COL 202: DISCRETE MATHEMATICAL STRUCTURES

LECTURE 39

INDEPENDENCE (CONTD.), BIRTHDAY PARADOX, AND RANDOM VARIABLES

APR 23, 2024 | ROHIT VAISH

INDEPENDENCE

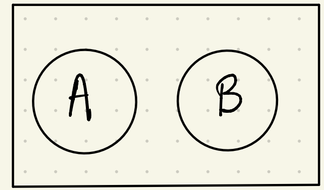
Two events A and B are independent if

$$\text{either} \quad \Pr(A \mid B) = \Pr(A)$$

$$\text{or} \quad \Pr(B) = 0 .$$

INDEPENDENCE

Disjoint $\not\Rightarrow$ independent

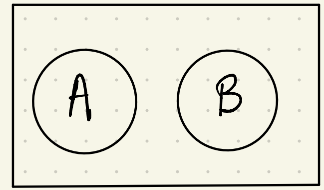


$P(A|B) = 0$

$P(A) \neq 0$

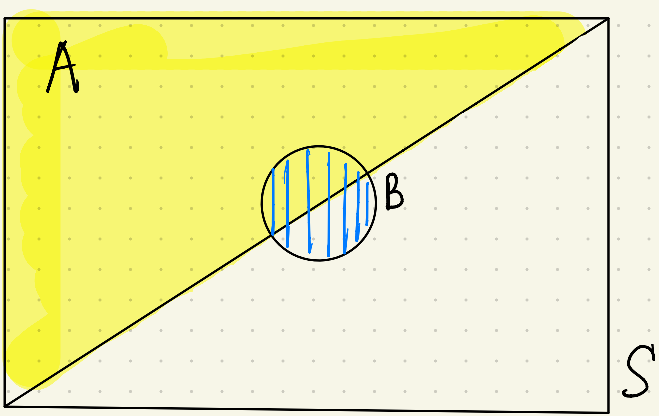
Independence

Disjoint $\not\Rightarrow$ independent



$$P(A|B) = 0$$

$$P(A) \neq 0$$



$$Pr(A|B) = Pr(A)$$

Product rule for independent events

If A and B are independent, then

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B).$$

BIAS $\longrightarrow$ FAIRNESS

Given two independent and biased coins

$$Pr(H) = p$$
$$Pr(T) = 1-p$$

How to simulate a fair coin toss ?

BIAS $\longrightarrow$ FAIRNESS

Given two independent and biased coins

$$Pr(H) = p$$
$$Pr(T) = 1-p$$

How to simulate a fair coin toss?

$$Pr(1^{st} \text{ coin is } H \mid \text{different outcomes})$$
$$=$$
$$Pr(1^{st} \text{ coin is } T \mid \text{different outcomes})$$
$$=$$
$$1/2.$$

BIAS $\longrightarrow$ FAIRNESS

Given two independent and biased coins

$$Pr(H) = p$$
$$Pr(T) = 1 - p$$

How to simulate a fair coin toss ?

1. Toss both coins

2. If same outcomes , repeat Step 1

3. If different outcomes, follow coin 1.

MUTUAL INDEPENDENCE

Events $A_1, A_2, \ldots, A_n$ are mutually independent if

$\forall i$ and $\forall J \subseteq \{1, 2, \ldots, n\} \setminus \{i\}$

$$\Pr\left(A_i \,\middle|\, \bigcap_{j \in J} A_j\right) = \Pr(A_i)$$

or $$\Pr\left(\bigcap_{j \in J} A_j\right) = 0 .$$

MUTUAL INDEPENDENCE

Events $A_1, A_2, \cdots, A_n$ are mutually independent if

$$\forall J \subseteq \{1, 2, \cdots, n\}$$

$$\Pr\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \Pr(A_j).$$

MUTUAL INDEPENDENCE

E.g., $n = 3$

A_1, A_2, A_3 are mutually independent if

$$\Pr(A_1 \cap A_2) = \Pr(A_1) \cdot \Pr(A_2)$$

$$\Pr(A_2 \cap A_3) = \Pr(A_2) \cdot \Pr(A_3)$$

$$\Pr(A_1 \cap A_3) = \Pr(A_1) \cdot \Pr(A_3)$$

$$\Pr(A_1 \cap A_2 \cap A_3) = \Pr(A_1) \cdot \Pr(A_2) \cdot \Pr(A_3).$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

A_1 : Event that coin 1 matches coin 2

A_2 : ———————— 2 ———————— 3

A_3 : ———————— 3 ———————— 1

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

A_1 : Event that coin 1 matches coin 2

A_2 : ———————— 2 ———————— 3

A_3 : ———————— 3 ———————— 1

Are A_1, A_2, A_3 mutually independent?

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) =$$

Mutual Independence

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\Pr(A_i \cap A_j) =$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\Pr(A_i \cap A_j) = \Pr(HHH) + \Pr(TTT)$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\Pr(A_i \cap A_j) = \Pr(HHH) + \Pr(TTT)$$
$$= \frac{1}{8} + \frac{1}{8}$$
$$= \frac{1}{4}$$

Mutual Independence

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\Pr(A_i \cap A_j) = \Pr(HHH) + \Pr(TTT)$$

$$= \frac{1}{8} + \frac{1}{8}$$

$$= \frac{1}{4} = \Pr(A_i) \cdot \Pr(A_j)$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_i) = \Pr(HH) + \Pr(TT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$\Pr(A_i \cap A_j) = \Pr(HHH) + \Pr(TTT)$$

$$= \frac{1}{8} + \frac{1}{8}$$

$$= \frac{1}{4} = \Pr(A_i) \cdot \Pr(A_j)$$

Not done yet!

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_1 \cap A_2 \cap A_3) =$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{4}$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{4}$$

$$\neq \frac{1}{8} = \Pr(A_1) \cdot \Pr(A_2) \cdot \Pr(A_3).$$

MUTUAL INDEPENDENCE

E.g., Flip 3 fair mutually independent coins

$$\Pr(A_1 \cap A_2 \cap A_3) = \frac{1}{4}$$

$$\neq \frac{1}{8} = \Pr(A_1) \cdot \Pr(A_2) \cdot \Pr(A_3).$$

NOT mutually independent

Pairwise Independence

Events $A_1, A_2, \ldots, A_n$ are pairwise independent if

$\forall\, i, j$ such that $i \neq j$

$$\Pr(A_i \cap A_j) = \Pr(A_i) \cdot \Pr(A_j).$$

PAIRWISE INDEPENDENCE

Events $A_1, A_2, \dots, A_n$ are pairwise independent if

$\forall\, i, j$ such that $i \neq j$

$$\Pr(A_i \cap A_j) = \Pr(A_i) \cdot \Pr(A_j).$$

Mutual independence $\Longrightarrow$ Pairwise independence

BIRTHDAY PARADOX

Birthday Paradox

Jan	24 23 11
Feb	5
Mar	11
Apr	7 2
May	
Jun	13 4 12 Yay!
Jul	15 21 2 5 6
Aug	17 12
Sep	6 21 19 22 29
Oct	5 24 20 25 21
Nov	8 24
Dec	18 24 10 28

BIRTHDAY PARADOX

N birthdays (e.g., $N = 365$)

M people (e.g., $M = 80$)

$M \leq N$

BIRTHDAY PARADOX

N birthdays (e.g., $N = 365$)

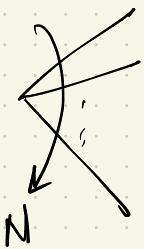
M people (e.g., $M = 80$)

$M \leq N$

What is the probability that 2 or more people have the same birthday?

