Lesson 011
Probability Distributions for Discrete Random Variables

Wednesday, October 4

There is no probabilistic difference between modelling a coin flip and a plane crash.

Probability Distributions

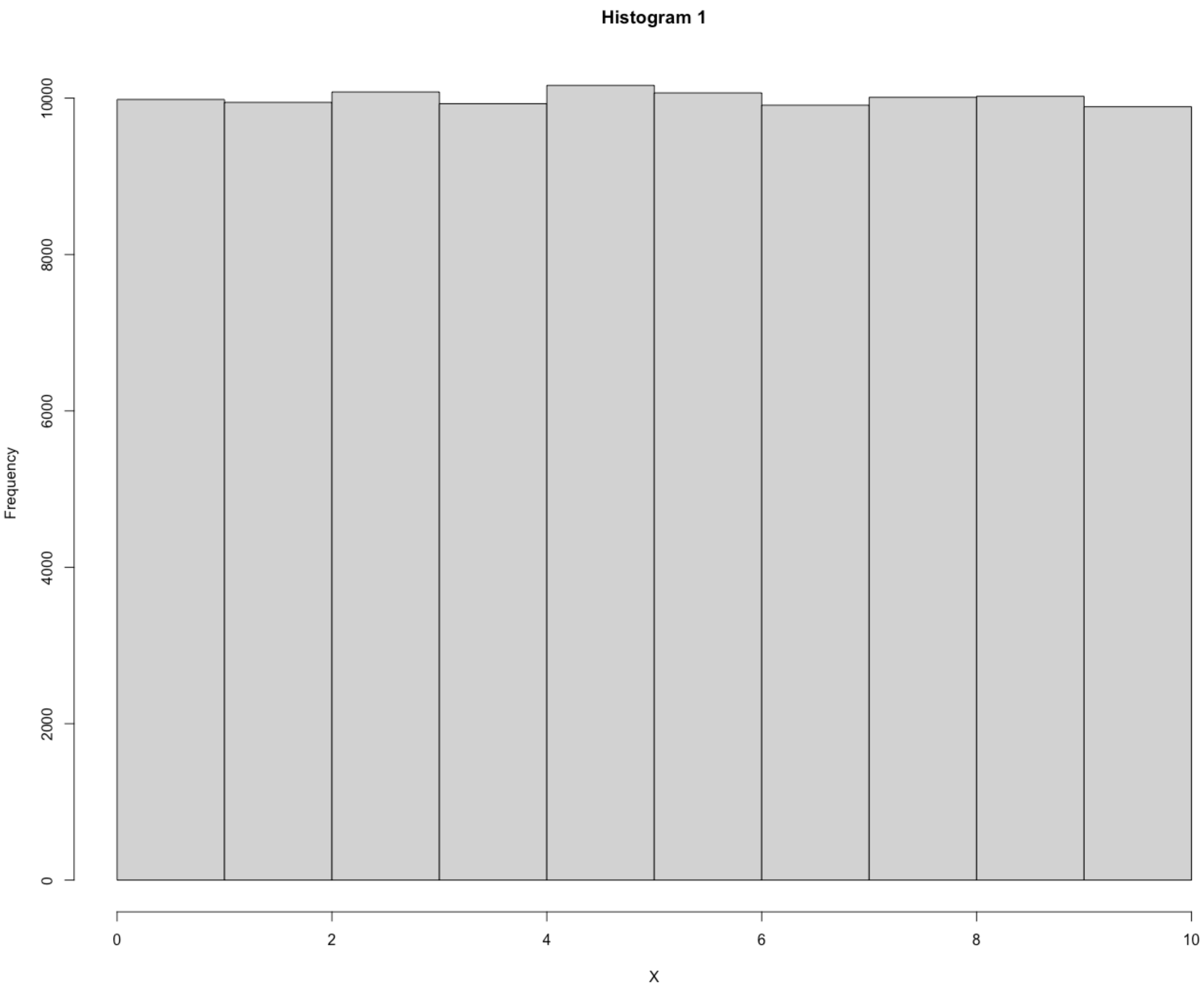
- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.

