STAT 2593

Lecture 011 - Probability Distributions for Discrete Random Variables

Dylan Spicker

Probability Distributions for Discrete Random Variables

Learning Objectives

1. Define the probability mass function and illustrate its use.
2. Understand the properties of a valid PMF.
3. Interpret the PMF, and its parameters.
4. Understand the cumulative distribution function.
5. Understand the Bernoulli and geometric distributions, and their use cases.

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.

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 $X \sim p(x)$.

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.

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.

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.

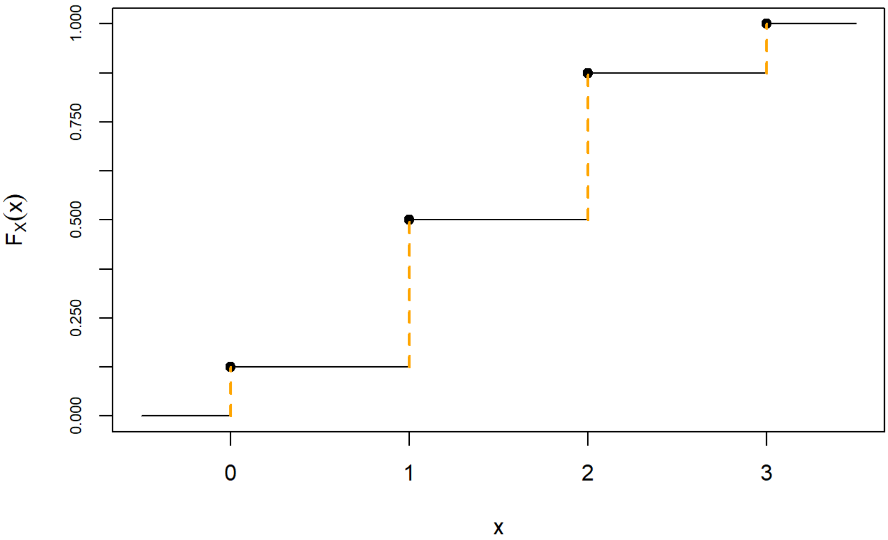
Cumulative Distribution Functions

- 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.



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

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 $F_X(x) = 1 - (1-p)^{\lfloor x \rfloor}$.

Summary

- Distributions summarize the breakdown of probability values.
- Discrete distributions are characterized by a PMF.
- We can also consider the cumulative distribution function, which will be a step function for discrete variables.
- The Bernoulli distribution models a single success-fail experiment.
- The geometric distribution models repeated success-fail experiments until a success is reached.