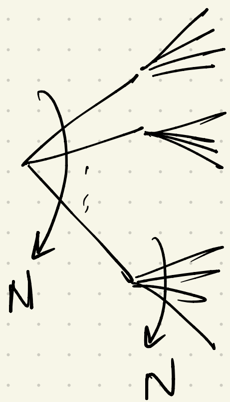
BIRTHDAY PARADOX

1^{st}
person

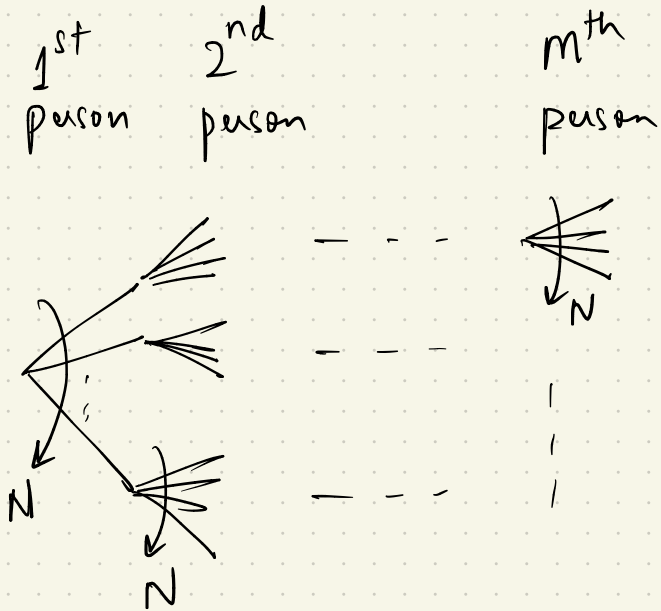


BIRTHDAY PARADOX

1^{st} person 2^{nd} person

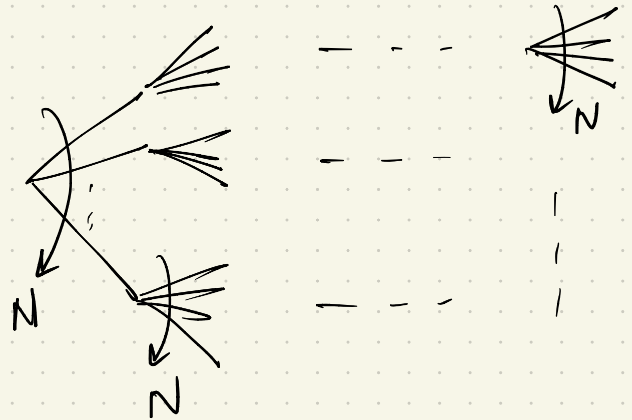


BIRTHDAY PARADOX



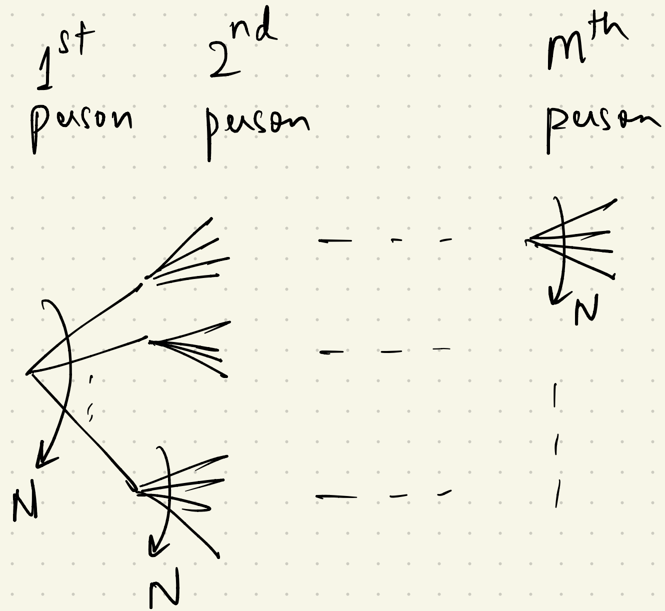
BIRTHDAY PARADOX

1st person 2nd person m^{th} person



$$S = \{(b_1, b_2, \dots, b_m) \mid 1 \le b_i \le N\}$$

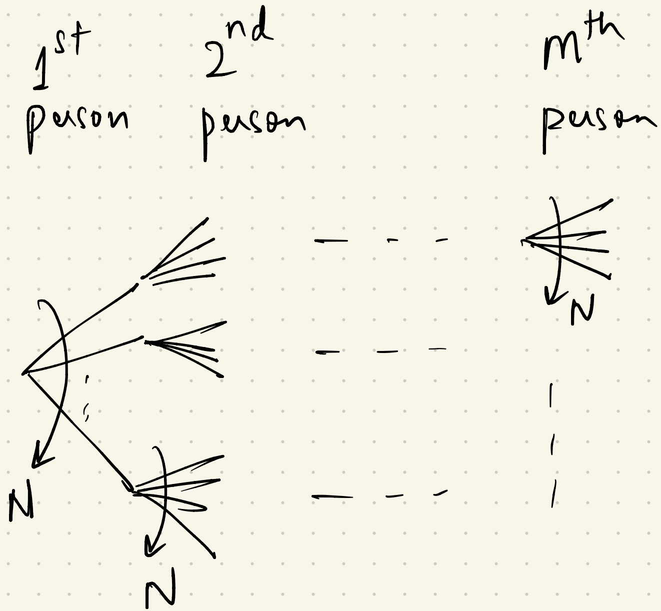
BIRTHDAY PARADOX



$$S = \{(b_1, b_2, \cdots, b_m) \mid 1 \leq b_i \leq N\}$$

$$|S| =$$

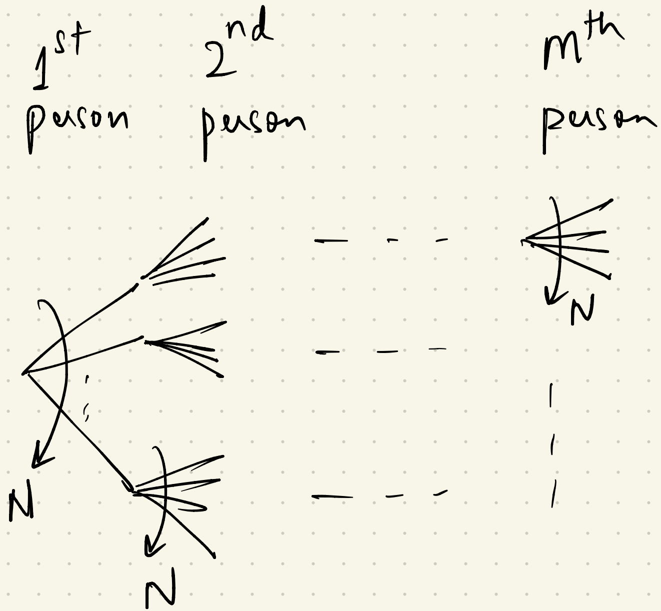
BIRTHDAY PARADOX



$$S = \{(b_1, b_2, \ldots, b_m) \mid 1 \le b_i \le N\}$$

$$|S| = N^m$$

BIRTHDAY PARADOX

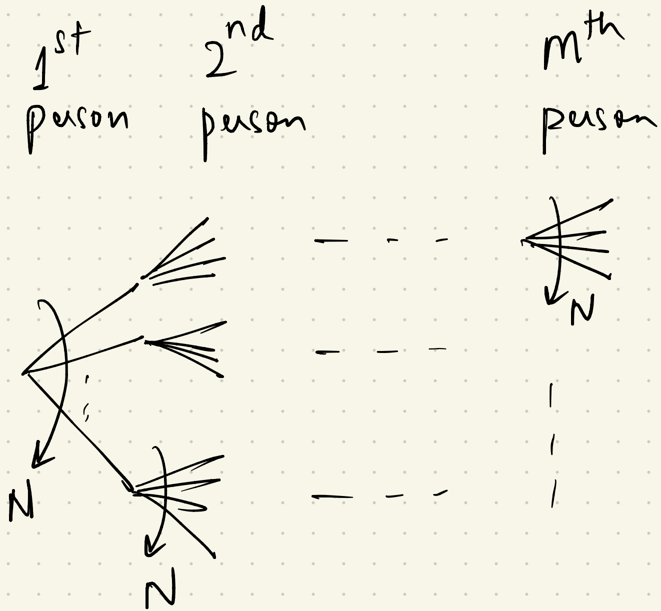


$$S = \{ (b_1, b_2, \cdots, b_m) \mid 1 \le b_i \le N \}$$

$$|S| = N^m$$

$$Pr((b_1, b_2, \ldots, b_m)) =$$

BIRTHDAY PARADOX



$$S = \{ (b_1, b_2, \dots, b_m) \mid 1 \le b_i \le N \}$$

$$|S| = N^m$$

$$Pr((b_1, b_2, \dots, b_m)) = \left(\frac{1}{N}\right)^m$$

assuming all birthdays are equally likely and mutually independent

BIRTHDAY PARADOX

No. of outcomes with all birthdays distinct

=

BIRTHDAY PARADOX

No. of outcomes with all birthdays distinct

$$= N \cdot (N-1) \cdot (N-2) \cdot \; \cdots \; \cdot (N-M+1)$$

BIRTHDAY PARADOX

No. of outcomes with all birthdays distinct

$$= N \cdot (N-1) \cdot (N-2) \cdot \;\cdots\; \cdot (N-M+1)$$

$$= \frac{N!}{(N-M)!}$$

BIRTHDAY PARADOX

No. of outcomes with all birthdays distinct

$$= N \cdot (N-1) \cdot (N-2) \cdot \; \cdots \; \cdot (N-M+1)$$

$$= \frac{N!}{(N-M)!}$$

Pr (all birthdays differ) =

BIRTHDAY PARADOX

No. of outcomes with all birthdays distinct

$$= N \cdot (N-1) \cdot (N-2) \cdot \;---\; \cdot (N-m+1)$$

$$= \frac{N!}{(N-m)!}$$

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

BIRTHDAY PARADOX

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

BIRTHDAY PARADOX

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

$$\left[\text{Stirling's formula} : \; N! \sim \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \right]$$

BIRTHDAY PARADOX

