Analysis of Algorithms

The Mathematics of Analysis of Algorithms

Andres Mendez-Vazquez

September 10, 2015

Outline

Cinvestav

Outline

Cinvestav

Basic Induction

Principle of Mathematical Induction

Let $P(n)$ be a property that is defined for integers n, and let a be a fixed integer.

Basic Induction

Principle of Mathematical Induction

Let $P(n)$ be a property that is defined for integers n, and let a be a fixed integer.

Suppose the following two statements are true

1. $P(a)$ is true.
2. For all integers $k \geq a$, if $P(k)$ is true then $P(k+1)$ is true.

Basic Induction

Principle of Mathematical Induction

Let $P(n)$ be a property that is defined for integers n, and let a be a fixed integer.

Suppose the following two statements are true

1. $P(a)$ is true.
2. For all integers $k \geq a$, if $P(k)$ is true then $P(k+1)$ is true.

Then the statement

$$\text{For all integers } n \geq a,\ P(n) \text{ is true.} \quad (1)$$

Basic Induction

Principle of Mathematical Induction

Let $P(n)$ be a property that is defined for integers n, and let a be a fixed integer.

Suppose the following two statements are true

1. $P(a)$ is true.
2. For all integers $k \geq a$, if $P(k)$ is true then $P(k+1)$ is true.

Then the statement

$$\text{For all integers } n \geq a, \ P(n) \ \text{is true.} \tag{1}$$

Cinvestav

We have the following method for Mathematical Induction

Consider a statement of the form

$$\text{For all integers } n \geq a,\ P(n) \text{ is true.} \tag{2}$$

To prove such a statement

Perform the following two steps

Step 1 (Basis step)

Show that $P(a)$ is true - we normally use $a = 1$.

We have the following method for Mathematical Induction

Consider a statement of the form

$$\text{For all integers } n \geq a, \; P(n) \; \text{ is true.} \tag{2}$$

To prove such a statement

Perform the following two steps

Step 1 (Basis step)

Show that $P(a)$ is true - we normally use $a = 1$.

We have the following method for Mathematical Induction

Consider a statement of the form

$$\text{For all integers } n \geq a, \; P(n) \text{ is true.} \tag{2}$$

To prove such a statement

Perform the following two steps

Step 1 (Basis step)

Show that $P(a)$ is true - we normally use $a = 1$.

Cinvestav

Then

Step 2 (Inductive step)

Show that for all integers $k \geq a$, if $P(k)$ is true, then $P(k+1)$ is true.

Inductive hypothesis

Suppose that $P(k)$ is true, where k is any particular but arbitrarily chosen integer with $k \geq a$.

Then, you can prove that

$P(k+1)$ is true.

Then

Step 2 (Inductive step)

Show that for all integers $k \geq a$, if $P(k)$ is true, then $P(k+1)$ is true.

Inductive hypothesis

Suppose that $P(k)$ is true, where k is any particular but arbitrarily chosen integer with $k \geq a$.

Then, you can prove that

$P(k+1)$ is true.

Then

Step 2 (Inductive step)

Show that for all integers $k \geq a$, if $P(k)$ is true, then $P(k+1)$ is true.

Inductive hypothesis

Suppose that $P(k)$ is true, where k is any particular but arbitrarily chosen integer with $k \geq a$.

Then, you can prove that

$P(k+1)$ is true.

Example

Proposition

For all integers $n \geq 8$, $n¢$ can be obtained using $3¢$ and $5¢$ coins.

Show that P(8) is true

$P(8)$ is true because $8¢$ can be obtained using one coin $3¢$ and another coin of $5¢$.

Show that for all integers $k \geq 8$, if $P(k)$ is true then $P(k+1)$ is also true

We can do the following.

Example

Proposition

For all integers $n \geq 8$, n¢ can be obtained using 3¢ and 5¢ coins.

Show that P(8) is true

$P(8)$ is true because 8¢ can be obtained using one coin 3¢ and another coin of 5¢.

Show that for all integers $k \geq 8$, if $P(k)$ is true then $P(k+1)$ is also true

We can do the following.

Example

Proposition

For all integers $n \geq 8$, $n¢$ can be obtained using $3¢$ and $5¢$ coins.

Show that P(8) is true

$P(8)$ is true because $8¢$ can be obtained using one coin $3¢$ and another coin of $5¢$.

Show that for all integers $k \geq 8$, if $P(k)$ is true then $P(k+1)$ is also true

We can do the following.

Cinvestav

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- k¢ can be obtained using 3¢ and 5¢ coins.

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- $k¢$ can be obtained using $3¢$ and $5¢$ coins.

Case 1 - There is a $5¢$ among those making the change for $k¢$

In this case, replace $5¢$ by two $3¢$. Thus, we get the change for $(k+1)¢$

- At least three $3¢$ coins must have been used
- Remove those coins and replaced them using two $5¢$
- The result will be $(k+1)¢$

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- k¢ can be obtained using 3¢ and 5¢ coins.

Case 1 - There is a 5¢ among those making the change for k¢

In this case, replace 5¢ by two 3¢. Thus, we get the change for $(k+1)$ ¢

Case 2 - There is not a 5¢ among those making the change for k¢

- Because $k \geq 8$, at leat three coins must have been used.
- At least three 3¢ coins must have been used.
- Remove those coins and replaced them using two 5¢.
- The result will be $(k+1)$ ¢.

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- k¢ can be obtained using 3¢ and 5¢ coins.

Case 1 - There is a 5¢ among those making the change for k¢

In this case, replace 5¢ by two 3¢. Thus, we get the change for $(k+1)$ ¢

Case 2 - There is not a 5¢ among those making the change for k¢

- Because $k \geq 8$, at leat three coins must have been used.
- At least three 3¢ coins must have been used.
- Remove those coins and replaced them using two 5¢.
- The result will be $(k+1)$ ¢.

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- k¢ can be obtained using 3¢ and 5¢ coins.

Case 1 - There is a 5¢ among those making the change for k¢

In this case, replace 5¢ by two 3¢. Thus, we get the change for $(k+1)$ ¢

Case 2 - There is not a 5¢ among those making the change for k¢

- Because $k \geq 8$, at leat three coins must have been used.
- At least three 3¢ coins must have been used.
- Remove those coins and replaced them using two 5¢.
- The result will be $(k+1)$ ¢.

Cinvestav

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- k¢ can be obtained using 3¢ and 5¢ coins.

Case 1 - There is a 5¢ among those making the change for k¢

In this case, replace 5¢ by two 3¢. Thus, we get the change for $(k+1)$ ¢

Case 2 - There is not a 5¢ among those making the change for k¢

- Because $k \geq 8$, at leat three coins must have been used.
- At least three 3¢ coins must have been used.
- Remove those coins and replaced them using two 5¢.
- The result will be $(k+1)$ ¢.

Example

Inductive hypothesis

Suppose that k is any integer with $k \geq 8$ such that

- $k¢$ can be obtained using $3¢$ and $5¢$ coins.

Case 1 - There is a $5¢$ among those making the change for $k¢$

In this case, replace $5¢$ by two $3¢$. Thus, we get the change for $(k+1)\,¢$

Case 2 - There is not a $5¢$ among those making the change for $k¢$

- Because $k \geq 8$, at leat three coins must have been used.
- At least three $3¢$ coins must have been used.
- Remove those coins and replaced them using two $5¢$.
- The result will be $(k+1)\,¢$.

Outline

We can go further...

Recursively defined sets and structures

Assume S is a set. We use two steps to define the elements of S.

Basis Step

Specify an initial collection of elements.

Recursive Step

Give a rule for forming new elements from those already known to be in S.

We can go further...

Recursively defined sets and structures

Assume S is a set. We use two steps to define the elements of S.

Basis Step

Specify an initial collection of elements.

Recursive Step

Give a rule for forming new elements from those already known to be in S.

We can go further...

Recursively defined sets and structures

Assume S is a set. We use two steps to define the elements of S.

Basis Step

Specify an initial collection of elements.

Recursive Step

Give a rule for forming new elements from those already known to be in S.

Thus

Let S be a set that has been defined recursively

To prove that every object in S satisfies a certain property:

1. Show that each object in the BASE for S satisfies the property.
2. Show that for each rule in the RECURSION, if the rule is applied to objects in S that satisfy the property, then the objects defined by the rule also satisfy the property.

Cinvestav

Thus

Let S be a set that has been defined recursively

To prove that every object in S satisfies a certain property:

1. Show that each object in the BASE for S satisfies the property.
2. Show that for each rule in the RECURSION, if the rule is applied to objects in S that satisfy the property, then the objects defined by the rule also satisfy the property.

Thus

Let S be a set that has been defined recursively

To prove that every object in S satisfies a certain property:

1. Show that each object in the BASE for S satisfies the property.
2. Show that for each rule in the RECURSION, if the rule is applied to objects in S that satisfy the property, then the objects defined by the rule also satisfy the property.

Example: Binary trees recursive definition

Recall that the set B of binary trees over an alphabet A is defined as follows

1. Basis: $\langle\rangle \in B$
2. Recursive definition: If $L, R \in B$ and $x \in A$ then $< L, x, R >\in B$.

Now define the function $f : B \to \mathbb{N}$ defined as

$$f(\langle\rangle) = 0$$

$$f(\langle L, x, R\rangle) = \begin{cases} 1 & \text{if } L = R = \langle\rangle \\ f(L) + f(R) & \text{otherwise} \end{cases}$$

Theorem

Let T in B be a binary tree. Then $f(T)$ yields the number of leaves of T.

Example: Binary trees recursive definition

Recall that the set B of binary trees over an alphabet A is defined as follows

1. Basis: $\langle\rangle \in B$
2. Recursive definition: If $L, R \in B$ and $x \in A$ then $< L, x, R > \in B$.

Now define the function $f : B \to \mathbb{N}$ defined as

$$f(\langle\rangle) = 0$$

$$f(\langle L, x, R\rangle) = \begin{cases} 1 & \text{if } L = R = \langle\rangle \\ f(L) + f(R) & \text{otherwise} \end{cases}$$

Theorem

Let T in B be a binary tree. Then $f(T)$ yields the number of leaves of T.

Example: Binary trees recursive definition

Recall that the set B of binary trees over an alphabet A is defined as follows

1. Basis: $\langle\rangle \in B$
2. Recursive definition: If $L, R \in B$ and $x \in A$ then $< L, x, R > \in B$.

Now define the function $f : B \to \mathbb{N}$ defined as

$$f\left(\langle\rangle\right) = 0$$

$$f\left(\langle L, x, R\rangle\right) = \begin{cases} 1 & \text{if } L = R = \langle\rangle \\ f\left(L\right) + f\left(R\right) & \text{otherwise} \end{cases}$$

Theorem

Let T in B be a binary tree. Then $f\left(T\right)$ yields the number of leaves of T.