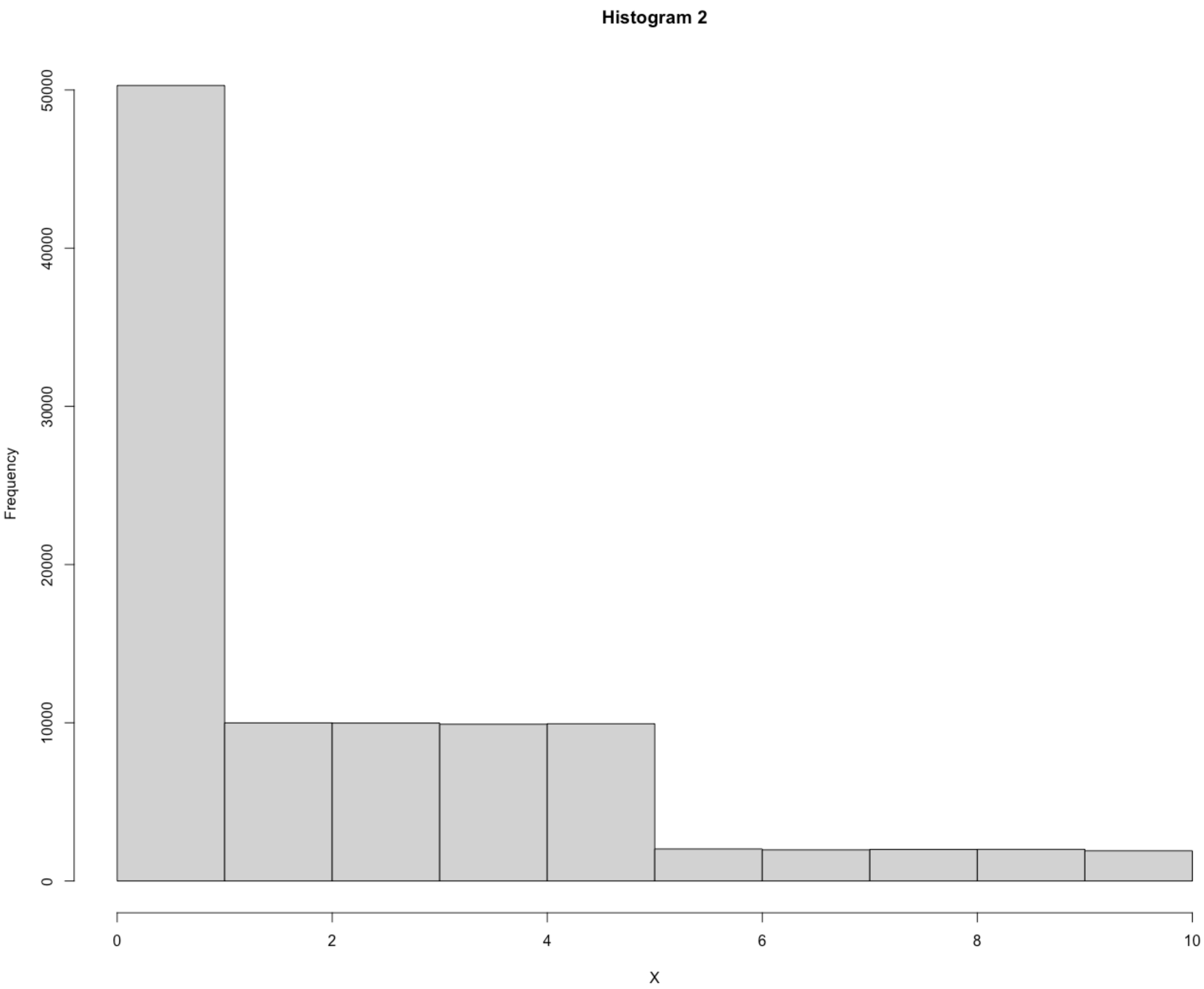
Probability Distributions

- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.
- A probability distribution demonstrates how the total probability is allocated to each outcome.

Probability Distributions

- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.
- A probability distribution demonstrates how the total probability is allocated to each outcome.
- Can think of these being represented, visually, as a histogram.


Histogram 1
Frequency
0
2000
4000
6000
8000
10000
0
2
4
6
8
10
X


Histogram 2
Frequency
0
10000
20000
30000
40000
50000
0
2
4
6
8
10
X

Probability Distributions

- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.
- A probability distribution demonstrates how the total probability is allocated to each outcome.
- Can think of these being represented, visually, as a histogram.
- Normally, we discuss the distribution of a random variable.

