Probability and Statistics

UG2, Core course, IIIT,H

Pawan Kumar

IIIT, Hyderabad

September 14, 2021

1. Conditional Probability, Bayes Theorem
2. Independence
3. Conditional Independence

1 Outline

Random Monty Hall Problem

This result depends crucially on the fact that Monty was always guaranteed to open a door with a goat behind it, regardless of what door you picked initially. That is, $P(E \mid H) = P(E \mid H^c)P(E \mid H)$. Now consider what would happen if Monty randomly opened a door we did not pick and it contained a goat. What is the probability that our first pick is correct, regardless of which specific door we picked?

Solution:

H : door 1 has a car

E : Monty reveals a goat door

$P(H) = \frac{1}{3}$, $P(H^c) = \frac{2}{3}$, $P(E|H) = 1$

$P(E|H^c) = \frac{1}{2}$

$P(H|E) = \frac{1}{2}$ ← Either you switch or not, probab. does not change.

1 Problem Similar to Monty Hall Problem...

Problem Similar to Monty Hall...

Suppose we have 3 cards identical in form except that both sides of the first card are colored red, both sides of the second card are colored black, and one side of the third card is colored red and the other side is colored black. The 3 cards are mixed up in a hat, and 1 card is randomly selected and put down on the ground. If the upper side of the chosen card is colored red, what is the probability that the other side is colored black?

R R B B R B

Pick one randomly

RR : Chose Card is Red-Red

BB : Chosen card is black-black

RB : Chosen card is red-black

R : upper side of chosen card is red

$$P(RB \mid R) = \frac{P(RB \cap R)}{P(R)}$$

$$= \frac{P(R \mid RB)\,P(RB)}{P(R \mid RR)\,P(RR) + P(R \mid RB)\,P(RB) + P(R \mid BB)\,P(BB)}$$

$$= \frac{\frac{1}{2}\cdot\frac{1}{3}}{1\cdot\frac{1}{3} + \frac{1}{2}\cdot\frac{1}{3} + 0\cdot\frac{1}{3}} = \frac{1}{3}$$

1 Some Problems in Probability...

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls.

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls. That night, a woman gives birth to a child, and the child is placed in the nursery.

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls. That night, a woman gives birth to a child, and the child is placed in the nursery. On January 2, a statistician conducts a survey

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls. That night, a woman gives birth to a child, and the child is placed in the nursery. On January 2, a statistician conducts a survey and selects a child at random from the nursery (including the newborn and every child from January 1).

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls. That night, a woman gives birth to a child, and the child is placed in the nursery. On January 2, a statistician conducts a survey and selects a child at random from the nursery (including the newborn and every child from January 1). The selected child is a boy.

Problem: Boy/Girl Problem

On the morning of January 1, a hospital nursery has 3 boys and some number of girls. That night, a woman gives birth to a child, and the child is placed in the nursery. On January 2, a statistician conducts a survey and selects a child at random from the nursery (including the newborn and every child from January 1). The selected child is a boy. What is the probability the child born on January 1 was a boy?

Solution:

Jan-1 : 3 boys + k girls

$\downarrow$ $+1$ ← what is the prob that this is a boy. what is the incr. is

Jan-2 : Statistician randomly selects a child, which happens to be a boy.

Example

3 boys, 3 girls $\xrightarrow{+1}$

- $\frac{1}{2}$ → 3 girls, 4 boys $\quad \frac{1}{2} \cdot \frac{4}{7}$
- $\frac{1}{2}$ → 4 girls, 3 boys $\quad \frac{1}{2} \cdot \frac{3}{7}$

Define event :

B : select boy

N : new born is a boy

$$B = \underbrace{(B \cap N)}\ \cup\ \underbrace{(B \cap N^c)}$$

mutually exclusive

$$P(B) = P(B \cap N) + P(B \cap N^c) = \frac{1}{2} \cdot \frac{4}{7} + \frac{1}{2} \cdot \frac{3}{7} = \frac{7}{14} = \frac{1}{2}.$$

$P(N|B) \leftarrow$ we want.