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

$$\left[\text{Stirling's formula}: \; N! \sim \sqrt{2\pi N}\left(\frac{N}{e}\right)^N \right]$$

$$\approx e^{\left(N-m+\frac{1}{2}\right)\ln\left(\frac{N}{N-m}\right) - m}$$

BIRTHDAY PARADOX

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

$$\left[\text{Stirling's formula} : \; N! \sim \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \right]$$

$$\approx e^{\left(N-m+\frac{1}{2}\right)\ln\left(\frac{N}{N-m}\right) - m}$$

$N = 365 \qquad m = 23 \qquad \Pr(\text{all diff}) = 0.49\ldots$

BIRTHDAY PARADOX

$$\Pr(\text{all birthdays differ}) = \frac{N!}{(N-m)!} \cdot \frac{1}{N^m}$$

$$\left[\text{Stirling's formula : } N! \sim \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \right]$$

$$\approx e^{(N-m+\frac{1}{2}) \ln\left(\frac{N}{N-m}\right) - m}$$

$N = 365$ $\quad m = 23$ $\quad \Pr(\text{all diff}) = 0.49\ldots$

$N = 365$ $\quad m = 85$ $\quad \Pr(\text{all diff}) = 0.00006$

$\rightarrow$ |COL 202| in Spring 2024

BIRTHDAY PARADOX

What about M and N that can grow arbitrarily?

Birthday Paradox

What about M and N that can grow arbitrarily?

$$\Pr(\text{all differ}) \text{ equals } \frac{1}{2} \text{ when } M = \sqrt{2N \ln 2}$$

$$= 1.177\sqrt{N}$$

BIRTHDAY PARADOX

What about m and N that can grow arbitrarily?

$$\Pr(\text{all differ}) \text{ equals } \frac{1}{2} \text{ when } m = \sqrt{2N\ln 2}$$

$$= 1.177\sqrt{N}$$

Birthday Principle: Allocating $1.177\sqrt{N}$ objects to N bins uniformly at random has at least 50% chance of a "collision".

BIRTHDAY PARADOX

What about M and N that can grow arbitrarily?

$$\Pr(\text{all differ}) \text{ equals } \frac{1}{2} \text{ when } M = \sqrt{2N \ln 2}$$

$$= 1.177\sqrt{N}$$

Birthday Principle: Allocating $1.177\sqrt{N}$ objects to N bins uniformly at random has at least 50% chance of a "collision".

Applications in CS: hashing, cryptography, error correction, ...

RANDOM VARIABLES

Events v/s Random Variables

$\downarrow$ either happens or doesn't

$\downarrow$ how much, how many ...