Proof

By structural induction on T

Basis: The empty tree has no leaves, so $f(\langle\rangle) = 0$ is correct.

Induction

Let L, R be trees in B, $x \in A$.

Now

Suppose that $f(L)$ and $f(R)$ denotes the number of leaves of L and R, respectively.

Proof

By structural induction on T

Basis: The empty tree has no leaves, so $f(\langle\rangle) = 0$ is correct.

Induction

Let L, R be trees in B, $x \in A$.

Now

Suppose that $f(L)$ and $f(R)$ denotes the number of leaves of L and R, respectively.

Proof

By structural induction on T

Basis: The empty tree has no leaves, so $f(\langle\rangle) = 0$ is correct.

Induction

Let L, R be trees in B, $x \in A$.

Now

Suppose that $f(L)$ and $f(R)$ denotes the number of leaves of L and R, respectively.

Proof

Case 1

If $L = R = \langle\rangle$, then $\langle L, x, R\rangle = \langle\langle\rangle, x, \langle\rangle\rangle$ has one leaf, namely x, so $f(\langle\langle\rangle, x, \langle\rangle\rangle) = 1$ is correct.

Case 2

If L and R are both not empty, then the number of leaves of the tree $\langle L, x, R\rangle$ is equal to the number of leaves of L plus the number of leaves of R.

$$f(\langle L, x, R\rangle) = f(L) + f(R) \tag{3}$$

Q.E.D.

Proof

Case 1

If $L = R = \langle\rangle$, then $\langle L, x, R\rangle = \langle\langle\rangle, x, \langle\rangle\rangle$ has one leaf, namely x, so $f(\langle\langle\rangle, x, \langle\rangle\rangle) = 1$ is correct.

Case 2

If L and R are both not empty, then the number of leaves of the tree $\langle L, x, R\rangle$ is equal to the number of leaves of L plus the number of leaves of R.

$$f(\langle L, x, R\rangle) = f(L) + f(R) \tag{3}$$

Q.E.D.

Cinvestav

Introduction

When an algorithm contains an iterative control construct such as a while or for loop

It is possible to express its running time as a series:

$$\sum_{j=1}^{n} j \tag{4}$$

Thus, what is the objective of using these series

To be able to find bounds for the complexities of algorithms

Introduction

When an algorithm contains an iterative control construct such as a while or for loop

It is possible to express its running time as a series:

$$\sum_{j=1}^{n} j \tag{4}$$

Thus, what is the objective of using these series

To be able to find bounds for the complexities of algorithms

Cinvestav

Definition of Series

Definition

Given a sequence $a_1, a_2, ..., a_n$ of numbers, where n is a no-negative integer, we can say that

$$a_1 + a_2 + ... + a_n = \sum_{k=1}^{n} a_k \tag{5}$$

Cinvestav

Definition of Series

Definition

Given a sequence $a_1, a_2, ..., a_n$ of numbers, where n is a no-negative integer, we can say that

$$a_1 + a_2 + ... + a_n = \sum_{k=1}^{n} a_k \tag{5}$$

In the case of infinite series

$$a_1 + a_2 + ... = \sum_{k=1}^{\infty} a_k = \lim_{n \to \infty} \sum_{k=1}^{n} a_k \tag{6}$$

Here, we have concepts of convergence and divergence that I will allow you to study.

Definition of Series

Definition

Given a sequence $a_1, a_2, ..., a_n$ of numbers, where n is a no-negative integer, we can say that

$$a_1 + a_2 + ... + a_n = \sum_{k=1}^{n} a_k \tag{5}$$

In the case of infinite series

$$a_1 + a_2 + ... = \sum_{k=1}^{\infty} a_k = \lim_{n \to \infty} \sum_{k=1}^{n} a_k \tag{6}$$

Here, we have concepts of convergence and divergence that I will allow you to study.

Definition of Series

Definition

Given a sequence $a_1, a_2, ..., a_n$ of numbers, where n is a no-negative integer, we can say that

$$a_1 + a_2 + ... + a_n = \sum_{k=1}^{n} a_k \tag{5}$$

In the case of infinite series

$$a_1 + a_2 + ... = \sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=1}^{n} a_k \tag{6}$$

Here, we have concepts of convergence and divergence that I will allow you to study.

Outline

Linearity

For any real number c and any finite sequences $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$

$$\sum_{k=1}^{n} [ca_k + b_k] = c \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k \tag{7}$$

For More

Please take a look at page 1146 of Cormen's book.

Linearity

For any real number c and any finite sequences $a_1, a_2, ..., a_n$ and $b_1, b_2, ..., b_n$

$$\sum_{k=1}^{n} [ca_k + b_k] = c\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k \quad (7)$$

For More

Please take a look at page 1146 of Cormen's book.

Outline

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k} - \frac{1}{k+1} = \frac{(k+1)-k}{k(k+1)} = \frac{1}{k(k+1)} \tag{8}$$

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} \tag{9}$$

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k} - \frac{1}{k+1} = \frac{(k+1) - k}{k\,(k+1)} = \frac{1}{k\,(k+1)} \quad (8)$$

What is the result of the following sum?

$$\sum_{k=1}^{n} \frac{1}{k\,(k+1)} \quad (9)$$

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k}-\frac{1}{k+1}=\frac{(k+1)-k}{k(k+1)}=\frac{1}{k(k+1)} \tag{8}$$

What is the result of the following sum?

$$\sum_{k=1}^{n}\frac{1}{k(k+1)} \tag{9}$$

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k} - \frac{1}{k+1} = \frac{(k+1) - k}{k\,(k+1)} = \frac{1}{k\,(k+1)} \quad (8)$$

What is the result of the following sum?

$$\sum_{k=1}^{n} \frac{1}{k\,(k+1)} \quad (9)$$

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k} - \frac{1}{k+1} = \frac{(k+1) - k}{k\,(k+1)} = \frac{1}{k\,(k+1)} \tag{8}$$

What is the result of the following sum?

$$\sum_{k=1}^{n} \frac{1}{k\,(k+1)} \tag{9}$$

Telescopic Sum

Observation

- In certain sums each term is a difference of two quantities.
- You sometimes see that all the terms cancel except the first and the last.

Example

Imagine that each term in the sum has the following structure:

$$\frac{1}{k} - \frac{1}{k+1} = \frac{(k+1) - k}{k(k+1)} = \frac{1}{k(k+1)} \quad (8)$$

What is the result of the following sum?

$$\sum_{k=1}^{n} \frac{1}{k(k+1)} \quad (9)$$

Telescopic Sum

Definition

For any sequence $a_0, a_1, ..., a_n$,

$$\sum_{k=1}^{n} (a_k - a_{k-1}) = a_n - a_0 \tag{10}$$

Similarly

$$\sum_{k=0}^{n-1} (a_k - a_{k+1}) = a_0 - a_n \tag{11}$$

Cinvestav

Telescopic Sum

Definition

For any sequence $a_0, a_1, ..., a_n$,

$$\sum_{k=1}^{n} (a_k - a_{k-1}) = a_n - a_0 \tag{10}$$

Similarly

$$\sum_{k=0}^{n-1} (a_k - a_{k+1}) = a_0 - a_n \tag{11}$$

Cinvestav

Arithmetic series

Summing over the set $\{1, 2, 3, ..., n\}$

We can prove that

$$\sum_{i=1}^{n} i = \frac{n\left(n+1\right)}{2} \tag{12}$$

Proof

Basis. If $n = 1$ then $\frac{1\times 2}{2} = 1$

Arithmetic series

Summing over the set $\{1, 2, 3, ..., n\}$

We can prove that

$$\sum_{i=1}^{n} i = \frac{n\,(n+1)}{2} \tag{12}$$

Proof

Basis: If $n = 1$ then $\frac{1 \times 2}{2} = 1$

Proof

Induction, assume that is true for n

$$\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + n + 1$$

$$= \frac{n(n+1)}{2} + n + 1$$

$$= \frac{n(n+1) + 2(n+1)}{2}$$

$$= \frac{(n+1)(n+2)}{2}$$

Proof

Induction, assume that is true for n

$$
\begin{aligned}
\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + n + 1 \\
&= \frac{n\,(n+1)}{2} + n + 1 \\
&= \frac{n\,(n+1) + 2\,(n+1)}{2} \\
&= \frac{(n+1)\,(n+2)}{2}
\end{aligned}
$$

Proof

Induction, assume that is true for n

$$
\begin{aligned}
\sum_{i=1}^{n+1} i &= \sum_{i=1}^{n} i + n + 1 \\
&= \frac{n\left(n+1\right)}{2} + n + 1 \\
&= \frac{n\left(n+1\right) + 2\left(n+1\right)}{2} \\
&= \frac{\left(n+1\right)\left(n+2\right)}{2}
\end{aligned}
$$

Proof

Induction, assume that is true for n

$$\sum_{i=1}^{n+1} i = \sum_{i=1}^{n} i + n + 1$$

$$= \frac{n(n+1)}{2} + n + 1$$

$$= \frac{n(n+1) + 2(n+1)}{2}$$

$$= \frac{(n+1)(n+2)}{2}$$

Cinvestav

Series of Squares and Sums

Series of Squares

$$\sum_{k=0}^{n} k^2 = \frac{n\left(n+1\right)\left(2n+1\right)}{6} \quad (13)$$

Series of Cubes

$$\sum_{k=0}^{n} k^2 = \frac{n^2\left(n+1\right)^2}{4} \quad (14)$$

Cinvestav

Series of Squares and Sums

Series of Squares

$$\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \quad (13)$$

Series of Cubes

$$\sum_{k=0}^{n} k^2 = \frac{n^2 (n+1)^2}{4} \quad (14)$$

Cinvestav

Geometric Series

Definition

For a real $x \neq 1$, we have that

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + ... + x^n \tag{15}$$

It is called the geometric series

It is possible to prove that

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1} \tag{16}$$

Proof

Now multiply both sides by of (Eq. 15) by x

$$x \left[\sum_{k=0}^{n} x^k \right] = x + x^2 + x^3 + ... + x^{n+1} \tag{17}$$

Geometric Series

Definition

For a real $x \neq 1$, we have that

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + ... + x^n \tag{15}$$

It is called the geometric series

It is possible to prove that

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1} \tag{16}$$

Proof

Now multiply both sides by of (Eq. 15) by x

$$x\left[\sum_{k=0}^{n} x^k\right] = x + x^2 + x^3 + ... + x^{n+1} \tag{17}$$

Geometric Series

Definition

For a real $x \neq 1$, we have that

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + ... + x^n \tag{15}$$

It is called the geometric series

It is possible to prove that

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1} - 1}{x - 1} \tag{16}$$

Proof

Now multiply both sides by of (Eq. 15) by x

$$x \left[\sum_{k=0}^{n} x^k \right] = x + x^2 + x^3 + ... + x^{n+1} \tag{17}$$

Proof

Subtract (Eq. 17) from (Eq. 15)

$$\sum_{k=0}^{n} x^k - x\left[\sum_{k=0}^{n} x^k\right] = 1 - x^{n+1} \tag{18}$$

Finally

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} = \frac{x^{n+1} - 1}{x - 1} \tag{19}$$

