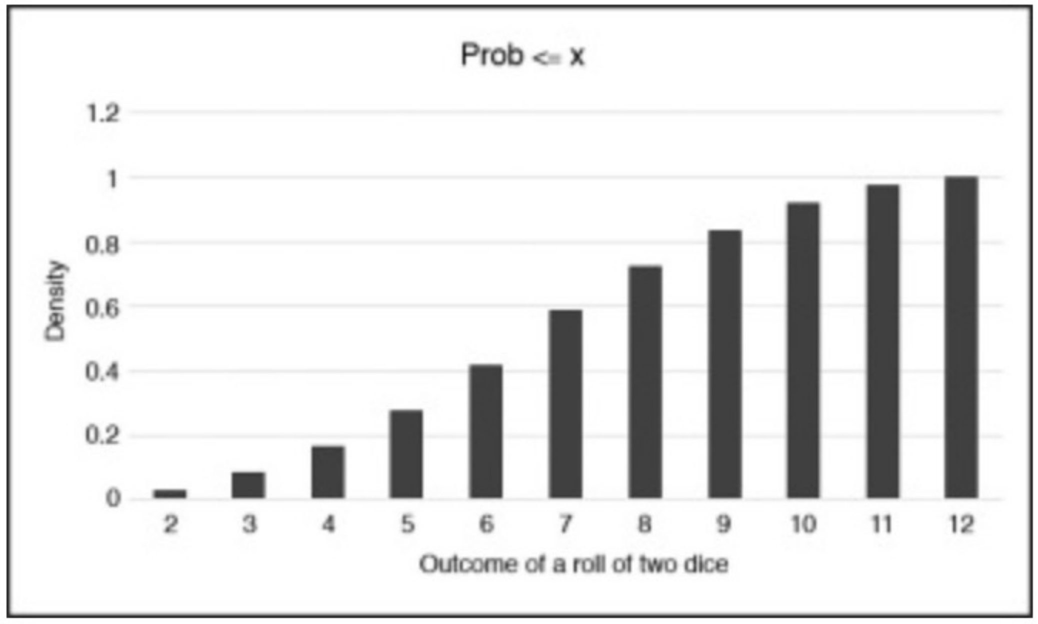
PMF



Notice three properties of the CDF of S:

(The cumulative distribution function of the discrete random variable)

1] $F(a) \leq F(b)$.
This is because $a \leq b$ implies that the event $\{S \leq a\}$ is contained in the event $\{S \leq b\}$.



2] $F(a)$ is a probability.
Therefore, $0 \leq F(a) \leq 1$, and

$$\lim_{a \to +\infty} F(a) = \lim_{a \to +\infty} P(S \leq a) = 1$$

$$\lim_{a \to -\infty} F(a) = \lim_{a \to -\infty} P(S \leq a) = 0$$

3] F is right-continuous: $\lim_{\varepsilon \downarrow 0} F(a + \varepsilon) = F(a)$

Now let's repeat that whole discussion with a different experiment

Experiment: Roll two dice

$\Omega = \{(\omega_1, \omega_2) : \omega_1, \omega_2 \ \in \{1, 2, 3, 4, 5, 6\} \}$

$\Omega = \{(1, 1), (1, 2), (1, 3), \dots, (2, 1), \dots, (6, 5), (6, 6)\}$

We want the maximum of the two dice: $\{1, 2, 3, 4, 5, 6\}$.

So, we are interested in the value of the function M: $\Omega \to \mathbb{R}$ given by:

$M(\omega_1, \omega_2) = \omega^{max} = k$ for $(\omega_1, \omega_2) \in \Omega$. $\quad k \in \{1, 2, 3, 4, 5, 6\}$

Discrete Random Variable aka DRV

The function $M(\omega_1, \omega_2) = \omega^{max} = k$ for $(\omega_1, \omega_2) \in \Omega$
is a discrete random variable.

$M: \Omega \longrightarrow \mathbb{R}.$

That is to say M transforms
$\Omega = \{(1,1), (1,2), (1,3), \ldots, (2,1), \ldots, (6,5), (6,6)\}$ to

$$\widetilde{\Omega} = \{1,2,3,4,5,6\}$$

The DRV is a function.

Practice:

What are the **outcomes** in the event $\{M = 3\}$?

Practice:

What are the **outcomes** in the event $\{M = 3\}$?

Ans: $\{M = 3\} = \{(1,3),(2,3),(3,3),(3,1),(3,2)\}$

$\{M = 3\}$ refers to the subset of ordered pairs $(\omega_1, \omega_2) \in \Omega$ for which $M(\omega_1, \omega_2) = \omega^{max} = 3$

Now we can talk about the probability of the event $\{M = k\}$, or $P(M = k)$.

$$P(M = 1) = P((1,1)) = \frac{1}{36}$$

$$P(M = 2) = P\big((\{(1,2),(2,1),(2,2)\})\big) = \frac{3}{36}$$

$$P(M = 3) = P(\{(1,3),(2,3),(3,3),(3,1),(3,2)\}) = \frac{5}{36}$$

$$P(M = 7) = P(\emptyset) = 0$$

Now we can in essence ignore Ω.

It is sufficient to merely list the values of M and the corresponding probabilities.

This information is contained in the Probability Mass Function, p of M.

$p: \mathbb{R} \to [0, 1]$ such that $p(a) = P(M = a)$.

In this two-dice maximum-value example, we would have:

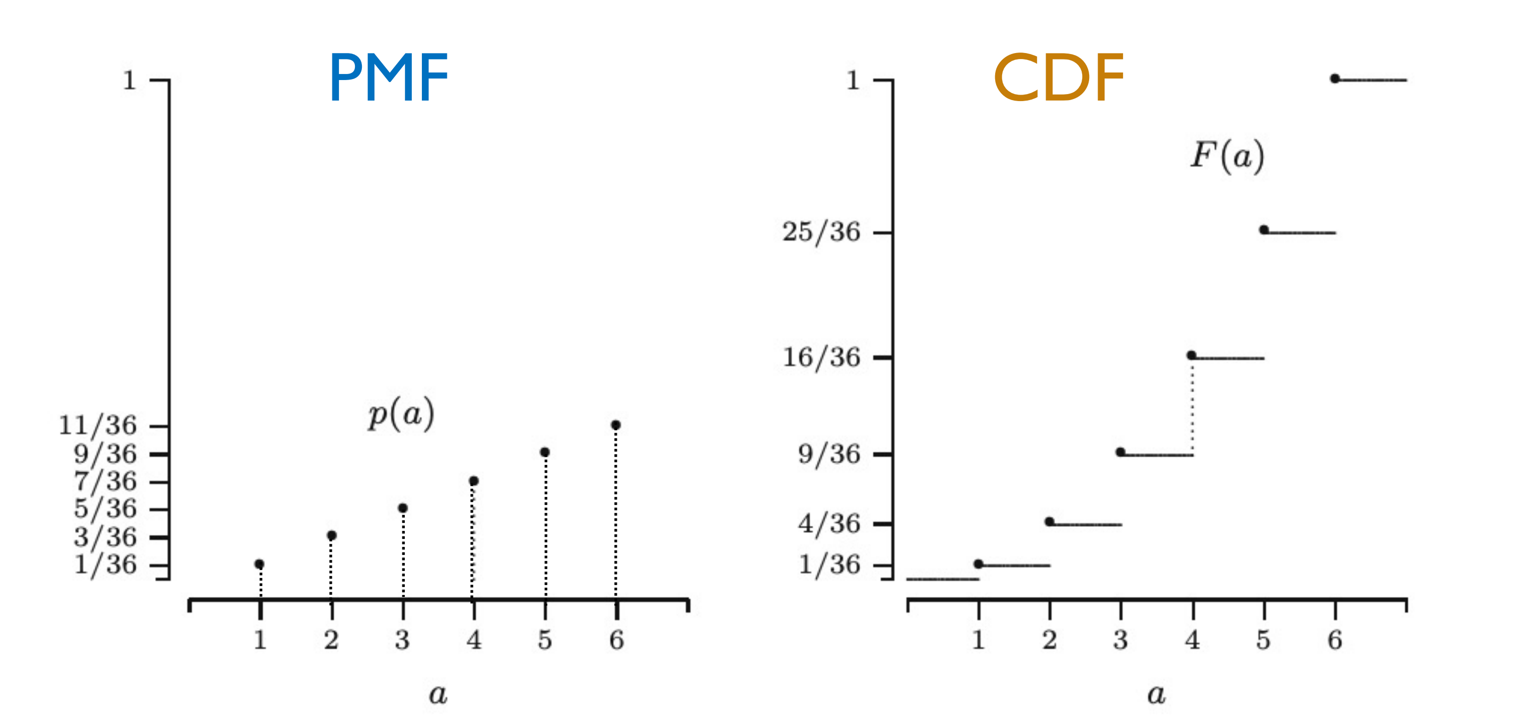
$$p(1) = \frac{1}{36}, p(2) = \frac{3}{36}, p(3) = \frac{5}{36}, \ldots, p(6) = \frac{11}{36}.$$

Notice that $p(1) + p(2) + \cdots + p(6) = 1$.

This brings us to the Cumulative Distribution Function (or CDF) of M.

$F\colon \mathbb{R} \to [0,1]$, defined by $F(a) = P(M \leq a)$.

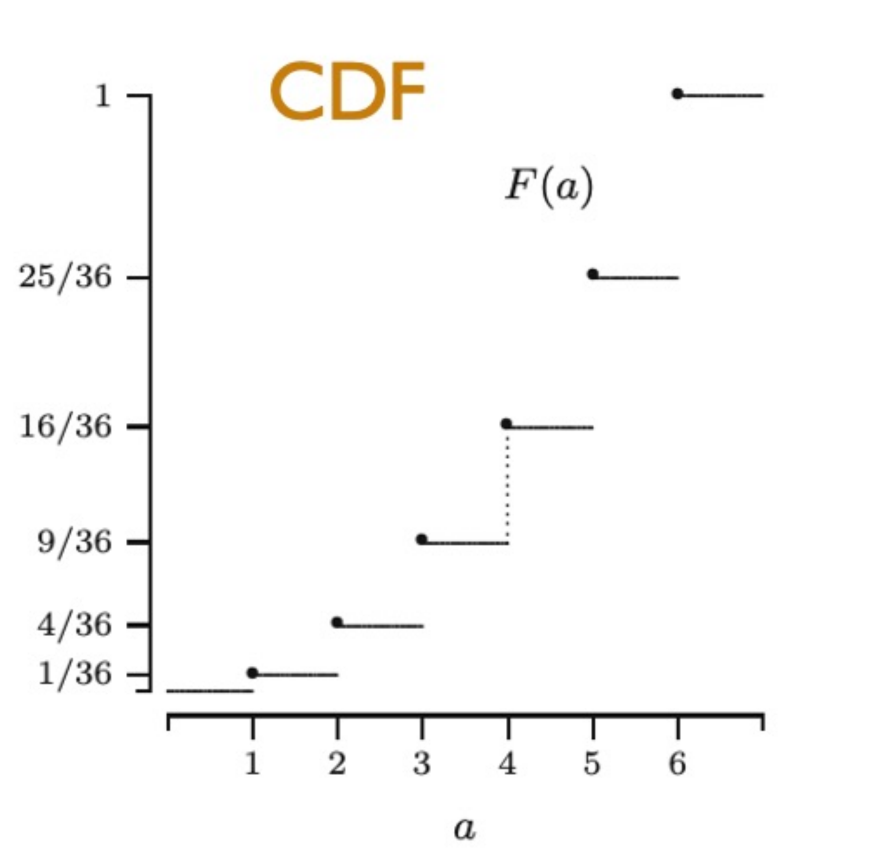
The probability distribution of M is determined by the PMF or the CDF.



