Probability and Statistics

UG2, Core course, IIIT,H

Pawan Kumar

IIIT, Hyderabad

September 28, 2021

Outline

Examples involving independent trials...

Problem: coin toss

Suppose we flip a coin n times. Each coin flip is an independent trial with probability p of coming up heads. Write an expression for the following:

Examples involving independent trials...

Problem: coin toss

Suppose we flip a coin n times. Each coin flip is an independent trial with probability p of coming up heads. Write an expression for the following:

- P(n heads on n coin flips) $P(\underbrace{H\ H\ \ldots\ H}_{n \text{ times}})$

Examples involving independent trials...

Problem: coin toss

Suppose we flip a coin n times. Each coin flip is an independent trial with probability p of coming up heads. Write an expression for the following:

- P(n heads on n coin flips)
- P(n tails on n coin flips) $P(TT\cdots T)$

Examples involving independent trials...

Problem: coin toss

Suppose we flip a coin n times. Each coin flip is an independent trial with probability p of coming up heads. Write an expression for the following:

- P(n heads on n coin flips)
- P(n tails on n coin flips)
- P(first k heads, then $n-k$ tails) $P(\underbrace{H \cdots H}_{k}\underbrace{T \cdots T}_{n-k})$

Examples involving independent trials...

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

Problem: coin toss

Suppose we flip a coin n times. Each coin flip is an independent trial with probability p of coming up heads. Write an expression for the following:

- P(n heads on n coin flips)
- P(n tails on n coin flips)
- P(first k heads, then $n-k$ tails)
- P(exactly k heads on n coin flips)

$$= P(H_1 \cap H_2 \cap H_3 \cdots \cap H_n) = P(H_1)P(H_2)\cdots P(H_n)$$

because H_i's are independ.

$$= \underbrace{p \cdot p \cdots p}_{n \text{ times}} = p^n$$

$$P(T_1 \cap T_2 \cap \cdots \cap T_n) = P(T_1)P(T_2)\cdots P(T_n)$$

$$= P(H_1^c)P(H_2^c)\cdots P(H_n^c)$$

$$= (1-P(H_1))(1-P(H_2))\cdots(1-P(H_n))$$

$$= (1-p)^n$$

$$P(H_1 \cap \cdots \cap H_k \cap T_{k+1} \cap \cdots \cap T_n)$$

$$= P(H_1)\cdots P(H_k)P(T_{k+1})\cdots P(T_n)$$

$$= p^k (1-p)^{n-k}$$

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

2, 4, 6, 8 . . .

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution: Consider the partition of E into $E_1, E_2, \ldots E_k, \ldots$

where E_k = event that 1st head occurs on the $\underbrace{2k^{th}}_{\text{even!}}$ toss

$E = \bigcup_{k=1}^{\infty} E_k$, E_k's are mutually exclusive!

$P(E) = \sum_{k=1}^{\infty} P(E_k)$, $P(E_k) = (1-p)^{2k-1} \cdot p$

$$P(E) = \sum_{k=1}^{\infty} (1-p)^{2k-1} \cdot p = \frac{1-p}{2-p}$$

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

What are sample space and events in this problem?

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

What are sample space and events in this problem?

1. Sample space, $S =$ all possible infinite binary sequences of coin toss.

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

What are sample space and events in this problem?

1. Sample space, $S =$ all possible infinite binary sequences of coin toss.
2. Consider event H_1 : head on first toss.

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

What are sample space and events in this problem?

1. Sample space, $S =$ all possible infinite binary sequences of coin toss.
2. Consider event H_1 : head on first toss.
3. Consider event E : first head on even numbered toss.

We want to compute $P(E)$.

Biased Coin, Independence, Infinite Sample Space, Total Probability, Bayes Theorem...

Problem

A biased coin (with probability of obtaining a Head equal to $p > 0$ is tossed repeatedly and independently until the first head is observed. Compute the probability that the first head appears at an even numbered toss.

Solution to the problem...

What are sample space and events in this problem?

1. Sample space, $S =$ all possible infinite binary sequences of coin toss.
2. Consider event H_1 : head on first toss.
3. Consider event E : first head on even numbered toss.

We want to compute $P(E)$.
How do we solve problems like this?

Solution...

Solution...

Coin Toss Example...

Coin Toss Example...

H | T | H H T T H H H T . . . . T T ← experiment ends!