Probability Distributions

- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.
- A probability distribution demonstrates how the total probability is allocated to each outcome.
- Can think of these being represented, visually, as a histogram.
- Normally, we discuss the distribution of a random variable.
  - Discrete random variables have discrete probability distributions.

Probability Distributions

- We know that the probability over the whole sample space is one, $P(\mathcal{S}) = 1$.
- A probability distribution demonstrates how the total probability is allocated to each outcome.
- Can think of these being represented, visually, as a histogram.
- Normally, we discuss the distribution of a random variable.
  - Discrete random variables have discrete probability distributions.
  - Discrete probability distributions are characterized by **probability mass functions**.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots$, then the PMF is a function where $p(x_i)$ takes a value.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots,$ then the PMF is a function where $p(x_i)$ takes a value.
  - For the PMF to be valid we require $0 \leq p(x_i) \leq 1$ and $\sum_i p(x_i) = 1$.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots$, then the PMF is a function where $p(x_i)$ takes a value.
  - For the PMF to be valid we require $0 \leq p(x_i) \leq 1$ and $\sum_i p(x_i) = 1$.
- In the long term, we would expect that in $p(x_i)$ proportion of experiments, we observe the event corresponding to $X = x_i$.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots$, then the PMF is a function where $p(x_i)$ takes a value.
  - For the PMF to be valid we require $0 \leq p(x_i) \leq 1$ and $\sum_i p(x_i) = 1$.
- In the long term, we would expect that in $p(x_i)$ proportion of experiments, we observe the event corresponding to $X = x_i$.
- Often, distributions are indexed by parameters controlling the values the function takes on.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots$, then the PMF is a function where $p(x_i)$ takes a value.
  - For the PMF to be valid we require $0 \leq p(x_i) \leq 1$ and $\sum_i p(x_i) = 1$.
- In the long term, we would expect that in $p(x_i)$ proportion of experiments, we observe the event corresponding to $X = x_i$.
- Often, distributions are indexed by parameters controlling the values the function takes on.
  - Any distributions which differ only in the value of the parameter are from the same *family*.

Probability Mass Functions

- A PMF is a mathematical function that assigns a numeric value to each outcome that a random variable X can take.
- Suppose X can take on values $x_1, x_2, \ldots$, then the PMF is a function where $p(x_i)$ takes a value.
  - For the PMF to be valid we require $0 \leq p(x_i) \leq 1$ and $\sum_i p(x_i) = 1$.
- In the long term, we would expect that in $p(x_i)$ proportion of experiments, we observe the event corresponding to $X = x_i$.
- Often, distributions are indexed by parameters controlling the values the function takes on.
  - Any distributions which differ only in the value of the parameter are from the same *family*.
- We will often write $X \sim p(x)$.

Example

$$p(x) = \begin{cases} 0.5 & x = 0 \\ 0.25 & x = 1 \\ 0.1 & x = 2 \\ 0.15 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

Example

$$p(x) = p^x(1-p)^{1-x} \quad x \in \{0,1\}$$

$$p \in (0,1)$$

$$0 \quad \text{otherwise}$$

Consider the following PMF: $p(x) = \begin{cases} 0.5 & x = 0 \\ z & x = 1 \\ 0.125 & x \in \{2, 3\} \\ 0 & \text{otherwise} \end{cases}$. What is the value of z?

$z = 0.5$

0%

$z = 0.375$

0%

$z = 0.25$

0%

$z = 0$

SEE MORE

Consider the following PMF: $p(x) = \begin{cases} 0.5 & x = 0 \\ 0.25 & x = 1 \\ 0.125 & x \in \{2, 3\} \\ 0 & \text{otherwise} \end{cases}$. What is $P(X = \{0\} \cup \{2\})$?

$z = 0.625$

0%

$z = 0.5$

0%

$z = 0.25$

0%

$z = 0.125$

SEE MORE

Bernoulli Distribution

- A random variable which can take on only values 0 or 1 always follows a **Bernoulli distribution**.

Bernoulli Distribution

- A random variable which can take on only values 0 or 1 always follows a **Bernoulli distribution**.
- The PMF of a Bernoulli random variable is given by $p(x) = p^x(1-p)^{1-x}$, for some **parameter** p.

Bernoulli Distribution

- A random variable which can take on only values 0 or 1 always follows a **Bernoulli distribution**.
- The PMF of a Bernoulli random variable is given by $p(x) = p^x(1-p)^{1-x}$, for some **parameter** p.
- We typically refer to a 1 as a success and a 0 as a failure.