Proof

Subtract (Eq. 17) from (Eq. 15)

$$\sum_{k=0}^{n} x^k - x\left[\sum_{k=0}^{n} x^k\right] = 1 - x^{n+1} \quad (18)$$

Finally

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} = \frac{x^{n+1} - 1}{x - 1} \quad (19)$$

Cinvestav

Infinite Geometric Series

When the summation is infinite and $|x| < 1$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{20}$$

Proof

Given that

Infinite Geometric Series

When the summation is infinite and $|x| < 1$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{20}$$

Proof

Given that

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} \tag{21}$$

Infinite Geometric Series

When the summation is infinite and $|x| < 1$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{20}$$

Proof

Given that

$$\sum_{k=0}^{\infty} x^k = \lim_{n \to \infty} \sum_{k=0}^{n} x^k = \lim_{n \to \infty} \frac{1 - x^{n+1}}{1-x} = \frac{1}{1-x} \tag{21}$$

Infinite Geometric Series

When the summation is infinite and $|x| < 1$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{20}$$

Proof

Given that

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} \tag{21}$$

Cinvestav

Infinite Geometric Series

When the summation is infinite and $|x| < 1$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \tag{20}$$

Proof

Given that

$$\sum_{k=0}^{\infty} x^k = \lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1-x^{n+1}}{1-x} = \frac{1}{1-x} \tag{21}$$

For more on the series

Please take a look to

Cormen's book - Appendix A.

Outline

Cinvestav

This is quite useful for Analysis of Algorithms

Important

The most basic way to evaluate a series is to use mathematical induction.

Example

Prove that

$$\sum_{k=0}^{n} 3^k \leq c3^n \qquad (22)$$

This is quite useful for Analysis of Algorithms

Important

The most basic way to evaluate a series is to use mathematical induction.

Example

Prove that

$$\sum_{k=0}^{n} 3^k \leq c3^n \tag{22}$$

Fast Bounding of Series

A quick upper bound on the arithmetic series

$$\sum_{k=1}^{n} k \leq \sum_{k=1}^{n} n = n^2 \tag{23}$$

In general, for a series $\sum_{k=1}^{n} a_k$

If $a_{\max} = \max_{1 \leq k \leq n} a_k$ then

$$\sum_{k=1}^{n} a_k \leq n \cdot a_{\max} \tag{24}$$

Another fast way of bounding finite series is

Suppose that $\frac{a_{k+1}}{a_k} \leq r$ for all $k \geq 0$ where $0 < r < 1$.

Fast Bounding of Series

A quick upper bound on the arithmetic series

$$\sum_{k=1}^{n} k \leq \sum_{k=1}^{n} n = n^2 \quad (23)$$

In general, for a series $\sum_{k=1}^{n} a_k$

If $a_{\max} = \max_{1\leq k\leq n} a_k$ then

$$\sum_{k=1}^{n} a_k \leq n \cdot a_{\max} \quad (24)$$

Another fast way of bounding finite series is

Suppose that $\frac{a_{k+1}}{a_k} \leq r$ for all $k \geq 0$ where $0 < r < 1$.

Fast Bounding of Series

A quick upper bound on the arithmetic series

$$\sum_{k=1}^{n} k \leq \sum_{k=1}^{n} n = n^2 \quad (23)$$

In general, for a series $\sum_{k=1}^{n} a_k$

If $a_{\max} = \max_{1 \leq k \leq n} a_k$ then

$$\sum_{k=1}^{n} a_k \leq n \cdot a_{\max} \quad (24)$$

Another fast way of bounding finite series is

Suppose that $\frac{a_{k+1}}{a_k} \leq r$ for all $k \geq 0$ where $0 < r < 1$.

A More Elegant Method

Thus, we have

$$a_k \leq a_0 r^k \tag{25}$$

Thus, we can use a infinite decreasing geometric series

A More Elegant Method

Thus, we have

$$a_k \leq a_0 r^k \tag{25}$$

Thus, we can use a infinite decreasing geometric series

$$\begin{aligned}\sum_{k=0}^{n} a_k &\leq \sum_{k=0}^{\infty} a_0 r^k \\ &= a_0 \sum_{k=0}^{\infty} r^k \\ &= a_0 \frac{1}{1-r}\end{aligned}$$

A More Elegant Method

Thus, we have

$$a_k \leq a_0 r^k \tag{25}$$

Thus, we can use a infinite decreasing geometric series

$$\begin{aligned}\sum_{k=0}^{n} a_k \leq & \sum_{k=0}^{\infty} a_0 r^k \\ = & a_0 \sum_{k=0}^{\infty} r^k \\ = & a_0 \frac{1}{1-r}\end{aligned}$$

Cinvestav

A More Elegant Method

Thus, we have

$$a_k \leq a_0 r^k \tag{25}$$

Thus, we can use a infinite decreasing geometric series

$$\begin{aligned}\sum_{k=0}^{n} a_k \leq & \sum_{k=0}^{\infty} a_0 r^k \\ = & a_0 \sum_{k=0}^{\infty} r^k \\ = & a_0 \frac{1}{1-r}\end{aligned}$$

Approximation by integrals

When a summation has the from $\sum_{k=m}^{n} f(k)$, where $f(k)$ is a monotonically increasing function

$$\int_{m-1}^{n} f(x)\, dx \leq \sum_{k=m}^{n} f(k) \leq \int_{m}^{n+1} f(x)\, dx \quad (26)$$

Cinvestav

For example

Given

$$\ln(n+1) = \int_1^{n+1} \frac{1}{x}dx \leq \sum_{k=1}^{n} \frac{1}{k} \tag{27}$$

In addition

$$\sum_{k=1}^{n} \frac{1}{k} \leq \int_1^{n} \frac{1}{x}dx = \ln n \tag{28}$$

For example

Given

$$\ln(n+1) = \int_1^{n+1} \frac{1}{x} dx \leq \sum_{k=1}^{n} \frac{1}{k} \quad (27)$$

In addition

$$\sum_{k=1}^{n} \frac{1}{k} \leq \int_1^{n} \frac{1}{x} dx = \ln n \quad (28)$$

Cinvestav

Outline

Cinvestav

Gerolamo Cardano: Gambling out of Darkness

Gambling

Gambling shows our interest in quantifying the ideas of probability for millennia, but exact mathematical descriptions arose much later.

Gerolamo Cardano (16th century)

While gambling he developed the following rule!!!

Equal conditions

“The most fundamental principle of all in gambling is simply equal conditions, e.g. of opponents, of bystanders, of money, of situation, of the dice box and of the dice itself. To the extent to which you depart from that equity, if it is in your opponent’s favour, you are a fool, and if in your own, you are unjust.”

Gerolamo Cardano: Gambling out of Darkness

Gambling

Gambling shows our interest in quantifying the ideas of probability for millennia, but exact mathematical descriptions arose much later.

Gerolamo Cardano (16th century)

While gambling he developed the following rule!!!

Equal conditions

"The most fundamental principle of all in gambling is simply equal conditions, e.g. of opponents, of bystanders, of money, of situation, of the dice box and of the dice itself. To the extent to which you depart from that equity, if it is in your opponent's favour, you are a fool, and if in your own, you are unjust."

Gerolamo Cardano: Gambling out of Darkness

Gambling

Gambling shows our interest in quantifying the ideas of probability for millennia, but exact mathematical descriptions arose much later.

Gerolamo Cardano (16th century)

While gambling he developed the following rule!!!

Equal conditions

"The most fundamental principle of all in gambling is simply equal conditions, e.g. of opponents, of bystanders, of money, of situation, of the dice box and of the dice itself. To the extent to which you depart from that equity, if it is in your opponent's favour, you are a fool, and if in your own, you are unjust."

Gerolamo Cardano's Definition

Probability

"If therefore, someone should say, I want an ace, a deuce, or a trey, you know that there are 27 favourable throws, and since the circuit is 36, the rest of the throws in which these points will not turn up will be 9; the odds will therefore be 3 to 1."

Meaning

Probability as a ratio of favorable to all possible outcomes!!! As long all events are equiprobable...

Thus, we get

$$P(\text{All favourable throws}) = \frac{\text{Number All favourable throws}}{\text{Number of All throws}} \tag{29}$$

Gerolamo Cardano's Definition

Probability

"If therefore, someone should say, I want an ace, a deuce, or a trey, you know that there are 27 favourable throws, and since the circuit is 36, the rest of the throws in which these points will not turn up will be 9; the odds will therefore be 3 to 1."

Meaning

Probability as a ratio of favorable to all possible outcomes!!! As long all events are equiprobable...

Thus, we get

$$P(\text{All favourable throws}) = \frac{\text{Number All favourable throws}}{\text{Number of All throws}} \quad (29)$$

Gerolamo Cardano's Definition

Probability

"If therefore, someone should say, I want an ace, a deuce, or a trey, you know that there are 27 favourable throws, and since the circuit is 36, the rest of the throws in which these points will not turn up will be 9; the odds will therefore be 3 to 1."

Meaning

Probability as a ratio of favorable to all possible outcomes!!! As long all events are equiprobable...

Thus, we get

$$P(\text{All favourable throws}) = \frac{\text{Number All favourable throws}}{\text{Number of All throws}} \tag{29}$$

Intuitive Formulation

Empiric Definition

Intuitively, the probability of an event A could be defined as:

$$P(A) = \lim_{n \to \infty} \frac{N(A)}{n}$$

Where $N(A)$ is the number that event a happens in n trials.

Example

Imagine you have three dices, then

Intuitive Formulation

Empiric Definition

Intuitively, the probability of an event A could be defined as:

$$P(A) = \lim_{n \to \infty} \frac{N(A)}{n}$$

Where $N(A)$ is the number that event a happens in n trials.

Example

Imagine you have three dices, then

- The total number of outcomes is 6^3
- If we have event A = all numbers are equal, $|A| = 6$
- Then, we have that $P(A) = \frac{6}{6^3} = \frac{1}{36}$

Intuitive Formulation

Empiric Definition

Intuitively, the probability of an event A could be defined as:

$$P(A) = \lim_{n \to \infty} \frac{N(A)}{n}$$

Where $N(A)$ is the number that event a happens in n trials.