$n-1$ n

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

E_1: 1st toss is H

E_2: first two tosses are TH

E_3: first two tosses are TT

F_n: experiment completes on nth toss.

E_1, E_2, E_3 partition Ω (Sample sp.)

H _ _ _ _ _ _

T H

T T

Must have n atleast 2

For $n=2$: $p_2 = P(F_2) = (1-p)(1-p) = (1-p)^2$

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

1. Sample space, S: all possible infinite sequences of tosses

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

1. Sample space, S: all possible infinite sequences of tosses
2. event E_1 : first toss is H

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

1. Sample space, S: all possible infinite sequences of tosses
2. event E_1 : first toss is H
3. event E_2 : first two tosses are TH

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

1. Sample space, S: all possible infinite sequences of tosses
2. event E_1 : first toss is H
3. event E_2 : first two tosses are TH
4. event E_3 : first two tosses are TT

Coin Toss Example...

Problem

A coin for which P(Heads) $= p$ is tossed until two successive Tails are obtained. Find the probability that the experiment is completed on the nth toss.

Solution

What are sample space and events?

1. Sample space, S: all possible infinite sequences of tosses
2. event E_1 : first toss is H
3. event E_2 : first two tosses are TH
4. event E_3 : first two tosses are TT
5. event F_n : experiment completed on the nth toss.

Solution to problem in previous slide...part-1

For $n > 2$: $P(F_n | E_1) = P(F_{n-1})$

H | ... TT

as if exp. started from $n = 2$

TH | ... TT

TT ...

$P(F_n | E_2) = P(F_{n-2})$

$P(F_n | E_3) = 0$

Notation: $P(F_n) = p_n$

Using total probab. $n > 2$,

$$P(F_n) = P(F_n | E_1) P(E_1) + P(F_n | E_2) P(E_2) + \cancel{P(F_n | E_3)} P(E_3)$$

$$p_n = p_{n-1} \cdot p + p_{n-2} \cdot (1-p) p + 0$$

$$\boxed{p_n = p \, p_{n-1} + p(1-p) \, p_{n-2},}$$

$$\boxed{\begin{aligned} p_1 &= 0 \\ p_2 &= (1-p)^2 \end{aligned}}$$

Solution to problem in previous slide...part-2

Solution to problem in previous slide...part-3

Properties of conditional probabilities...

Properties of conditional probabilities...

Properties

For any events A,B, and E we have the following:

Properties of conditional probabilities...

Properties

For any events A,B, and E we have the following:

- $0 \leq P(A \cap E) \leq 1$

Properties of conditional probabilities...

$$P(A|E) = \frac{P(A \cap E)}{P(E)}$$

Properties

For any events A,B,and E we have the following:

- $0 \leq P(A \cap E) \leq 1$
- $P(A \mid E) = 1 - P(A^c \mid E)$

$$P(A^c|E) = \frac{P(A^c \cap E)}{P(E)}$$

$$P(A|E) + P(A^c|E) = \frac{1}{\cancel{P(E)}}\left[\underbrace{P(A \cap E) + P(A^c \cap E)}_{= \cancel{P(E)}}\right]$$

$$E = (A \cap E) \cup (A^c \cap E)$$

mutually exc.

$$P(E) = P(A \cap E) + P(A^c \cap E)$$

Properties of conditional probabilities...

Properties

For any events A,B,and E we have the following:

- $0 \leq P(A \cap E) \leq 1$
- $P(A \mid E) = 1 - P(A^c \mid E)$
- $P(A \cap B \mid E) = P(B \cap A \mid E)$

Properties of conditional probabilities...

Properties

For any events A,B,and E we have the following:

- $0 \leq P(A \cap E) \leq 1$
- $P(A \mid E) = 1 - P(A^c \mid E)$
- $P(A \cap B \mid E) = P(B \cap A \mid E)$
- $P(A \cap B \mid E) = P(B \mid E)P(A \mid B \cap E)$

Properties of conditional probabilities...

Properties

For any events A,B,and E we have the following:

- $0 \leq P(A \cap E) \leq 1$
- $P(A \mid E) = 1 - P(A^c \mid E)$
- $P(A \cap B \mid E) = P(B \cap A \mid E)$
- $P(A \cap B \mid E) = P(B \mid E)P(A \mid B \cap E)$
- $P(A \mid B \cap E) = \dfrac{P(B \mid A \cap E)P(A \mid E)}{P(B \mid E)}$

Scratch Space for Proving Conditional Probabilities...