RANDOM VARIABLES

A random variable R is a function that maps each outcome to a real number.

$$R : \underbrace{S}_{\text{Sample space}} \longrightarrow \mathbb{R}$$

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

$R :=$ Number of heads

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

$R :=$ Number of heads

$R(H, T, H) = 2$, $R(T, T, T) = 0$

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

$R :=$ Number of heads

$R(H,T,H) = 2$, $R(T,T,T) = 0$

$$M = \begin{cases} 1 & \text{if all coins match} \\ 0 & \text{otherwise} \end{cases}$$

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

$R :=$ Number of heads

$R(H,T,H) = 2$, $R(T,T,T) = 0$

$$M = \begin{cases} 1 & \text{if all coins match} \\ 0 & \text{otherwise} \end{cases}$$

$M(H,T,H) = 0$, $M(T,T,T) = 1$

RANDOM VARIABLES

E.g., Toss 3 coins (fair and mutually independent)

$R :=$ Number of heads

$R(H,T,H) = 2$, $R(T,T,T) = 0$

$$M = \begin{cases} 1 & \text{if all coins match} \\ 0 & \text{otherwise} \end{cases}$$

range $\{0,1\}$

indicator/Bernoulli/characteristic r.v.

$M(H,T,H) = 0$, $M(T,T,T) = 1$