Example

Imagine you have three dices, then

- The total number of outcomes is 6^3
- If we have event $A =$ all numbers are equal, $|A| = 6$
- Then, we have that $P(A) = \frac{6}{6^3} = \frac{1}{36}$

Intuitive Formulation

Empiric Definition

Intuitively, the probability of an event A could be defined as:

$$P(A) = \lim_{n\to\infty} \frac{N(A)}{n}$$

Where $N(A)$ is the number that event a happens in n trials.

Example

Imagine you have three dices, then

- The total number of outcomes is 6^3
- If we have event $A =$ all numbers are equal, $|A| = 6$
- Then, we have that $P(A) = \frac{6}{6^3} = \frac{1}{36}$

Cinvestav

Outline

Cinvestav

Axioms of Probability

Axioms

Given a sample space S of events, we have that

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. If $A_1, A_2, ..., A_n$ are mutually exclusive events (i.e. $P(A_i \cap A_j) = 0$), then:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

Cinvestav

Axioms of Probability

Axioms

Given a sample space S of events, we have that

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. If $A_1, A_2, ..., A_n$ are mutually exclusive events (i.e. $P(A_i \cap A_j) = 0$), then:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

Cinvestav

Axioms of Probability

Axioms

Given a sample space S of events, we have that

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. If $A_1, A_2, ..., A_n$ are mutually exclusive events (i.e. $P(A_i \cap A_j) = 0$), then:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

Axioms of Probability

Axioms

Given a sample space S of events, we have that

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. If $A_1, A_2, ..., A_n$ are mutually exclusive events (i.e. $P(A_i \cap A_j) = 0$), then:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

Cinvestav

Set Operations

We are using

Set Notation

Thus

What Operations?

Set Operations

We are using

Set Notation

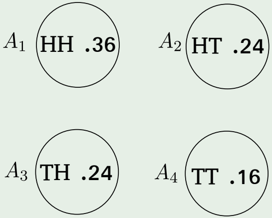
Thus

What Operations?

Example

Setup

Throw a biased coin twice



We have the following event

At least one head!!! Can you tell me which events are part of it?

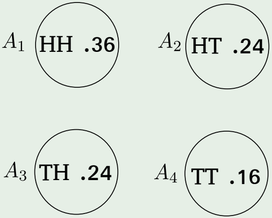
What about this one?

Tail on first toss.

Example

Setup

Throw a biased coin twice



We have the following event

At least one head!!! Can you tell me which events are part of it?

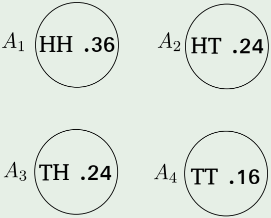
What about this one?

Tail on first toss.

Example

Setup

Throw a biased coin twice



We have the following event

At least one head!!! Can you tell me which events are part of it?

What about this one?

Tail on first toss.

We need to count!!!

We have four main methods of counting

1. Ordered samples of size r with replacement
2. Ordered samples of size r without replacement
3. Unordered samples of size r without replacement
4. Unordered samples of size r with replacement

Cinvestav

We need to count!!!

We have four main methods of counting

1. Ordered samples of size r with replacement
2. Ordered samples of size r without replacement
3. Unordered samples of size r without replacement
4. Unordered samples of size r with replacement

We need to count!!!

We have four main methods of counting

1. Ordered samples of size r with replacement
2. Ordered samples of size r without replacement
3. Unordered samples of size r without replacement
4. Unordered samples of size r with replacement

We need to count!!!

We have four main methods of counting

1. Ordered samples of size r with replacement
2. Ordered samples of size r without replacement
3. Unordered samples of size r without replacement
4. Unordered samples of size r with replacement

Ordered samples of size r with replacement

Definition

The number of possible sequences $(a_{i_1}, ..., a_{i_r})$ for n different numbers is

$$n \times n \times ... \times n = n^r \quad (30)$$

Example

If you throw three dices you have $6 \times 6 \times 6 = 216$

Ordered samples of size r with replacement

Definition

The number of possible sequences $(a_{i_1}, ..., a_{i_r})$ for n different numbers is

$$n \times n \times ... \times n = n^r \quad (30)$$

Example

If you throw three dices you have $6 \times 6 \times 6 = 216$

Ordered samples of size r without replacement

Definition

The number of possible sequences $(a_{i_1}, ..., a_{i_r})$ for n different numbers is

$$n \times n-1 \times ... \times (n-(r-1)) = \frac{n!}{(n-r)!} \tag{31}$$

Example

The number of different numbers that can be formed if no digit can be repeated. For example, if you have 4 digits and you want numbers of size 3.

Ordered samples of size r without replacement

Definition

The number of possible sequences $(a_{i_1}, ..., a_{i_r})$ for n different numbers is

$$n \times n-1 \times ... \times (n-(r-1)) = \frac{n!}{(n-r)!} \tag{31}$$

Example

The number of different numbers that can be formed if no digit can be repeated. For example, if you have 4 digits and you want numbers of size 3.

Unordered samples of size r without replacement

Definition

Actually, we want the number of possible unordered sets.

However

We have $\frac{n!}{(n-r)!}$ collections where we care about the order. Thus

$$\frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{r!\,(n-r)!} = \binom{n}{r} \tag{32}$$

Cinvestav

Unordered samples of size r without replacement

Definition

Actually, we want the number of possible unordered sets.

However

We have $\frac{n!}{(n-r)!}$ collections where we care about the order. Thus

$$\frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{r!\,(n-r)!} = \binom{n}{r} \tag{32}$$

Unordered samples of size r with replacement

Definition

We want to find an unordered set $\{a_{i_1}, ..., a_{i_r}\}$ with replacement

Use a digit trick for that

Look at the Board

Thus

$$\binom{n+r-1}{r} \tag{33}$$

Unordered samples of size r with replacement

Definition

We want to find an unordered set $\{a_{i_1}, ..., a_{i_r}\}$ with replacement

Use a digit trick for that

Look at the Board

Thus

$$\binom{n+r-1}{r} \tag{33}$$

Cinvestav

Unordered samples of size r with replacement

Definition

We want to find an unordered set $\{a_{i_1}, ..., a_{i_r}\}$ with replacement

Use a digit trick for that

Look at the Board

Thus

$$\binom{n+r-1}{r} \tag{33}$$

How?

Change encoding by adding more signs

Imagine all the strings of three numbers with $\{1, 2, 3\}$

We have

Old String	New String
111	1+0,1+1,1+2=123
112	1+0,1+1,2+2=124
113	1+0,1+1,3+2=125
122	1+0,2+1,2+2=134
123	1+0,2+1,3+2=135
133	1+0,3+1,3+2=145
222	2+0,2+1,2+2=234
223	2+0,2+1,3+2=235
233	1+0,3+1,3+2=245
333	3+0,3+1,3+2=345

How?

Change encoding by adding more signs

Imagine all the strings of three numbers with $\{1, 2, 3\}$

We have

Old String	New String
111	1+0,1+1,1+2=123
112	1+0,1+1,2+2=124
113	1+0,1+1,3+2=125
122	1+0,2+1,2+2=134
123	1+0,2+1,3+2=135
133	1+0,3+1,3+2=145
222	2+0,2+1,2+2=234
223	2+0,2+1,3+2=235
233	1+0,3+1,3+2=245
333	3+0,3+1,3+2=345

Outline

Cinvestav

Independence

Definition

Two events A and B are independent if and only if
$P(A, B) = P(A \cap B) = P(A)P(B)$

Do you have any example?

Any idea?

Independence

Definition

Two events A and B are independent if and only if
$P(A, B) = P(A \cap B) = P(A)P(B)$

Do you have any example?

Any idea?

Example

We have two dices

Thus, we have all pairs (i, j) such that $i, j = 1, 2, 3, ..., 6$

Example

We have two dices

Thus, we have all pairs (i, j) such that $i, j = 1, 2, 3, ..., 6$

We have the following events

- A ={First dice 1,2 or 3}
- B = {First dice 3, 4 or 5}
- C = {The sum of two faces is 9}

Example

We have two dices

Thus, we have all pairs (i, j) such that $i, j = 1, 2, 3, ..., 6$

We have the following events

- A ={First dice 1,2 or 3}
- B = {First dice 3, 4 or 5}
- C = {The sum of two faces is 9}

So, we can do

Look at the board!!! Independence between A, B, C

Cinvestav

Example

We have two dices

Thus, we have all pairs (i, j) such that $i, j = 1, 2, 3, ..., 6$

We have the following events

- A ={First dice 1,2 or 3}
- B = {First dice 3, 4 or 5}
- C = {The sum of two faces is 9}

So, we can do

Look at the board!!! Independence between A, B, C

Example

We have two dices

Thus, we have all pairs (i, j) such that $i, j = 1, 2, 3, ..., 6$

We have the following events

- A ={First dice 1,2 or 3}
- B = {First dice 3, 4 or 5}
- C = {The sum of two faces is 9}

So, we can do

Look at the board!!! Independence between A, B, C

We can use it to derive the Binomial Distribution

WHAT?????

First, we use a sequence of n Bernoulli Trials

We have this

- "Success" has a probability p.
- "Failure" has a probability $1 - p$.

First, we use a sequence of n Bernoulli Trials

We have this

- "Success" has a probability p.
- "Failure" has a probability $1 - p$.

Examples

- Toss a coin independently n times.
- Examine components produced on an assembly line.

Cinvestav

First, we use a sequence of n Bernoulli Trials

We have this

- "Success" has a probability p.
- "Failure" has a probability $1 - p$.

Examples

- Toss a coin independently n times.
- Examine components produced on an assembly line.

Now

We take S =all 2^n ordered sequences of length n, with components 0(failure) and 1(success).

Cinvestav

First, we use a sequence of n Bernoulli Trials

We have this

- "Success" has a probability p.
- "Failure" has a probability $1-p$.

Examples

- Toss a coin independently n times.
- Examine components produced on an assembly line.

Now

We take S =all 2^n ordered sequences of length n, with components 0(failure) and 1(success).

First, we use a sequence of n Bernoulli Trials

We have this

- "Success" has a probability p.
- "Failure" has a probability $1-p$.

Examples

- Toss a coin independently n times.
- Examine components produced on an assembly line.

Now

We take S =all 2^n ordered sequences of length n, with components 0(failure) and 1(success).

Thus, taking a sample ω

$\omega = 11 \cdots 10 \cdots 0$

k 1's followed by $n - k$ 0's.