- A biased coin is flipped, which turns up heads 70% of the time.

$$X \sim \text{Bern}(0.7)$$

$$p(x) = (0.7)^x (0.3)^{1-x} \quad x \in \{0,1\}$$

- A commercial plane takes flight. Of all commercial flights, 0.000414% end up crashing.

$$X \sim \text{Bern}(0.99999586)$$

$$p(x) = (0.99999586)^x (0.00000414)^{1-x} \quad x \in \{0,1\}$$

Which of the following random variables is represented by a Bernoulli distribution?

Counting the number of hearts (successes) on 5 draws from a deck of cards, with replacement.

0%

Counting the number of hearts (successes) on 5 draws from a deck of cards, without replacement.

0%

An indicator variable as to whether 5 hearts are observed in 5 draws from a deck of cards, without replacement.

0%

All of the above

0%

Cumulative Distribution Functions

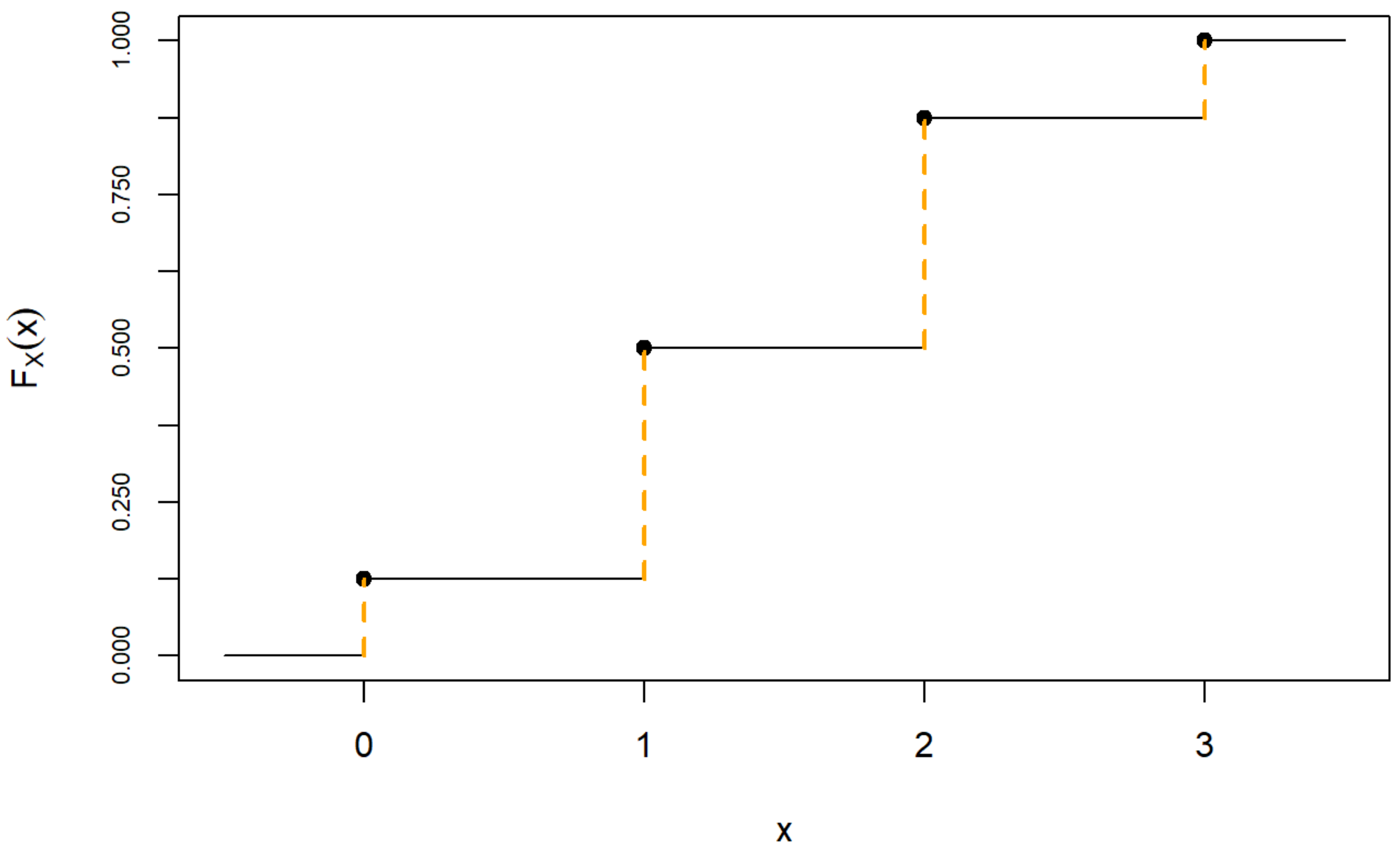
- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.