These are the same types of graphs seen in experiment *1*, they are just a different style.

Notice three properties of the CDF of M:
(The cumulative distribution function of the discrete random variable)

1] $F(a) \leq F(b)$.
This is because $a \leq b$ implies that the event $\{M \leq a\}$ is contained in the event $\{M \leq b\}$.



2] $F(a)$ is a probability.
Therefore, $0 \leq F(a) \leq 1$, and

$$\lim_{a \to +\infty} F(a) = \lim_{a \to +\infty} P(M \leq a) = 1$$

$$\lim_{a \to -\infty} F(a) = \lim_{a \to -\infty} P(M \leq a) = 0$$

3] F is right-continuous: $\lim_{\varepsilon \downarrow 0} F(a+\varepsilon) = F(a)$

Let's try another experiment

Experiment: Rolling two dice.

You are interested in the magnitude of the difference shown on each die.

What is the DRV? Call it D.
What is Ω ?
What are the outcomes in the event $\{D = 4\}$?
What is the PMF of D ?
What is the CDF of D ?

Find the number of outcomes, the actual outcomes, and their frequencies.

$D: \Omega \to \mathbb{R}$

$\Omega = \{(1,1),(1,2),(1,3),\dots,(6,6)\}$

$D(d_1, d_2) = |d_1 - d_2|$

$\widetilde{\Omega} = \{0,1,2,3,4,5\}$

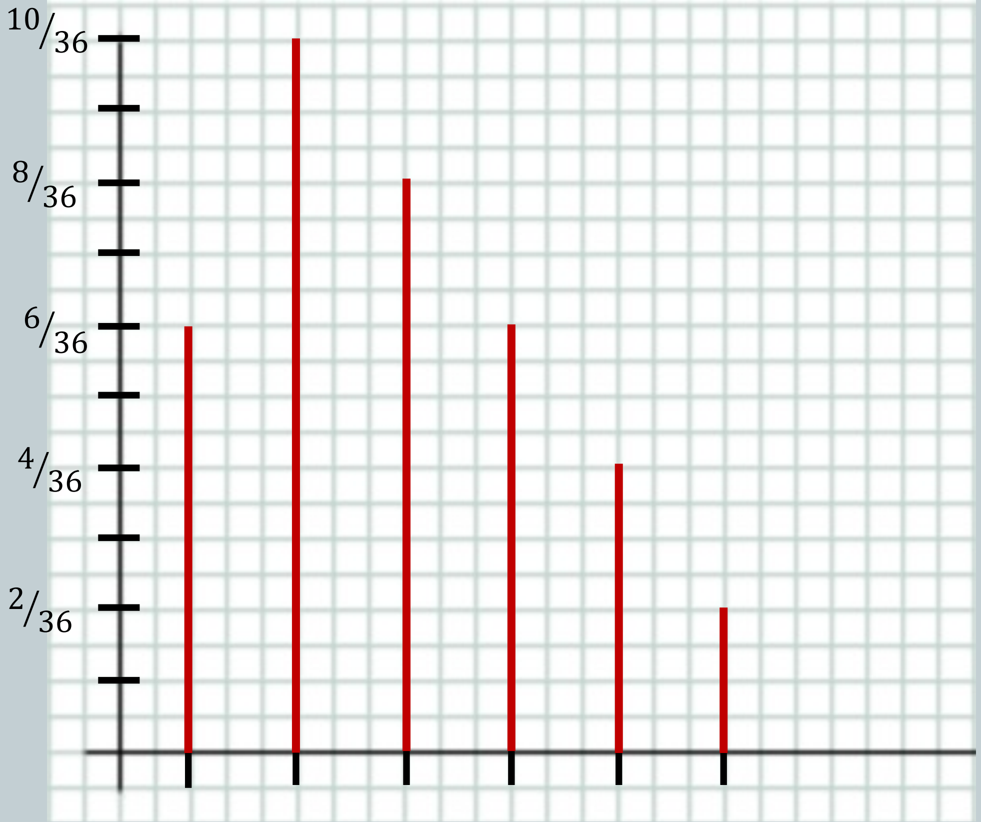
	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

outcome	Count
0	6
1	10
2	8
3	6
4	4
5	2

$\{D = 4\} = \{(5,1),(6,2),(1,5),(2,6)\}$

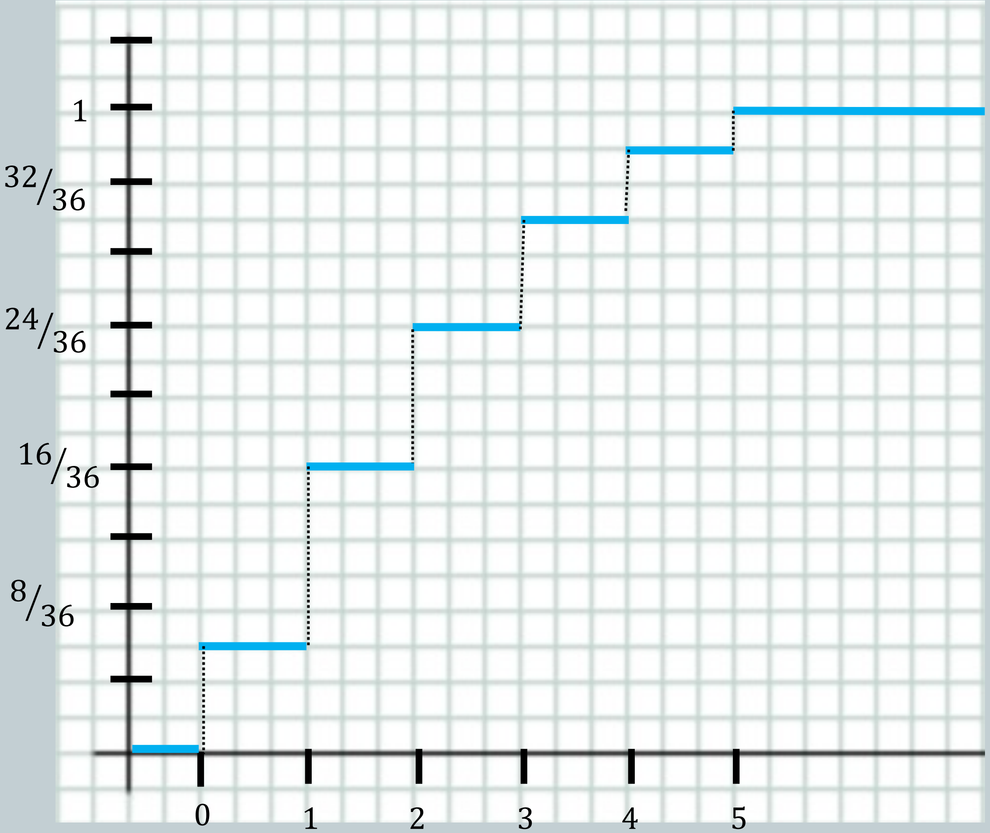
	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

outcome	Count
0	6
1	10
2	8
3	6
4	4
5	2



PMF

CDF



So, in general… (now that we've seen three experiments)

Discrete Random Variable

Let Ω be a sample space.

A discrete random variable is a function $X: \Omega \to \mathbb{R}$ that takes a finite number of values $a_1, a_2, a_3, \ldots, a_n$ (or an infinite number of values $a_1, a_2, a_3, \ldots,$)

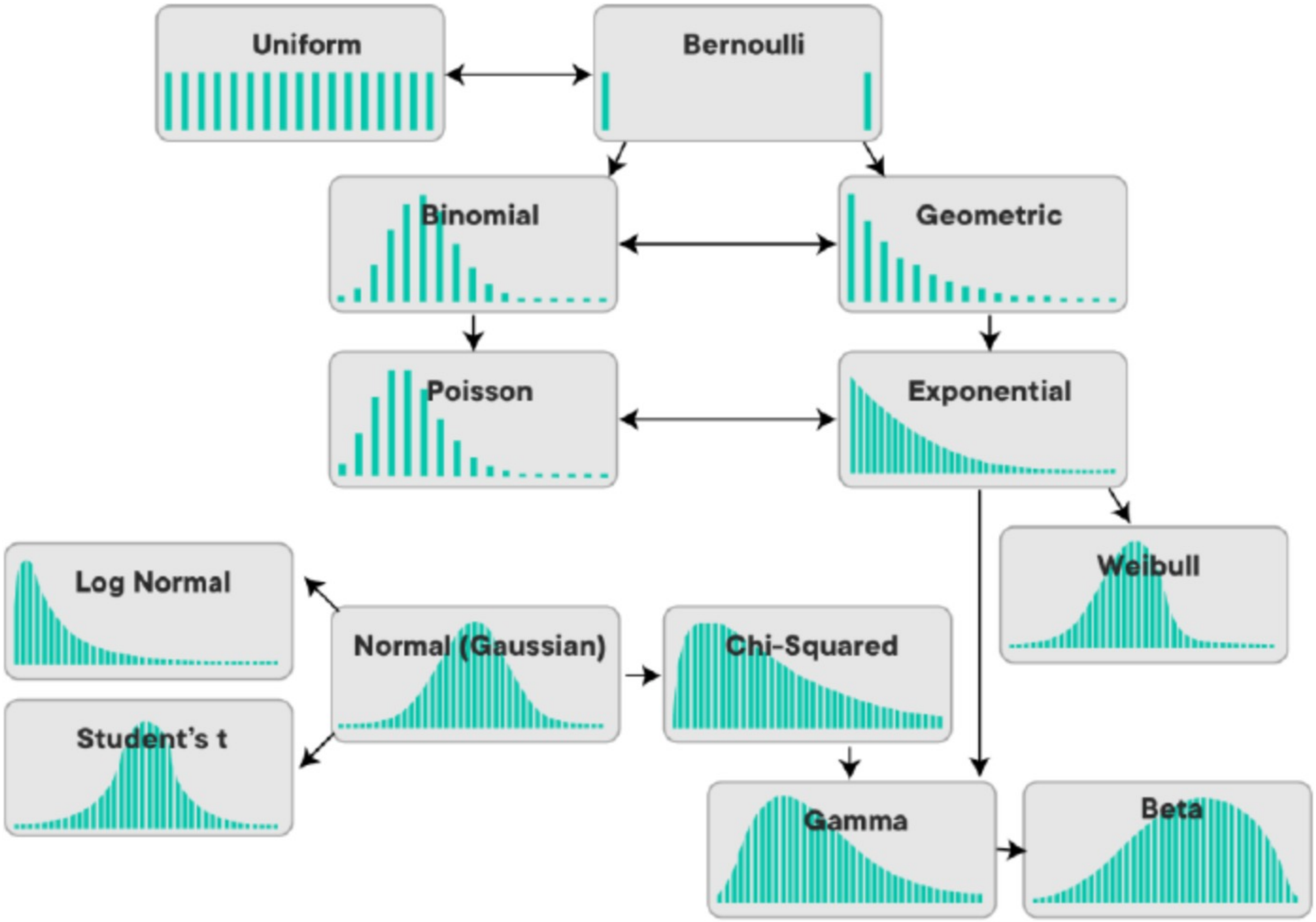
A DRV X transforms a sample space Ω to a more tangible sample space $\widetilde{\Omega}$, whose events are more directly related to what you are interested in.