We have then

$$
\begin{aligned}
P(\omega) &= P\left(A_1 \cap A_2 \cap \ldots \cap A_k \cap A_{k+1}^c \cap \ldots \cap A_n^c\right) \\
&= P(A_1)\, P(A_2) \cdots P(A_k)\, P\left(A_{k+1}^c\right) \cdots P\left(A_n^c\right) \\
&= p^k (1-p)^{n-k}
\end{aligned}
$$

Important

The number of such sample is the number of sets with k elements.... or...

$$
\binom{n}{k}
$$

Thus, taking a sample ω

$\omega = 11\cdots10\cdots0$

k 1's followed by $n-k$ 0's.

We have then

$$
\begin{aligned}
P(\omega) &= P\left(A_1 \cap A_2 \cap \ldots \cap A_k \cap A_{k+1}^c \cap \ldots \cap A_n^c\right) \\
&= P(A_1)\, P(A_2) \cdots P(A_k)\, P\left(A_{k+1}^c\right) \cdots P(A_n^c) \\
&= p^k (1-p)^{n-k}
\end{aligned}
$$

Important

The number of such sample is the number of sets with k elements.... or...

$$\binom{n}{k}$$

Thus, taking a sample ω

$\omega = 11\cdots 10\cdots 0$

k 1's followed by $n-k$ 0's.

We have then

$$\begin{aligned} P(\omega) &= P\left(A_1 \cap A_2 \cap \ldots \cap A_k \cap A_{k+1}^c \cap \ldots \cap A_n^c\right) \\ &= P(A_1)\, P(A_2) \cdots P(A_k)\, P\left(A_{k+1}^c\right) \cdots P(A_n^c) \\ &= p^k (1-p)^{n-k} \end{aligned}$$

Important

The number of such sample is the number of sets with k elements.... or...

$$\binom{n}{k}$$

Did you notice?

We do not care where the 1's and 0's are

Thus all the probabilities are equal to $p^k (1-p)^k$

Thus, we are looking to sum all those probabilities of all those combinations of 1's and 0's

$$\sum_{k \text{ 1's}} p\left(\omega^k\right)$$

Then

$$\sum_{k \text{ 1's}} p\left(\omega^k\right) = \begin{pmatrix} n \\ k \end{pmatrix} p\,(1-p)^{n-k}$$

Did you notice?

We do not care where the 1's and 0's are

Thus all the probabilities are equal to $p^k (1-p)^k$

Thus, we are looking to sum all those probabilities of all those combinations of 1's and 0's

$$\sum_{k \text{ 1's}} p\left(\omega^k\right)$$

Then

$$\sum_{k \text{ 1's}} p\left(\omega^k\right) = \binom{n}{k} p (1-p)^{n-k}$$

Did you notice?

We do not care where the 1's and 0's are

Thus all the probabilities are equal to $p^k (1-p)^k$

Thus, we are looking to sum all those probabilities of all those combinations of 1's and 0's

$$\sum_{k \text{ 1's}} p\left(\omega^k\right)$$

Then

$$\sum_{k \text{ 1's}} p\left(\omega^k\right) = \begin{pmatrix} n \\ k \end{pmatrix} p (1-p)^{n-k}$$

Proving this is a probability

Sum of these probabilities is equal to 1

$$\sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p\,(1-p)^{n-k} = (p + (1-p))^{n} = 1$$

The other is simple

$$0 \leq \left(\begin{array}{c} n \\ k \end{array} \right) p\,(1-p)^{n-k} \leq 1 \;\forall k$$

This is know as

The Binomial probability function!!!

Proving this is a probability

Sum of these probabilities is equal to 1

$$\sum_{k=0}^{n} \binom{n}{k} p\left(1-p\right)^{n-k} = \left(p + \left(1-p\right)\right)^{n} = 1$$

The other is simple

$$0 \leq \binom{n}{k} p\left(1-p\right)^{n-k} \leq 1 \ \forall k$$

This is know as

The Binomial probability function!!!

Proving this is a probability

Sum of these probabilities is equal to 1

$$\sum_{k=0}^{n} \binom{n}{k} p(1-p)^{n-k} = (p + (1-p))^n = 1$$

The other is simple

$$0 \leq \binom{n}{k} p(1-p)^{n-k} \leq 1 \ \forall k$$

This is know as

The Binomial probability function!!!

Cinvestav

Outline

Cinvestav

Different Probabilities

Unconditional

This is the probability of an event A prior to arrival of any evidence, it is denoted by $P(A)$. For example:

- P(Cavity)=0.1 means that "in the absence of any other information, there is a 10% chance that the patient is having a cavity".

- P(Cavity/Toothache)=0.8 means that "there is an 80% chance that the patient is having a cavity given that he is having a toothache."

Different Probabilities

Unconditional

This is the probability of an event A prior to arrival of any evidence, it is denoted by $P(A)$. For example:

- P(Cavity)=0.1 means that "in the absence of any other information, there is a 10% chance that the patient is having a cavity".

Conditional

This is the probability of an event A given some evidence B, it is denoted $P(A|B)$. For example:

- P(Cavity/Toothache)=0.8 means that "there is an 80% chance that the patient is having a cavity given that he is having a toothache"

Different Probabilities

Unconditional

This is the probability of an event A prior to arrival of any evidence, it is denoted by $P(A)$. For example:

- P(Cavity)=0.1 means that "in the absence of any other information, there is a 10% chance that the patient is having a cavity".

Conditional

This is the probability of an event A given some evidence B, it is denoted $P(A|B)$. For example:

- P(Cavity/Toothache)=0.8 means that "there is an 80% chance that the patient is having a cavity given that he is having a toothache"

Cinvestav

Different Probabilities

Unconditional

This is the probability of an event A prior to arrival of any evidence, it is denoted by $P(A)$. For example:

- P(Cavity)=0.1 means that "in the absence of any other information, there is a 10% chance that the patient is having a cavity".

Conditional

This is the probability of an event A given some evidence B, it is denoted $P(A|B)$. For example:

- P(Cavity/Toothache)=0.8 means that "there is an 80% chance that the patient is having a cavity given that he is having a toothache"

Outline

Posterior Probabilities

Relation between conditional and unconditional probabilities

- Conditional probabilities can be defined in terms of unconditional probabilities:

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

which generalizes to the **chain rule** $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$.

Posterior Probabilities

Relation between conditional and unconditional probabilities

- Conditional probabilities can be defined in terms of unconditional probabilities:

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

which generalizes to the **chain rule** $P(A,B) = P(B)P(A|B) = P(A)P(B|A)$.

Law of Total Probabilities

- if $B_1, B_2, ..., B_n$ is a partition of mutually exclusive events and A is an event, then $P(A) = \sum_{i=1}^{n} P(A \cap B_i)$. An special case $P(A) = P(A,B) + P(A,\overline{B})$.
- In addition, this can be rewritten into $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$.

Cinvestav

Posterior Probabilities

Relation between conditional and unconditional probabilities

- Conditional probabilities can be defined in terms of unconditional probabilities:

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

which generalizes to the **chain rule** $P(A, B) = P(B)P(A|B) = P(A)P(B|A)$.

Law of Total Probabilities

- if $B_1, B_2, ..., B_n$is a partition of mutually exclusive events and $A$is an event, then $P(A) = \sum_{i=1}^{n} P(A \cap B_i)$. An special case $P(A) = P(A, B) + P(A, \overline{B})$.
- In addition, this can be rewritten into $P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i)$.

Example

Three cards are drawn from a deck

Find the probability of no obtaining a heart

We have

- 52 cards
- 39 of them not a heart

Define

A_i ={Card i is not a heart} Then?

Example

Three cards are drawn from a deck

Find the probability of no obtaining a heart

We have

- 52 cards
- 39 of them not a heart

Define

A_i ={Card i is not a heart} Then?

Example

Three cards are drawn from a deck

Find the probability of no obtaining a heart

We have

- 52 cards
- 39 of them not a heart

Define

$A_i =${Card i is not a heart} Then?

Independence and Conditional

From here, we have that...

$P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Conditional independence

A and B are conditionally independent given C if and only if

$$P(A|B, C) = P(A|C)$$

Example: $P(WetGrass|Season, Rain) = P(WetGrass|Rain)$.

Independence and Conditional

From here, we have that...

$P(A|B) = P(A)$ and $P(B|A) = P(B)$.

Conditional independence

A and B are conditionally independent given C if and only if

$$P(A|B, C) = P(A|C)$$

Example: $P(WetGrass|Season, Rain) = P(WetGrass|Rain)$.

Bayes Theorem

One Version

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

- $P(A|B)$ is the conditional probability of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
- $P(B|A)$ is the conditional probability of B given A. It is also called the likelihood.
- $P(B)$ is the prior or marginal probability of B, and acts as a normalizing constant.

Bayes Theorem

One Version

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where

- $P(A)$ is the **prior probability** or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
- $P(A|B)$ is the **conditional probability** of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
- $P(B|A)$ is the **conditional probability** of B given A. It is also called the likelihood.
- $P(B)$ is the **prior or marginal probability** of B, and acts as a normalizing constant.

Bayes Theorem

One Version

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where

- $P(A)$ is the **prior probability** or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
- $P(A|B)$ is the **conditional probability** of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
- $P(B|A)$ is the **conditional probability** of B given A. It is also called the likelihood.
- $P(B)$ is the **prior or marginal probability** of B, and acts as a normalizing constant.

Bayes Theorem

One Version

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where

- $P(A)$ is the **prior probability** or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
- $P(A|B)$ is the **conditional probability** of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
- $P(B|A)$ is the **conditional probability** of B given A. It is also called the likelihood.
- $P(B)$ is the **prior or marginal probability** of B, and acts as a normalizing constant.

Bayes Theorem

One Version

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Where

- $P(A)$ is the **prior probability** or marginal probability of A. It is "prior" in the sense that it does not take into account any information about B.
- $P(A|B)$ is the **conditional probability** of A, given B. It is also called the posterior probability because it is derived from or depends upon the specified value of B.
- $P(B|A)$ is the **conditional probability** of B given A. It is also called the likelihood.
- $P(B)$ is the **prior or marginal probability** of B, and acts as a normalizing constant.

General Form of the Bayes Rule

Definition

If $A_1, A_2, ..., A_n$ is a partition of mutually exclusive events and B any event, then:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

where

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

General Form of the Bayes Rule

Definition

If $A_1, A_2, ..., A_n$ is a partition of mutually exclusive events and B any event, then:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

where

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Example

Setup

Throw two unbiased dice independently.

Let

1. A ={sum of the faces =8}
2. B ={faces are equal}

Then calculate $P(B|A)$

Look at the board

Example

Setup

Throw two unbiased dice independently.

Let

1. A ={sum of the faces $=8$}
2. B ={faces are equal}

Then calculate $P(B|A)$

Look at the board

Example

Setup

Throw two unbiased dice independently.