$$P(N|B) = \frac{P(N \cap B)}{P(B)} = \frac{\frac{1}{2} \cdot 4/7}{1/2} = \frac{4}{7} = 57.1$$

3

g girls, 3 boys

6

10

+1

$\xrightarrow{1/2}$ g girls, 4 boys, $\frac{1}{2} \cdot \frac{4}{g+4}$

$\xrightarrow{1/2}$ g+1 girls, 3 boys, $\frac{1}{2} \cdot \frac{3}{g+4}$

$$P(N \mid B) = \frac{P(N \cap B)}{P(B)} = \frac{\frac{1}{2} \cdot \frac{4}{\cancel{g+4}}}{\frac{3.5}{\cancel{g+4}}} = \frac{4}{7} = 57\%$$

same

Note

$$P(B) = \frac{7}{2(g+4)} = \frac{3.5}{g+4}$$

1 Probability of Seeing a Car...

Problem: Probability of seeing a car

If the probability of seeing a car on the highway in 30 minutes is 0.95,

Problem: Probability of seeing a car

If the probability of seeing a car on the highway in 30 minutes is 0.95, what is the probability of seeing a car on the highway in 10 minutes? (assume a constant default probability)

Solution:

$p = P(\text{no car in 10 min})$

$P(\text{car in 10 min}) = 1 - p = 1 - \sqrt[3]{0.05}$

$P(\text{no car in 20 min}) = p \cdot p = p^2$

$P(\text{" " 30 min}) = p \cdot p \cdot p = p^3$

$P(\text{car in 30 min}) = 1 - p^3 = 0.95$

$\Rightarrow p = \sqrt[3]{1 - 0.95} = \sqrt[3]{0.05}$

1 Skewed Die Problem...

Problem: Skewed Die Problem

A standard dice has the 6 showing with a probability of 1/6, which is the same as every other number. A loaded dice has the 6 showing with 1/2 the probability of the other numbers. What is the probability of rolling a 6?

skewed

Solution:

$p(\{6\}) = \frac{p}{2}$,

p = prob. of ~~other~~ numbers other than 6.

$p(\{1\}) = p$

$p(\{2\}) = p$

$\vdots$

$p(\{5\}) = p$

$\frac{p}{2} + 5 \cdot p = 1$

solve for p.

1 Spam Email Problems...

1 Spam Email Problems...

Problem

It is estimated that 50% of emails are spam emails.

1 Spam Email Problems...

Problem

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox.

Problem

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails,

Problem

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%.

Problem

It is estimated that 50% of emails are spam emails. Some software has been applied to filter these spam emails before they reach your inbox. A certain brand of software claims that it can detect 99% of spam emails, and the probability for a false positive (a non-spam email detected as spam) is 5%. Now if an email is detected as spam, then what is the probability that it is in fact a non-spam email?

A = email detected as spam

B = email is spam

B^c = email is not spam

$P(B) = P(B^c) = \frac{1}{2}$

$P(A|B) = 0.99$

$P(A|B^c) = 0.05$

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

$$= \frac{0.05 \cdot \frac{1}{2}}{0.99 \times 0.5 + 0.05 \times 0.5}$$

1 Scratch Space for Spam Email Problems...

2 Outline

2 Definition of Independence

Definition of independent events

Two events E and F are defined to be independent if

$P(E,F)$ "Ross"

$P(E,F) = P(E)P(F)$

Definition of independent events

Two events E and F are defined to be independent if

$$P(E \cap F) = P(E)P(F).$$

Definition of independent events

Two events E and F are defined to be independent if

$$P(E \cap F) = P(E)P(F).$$

Otherwise, E and F are called dependent events.