Scratch Space for Proving Conditional Probabilities...

Outline

1. Conditional Probability, Bayes Theorem
2. Conditional Independence
3. Random Variables
4. Expectation
5. Saint Petersberg Paradox

Conditional Independence...

Conditional Independence...

Definition of conditional independence

Two events A and B are conditionally independent given E if

$$P(A \cap B|E) = P(A \mid E)P(B \mid E)$$

Conditional Independence...

Definition of conditional independence

Two events A and B are conditionally independent given E if

$$P(A \cap B|E) = P(A \mid E)P(B \mid E)$$

Fact on Conditional Independence

A and B independent does not mean that A and B are independent given E. That is,

$$\underbrace{P(A \cap B) = P(A)P(B)}_{} \not\Longrightarrow P(A \cap B \mid E) = P(A \mid E)P(B \mid E)$$

Quiz on independence...

Quiz on independence...

Quiz-1

Two events E and F are independent if

1. Knowing that F happens means that E can't happen

Quiz on independence...

Quiz-1

Two events E and F are independent if

1. Knowing that F happens means that E can't happen
2. Knowing that F happens doesn't change probability that E happened.

Quiz on independence...

Quiz-1

Two events E and F are independent if

1. Knowing that F happens means that E can't happen
2. Knowing that F happens doesn't change probability that E happened.

What is your answer?

Quiz on independence...

Quiz-1

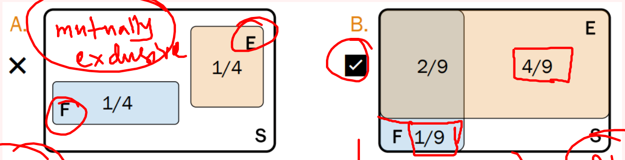
Two events E and F are independent if

1. Knowing that F happens means that E can't happen
2. Knowing that F happens doesn't change probability that E happened.

What is your answer?

Quiz-1

Are E and F independent in the following pictures (not to scale)?



mutually exclusive

F, E ind. $\Rightarrow P(E \cap F) = P(E)P(F) \Rightarrow E, F$ not ind.

$0 = \frac{1}{4} \cdot \frac{1}{4}$ ✗

$P(E \cap F) = 2/9$

$P(E) = 6/9, \; P(F) = \frac{3}{9}$

$P(E)P(F)$

$= \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$

$\Rightarrow E, F$ are ind.

Mutually Exclusive and Independent Events...

Mutually Exclusive and Independent Events...

Quiz

When are two events both mutually exclusive and independent?

More Problems on Independent Trials...

Problem: String-part 1

There are m strings that are hashed unequally into a hash table with n buckets. Each string hashed is an independent trial with probability p_i of getting hashed into bucket i.What is $P(E)$ if

- $E =$ bucket 1 has ≥ 1 string hashed into it?

More Problems on Independent Trials...

Problem: String-part 2

There are m strings that are hashed unequally into a hash table with n buckets. Each string hashed is an independent trial with probability p_i of getting hashed into bucket i.What is $P(E)$ if

- $E =$ at least 1 of buckets 1 to k has ≥ 1 string hashed into it?

Outline

Random Variables

Python $\leftarrow$ JIT

Random Variables

Examples of typed variables in C

In some languages, such as, C/C++. We have the concept of a typed variable:

Random Variables

Examples of typed variables in C

In some languages, such as, C/C++. we have the concept of a typed variable:

- int $i = 4$;
- float $x = 10$;
- char $y =$ 'x';

Random Variables

Examples of typed variables in C

In some languages, such as, C/C++. we have the concept of a typed variable:

- int i = 4;
- float x = 10;
- char y = 'x';

Examples of random variable

Let X denote the outputs after we roll a die, then

$$X = 3$$

$P(\{X = 3\})$

means that after rolling a die, we obtain 3 as output.

Random Variables

Examples of typed variables in C

In some languages, such as, C/C++. we have the concept of a typed variable:

- int i = 4;
- float x = 10;
- char y = 'x';

Examples of random variable

Let X denote the outputs after we roll a die, then

$$X = 3$$

means that after rolling a die, we obtain 3 as output.
Since the number that is going to be assigned to variable X is going to be random, it is called random variable.

Define Random Variable

Define Random Variable

$X : \{1, \ldots\} - 1, 2 \ldots$ $x = 1 \Leftarrow$ $x = 2 \Leftarrow$ $P\{x = -2\} \in [0, 1]$

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

P

Define Random Variable

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

- We usually denote random variables by capital letters: X, Y, ...