Let

1. $A =$ {sum of the faces $=8$}
2. $B =$ {faces are equal}

Then calculate $P(B|A)$

Look at the board

Another Example

We have the following

Two coins are available, one unbiased and the other two headed

Assume

That you have a probability of $\frac{3}{4}$ to choose the unbiased

- $B_1 = \{\text{Unbiased coin chosen}\}$
- $B_2 = \{\text{Biased coin chosen}\}$
  - Find that if a head come up, find the probability that the two headed coin was chosen

Another Example

We have the following

Two coins are available, one unbiased and the other two headed

Assume

That you have a probability of $\frac{3}{4}$ to choose the unbiased

Events

- $A=$ {head comes up}
- $B_1=$ {Unbiased coin chosen}
- $B_2=$ {Biased coin chosen}
  - Find that if a head come up, find the probability that the two headed coin was chosen

Another Example

We have the following

Two coins are available, one unbiased and the other two headed

Assume

That you have a probability of $\frac{3}{4}$ to choose the unbiased

Events

- $A=$ {head comes up}
- $B_1=$ {Unbiased coin chosen}
- $B_2=$ {Biased coin chosen}
  - Find that if a head come up, find the probability that the two headed coin was chosen

Cinvestav

Another Example

We have the following

Two coins are available, one unbiased and the other two headed

Assume

That you have a probability of $\frac{3}{4}$ to choose the unbiased

Events

- $A=$ {head comes up}
- $B_1=$ {Unbiased coin chosen}
- $B_2=$ {Biased coin chosen}
    - Find that if a head come up, find the probability that the two headed coin was chosen

Another Example

We have the following

Two coins are available, one unbiased and the other two headed

Assume

That you have a probability of $\frac{3}{4}$ to choose the unbiased

Events

- $A=$ {head comes up}
- $B_1=$ {Unbiased coin chosen}
- $B_2=$ {Biased coin chosen}
  - Find that if a head come up, find the probability that the two headed coin was chosen

Outline

Cinvestav

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable X = number of "1"'s recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Cinvestav

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable X =number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Cinvestav

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable X =number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Cinvestav

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable X =number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable X =number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable $X =$number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Random Variables

Definition

In many experiments, it is easier to deal with a summary variable than with the original probability structure.

Example

In an opinion poll, we ask 50 people whether agree or disagree with a certain issue.

- Suppose we record a "1" for agree and "0" for disagree.
- The sample space for this experiment has 2^{50} elements. Why?
- Suppose we are only interested in the number of people who agree.
- Define the variable $X =$number of "1" 's recorded out of 50.
- Easier to deal with this sample space (has only 51 elements).

Thus...

It is necessary to define a function random variable as follow

$$X : S \to \mathbb{R}$$

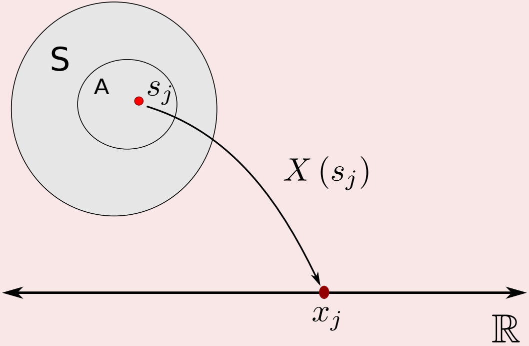
Graphically

Thus...

It is necessary to define a function random variable as follow

$$X : S \to \mathbb{R}$$

Graphically



Random Variables

How?

What is the probability function of the random variable is being defined from the probability function of the original sample space?

- Suppose the sample space is $S = \{s_1, s_2, ..., s_n\}$
- Suppose the range of the random variable $X =< x_1, x_2, ..., x_m >$
- Then, we observe $X = x_i$ if and only if the outcome of the random experiment is an $s_j \in S$ s.t. $X(s_j) = x_j$ or

Random Variables

How?

What is the probability function of the random variable is being defined from the probability function of the original sample space?

- Suppose the sample space is $S = \{s_1, s_2, ..., s_n\}$
- Suppose the range of the random variable $X =< x_1, x_2, ..., x_m >$
- Then, we observe $X = x_i$ if and only if the outcome of the random experiment is an $s_j \in S$ s.t. $X(s_j) = x_j$ or

Random Variables

How?

What is the probability function of the random variable is being defined from the probability function of the original sample space?

- Suppose the sample space is $S = \{s_1, s_2, ..., s_n\}$
- Suppose the range of the random variable $X =< x_1, x_2, ..., x_m >$
- Then, we observe $X = x_i$ if and only if the outcome of the random experiment is an $s_j \in S$ s.t. $X(s_j) = x_j$ or

Random Variables

How?

What is the probability function of the random variable is being defined from the probability function of the original sample space?

- Suppose the sample space is $S = \{s_1, s_2, ..., s_n\}$
- Suppose the range of the random variable $X =< x_1, x_2, ..., x_m >$
- Then, we observe $X = x_i$ if and only if the outcome of the random experiment is an $s_j \in S$ s.t. $X(s_j) = x_j$ or

$$P(X = x_j) = P(s_j \in S | X(s_j) = x_j)$$

Example

Setup

Throw a coin 10 times, and let R be the number of heads.

Then

S = all sequences of length 10 with components H and T

We have for

$\omega =$HHHHTTHTTH $\Rightarrow R(\omega) = 6$

Example

Setup

Throw a coin 10 times, and let R be the number of heads.

Then

$S =$ all sequences of length 10 with components H and T

We have for

$\omega =$ HHHHTTHTTH $\Rightarrow R(\omega) = 6$

Example

Setup

Throw a coin 10 times, and let R be the number of heads.

Then

$S =$ all sequences of length 10 with components H and T

We have for

$\omega =$HHHHTTHTTH $\Rightarrow R(\omega) = 6$

Example

Setup

Let R be the number of heads in two independent tosses of a coin.

- Probability of head is .6

What are the probabilities?

$\Omega = \{HH, HT, TH, TT\}$

Thus, we can calculate

$P(R=0), P(R=1), P(R=2)$

Cinvestav

Example

Setup

Let R be the number of heads in two independent tosses of a coin.

- Probability of head is .6

What are the probabilities?

$\Omega = \{HH, HT, TH, TT\}$

Thus, we can calculate

$P(R=0), P(R=1), P(R=2)$

Cinvestav

Example

Setup

Let R be the number of heads in two independent tosses of a coin.

- Probability of head is .6

What are the probabilities?

Ω ={HH,HT,TH,TT}

Thus, we can calculate

$P(R=0), P(R=1), P(R=2)$

Outline

Cinvestav

Types of Random Variables

Discrete

A discrete random variable can assume only a countable number of values.

Continuous

A continuous random variable can assume a continuous range of values.

Types of Random Variables

Discrete

A discrete random variable can assume only a countable number of values.

Continuous

A continuous random variable can assume a continuous range of values.

Properties

Probability Mass Function (PMF) and Probability Density Function (PDF)

The pmf /pdf of a random variable X assigns a probability for each possible value of X.

- $\sum_x p(x) = 1$ and $P(a < X < b) = \sum_{k=a}^{b} p(k)$
- In a similar way for the pdf:
  - $\int_{-\infty}^{\infty} p(x)dx = 1$ and $P(a < X < b) = \int_a^b p(t)dt$

Properties

Probability Mass Function (PMF) and Probability Density Function (PDF)

The pmf /pdf of a random variable X assigns a probability for each possible value of X.

Properties of the pmf and pdf

- Some properties of the pmf:
  - $\sum_x p(x) = 1$ and $P(a < X < b) = \sum_{k=a}^{b} p(k)$.
- In a similar way for the pdf:
  - $\int_{-\infty}^{\infty} p(x)dx = 1$ and $P(a < X < b) = \int_{a}^{b} p(t)dt$.

Cinvestav

Properties

Probability Mass Function (PMF) and Probability Density Function (PDF)

The pmf /pdf of a random variable X assigns a probability for each possible value of X.

Properties of the pmf and pdf

- Some properties of the pmf:
  - $\sum_x p(x) = 1$ and $P(a < X < b) = \sum_{k=a}^{b} p(k)$.
- In a similar way for the pdf:
  - $\int_{-\infty}^{\infty} p(x)dx = 1$ and $P(a < X < b) = \int_a^b p(t)dt$.

Cinvestav

Properties

Probability Mass Function (PMF) and Probability Density Function (PDF)

The pmf /pdf of a random variable X assigns a probability for each possible value of X.

Properties of the pmf and pdf

- Some properties of the pmf:
  - $\sum_x p(x) = 1$ and $P(a < X < b) = \sum_{k=a}^{b} p(k)$.
- In a similar way for the pdf:
  - $\int_{-\infty}^{\infty} p(x)dx = 1$ and $P(a < X < b) = \int_a^b p(t)dt$.

Cinvestav

Properties

Probability Mass Function (PMF) and Probability Density Function (PDF)

The pmf /pdf of a random variable X assigns a probability for each possible value of X.

Properties of the pmf and pdf

- Some properties of the pmf:
  - $\sum_x p(x) = 1$ and $P(a < X < b) = \sum_{k=a}^{b} p(k)$.
- In a similar way for the pdf:
  - $\int_{-\infty}^{\infty} p(x)dx = 1$ and $P(a < X < b) = \int_a^b p(t)dt$.

Cinvestav

Outline

Cumulative Distributive Function I

Cumulative Distribution Function

- With every random variable, we associate a function called Cumulative Distribution Function (CDF) which is defined as follows:

$$F_X(x) = P(f(X) \leq x)$$

- With properties:
  - $F_X(x) \geq 0$
  - $F_X(x)$ in a non-decreasing function of X.

Example

- If X is discrete, its CDF can be computed as follows:

$$F_X(x) = P(f(X) \leq x) = \sum_{k=1}^{N} P(X_k = p_k).$$

Cumulative Distributive Function I

Cumulative Distribution Function

- With every random variable, we associate a function called Cumulative Distribution Function (CDF) which is defined as follows:

$$F_X(x) = P(f(X) \leq x)$$

- With properties:
  - $F_X(x) \geq 0$
  - $F_X(x)$ in a non-decreasing function of X.

Example

- If X is discrete, its CDF can be computed as follows:

$$F_X(x) = P(f(X) \leq x) = \sum_{k=1}^{N} P(X_k = p_k).$$

Cumulative Distributive Function I

Cumulative Distribution Function

- With every random variable, we associate a function called Cumulative Distribution Function (CDF) which is defined as follows:

$$F_X(x) = P(f(X) \leq x)$$

- With properties:
  - $F_X(x) \geq 0$
  - $F_X(x)$ in a non-decreasing function of X.