2 Definition of Independence

Definition of independent events

Two events E and F are defined to be independent if

$$P(E \cap F) = P(E)P(F).$$

Otherwise, E and F are called dependent events.
If E and F are independent, then

Definition of independent events

Two events E and F are defined to be independent if

$$P(E \cap F) = P(E)P(F).$$

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

Otherwise, E and F are called dependent events.
If E and F are independent, then

$$P(E \mid F) = P(E).$$

$$= \frac{P(E)\cancel{P(F)}}{\cancel{P(F)}}$$

$$= P(E)$$

Solution:

Example

Roll two 6-sided dice, yielding values D_1 and D_2. Let us consider the following events:

- $E : D_1 = 1$
- $F : D_2 = 6$
- $G : D_1 + D_2 = 5$
- That is, $G = \{(1,4), (2,3), (3,2), (4,1)\}$

1. Are E and F independent?

$E \cap G = \{(1,4)\}$

Example

Roll two 6-sided fair dice, yielding values $(D_1$ and D_2. Let us consider the following events:

- $E : D_1 = 1$
- $F : D_2 = 6$
- $G : D_1 + D_2 = 5$
- That is, $G = \{(1,4), (2,3), (3,2), (4,1)\}$

1. Are E and F independent?
2. Are E and G independent?

$$P(E \cap G) = P(E)\, P(G).$$

Can two events A, B be both mutually exclusive & independent?

$$P(E \cap G) = \frac{1}{36}$$

$$P(E) = \frac{1}{6}$$

$$P(G) = \frac{4}{36} = \frac{1}{9}$$

$$P(E) \cdot P(G) = \frac{1}{6} \cdot \frac{1}{9} = \frac{1}{54} \neq \frac{1}{36} = P(E \cap G)$$

$\Rightarrow$ E & G are not independent events.

2 General Definition of Independence...

E, F

Definition of independence for 3 events

Three events E, F, and G are independent if

- $P(E \cap F \cap G) = P(E)P(F)P(G)$

Definition of independence for 3 events

Three events E, F, and G are independent if

- $P(E \cap F \cap G) = P(E)P(F)P(G)$ $\leftarrow$ not enough
- $P(E \cap F) = P(E)P(F)$ $\leftarrow$

Definition of independence for 3 events

Three events E, F, and G are independent if

- $P(E \cap F \cap G) = P(E)P(F)P(G)$
- $P(E \cap F) = P(E)P(F)$
- $P(E \cap G) = P(E)P(G)$

2 General Definition of Independence...

Definition of independence for 3 events

Three events E, F, and G are independent if

- $P(E \cap F \cap G) = P(E)P(F)P(G)$
- $P(E \cap F) = P(E)P(F)$
- $P(E \cap G) = P(E)P(G)$
- $P(F \cap G) = P(F)P(G)$

General Definition for many events

The n events $E_1, E_2, \ldots, E_n$ are independent if

for $r = 1, \ldots, n$:

for every subset $E_1, E_2, \ldots, E_r$:

$$P(E_1 \cap E_2 \cap \cdots \cap E_r) = P(E_1)P(E_2) \cdots P(E_r)$$

2 Example of general independence...

Question

Each roll of 6-sided die is an independent trial. Two rolls with output D_1 and D_2. Consider the following events:

- $E : D_1 = 1$
- $F : D_2 = 6$
- $G : D_1 + D_2 = 7$
- $G = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

Answer the following:

1. Are E and F independent?
2. Are E and G independent?
3. Are F and G independent?
4. Are E, F, G independent?

2 Solution to problem on previous slide...

We have $E : D_1 = 1 \quad F : D_2 = 6 \quad G : D_1 + D_2 = 7$

2 Independent Trials...

Definition of Independent Trials

A set of n trials are called independent trials if

1st trial 2nd trial . . . .

{H, T} {H, T}

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

2 Independent Trials...

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

Examples of Independent Trials...

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

Examples of Independent Trials...

- Flip a coin n times

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

Examples of Independent Trials...

- Flip a coin n times
- Roll a die n times

2 Independent Trials...

Definition of Independent Trials

A set of n trials are called independent trials if

1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

Examples of Independent Trials...