Cumulative Distribution Functions

- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.
  - $$P(X \leq x) = F_X(x) = \sum_{k=-\infty}^{x} p(k).$$

Cumulative Distribution Functions

- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.
  - $$P(X \leq x) = F_X(x) = \sum_{k=-\infty}^{x} p(k).$$
  - For discrete random variables, the CDF will always be a step function.



1.000
0.750
0.500
0.250
0.000
$F_X(x)$
0
1
2
3
x

Cumulative Distribution Functions

- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.
  - $$P(X \leq x) = F_X(x) = \sum_{k=-\infty}^{x} p(k).$$
  - For discrete random variables, the CDF will always be a step function.
  - We can interchange between a PMF and a CDF.

Cumulative Distribution Functions

- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.
  - $$P(X \leq x) = F_X(x) = \sum_{k=-\infty}^{x} p(k).$$
  - For discrete random variables, the CDF will always be a step function.
  - We can interchange between a PMF and a CDF.
  - To find the probability of falling in an interval,

$$P(a < X \leq b) = F(b) - F(a).$$

Cumulative Distribution Functions

- The **cumulative distribution function** is a related quantity which gives the probability that the random variable takes on a value **less than** some threshold.
  - $$P(X \leq x) = F_X(x) = \sum_{k=-\infty}^{x} p(k).$$
  - For discrete random variables, the CDF will always be a step function.
  - We can interchange between a PMF and a CDF.
  - To find the probability of falling in an interval,
    $$P(a < X \leq b) = F(b) - F(a).$$
- It will often be easier to work with a CDF rather than a PMF.

Suppose that $P(Y=8)=0.40$. Moreover, assume that the CDF of Y is given by

$$\begin{cases} 0 & y<1 \\ 0.05 & 1\leq y<2 \\ 0.15 & 2\leq y<4 \\ 0.5 & 4\leq y<8 \\ z & 8\leq y<16 \\ 1 & 16\leq y \end{cases}$$

. What is the value of z?

$z=0.5$

0%

$z=0.75$

0%

SEE MORE

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.
- We call this distribution the **geometric distribution**, and it is also governed by a parameter p.

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.
- We call this distribution the **geometric distribution**, and it is also governed by a parameter p.
  - The PMF of the distribution is given by $p(x) = (1-p)^{x-1}p$.

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.
- We call this distribution the **geometric distribution**, and it is also governed by a parameter p.
  - The PMF of the distribution is given by $p(x) = (1 - p)^{x-1}p$.
  - X represents the number of trials until the first success is observed.

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.
- We call this distribution the **geometric distribution**, and it is also governed by a parameter p.
  - The PMF of the distribution is given by $p(x) = (1 - p)^{x-1}p$.
  - X represents the number of trials until the first success is observed.
  - Sometimes you will count non-inclusively, giving $p(x) = (1 - p)^{x}p$.

The Geometric Distribution

- What would the distribution be if you wanted to count how many flips of a coin it took until you got a heads?
- This is related to the Bernoulli distribution, but is not exactly the same.
- We call this distribution the **geometric distribution**, and it is also governed by a parameter p.
  - The PMF of the distribution is given by $p(x) = (1-p)^{x-1}p$.
  - X represents the number of trials until the first success is observed.
  - Sometimes you will count non-inclusively, giving $p(x) = (1-p)^{x}p$.
  - If we right $\lfloor x \rfloor$ as the lowest integer less than or equal to x, then $F_X(x) = 1 - (1-p)^{\lfloor x \rfloor}$.

Which of the following random variables is represented by a Geometric distribution?

Counting the number of draws until a heart is seen in a deck of cards, without replacing them each draw.

0%

Counting the number of draws until a heart is seen in a deck of cards, replacing them each draw.

0%

Counting the number of hearts (successes) in a set number of draws with replacement.

0%

Two or more of the above.

0%