Example

- If X is discrete, its CDF can be computed as follows:

$$F_X(x) = P(f(X) \leq x) = \sum_{k=1}^{N} P(X_k = p_k).$$

Cumulative Distributive Function I

Cumulative Distribution Function

- With every random variable, we associate a function called Cumulative Distribution Function (CDF) which is defined as follows:

$$F_X(x) = P(f(X) \leq x)$$

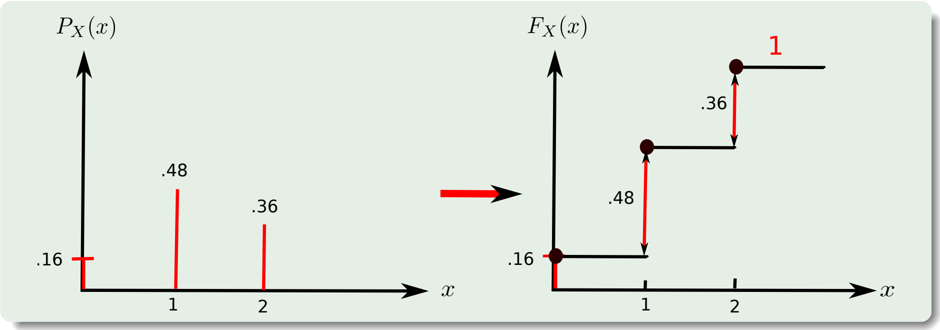
- With properties:
  - $F_X(x) \geq 0$
  - $F_X(x)$ in a non-decreasing function of X.

Example

- If X is discrete, its CDF can be computed as follows:

$$F_X(x) = P(f(X) \leq x) = \sum_{k=1}^{N} P(X_k = p_k).$$

Example: Discrete Function



Cumulative Distributive Function II

Continuous Function

If X is continuous, its CDF can be computed as follows:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt.$$

Remark

Based in the fundamental theorem of calculus, we have the following equality.

$$p(x) = \frac{dF}{dx}(x)$$

Note

This particular $p(x)$ is known as the Probability Mass Function (PMF) or Probability Distribution Function (PDF).

Cumulative Distributive Function II

Continuous Function

If X is continuous, its CDF can be computed as follows:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt.$$

Remark

Based in the fundamental theorem of calculus, we have the following equality.

$$p(x) = \frac{dF}{dx}(x)$$

Note

This particular $p(x)$ is known as the Probability Mass Function (PMF) or Probability Distribution Function (PDF).

Cumulative Distributive Function II

Continuous Function

If X is continuous, its CDF can be computed as follows:

$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

Remark

Based in the fundamental theorem of calculus, we have the following equality.

$$p(x) = \frac{dF}{dx}(x)$$

Note

This particular $p(x)$ is known as the Probability Mass Function (PMF) or Probability Distribution Function (PDF).

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- X has a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- Xhas a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

We have

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(t)\,dt \quad (34)$$

$$P\{a < X \leq b\} = \int_{a}^{b} f_X(t)\,dt \quad (35)$$

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- Xhas a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

We have

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(t)\, dt \quad (34)$$

$$P\{a < X \leq b\} = \int_{a}^{b} f_X(t)\, dt \quad (35)$$

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- Xhas a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

We have

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(t)\, dt \quad (34)$$

$$P\{a < X \leq b\} = \int_{a}^{b} f_X(t)\, dt \quad (35)$$

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- Xhas a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

We have

$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(t)\, dt \tag{34}$$

$$P\{a < X \leq b\} = \int_{a}^{b} f_X(t)\, dt \tag{35}$$

Cinvestav

Example: Continuous Function

Setup

- A number X is chosen at random between a and b
- Xhas a uniform distribution
  - $f_X(x) = \frac{1}{b-a}$ for $a \leq x \leq b$
  - $f_X(x) = 0$ for $x < a$ and $x > b$

We have

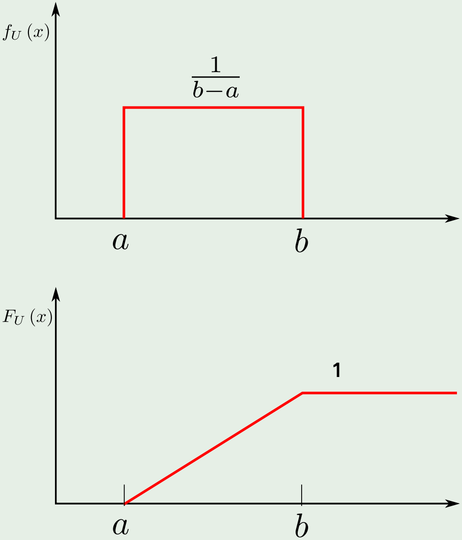
$$F_X(x) = P\{X \leq x\} = \int_{-\infty}^{x} f_X(t)\, dt \tag{34}$$

$$P\{a < X \leq b\} = \int_{a}^{b} f_X(t)\, dt \tag{35}$$

Cinvestav

Graphically

Example uniform distribution



Outline

Cinvestav

Properties of the PMF/PDF

Conditional PMF/PDF

We have the conditional pdf:

$$p(y|x) = \frac{p(x,y)}{p(x)}.$$

From this, we have the general chain rule

$$p(x_1, x_2, ..., x_n) = p(x_1|x_2, ..., x_n)p(x_2|x_3, ..., x_n)...p(x_n).$$

Independence

If X and Y are independent, then:

$$p(x,y) = p(x)p(y).$$

Properties of the PMF/PDF

Conditional PMF/PDF

We have the conditional pdf:

$$p(y|x) = \frac{p(x,y)}{p(x)}.$$

From this, we have the general chain rule

$$p(x_1, x_2, ..., x_n) = p(x_1|x_2, ..., x_n)p(x_2|x_3, ..., x_n)...p(x_n).$$

Independence

If X and Y are independent, then:

$$p(x,y) = p(x)p(y).$$

Properties of the PMF/PDF

Law of Total Probability

$$p(y) = \sum_x p(y|x)p(x).$$

Outline

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute) \;\; R_2 = 10(cents)$$

$$\text{if } 3 < R_1 \leq 6(minute) \;\; R_2 = 20(cents)$$

$$\text{if } 6 < R_1 \leq 9(minute) \;\; R_2 = 30(cents)$$

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute) \;\; R_2 = 10(cents)$$

$$\text{if } 3 < R_1 \leq 6(minute) \;\; R_2 = 20(cents)$$

$$\text{if } 6 < R_1 \leq 9(minute) \;\; R_2 = 30(cents)$$

We have then the probabilities

$P\{R_2 = 10\} = 0.6$, $P\{R_2 = 20\} = 0.25$, $P\{R_2 = 10\} = 0.15$.

We can say that we have $N \times 0.6$ calls and $10 \times N \times 0.6$ is the cost of those calls

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute) \;\; R_2 = 10(cents)$$

$$\text{if } 3 < R_1 \leq 6(minute) \;\; R_2 = 20(cents)$$

$$\text{if } 6 < R_1 \leq 9(minute) \;\; R_2 = 30(cents)$$

We have then the probabilities

$P\left\{R_2 = 10\right\} = 0.6$, $P\left\{R_2 = 20\right\} = 0.25$, $P\left\{R_2 = 10\right\} = 0.15$.

If we observe N calls and N is very large

We can say that we have $N \times 0.6$ calls and $10 \times N \times 0.6$ the cost of those calls.

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute) \;\; R_2 = 10(cents)$$
$$\text{if } 3 < R_1 \leq 6(minute) \;\; R_2 = 20(cents)$$
$$\text{if } 6 < R_1 \leq 9(minute) \;\; R_2 = 30(cents)$$

We have then the probabilities

$P\{R_2 = 10\} = 0.6$, $P\{R_2 = 20\} = 0.25$, $P\{R_2 = 10\} = 0.15$.

If we observe N calls and N is very large

We can say that we have $N \times 0.6$ calls and $10 \times N \times 0.6$ the cost of those calls.

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute)\ \ R_2 = 10(cents)$$
$$\text{if } 3 < R_1 \leq 6(minute)\ \ R_2 = 20(cents)$$
$$\text{if } 6 < R_1 \leq 9(minute)\ \ R_2 = 30(cents)$$

We have then the probabilities

$P\left\{R_2 = 10\right\} = 0.6,\ P\left\{R_2 = 20\right\} = 0.25,\ P\left\{R_2 = 10\right\} = 0.15.$

If we observe N calls and N is very large

We can say that we have $N \times 0.6$ calls and $10 \times N \times 0.6$ the cost of those calls.

Expectation

Something Notable

You have the random variables R_1, R_2 representing how long is a call and how much you pay for an international call:

$$\text{if } 0 \leq R_1 \leq 3(minute)\ \ R_2 = 10(cents)$$
$$\text{if } 3 < R_1 \leq 6(minute)\ \ R_2 = 20(cents)$$
$$\text{if } 6 < R_1 \leq 9(minute)\ \ R_2 = 30(cents)$$

We have then the probabilities

$P\left\{R_2 = 10\right\} = 0.6,\ P\left\{R_2 = 20\right\} = 0.25,\ P\left\{R_2 = 10\right\} = 0.15.$

If we observe N calls and N is very large

We can say that we have $N \times 0.6$ calls and $10 \times N \times 0.6$ the cost of those calls.

Expectation

Similarly

- $\{R_2 = 20\} \implies 0.25N$ and total cost $5N$.
- $\{R_2 = 20\} \implies 0.15N$ and total cost $4.5N$.

Expectation

Similarly

- $\{R_2 = 20\} \implies 0.25N$ and total cost $5N$.
- $\{R_2 = 20\} \implies 0.15N$ and total cost $4.5N$.

We have then the probabilities

The total cost is $6N + 5N + 4.5N = 15.5N$ or in average 15.5 cents per call.

Expectation

Similarly

- $\{R_2 = 20\} \Longrightarrow 0.25N$ and total cost $5N$.
- $\{R_2 = 20\} \Longrightarrow 0.15N$ and total cost $4.5N$.

We have then the probabilities

The total cost is $6N + 5N + 4.5N = 15.5N$ or in average 15.5 cents per call.

The average

$$\frac{10\,(0.6N) + 20\,(.25N) + 30\,(0.15N)}{N} = 10\,(0.6) + 20\,(.25) + 30\,(0.15) = \sum_y yP\{R_2 = y\}$$

Expectation

Similarly

- $\{R_2 = 20\} \implies 0.25N$ and total cost $5N$.
- $\{R_2 = 20\} \implies 0.15N$ and total cost $4.5N$.

We have then the probabilities

The total cost is $6N + 5N + 4.5N = 15.5N$ or in average 15.5 cents per call.