Then we must determine the probability distribution of X,

i.e., we describe how the probability mass is distributed over possible values of X.

There are numerous probability distributions to be familiar with.

We will discuss a few in more depth.



Once a DRV X is introduced, the sample space Ω is no longer needed.

We simply list the possible values of X and their corresponding probabilities.

This information is contained in the Probability Mass Function, or PMF.

The probability mass function p of a discrete random variable X is the function $p: \mathbb{R} \to [0, 1]$, defined by $p(a) = P(X = a)$ for $-\infty < a < \infty$.

If X is a DRV that takes on the values $a_1, a_2, \ldots$ then:

$p(a_i) > 0$ and

$p(a_1) + p(a_2) + \cdots = 1$ and

for all other a, we have $p(a) = 0$.

Which brings us to the Cumulative Distribution Function (or CDF) of X.

$$F\colon \mathbb{R} \to [0,1], \text{ defined by } F(a) = P(X \le a), \text{ for } -\infty < a < \infty.$$

Compare to the Probability Mass Function of X.

$$p\colon \mathbb{R} \to [0,1], \text{ defined by } p(a) = P(X = a), \text{ for } -\infty < a < \infty.$$

The probability distribution of X is determined by the PMF or the CDF.

Three properties of the CDF of X:

(The cumulative distribution function of the discrete random variable)

1] $F(a) \le F(b)$. This is because $a \le b$ implies that the event $\{X \le a\}$ is contained in the event $\{X \le b\}$.

2] $F(a)$ is a probability. Therefore, $0 \le F(a) \le 1$, and

$$\lim_{a\to+\infty} F(a) = \lim_{a\to+\infty} P(X \le a) = 1$$

$$\lim_{a\to-\infty} F(a) = \lim_{a\to-\infty} P(X \le a) = 0$$

3] F is right-continuous:

$$\lim_{\varepsilon\downarrow 0} F(a+\varepsilon) = F(a)$$

There are many discrete random variables that arise in a natural way.

Let’s now look at some of their distributions.

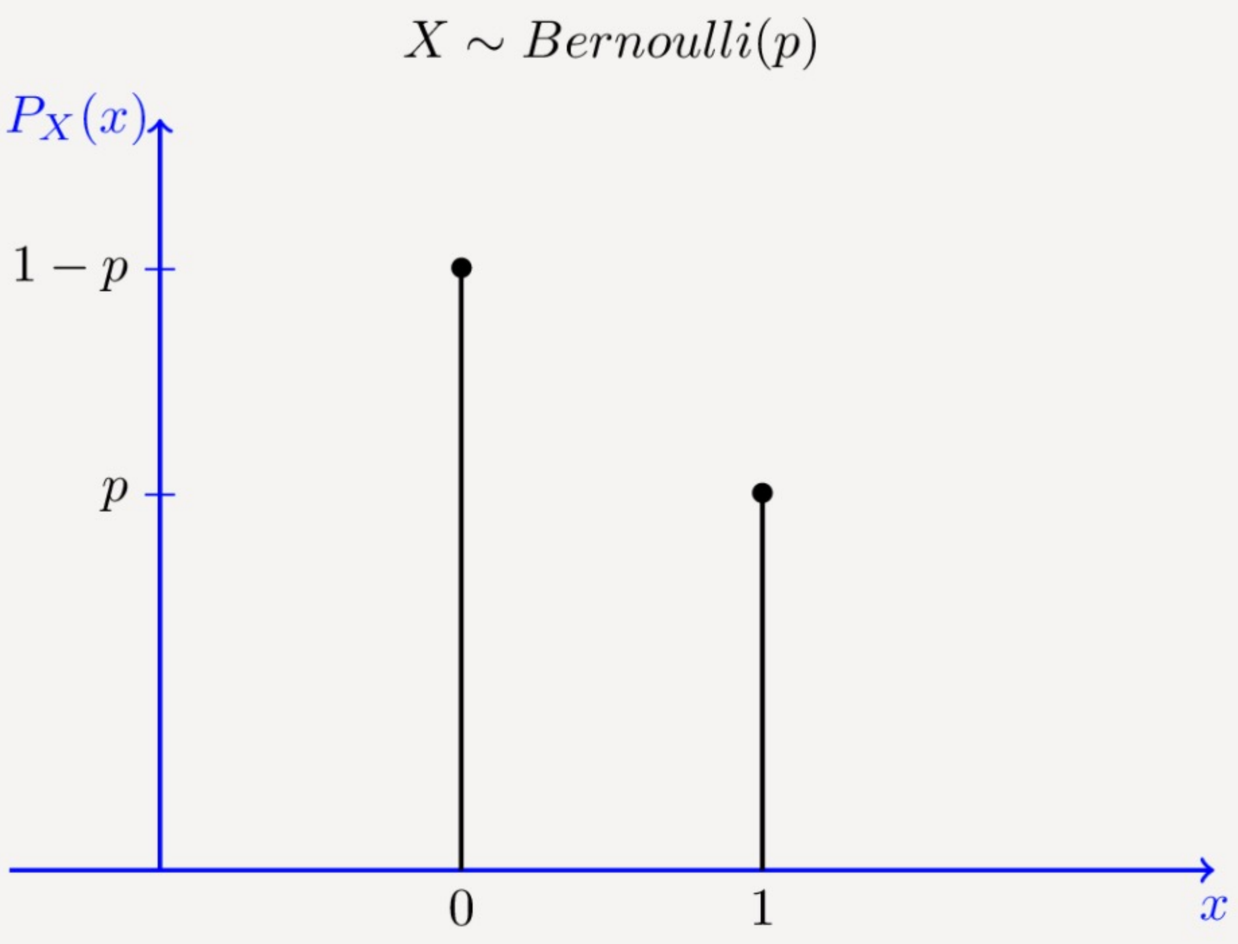
The Bernoulli distribution

- Consider an experiment with only *two possible outcomes:*

  "Success" and "Failure",

  typically encoded as *1* and *0*

- The experiment is modeled with the Bernoulli distribution.
- Denoted $\boldsymbol{Ber(p)}$



The Bernoulli distribution

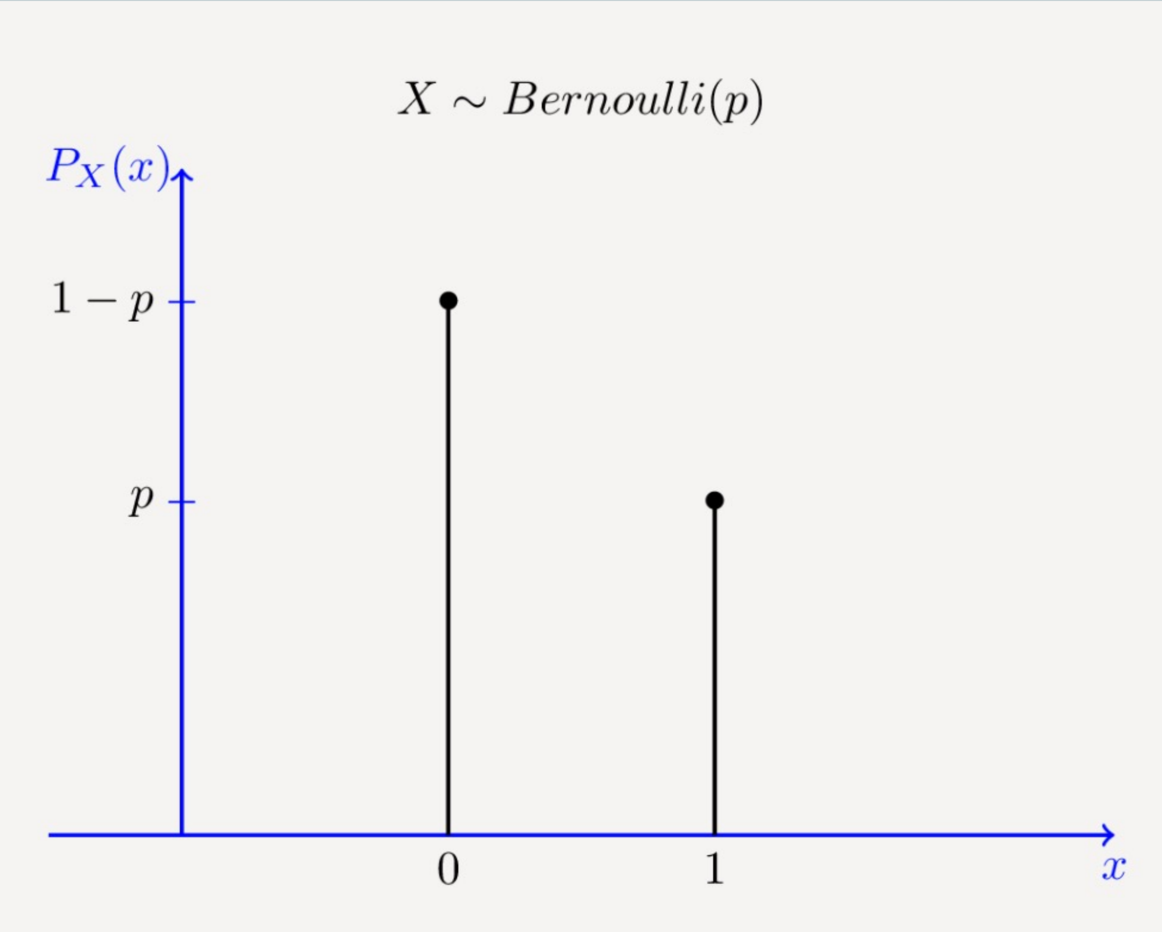
A discrete random variable X has a Bernoulli distribution if its probability mass function is given by:

$p_X(1) = P(X = 1) = p$ and

$p_X(0) = P(X = 0) = 1 - p$

with parameter p, where $0 \leq p \leq 1$.