RANDOM VARIABLES

HHH	TTT
HHT	HTT
HTH	THT
THH	TTH

RANDOM VARIABLES

HHH TTT $m=1$

HHT HTT

HTH THT $m=0$

THH TTH

RANDOM VARIABLES

$R = 0$

HHH TTT

$M = 1$

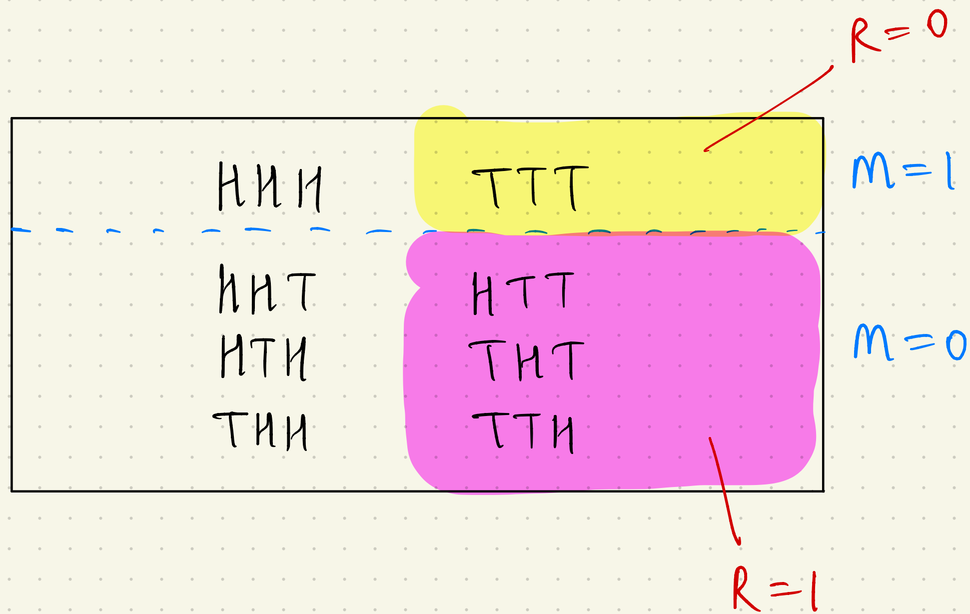
HHT HTT

HTH THT

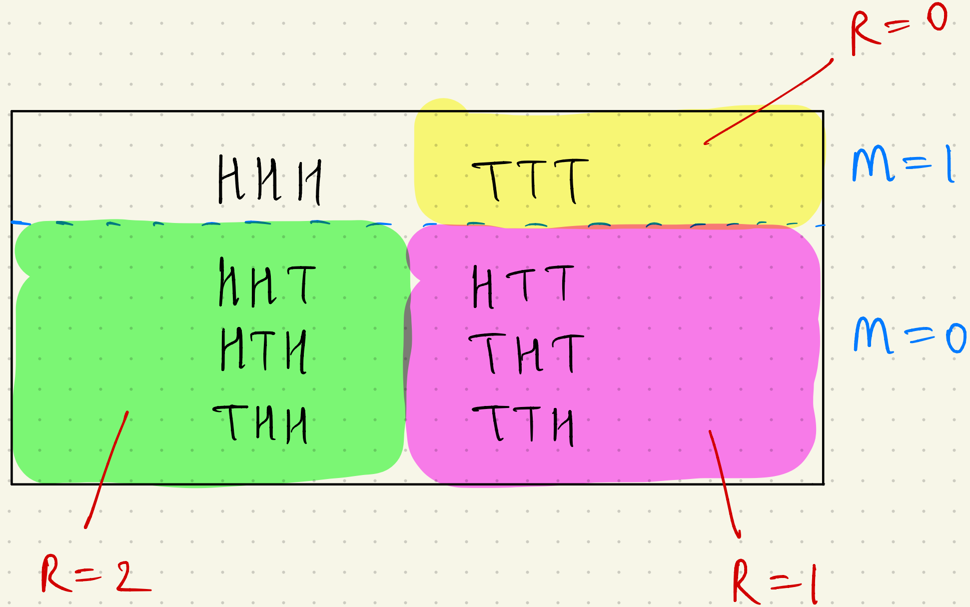
THH TTH

$M = 0$

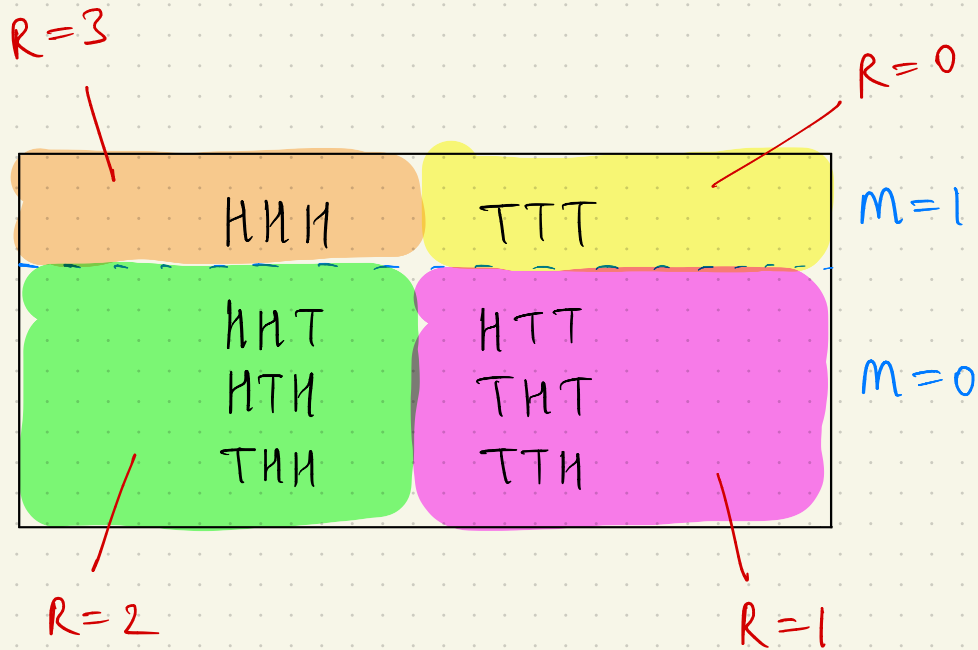
RANDOM VARIABLES



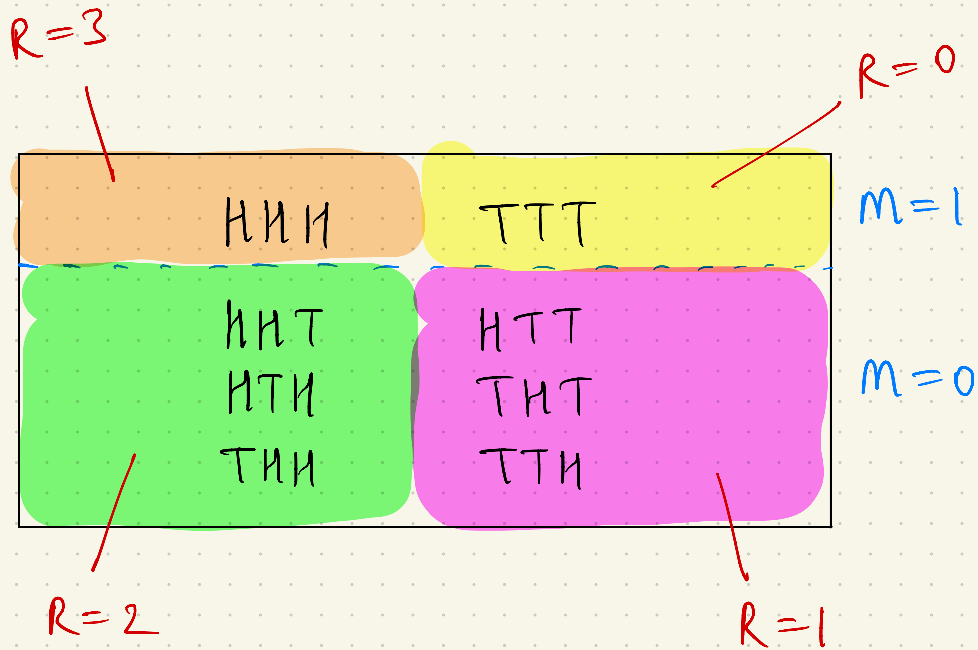
RANDOM VARIABLES



RANDOM VARIABLES



RANDOM VARIABLES



Each "block" of the partition is an event

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

RANDOM VARIABLES

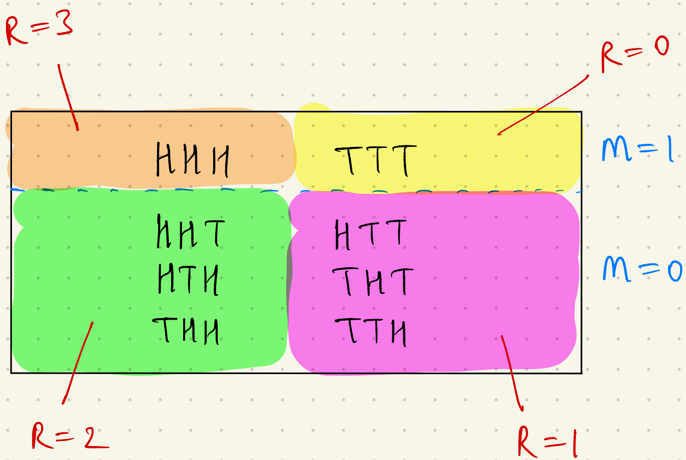
$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R = x) = \sum_{w \,:\, R(w) = x} \Pr(w)$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

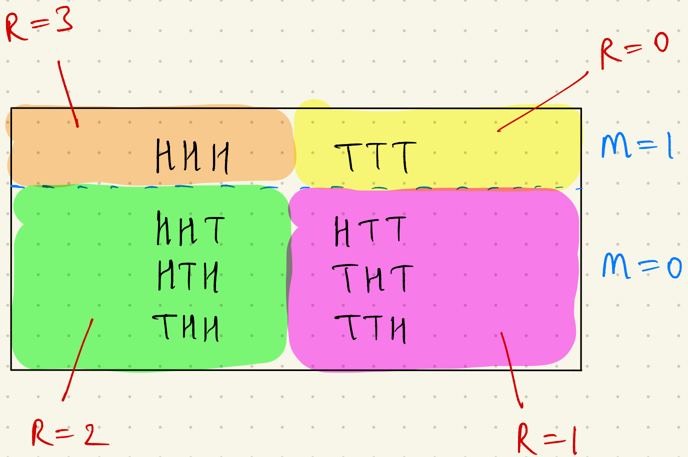
$$\Pr(R = x) = \sum_{w\,:\,R(w) = x} \Pr(w)$$



RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R = x) = \sum_{w \,:\, R(w) = x} \Pr(w)$$

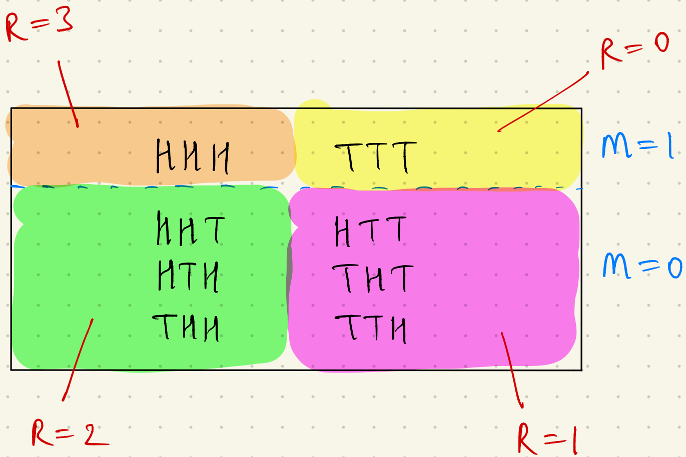


$\Pr(R = 1) =$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w\,:\,R(w)=x} \Pr(w)$$