- Flip a coin n times
- Roll a die n times
- Send a multiple choice survey to n people

2 Independent Trials...

Definition of Independent Trials

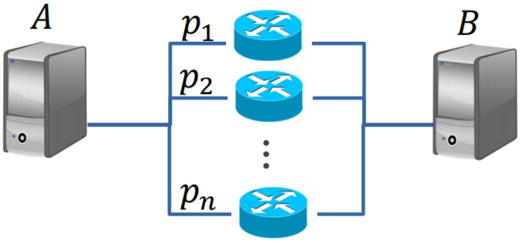
A set of n trials are called independent trials if

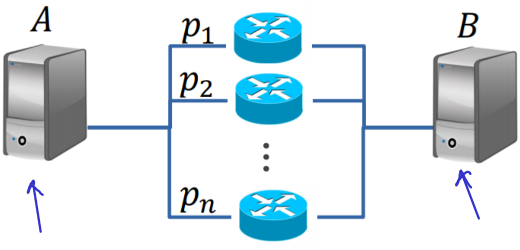
1. Each of the n trials have same set of possible outcomes
2. The trials are independent if an event in one subset of trials is independent of events in other subsets of trials

Examples of Independent Trials...

- Flip a coin n times
- Roll a die n times
- Send a multiple choice survey to n people
- Send n web requests to k different servers

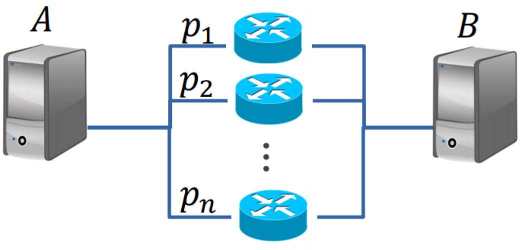
2 Examples involving independent trials...


A
p_1
p_2
p_n
B



Problem

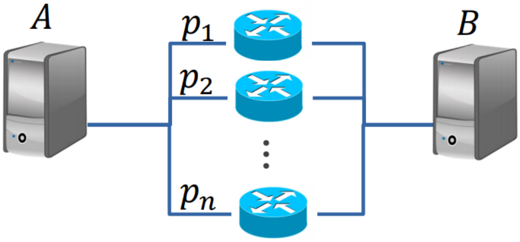
Consider the parallel network above:



Problem

Consider the parallel network above:

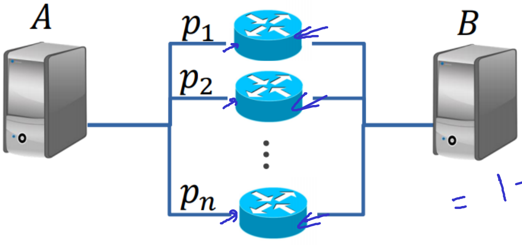
- n independent routers, each with probability p_i of functioning, where $1 \leq i \leq n$



Problem

Consider the parallel network above:

- n independent routers, each with probability p_i of functioning, where $1 \leq i \leq n$
- $E =$ functional path from A to B exists.



Problem

Consider the parallel network above:

- n independent routers, each with probability p_i of functioning, where $1 \leq i \leq n$
- E = functional path from A to B exists.

What is $P(E)$?

$P(E) = P(\text{at least one router works})$

$= 1 - P(\text{no router works})$

$= 1 - P(R_1 \text{ fails} \cap R_2 \text{ fails} \cap \dots \cap R_n \text{ fails})$

$= 1 - P(R_1 \text{ fails}) P(R_2 \text{ fails}) \dots P(R_n \text{ fails})$

$= 1 - (1-p_1)(1-p_2) \dots (1-p_n)$

$\{H, T\}$ $\{H, T\}$

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:

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)

H_i: obtaining a head in trial i

$$P(H_1 \cap H_2 \cap \cdots \cap H_n) = P(H_1)\,P(H_2)\cdots P(H_n) = p \cdot p \cdots p = p^n$$

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) $\leftarrow$ $(1-p)^n$.

$$P(H_1^c \cap H_2^c \cdots \cap H_n^c)$$
$$= P(H_1^c) \cdot P(H_2^c) \cdots P(H_n^c)$$
$$= (1-P(H_1))(1-P(H_2)) \cdots (1-P(H_n))$$
$$= (1-p) \cdots (1-p)$$
$$= (1-p)^n$$

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)

$$P(\underbrace{H_1 \cap H_2 \ldots \cap H_k}_{p^k} \cap \underbrace{H_{k+1}^c \cap H_{k+2}^c \ldots \cap H_n^c}_{})$$

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

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) ← Try