Each Bernoulli distribution is different, depending upon its parameter, p.



p_X instead of p for the probability mass function of X.
To emphasize its dependence on X and to avoid possible confusion with parameter p of the Bernoulli distribution.

Experiment: Rolling a single die

If you roll a multiple of 3, then you win, otherwise you lose.

As there are only two options (win/lose aka success/failure) this experiment can be modeled with a Bernoulli distribution.

We'll say a multiple of 3 is success, or *1*. Any other die outcome is failure, or *0*.

What is the DRV? Call it T.

What is Ω ?

What are the outcomes in the event $\{T = 1\}$?

What is the PMF of T ?

What is the CDF of T ?

Find the number of outcomes, the actual outcomes, and their frequencies.

$$T\colon \Omega \to \{0,1\}$$

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$T(f) = \begin{cases} 1, & f = 3n \\ 0, & f \neq 3n \end{cases}$$

$$\widetilde{\Omega} = \{0, 1\}$$

1	2	3	4	5	6
0	0	1	0	0	1

outcome	Count
0	4
1	2

$$\{T = 1\} = \{3, 6\}$$

•

1	2	3	4	5	6
0	0	1	0	0	1

outcome	Count
0	4
1	2

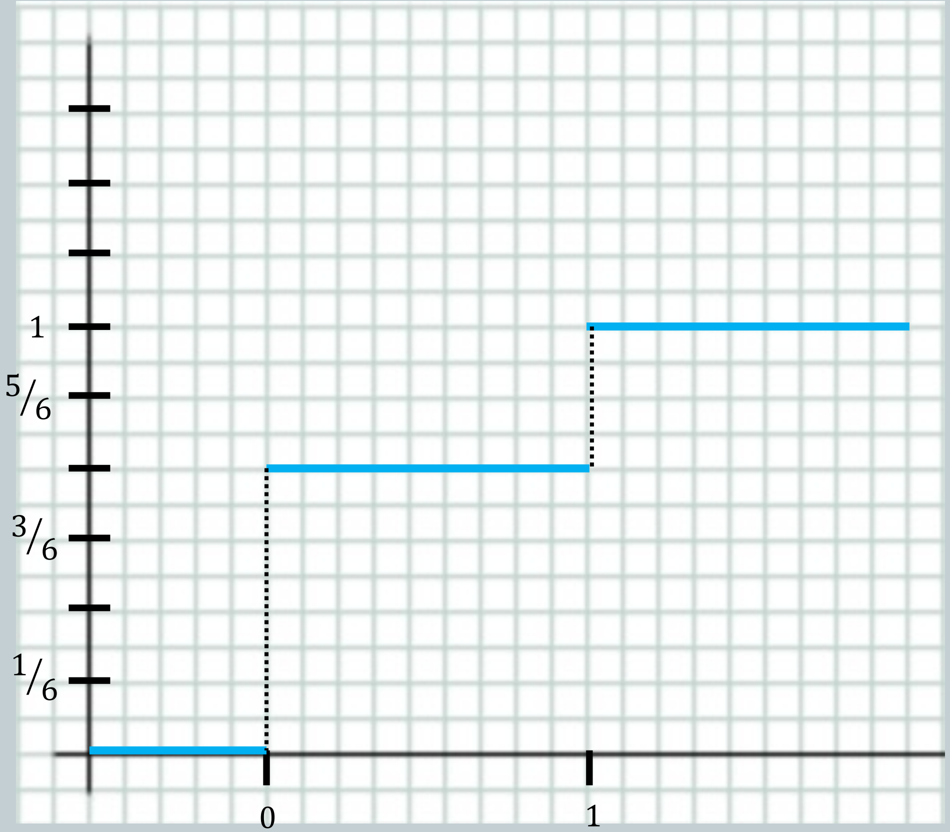
PMF

This is the Bernoulli distribution for:

$$p = \frac{1}{3}$$

or $$Ber\left(\frac{1}{3}\right)$$

CDF



Now consider an experiment with n repeated Bernoulli trials.

The discrete random variable X is then said to have a

Binomial Distribution

(with parameters n and p)

where n = 1, 2, 3, …

and $0 \leq p \leq 1$,

if its probability mass function is given by:

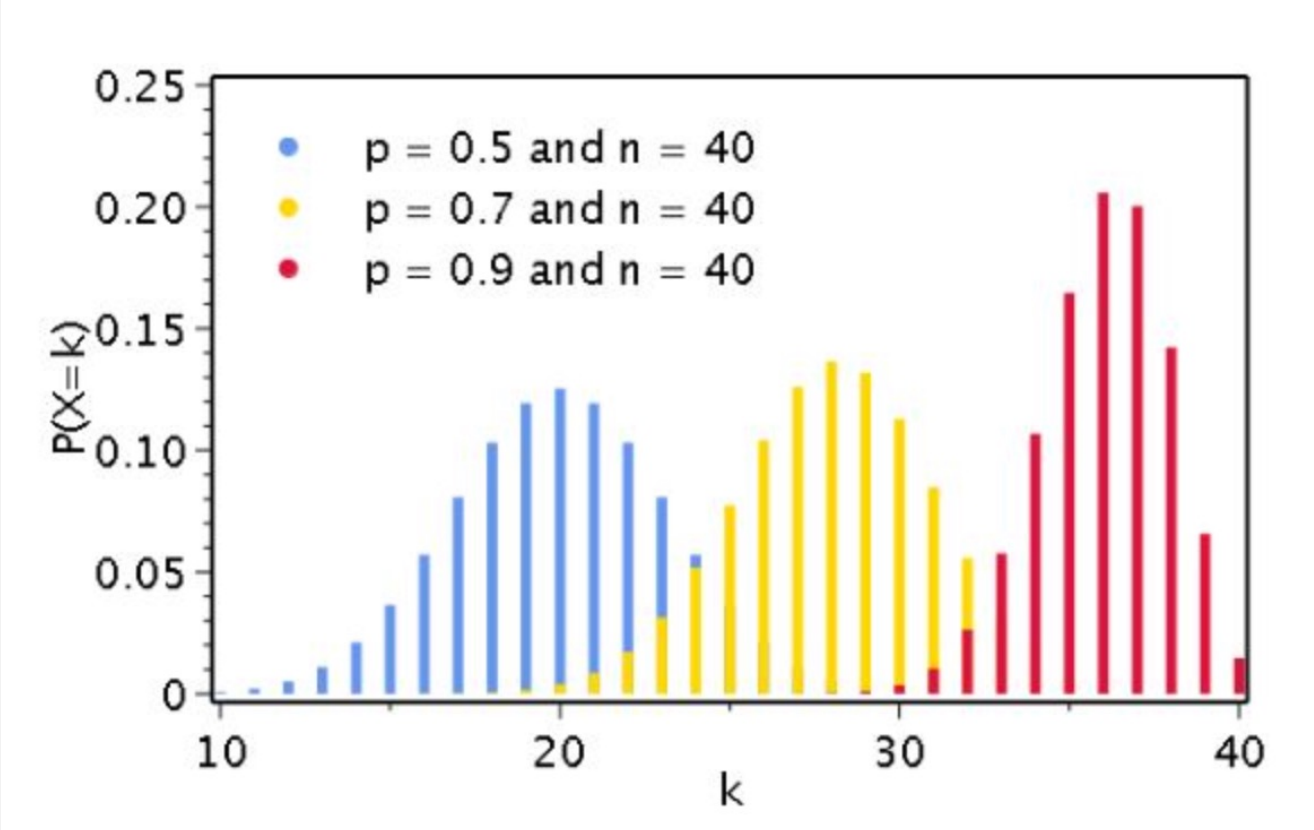
$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \dots, n$

- This is denoted $Bin(n, p)$.

The Binomial Distribution

(for 3 different probabilities)



Experiment: randomly picking 1 of 4 possible choices on a 5-question multiple choice quiz.

You need to get at least 3 of 5 correct to pass $\left(\frac{3}{5} = 60\%\right)$.

What is the probability, <u>with random guesses</u>, that you pass the quiz?

As there are multiple questions (5 total) with only two options on each individual question (success/failure) with the probability $\frac{1}{4}$ of successfully picking correctly, this experiment can be modeled with a Binomial distribution.

What is the DRV? Call it M.

What is Ω ?

What are the outcomes in the event $\{M = 3\}$?

What is the PMF of M ?

What is the CDF of M ?

There are five questions on the quiz: Q_1, Q_2, Q_3, Q_4, Q_5

$$Q_i = \begin{cases} 1, & if\ Q_i\ is\ correct \\ 0, & if\ Q_i\ is\ incorrect \end{cases}$$

$\Omega = \{\text{All the ways of choosing } 0, 1, 2, 3, 4, \text{or } 5\ of\ 5 \text{ questions correctly}\}$

The DRV is given by $M = Q_1 + Q_2 + Q_3 + Q_4 + Q_5$

and $M\colon \Omega \to \{0,1,2,3,4,5\}$ therefore $\widetilde{\Omega} = \{0,1,2,3,4,5\}$

$\{M = 3\} = \{The\ 10\ ways\ of\ choosing\ 3\ of\ 5\ questions\ correctly\}$

$M(The\ C_{5,k}\ ways) = k$

$$\boxed{\begin{aligned} C_{5,0} &= 1 \\ C_{5,1} &= 5 \\ C_{5,2} &= 10 \\ C_{5,3} &= 10 \\ C_{5,4} &= 5 \\ C_{5,5} &= 1 \end{aligned}}$$

$p(k) = P(M = k) = \left(\frac{1}{4}\right)^k \cdot \left(\frac{3}{4}\right)^{5-k} \cdot C_{5,k}$ Why?