The average

$$\frac{10\,(0.6N) + 20\,(.25N) + 30\,(0.15N)}{N} = 10\,(0.6) + 20\,(.25) + 30\,(0.15)$$
$$= \sum_y yP\,\{R_2 = y\}$$

Expected Value

Definition

- Discrete random variable X: $E(X) = \sum_x xp(x)$.
- Continuous random variable Y: $E(Y) = \int_x xp(x)dx$.

Expected Value

Definition

- Discrete random variable X: $E(X) = \sum_x xp(x)$.
- Continuous random variable Y: $E(Y) = \int_x xp(x)dx$.

Extension to a function $g(x)$

- $E(g(X)) = \sum_x g(x)p(x)$ (Discrete case).
- $E(g(X)) = \int_{-\infty}^{+\infty} g(x)p(x)dx$ (Continuous case).

Expected Value

Definition

- Discrete random variable X: $E(X) = \sum_x xp(x)$.
- Continuous random variable Y: $E(Y) = \int_x xp(x)dx$.

Extension to a function $g(x)$

- $E(g(X)) = \sum_x g(x)p(x)$ (Discrete case).
- $E(g(X)) = \int_{-\infty}^{+\infty} g(x)p(x)dx$ (Continuous case).

Cinvestav

Expected Value

Definition

- Discrete random variable X: $E(X) = \sum_x xp(x)$.
- Continuous random variable Y: $E(Y) = \int_x xp(x)dx$.

Extension to a function $g(x)$

- $E(g(X)) = \sum_x g(x)p(x)$ (Discrete case).
- $E(g(X)) = \int_{-\infty}^{+\infty} g(x)p(x)dx$ (Continuous case).

Cinvestav

Linear Property

Linearity property of the Expected Value

$$E(af(X) + bg(Y)) = aE(f(X)) + bE(g(Y)) \quad (36)$$

Example for a discrete distribution

Linear Property

Linearity property of the Expected Value

$$E(af(X)+bg(Y)) = aE(f(X)) + bE(g(Y)) \tag{36}$$

Example for a discrete distribution

$$E[aX+b] = \sum_x [ax+b]\, p(x|\theta)$$

$$= a\sum_x xp(x|\theta) + b\sum_x p(x|\theta)$$

$$= aE[X] + b$$

Cinvestav

Linear Property

Linearity property of the Expected Value

$$E(af(X)+bg(Y))=aE(f(X))+bE(g(Y)) \tag{36}$$

Example for a discrete distribution

$$\begin{aligned}E\left[aX+b\right] &= \sum_x \left[ax+b\right]p\left(x|\theta\right)\\ &= a\sum_x xp\left(x|\theta\right)+b\sum_x p\left(x|\theta\right)\\ &= aE\left[X\right]+b\end{aligned}$$

Linear Property

Linearity property of the Expected Value

$$E(af(X) + bg(Y)) = aE(f(X)) + bE(g(Y)) \tag{36}$$

Example for a discrete distribution

$$\begin{aligned} E[aX + b] &= \sum_x [ax + b]\, p(x|\theta) \\ &= a \sum_x x p(x|\theta) + b \sum_x p(x|\theta) \\ &= aE[X] + b \end{aligned}$$

Cinvestav

Example

Imagine the following

We have the following functions

1. $f(x) = e^{-x}, \; x \geq 0$
2. $g(x) = 0, \; x < 0$

Example

Imagine the following

We have the following functions

1. $f(x) = e^{-x},\ x \geq 0$
2. $g(x) = 0,\ x < 0$

Find

The expected Value

Example

Imagine the following

We have the following functions

1. $f(x) = e^{-x}$, $x \geq 0$
2. $g(x) = 0$, $x < 0$

Find

The expected Value

Example

Imagine the following

We have the following functions

1. $f(x) = e^{-x},\ x \geq 0$
2. $g(x) = 0,\ x < 0$

Find

The expected Value

Variance

Definition

- $Var(X) = E((X-\mu)^2)$ where $\mu = E(X)$

Standard Deviation

- The standard deviation is simply $\sigma = \sqrt{Var(X)}$.

Cinvestav

Variance

Definition

- $Var(X) = E((X-\mu)^2)$ where $\mu = E(X)$

Standard Deviation

- The standard deviation is simply $\sigma = \sqrt{Var(X)}$.

Outline

Cinvestav

Indicator Random Variable

Definition

The indicator of an event A is a random variable I_A defined as follows:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \quad (37)$$

Indicator Random Variable

Definition

The indicator of an event A is a random variable I_A defined as follows:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \tag{37}$$

Thus

- If $I_A = 1$ if A occurs.
- If $I_A = 0$ if A does not occur.

Indicator Random Variable

Definition

The indicator of an event A is a random variable I_A defined as follows:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \tag{37}$$

Thus

- If $I_A = 1$ if A occurs.
- If $I_A = 0$ if A does not occur.

Why is this useful?

Because we can count events!!!

Cinvestav

Indicator Random Variable

Definition

The indicator of an event A is a random variable I_A defined as follows:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \tag{37}$$

Thus

- If $I_A = 1$ if A occurs.
- If $I_A = 0$ if A does not occur.

Why is this useful?

Because we can count events!!!

Useful Property

The Expected value of an indicator function

$$E[I_A] = 0 \times P(I_A = 0) + 1 \times P(I_A = 1) = P(I_A = 1) = P(A) \quad (38)$$

Useful Property

The Expected value of an indicator function

$$E[I_A] = 0 \times P(I_A = 0) + 1 \times P(I_A = 1) = P(I_A = 1) = P(A) \quad (38)$$

Why is this useful?

Because we can use this to count things when a probability is involved!!!

Useful Property

The Expected value of an indicator function

$$E[I_A] = 0 \times P(I_A = 0) + 1 \times P(I_A = 1) = P(I_A = 1) = P(A) \quad (38)$$

Why is this useful?

Because we can use this to count things when a probability is involved!!!

Useful Property

The Expected value of an indicator function

$$E[I_A] = 0 \times P(I_A = 0) + 1 \times P(I_A = 1) = P(I_A = 1) = P(A) \quad (38)$$

Why is this useful?

Because we can use this to count things when a probability is involved!!!

Useful Property

The Expected value of an indicator function

$$E\left[I_A\right] = 0 \times P\left(I_A = 0\right) + 1 \times P\left(I_A = 1\right) = P\left(I_A = 1\right) = P\left(A\right) \quad (38)$$

Why is this useful?

Because we can use this to count things when a probability is involved!!!

Method of Indicators

It is possible to see random variables as a sum of indicators functions

$$R = I_{A_1} + ... + I_{A_n} \tag{39}$$

Then

$$E[R] = \sum_{j=1}^{n} E\left[I_{A_j}\right] = \sum_{j=1}^{n} P(A_j) \tag{40}$$

Hopefully

It is easier to calculate $P(A_j)$ than to evaluate $E[R]$ directly.

Method of Indicators

It is possible to see random variables as a sum of indicators functions

$$R = I_{A_1} + ... + I_{A_n} \tag{39}$$

Then

$$E\left[R\right] = \sum_{j=1}^{n} E\left[I_{A_j}\right] = \sum_{j=1}^{n} P\left(A_j\right) \tag{40}$$

Hopefully

It is easier to calculate $P\left(A_j\right)$ than to evaluate $E\left[R\right]$ directly.

Method of Indicators

It is possible to see random variables as a sum of indicators functions

$$R = I_{A_1} + ... + I_{A_n} \tag{39}$$

Then

$$E[R] = \sum_{j=1}^{n} E\left[I_{A_j}\right] = \sum_{j=1}^{n} P(A_j) \tag{40}$$

Hopefully

It is easier to calculate $P(A_j)$ than to evaluate $E[R]$ directly.

Cinvestav

Example

A single unbiased die is tossed independently n times

- Let R_1 be the numbers of 1's.
- Let R_2 be the numbers of 2's.

Find $E[R_1 R_2]$

We can express each variable as

$$R_1 = I_{A_1} + ... + I_{A_n}$$
$$R_2 = I_{B_1} + ... + I_{B_n}$$

Thus

$$E[R_1 R_2] = \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[I_{A_i} I_{B_j}\right] \quad (41)$$

Example

A single unbiased die is tossed independently n times

- Let R_1 be the numbers of 1's.
- Let R_2 be the numbers of 2's.

Find $E\left[R_1 R_2\right]$

We can express each variable as

$$
\begin{aligned}
R_1 &= I_{A_1} + ... + I_{A_n} \\
R_2 &= I_{B_1} + ... + I_{B_n}
\end{aligned}
$$

Thus

$$E\left[R_1 R_2\right] = \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[I_{A_i} I_{B_j}\right] \tag{41}$$

Example

A single unbiased die is tossed independently n times

- Let R_1 be the numbers of 1's.
- Let R_2 be the numbers of 2's.

Find $E\left[R_1 R_2\right]$

We can express each variable as

$$R_1 = I_{A_1} + ... + I_{A_n}$$
$$R_2 = I_{B_1} + ... + I_{B_n}$$

Thus

$$E\left[R_1 R_2\right] = \sum_{i=1}^{n} \sum_{j=1}^{n} E\left[I_{A_i} I_{B_j}\right] \quad (41)$$

Next

Case 1 $i \neq j$ I_{A_i} and I_{B_j} are independent

$$E\left[I_{A_i} I_{B_j}\right] = E\left[I_{A_i}\right] E\left[I_{B_j}\right] = P\left(A_i\right) P\left(B_j\right) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad (42)$$

Case 2 $i = j$ A_i and B_i are disjoint

Thus, $I_{A_i} I_{B_i} = I_{A_i \cap B_i} = 0$

Thus

$$E\left[R_1 R_2\right] = \frac{n\left(n-1\right)}{36} \quad (43)$$

Next

Case 1 $i \neq j$ I_{A_i} and I_{B_j} are independent

$$E\left[I_{A_i} I_{B_j}\right] = E\left[I_{A_i}\right] E\left[I_{B_j}\right] = P\left(A_i\right) P\left(B_j\right) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad (42)$$

Case 2 $i = j$ A_i and B_i are disjoint

Thus, $I_{A_i} I_{B_i} = I_{A_i \cap B_i} = 0$

Thus

$$E\left[R_1 R_2\right] = \frac{n(n-1)}{36} \quad (43)$$

Next

Case 1 $i \neq j$ I_{A_i} and I_{B_j} are independent

$$E\left[I_{A_i} I_{B_j}\right] = E\left[I_{A_i}\right] E\left[I_{B_j}\right] = P\left(A_i\right) P\left(B_j\right) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36} \quad (42)$$

Case 2 $i = j$ A_i and B_i are disjoint

Thus, $I_{A_i} I_{B_i} = I_{A_i \cap B_i} = 0$

Thus

$$E\left[R_1 R_2\right] = \frac{n\left(n-1\right)}{36} \quad (43)$$

Cinvestav