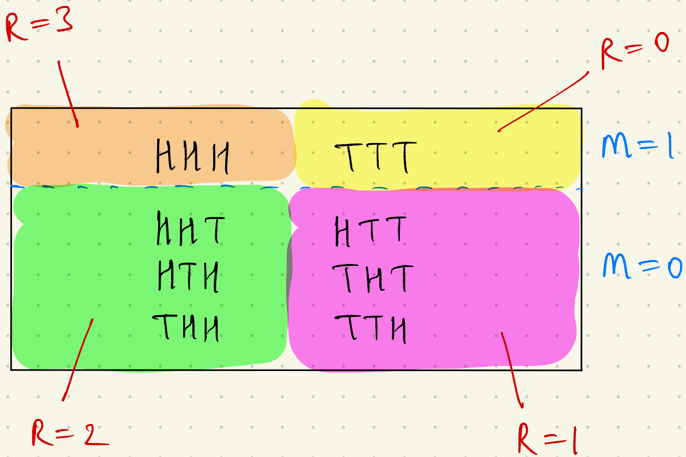


$$\Pr(R=1) = \Pr(HTT) + \Pr(THT) + \Pr(TTH)$$
$$= 3/8.$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w\,:\,R(w)=x} \Pr(w)$$



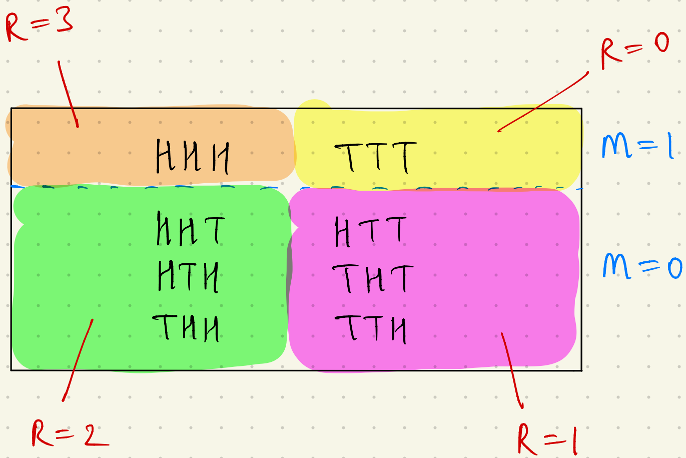
$$\Pr(R=1) = \Pr(HTT) + \Pr(THT) + \Pr(TTH)$$
$$= 3/8.$$

$\Pr(M=1) =$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R = x) = \sum_{w \,:\, R(w) = x} \Pr(w)$$



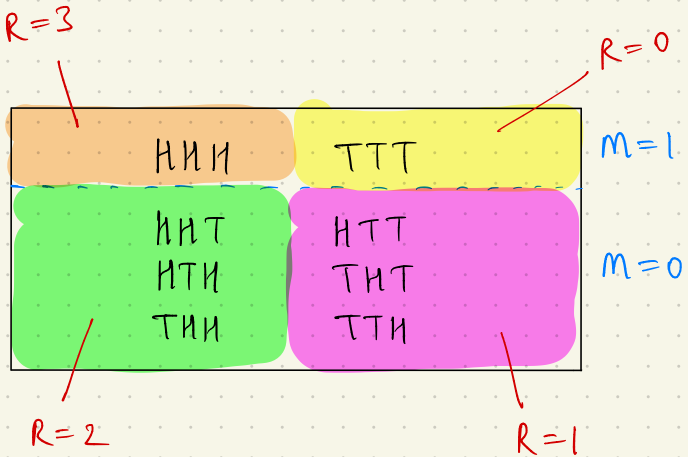
$$\Pr(R = 1) = \Pr(HTT) + \Pr(THT) + \Pr(TTH)$$
$$= 3/8.$$

$$\Pr(M = 1) = \Pr(HHH) + \Pr(TTT)$$
$$= 1/4$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R = x) = \sum_{w \,:\, R(w) = x} \Pr(w)$$

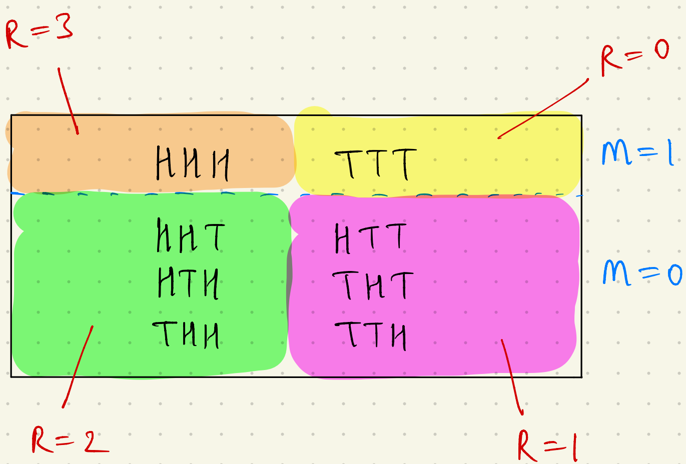


$$\Pr(R \geq 2) =$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w:\, R(w)=x} \Pr(w)$$



$$\Pr(R \geq 2) = \sum_{i=2}^{\infty} \Pr(R=i)$$

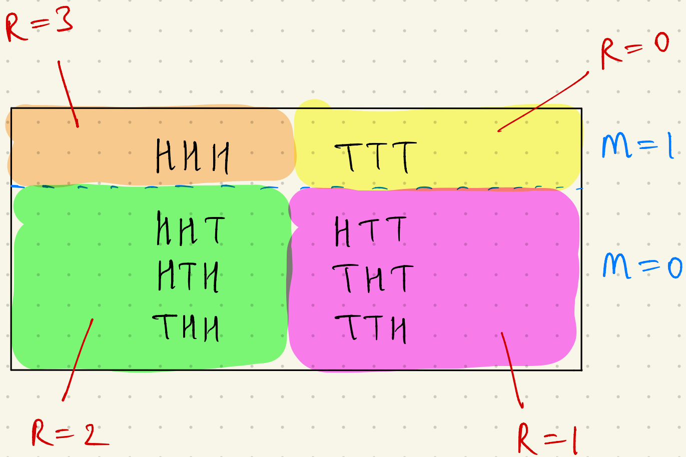
$$= \Pr(R=2) + \Pr(R=3)$$

$$= \frac{1}{2}$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w \,:\, R(w)=x} \Pr(w)$$