$p(3) = P(M = 3) = \left(\frac{1}{4}\right)^3 \cdot \left(\frac{3}{4}\right)^2 \cdot C_{5,3} \approx 8.8\%$ of getting 3 questions correct.

$$\boldsymbol{p(0)} = P(M = 0) = P(Not\ a\ single\ Q_i = 1)$$

$$= P(Q_1 = 0, Q_2 = 0, Q_3 = 0, Q_4 = 0, Q_5 = 0,)$$

$$= P(Q_1 = 0) \cdot P(Q_2 = 0) \cdot P(Q_3 = 0) \cdot P(Q_4 = 0) \cdot P(Q_5 = 0)$$

$$= \boxed{\left(\frac{\mathbf{3}}{\mathbf{4}}\right)^{\mathbf{5}}}$$

$$\boxed{\begin{array}{c} C_{5,0} = 1 \\ C_{5,1} = \mathbf{5} \\ C_{5,2} = 10 \\ C_{5,3} = 10 \\ C_{5,4} = 5 \\ C_{5,5} = 1 \end{array}}$$

$$\boldsymbol{p(1)} = P(M = 1) = \boxed{\frac{1}{4} \cdot \left(\frac{3}{4}\right)^{4} \cdot 5}$$

$$P(M = 1) =$$

$$P(Q_1 = 1) \cdot P(Q_2 = 0) P(Q_3 = 0) \cdot P(Q_4 = 0) \cdot P(Q_5 = 0) \qquad 1$$

$$+\ P(Q_1 = 0) \cdot P(Q_2 = 1) P(Q_3 = 0) \cdot P(Q_4 = 0) \cdot P(Q_5 = 0) \qquad 2$$

$$+\ P(Q_1 = 0) \cdot P(Q_2 = 0) P(Q_3 = 1) \cdot P(Q_4 = 0) \cdot P(Q_5 = 0) \qquad 3$$

$$+\ P(Q_1 = 0) \cdot P(Q_2 = 0) P(Q_3 = 0) \cdot P(Q_4 = 1) \cdot P(Q_5 = 0) \qquad 4$$

$$+\ P(Q_1 = 0) \cdot P(Q_2 = 0) P(Q_3 = 0) \cdot P(Q_4 = 0) \cdot P(Q_5 = 1) \qquad \text{5 ways}$$

In general, for this example of 5 questions with 4 choices:

$$p(k) = P(M = k) = \underline{\left(\frac{1}{4}\right)^{k}} \cdot \underline{\left(\frac{3}{4}\right)^{5-k}} \cdot \underline{C_{5,k}}$$

This is the probability that k questions were answered correctly,
times the probability that the other $5 - k$ answer are wrong,
times the number of ways this can occur.

In this example we noticed that:

$$\boxed{\begin{gathered} C_{5,0} = 1 \\ C_{5,1} = 5 \\ C_{5,2} = 10 \\ C_{5,3} = 10 \\ C_{5,4} = 5 \\ C_{5,5} = 1 \end{gathered}}$$

Is this symmetry always present?

i.e. is it always true that $\binom{n}{n-k} = \binom{n}{k}$?

What is the PMF and CDF for $Bin\left(5, \frac{1}{4}\right)$?

$$p(0) = \binom{5}{0}\left(\frac{1}{4}\right)^0\left(\frac{3}{4}\right)^5 \approx 0.237$$

$$p(1) = \binom{5}{1}\left(\frac{1}{4}\right)^1\left(\frac{3}{4}\right)^4 \approx 0.396$$

$$p(2) = \binom{5}{2}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^3 \approx 0.264$$

$$p(3) = \binom{5}{3}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2 \approx 0.088$$

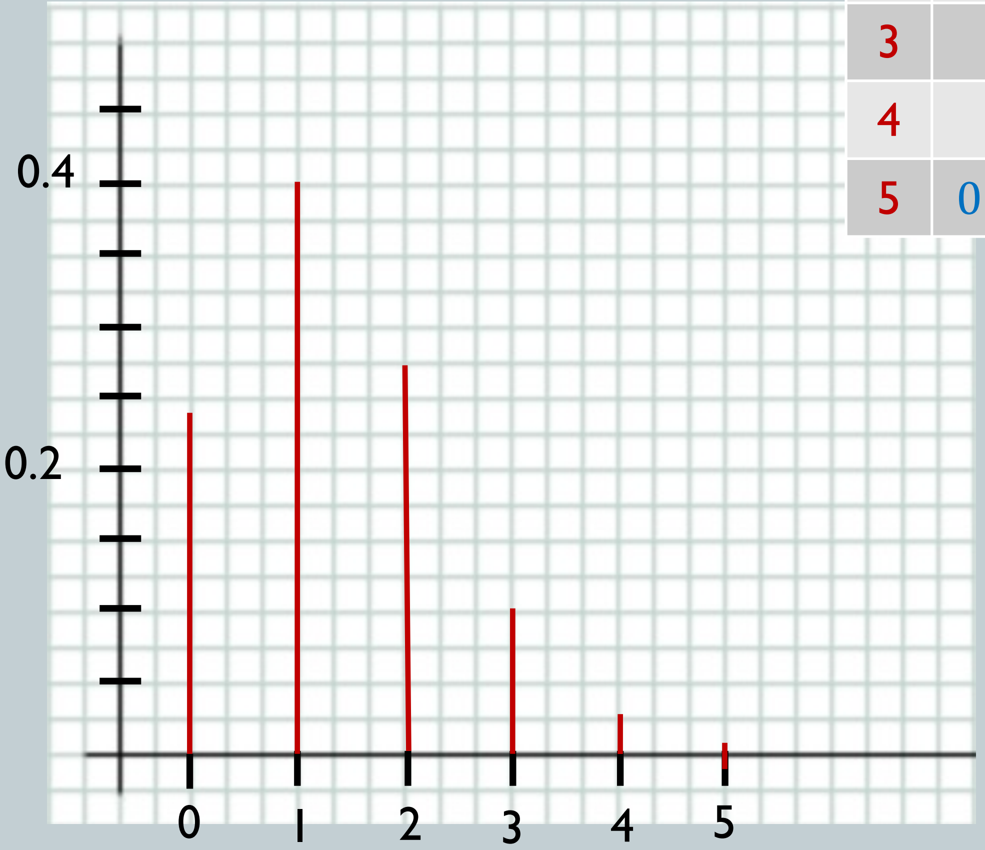
$$p(4) = \binom{5}{4}\left(\frac{1}{4}\right)^4\left(\frac{3}{4}\right)^1 \approx 0.015$$

$$p(5) = \binom{5}{5}\left(\frac{1}{4}\right)^5\left(\frac{3}{4}\right)^0 \approx 0.00098$$

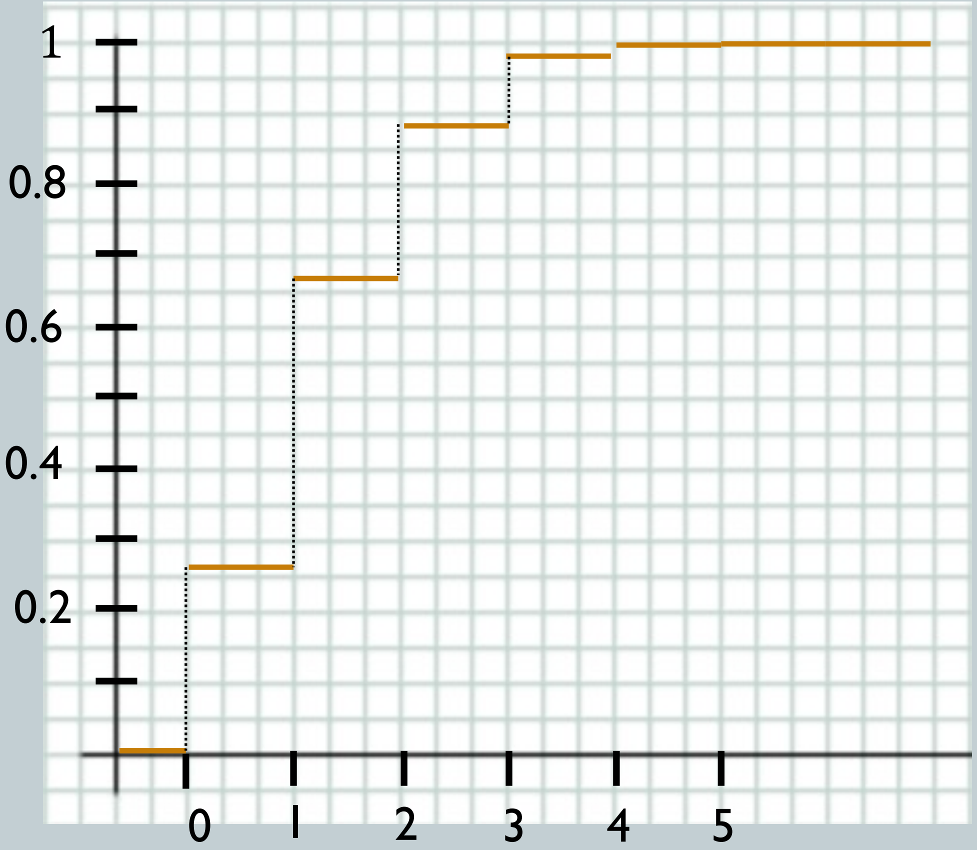
PMF and CDF for a Binomial distribution with parameters $n = 5$ and $p = \frac{1}{4}$

	PDF	CDF
0	0.237	0.237
1	0.396	0.633
2	0.264	0.897
3	0.088	0.985
4	0.015	.999
5	0.00098	1

PMF for $Bin\left(5, \frac{1}{4}\right)$



CDF for $Bin\left(5, \frac{1}{4}\right)$



We went over the Bernoulli and the Binomial, now we introduce the Geometric distribution.