Define Random Variable

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

- We usually denote random variables by capital letters: X,Y,...
- Random variable is a function with domain S and co-domain $\mathbb{R}$

Define Random Variable

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

- We usually denote random variables by capital letters: X,Y,...
- Random variable is a function with domain S and co-domain $\mathbb{R}$
- The range of a random variable is the set of possible values of X

Define Random Variable

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

- We usually denote random variables by capital letters: X,Y,...
- Random variable is a function with domain S and co-domain $\mathbb{R}$
- The range of a random variable is the set of possible values of X

$X = \{0, 1, 2, 3, ..., 10\}$

Examples of Random Variables...

Find the range of the following random variables:

- I toss a coin 10 times. Let X be the number of heads I observe

Define Random Variable

Q^c Q

Definition of Random Variable

A random variable X is a function from the sample space to the real numbers.

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

range of $Y = \mathbb{N}$

countable set

- We usually denote random variables by capital letters: X,Y,...
- Random variable is a function with domain S and co-domain $\mathbb{R}$
- The range of a random variable is the set of possible values of X

$Y = \{1, 2, 3, \ldots\} \subset \mathbb{R}$

Examples of Random Variables...

Find the range of the following random variables:

- I toss a coin 10 times. Let X be the number of heads I observe
- I toss a coin until the first tail appears. Let Y be the total number of coin tosses

Quiz on Random Variable...

Quiz on Random Variable...

Quiz on Random Variable

Consider and Experiment: 3 coins are flipped. Let X be the number of tails. Answer the following:

Quiz on Random Variable...

Quiz on Random Variable

Consider and Experiment: 3 coins are flipped. Let X be the number of tails. Answer the following:

- What is the value of X for the outcomes?

Quiz on Random Variable...

Quiz on Random Variable

Consider and Experiment: 3 coins are flipped. Let X be the number of tails. Answer the following:

- What is the value of X for the outcomes?
  - $(H, H, H) \leftarrow$ 0
  - $(T, T, H) \leftarrow$ 2

Quiz on Random Variable...

Quiz on Random Variable

Consider and Experiment: 3 coins are flipped. Let X be the number of tails. Answer the following:

- What is the value of X for the outcomes?
  - (H, H, H)
  - (T, T, H)
- What is the event when $X = 2$?

Quiz on Random Variable...

Quiz on Random Variable

Consider and Experiment: 3 coins are flipped. Let X be the number of tails. Answer the following:

- What is the value of X for the outcomes?
  - (H, H, H)
  - (T, T, H)
- What is the event when $X = 2$? = event that we get 2 tails when 3 coins are tossed.
- What is $P(X = 2)$?

$$\{X=2\} = \{\underbrace{(TTH), (HTT), (THT)}\}$$

$$2 \cdot 2 \cdot 2 = 8$$

$$P(X=2) = \frac{3}{8}$$

Random Variables are Not Events!

Random Variables are Not Events!

Remarks on Random variables

- random variables are not events!

Random Variables are Not Events!

Remarks on Random variables

- random variables are not events!
- when a random variable is assigned a value, then it becomes event

$X = x$	Set of Outcomes	$P(X = k)$
$X = 0$	$\{(T,T,T)\}$	1/8
$X = 1$	$\{(H,T,T),(T,H,T),(T,T,H)\}$	3/8
$X = 2$	$\{(H,H,T),(H,T,H),(T,H,T)\}$	3/8
$X = 3$	$\{(H,H,H)\}$	1/8
$X \geq 4$	$\{\ \}$	0

Table: Consider an experiment where 3 coins are flipped, and X denotes number of heads

Discrete Random Variables...

Discrete Random Variables...

$\Sigma \quad \longleftrightarrow \quad \int$

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Types of Random Variables...

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Types of Random Variables...

There are three types of random variables:

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Types of Random Variables...

There are three types of random variables:

1. discrete random variables

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Types of Random Variables...

There are three types of random variables:

1. discrete random variables
2. continuous random variables

Discrete Random Variables...

Recall: countable sets

A set A is countable if either it is a finite set, or it can be put in 1-1 correspondence with set of natural numbers.

Discrete Random Variables

A random variable X is called discrete random variable, if its range is countable.

Types of Random Variables...

There are three types of random variables:

1. discrete random variables
2. continuous random variables
3. mixed random variables