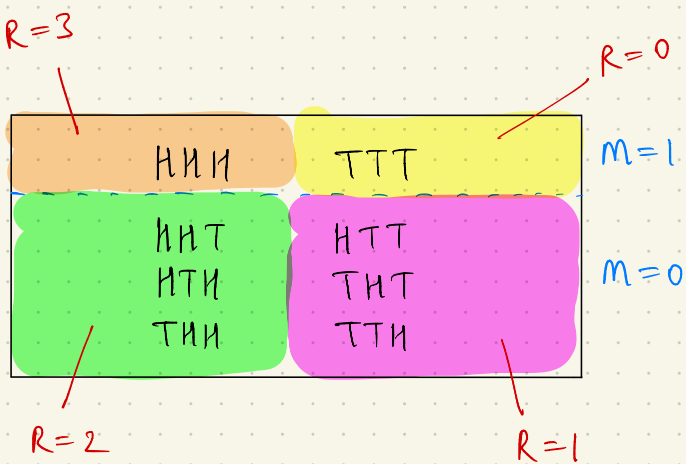
For $A \subseteq \mathbb{R}$,

$$\Pr(R \in A) = \sum_{a \in A} \Pr(R=a)$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w : R(w)=x} \Pr(w)$$



For $A \subseteq \mathbb{R}$,

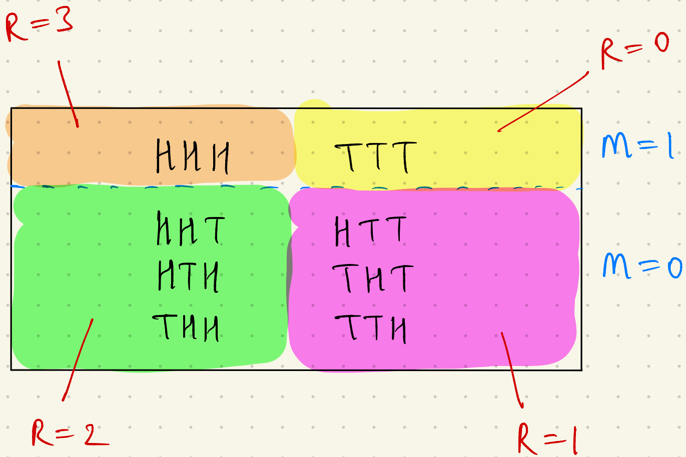
$$\Pr(R \in A) = \sum_{a \in A} \Pr(R=a)$$

$$\Pr(R \in \{1,3\}) = \frac{1}{2}.$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w \,:\, R(w)=x} \Pr(w)$$

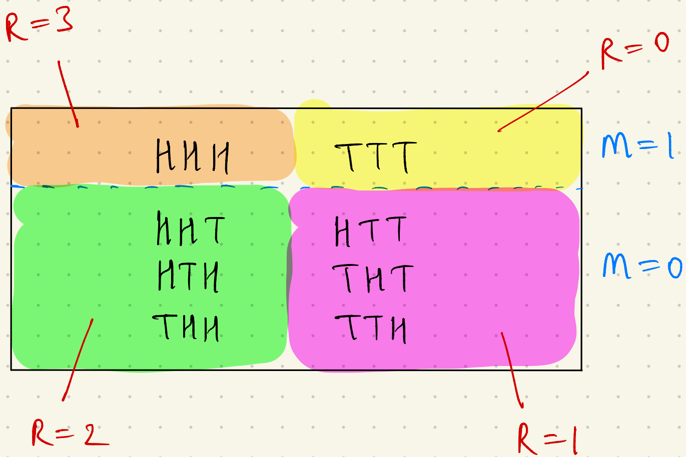


$$\Pr(R=1 \mid M=0) =$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R = x) = \sum_{w\,:\,R(w)=x} \Pr(w)$$

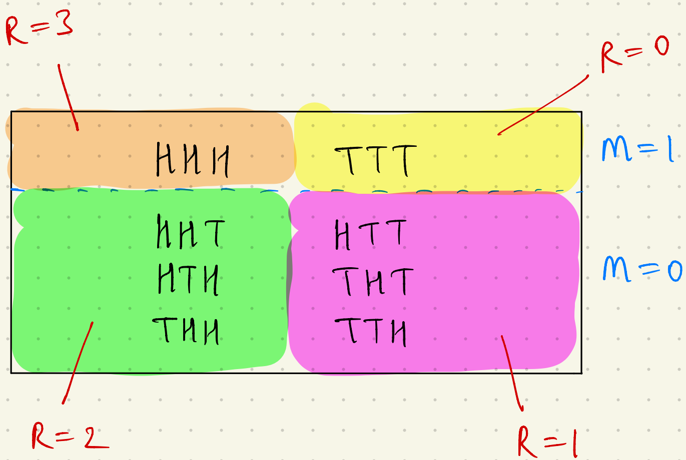


$$\Pr(R = 1 \mid M = 0) = 1/2$$

RANDOM VARIABLES

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w\,:\,R(w)=x} \Pr(w)$$



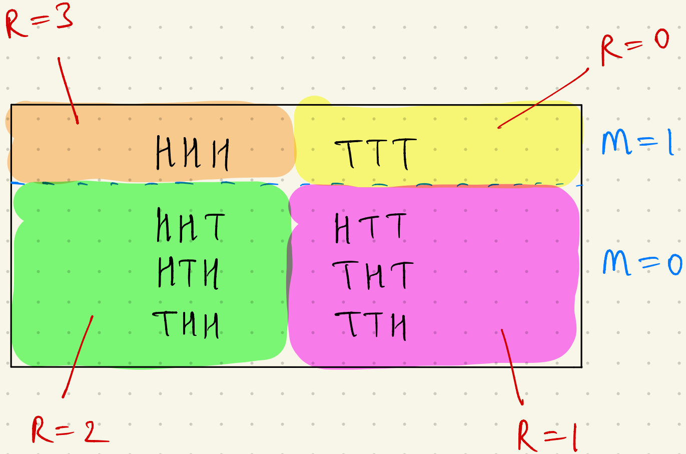
$$\Pr(R=1 \mid M=0) = 1/2$$

$$\Pr(R=2 \mid M=1) =$$

Random Variables

$\{w : R(w) = x\}$ is the event that $R = x$.

$$\Pr(R=x) = \sum_{w\,:\,R(w)=x} \Pr(w)$$



$$\Pr(R=1 \mid M=0) = 1/2$$

$$\Pr(R=2 \mid M=1) = 0$$

RANDOM VARIABLES

The random variables R_1, R_2 are independent if $\forall\, x_1, x_2 \in \mathbb{R}$

$$\Pr(R_1 = x_1 \mid R_2 = x_2) = \Pr(R_1 = x_1)$$

$$\text{or} \quad \Pr(R_2 = x_2) = 0.$$

RANDOM VARIABLES

The random variables R_1, R_2 are independent if $\forall\, x_1, x_2 \in \mathbb{R}$