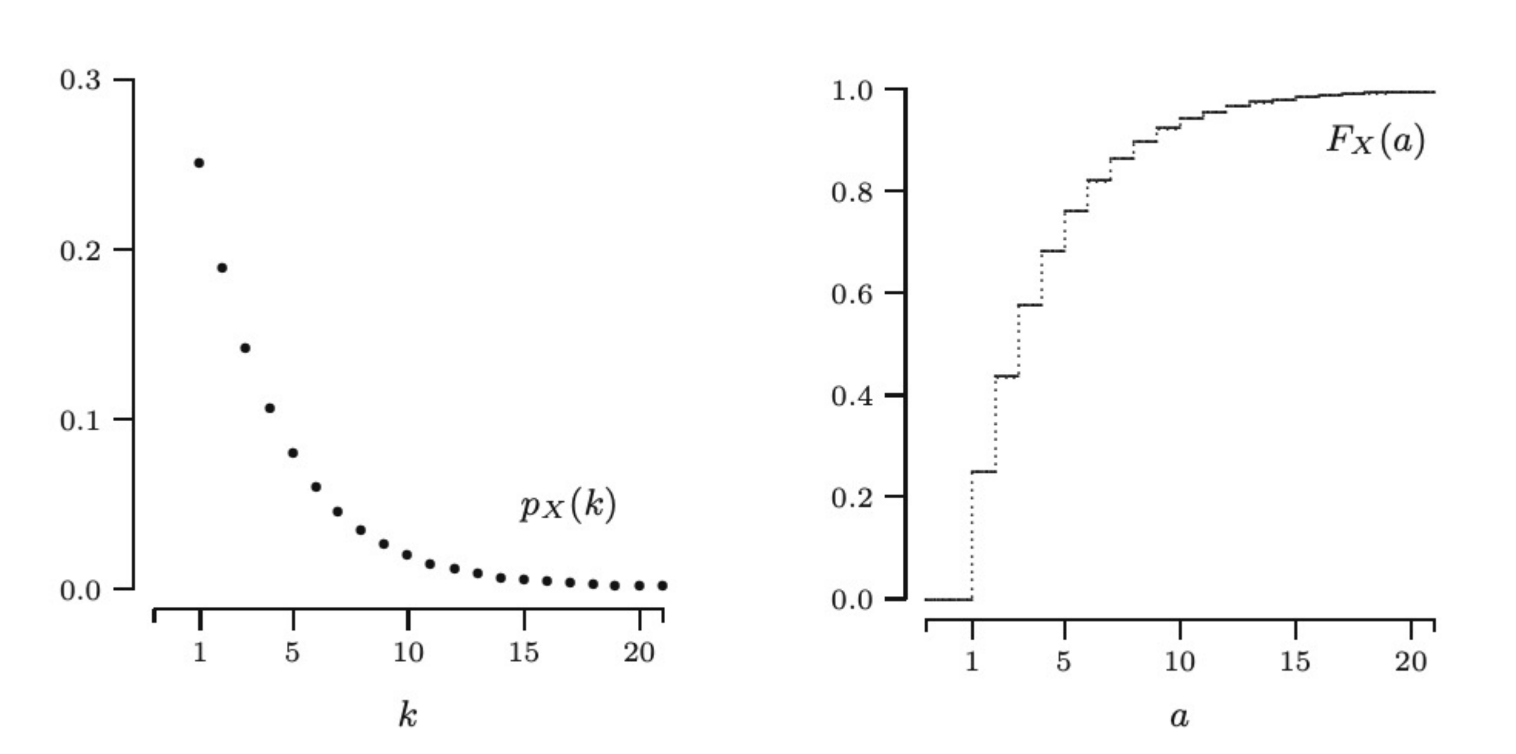
A discrete random variable X has a geometric distribution with parameter p, where $0 < p \leq 1$, if its probability mass function is given by:

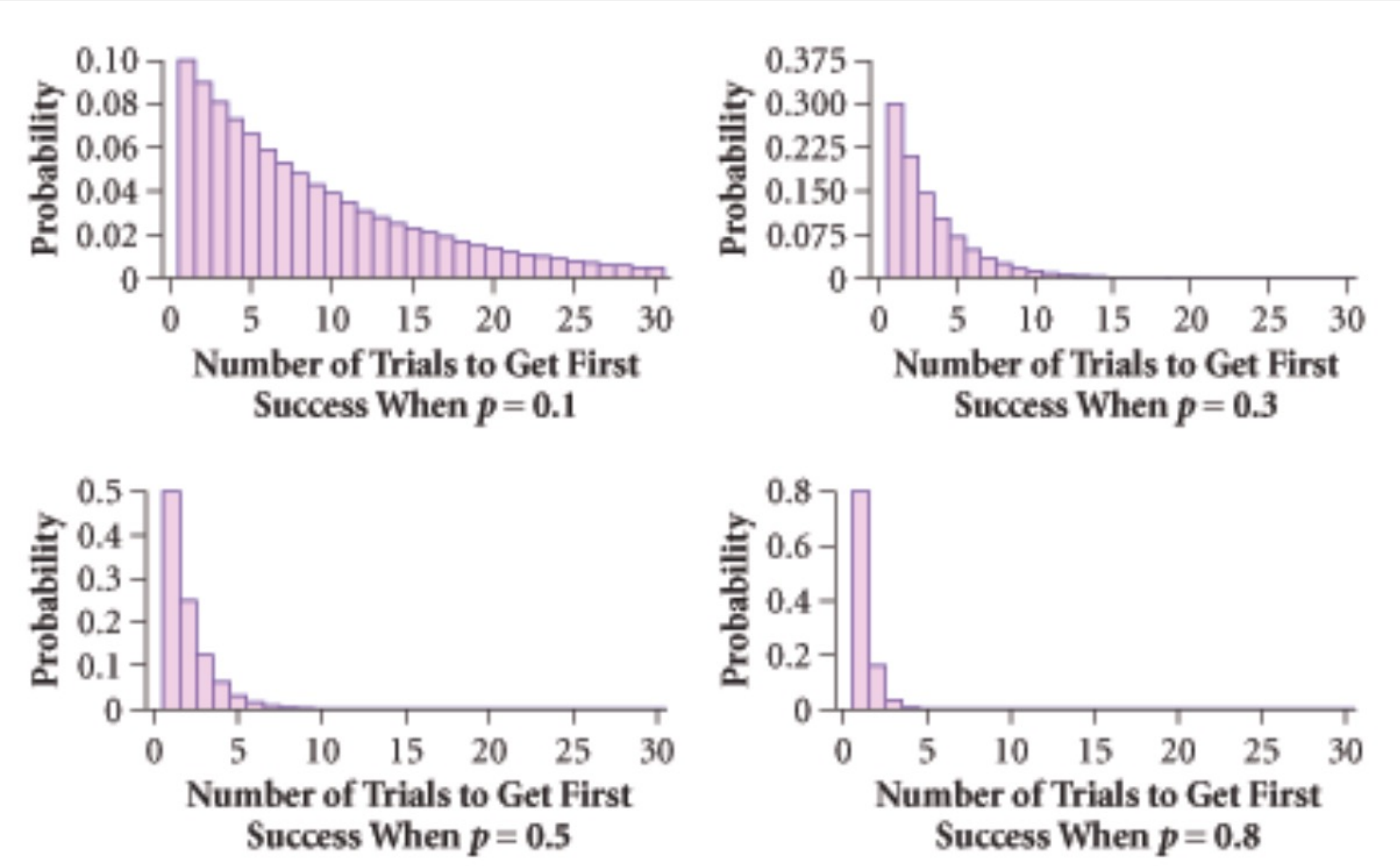
$$p_X(k) = P(X = k) = (1-p)^{k-1}p$$

for $k = 1, 2, \ldots$

This is denoted $\boldsymbol{Geo(p)}$

PMF and CDF for $\boldsymbol{Geo\left(\frac{1}{4}\right)}$





Assumptions:

- Each trial is independent
- Each trial is a Bernoulli random variable with probability of success p.

Suppose you have a coin that shows heads with probability p.

- So, if the coin is normal,

  then $p = 0.5$

- Or it could also be biased,

  $p = 0.3$ or maybe $p = 0.8$

For a given coin with probability p to show heads, how many times do you flip the coin before you first see heads?

$$p_X(k) = P(X = k) = (1-p)^{k-1}p$$

For a given coin with probability p to show heads, how many times do you flip the coin before you first see heads?

$$p(H) = p$$

'Heads' in 1 flip: p

'Heads' in 2 flips: $(1 - p) \cdot p$

'Heads in 3 flips: $(1 - p)^2 \cdot p$

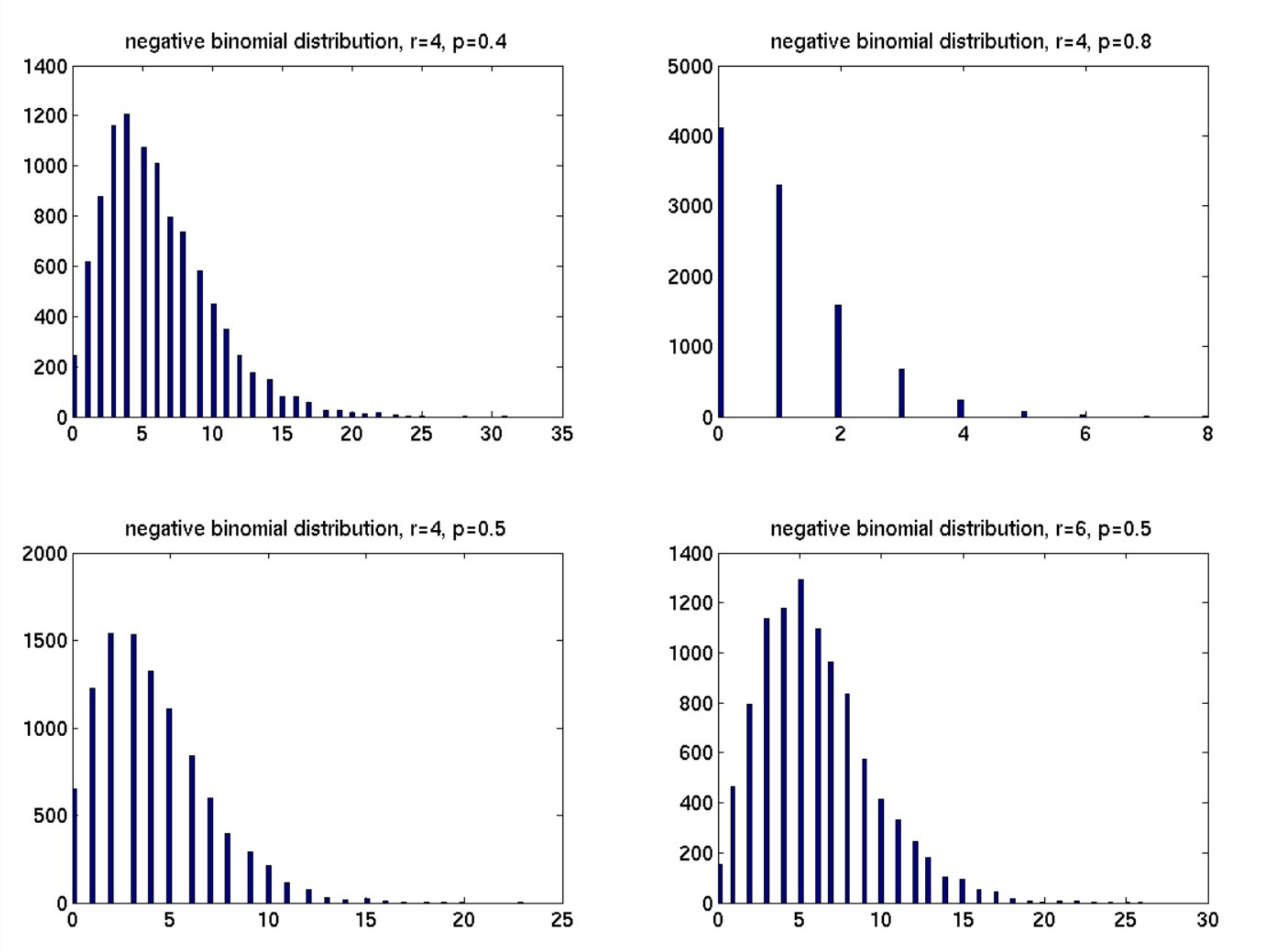
⋮

'Heads' in k flips: $p_x(k) = (1 - p)^{k-1} \cdot p$

The probability that you'll see 'heads' on the k^{th} flip.

Negative Binomial Distribution

The negative binomial distribution models the number of successes in a sequence of independent and identically distributed Bernoulli trials before a specified number of failures (r) occurs.



- A discrete random variable has a **negative binomial distribution** with parameters r and p where

$$r > 1 \text{ and } 0 \le p \le 1$$

if its probability mass function is:

$$p_X(k) = P(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot (1-p)^{k-r}$$

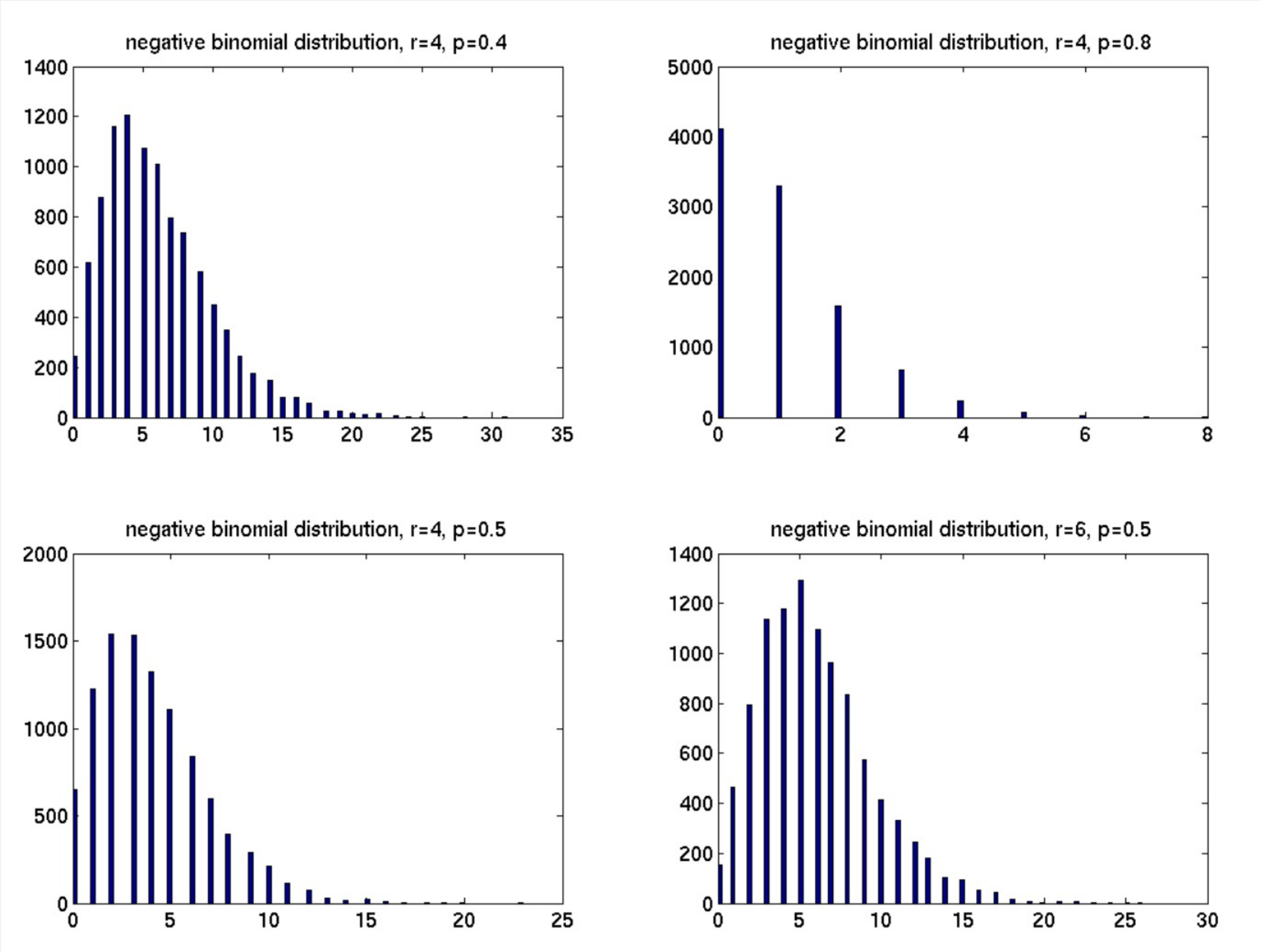
$p =$ probability of success for each trial

$r =$ number of successes we want to observe

$X =$ number of trials needed before we observe r successes.

Recall the Binomial Distribution:

$$p_X(k) = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$



"Success" and "failure" are arbitrary terms that are sometimes swapped. We could just as easily say that the negative binomial distribution is the distribution of the number of successes before r failures.

When applied to real-world problems, outcomes of success and failure may or may not be outcomes we ordinarily view as good and bad, respectively.

Experiment: Roll a die until a player 'loses'.

A loss is defined as the third time you roll a 1.

How many successful roles will we see before we see the third 1 get rolled ?

Therefore, we have $r = 3$ (i.e., the third failure).

Experiment: Roll a die until you 'lose'.

A loss is defined as the third time you roll a 1.

How many successful roles will we see before we see the third 1.

In this case, the probability distribution of the number of non-1s that appear will be a negative binomial distribution.

$$r = 3 \text{ and } p = \frac{1}{6}$$

$$\boldsymbol{p_X(k) = P(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot (1-p)^{k-r}}$$

Although no could possibly know the exact number of rolls until the game is over, we most certainly can determine the probability of various amounts of rolls until three 1's are rolled: $P(X = 3)$ or $P(X = 4)$ or $P(X = 5)$ … etc.

For instance, $P(X = 10) = \binom{9}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7 \approx 0.0465 \approx \binom{9}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^7 \cdot \left(\frac{1}{6}\right)$

$r = 3$ and $p = \frac{1}{6}$

$$\boldsymbol{p_X(k) = P(X = k) = \binom{k-1}{r-1} \cdot p^r \cdot (1-p)^{k-r}}$$

Although no could possibly know the exact number of rolls until the game is over, we most certainly can determine the probability of various amounts of rolls until there 1's are rolled: $P(X = 3)$ or $P(X = 4)$ or $P(X = 5)$ … etc.

For instance:

$$P(X = 3) = \binom{2}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^0 \approx 0.0046 \approx \binom{2}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^0 \cdot \left(\frac{1}{6}\right)$$

$$P(X = 5) = \binom{4}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 \approx 0.0193 \approx \binom{4}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)$$

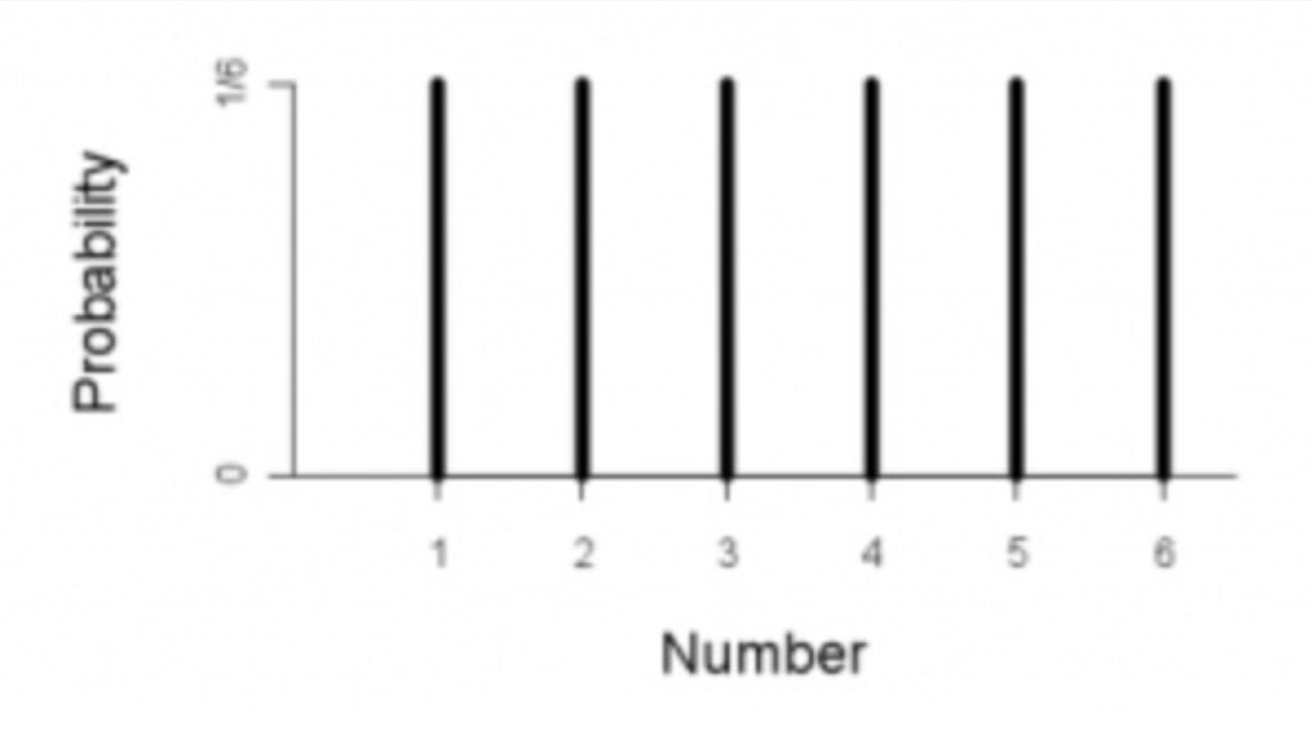
$$P(X = 10) = \binom{9}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^7 \approx 0.0465 \approx \binom{9}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^7 \cdot \left(\frac{1}{6}\right)$$

$$P(X = 20) = \binom{19}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^{17} \approx 0.0357 \approx \binom{19}{2}\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^{17} \cdot \left(\frac{1}{6}\right)$$

The Discrete Uniform Distribution

- A discrete random variable X has a uniform distribution with parameters a and b where $n = b - a + 1$ if $p_X(k) = \frac{1}{n}$ for $k = a, a+1, a+2, \dots, b.$

- e.g., The distribution of a fair 6-sided die is shown.



Poisson Distribution

The Poisson (Pwas - ah) Distribution expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate (μ) and independently of the time since the last event.

Examples that are modeled with a Poisson Distribution:

- The number of pieces of mail received in a day
- The number of people arriving at a restaurant during a shift
- The number of phone calls per hour received by a call center
- The number of meteors striking the earth each year

A discrete random variable has a

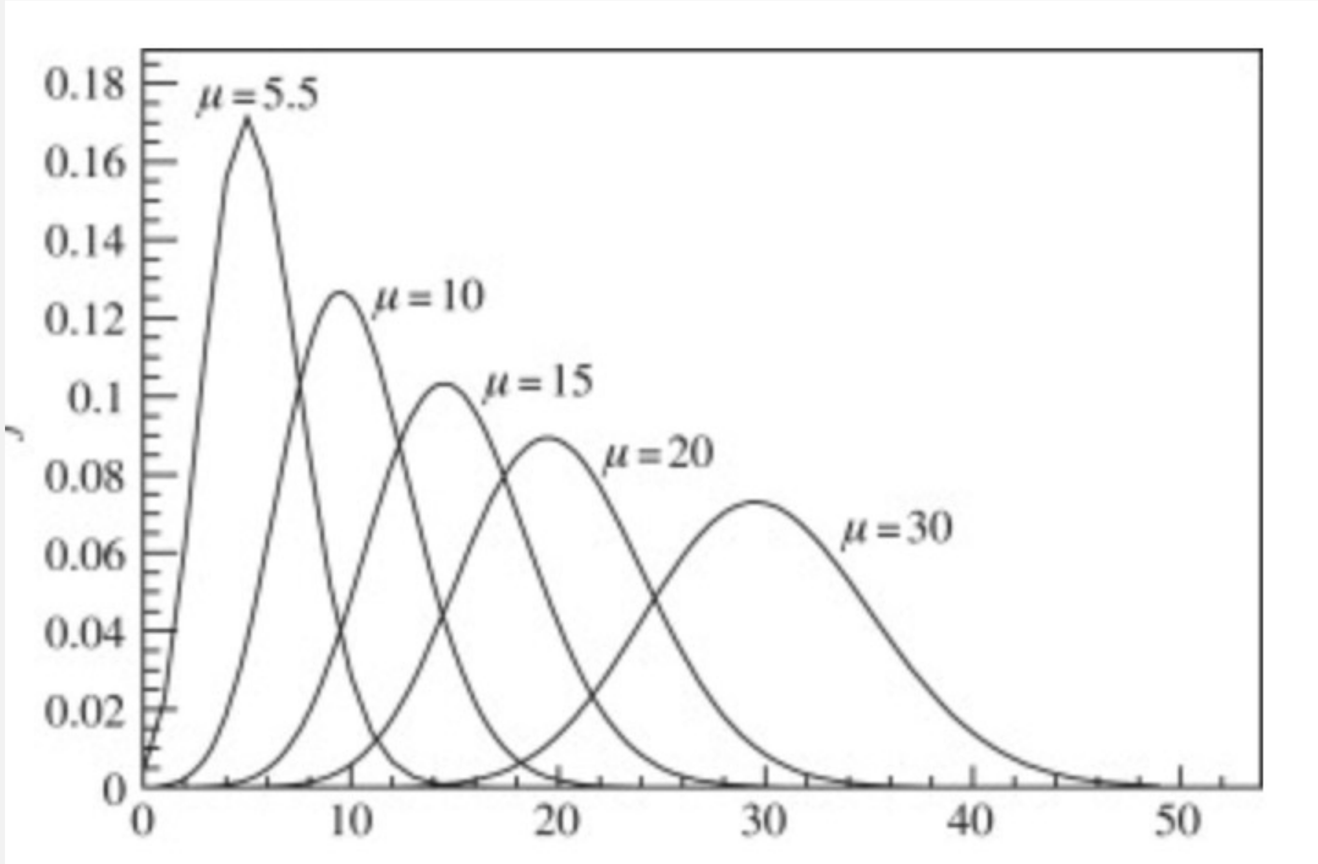
Poisson Distribution

with parameter μ, where $\mu > 0$, if its probability mass function is given by:

$$p_X(k) = P(X = k) = \frac{\mu^k e^{-\mu}}{k!}$$

for $k = 0, 1, 2, \ldots$

- By Poisson processes, we mean processes that are discrete, independent, and mutually exclusive



What does the PMF $\boldsymbol{p_X(k) = P(X = k) = \frac{\mu^k e^{-\mu}}{k!}}$ mean?

$\boldsymbol{\mu}$ is the expected number of events per... (time period, length, space, volume,...)

$$\boldsymbol{\mu} = \frac{\text{\# of events}}{\text{time}} \cdot \text{time period}$$

$$P(k \text{ events in a time period}) = \frac{\left(\frac{\text{\# of events}}{\text{time}} \cdot \text{time period}\right)^k \cdot e^{-\frac{\text{\# of events}}{\text{time}} \cdot \text{time period}}}{k!}$$

Experiment: A call center receives an average of 4320 calls per day. The calls are independent; receiving one does not change the probability of when the next one will arrive.

So, $\text{rate} = \dfrac{\text{events}}{\text{time}} = \dfrac{4320 \text{ calls}}{24 \text{ hours}} = \dfrac{3 \text{ calls}}{\text{minute}}$

The number of calls received during any minute has a Poisson probability distribution: the most likely numbers are 2 and 3 but 1 and 4 are also likely and there is a small probability of it being as low as zero and a very small probability it could be 10.

So $\mu = \dfrac{3 \text{ calls}}{\text{minute}} \cdot 1 \text{ minute}.$ $\quad \boldsymbol{p_X(k) = P(X = k) = \dfrac{\mu^k e^{-\mu}}{k!}}$

$$p_X(1) = P(X = 1) = \frac{3^1 e^{-3}}{1!} \approx 0.14936 \qquad p_X(4) = P(X = 4) = \frac{3^4 e^{-3}}{4!} \approx 0.16803$$

$$p_X(2) = P(X = 2) = \frac{3^2 e^{-3}}{2!} \approx 0.22404 \qquad p_X(5) = P(X = 5) = \frac{3^5 e^{-3}}{5!} \approx 0.10082$$

$$p_X(3) = P(X = 3) = \frac{3^3 e^{-3}}{3!} \approx 0.22404 \qquad p_X(10) = P(X = 10) = \frac{3^{10} e^{-3}}{10!} \approx 0.000810$$

The Poisson is bounded by 0 and ∞. Values become negligible after a point.

The Binomial Distribution tends to the Poisson Distribution as $n \to \infty$ and $p \to 0$, with a constant μ.

The PD approximates the BD if n is large and p is small.

$$\lim_{n\to\infty} (\text{Binomial}) = \text{Poisson}$$

$$\lim_{n\to\infty} \binom{n}{k} p^k (1 - p)^{n-k} = \frac{\mu^k e^{-\mu}}{k!} \qquad \text{for } \mu = np$$

The Binomial Distribution tends to the Poisson Distribution as $n \to \infty$ and $p \to 0$, with a constant μ.
The PD approximates the BD if n is large and p is small.

Example: Suppose a disease affects $\frac{1}{20{,}000}$ and you sample 1000 people.
How likely is it that your sample contains 2 people with the disease?
This is a BD since the person either has the disease, or they do not.

Binomial

$$\binom{1000}{2} \cdot \left(\frac{1}{20{,}000}\right)^2 \cdot \left(1 - \frac{1}{20{,}000}\right)^{998} = 0.001187965 \quad \text{(true value)}$$

Poisson

$$\mu = \frac{1}{20{,}000} \cdot 1000 = 0.05$$

$$p_X(2) = P(X = 2) = \frac{(0.05)^2 e^{-0.05}}{2!} = 0.0011890368 \quad \text{(Approximation)}$$

When should the Poisson be used instead of the Binomial?
There are different opinions. Some say when $n > 50$ and $\mu < 5$.
The Poisson works best when n is really big and p is close to 0.
Whether the estimate is reasonable or not depends on your needs.

Why do we approximate?
1] Some factorials and exponentials are problematic to calculate. There are round-off errors and magnitude will sometimes return no answer. The Poisson does not possess these problems.
2] The binomial distribution requires knowledge of n and p. The Poisson can be calculated with just μ.

Can you recognize <u>which distribution</u> is called for?

Experiment: You are at a football game in the student section. Freshmen make up 25% of the crowd, sophomores 25%, Juniors 25% and Seniors 25%. You are interested in finding seniors for giving them graduation information. You don't know anyone's class status. You stand in a door talking to people as they pass.

Let X be a random variable describing:

*The number of each class in a randomly chosen group of 100 passers-by.

*The number of people you have to talk to in order to speak with exactly 100 seniors.

*The number of seniors that pass by over a 15-minute period.

*The number of seniors you encounter after speaking with 100 people.

*The number of people you have talked to up to and including your first encountered senior.

*Whether or not a randomly chosen person was a senior.

Discrete Uniform Dist., NBD $r = 100, p = .25$, Poisson, $\text{Bin}(100, .25)$, $\text{Geo}(0.25)$, $\text{Ber}(0.25)$

In the next set of slides, we will move on from Discrete Random variables to Continuous Random variables!