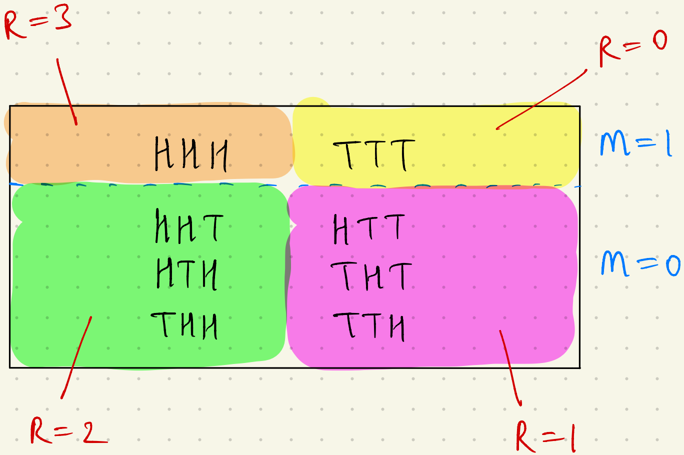
$$Pr(R_1 = x_1 \mid R_2 = x_2) = Pr(R_1 = x_1)$$

$$\text{or} \quad Pr(R_2 = x_2) = 0.$$

Equivalently,

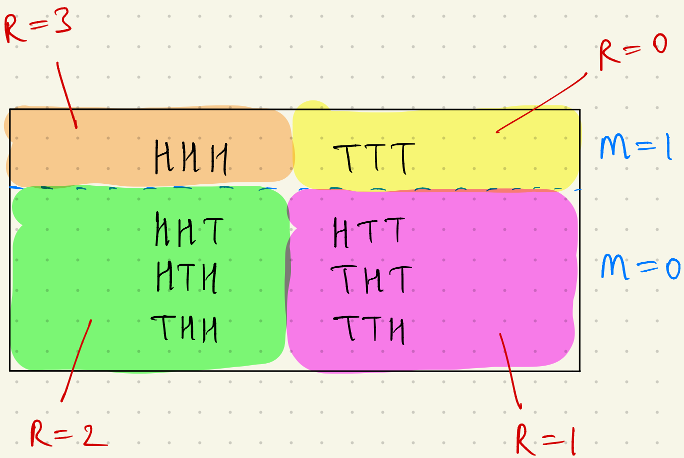
$$\forall\, x_1, x_2 \in \mathbb{R} \quad Pr(R_1 = x_1 \cap R_2 = x_2) = Pr(R_1 = x_1) \cdot Pr(R_2 = x_2)$$

RANDOM VARIABLES



Are R and M independent?

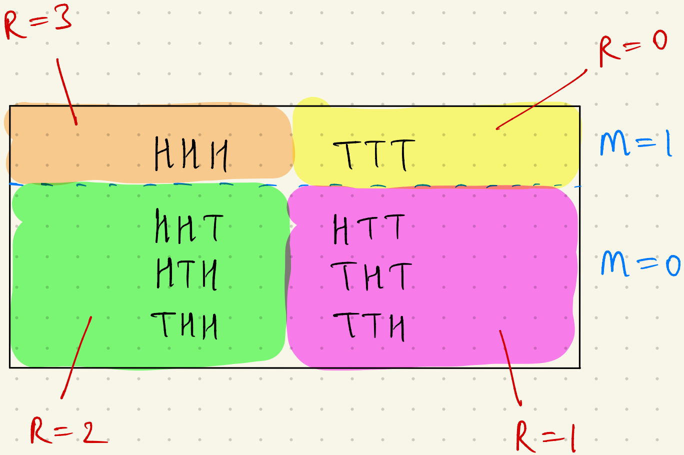
RANDOM VARIABLES



Are R and M independent?

No!

RANDOM VARIABLES

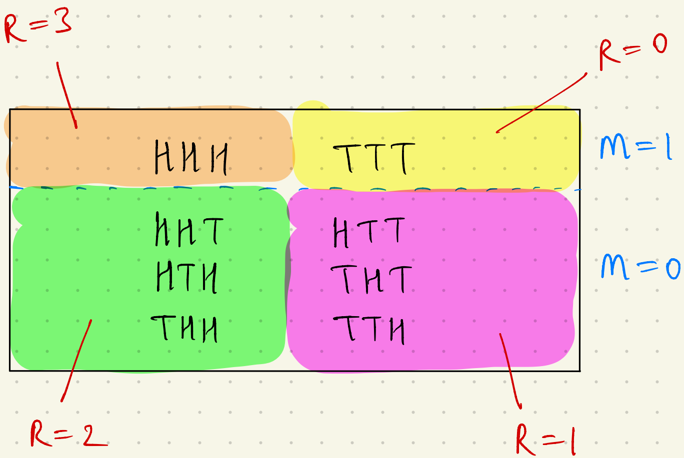


Are R and M independent?

No!

$$\Pr(R=2 \cap M=1) = 0$$

RANDOM VARIABLES



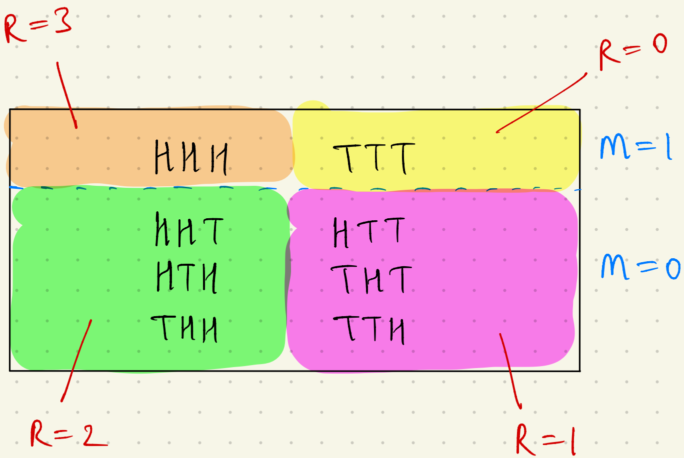
Are R and M independent?

No!

$$\Pr(R=2 \cap M=1) = 0$$

$$\Pr(R=2) \cdot \Pr(M=1) = \frac{3}{8} \cdot \frac{1}{4} \neq 0$$

RANDOM VARIABLES



Are R and M independent?

No!

$$\Pr(R=2 \cap M=1) = 0$$

$$\Pr(R=2) \cdot \Pr(M=1) = \frac{3}{8} \cdot \frac{1}{4} \neq 0$$

To show NOT independent $\rightarrow$ Find some x_1, x_2 that work

——— independent $\longrightarrow$ Prove for all x_1, x_2

E.g., 2 fair independent 6-sided die D_1, D_2

E.g., 2 fair independent 6-sided die D_1, D_2

Let $S := D_1 + D_2$

E.g., 2 fair independent 6-sided die D_1, D_2

Let $S := D_1 + D_2$ and $T := \begin{cases} 1 & \text{if } S = 7 \\ 0 & \text{otherwise} \end{cases}$

E.g., 2 fair independent 6-sided die D_1, D_2

Let $S := D_1 + D_2$ and $T := \begin{cases} 1 & \text{if } S = 7 \\ 0 & \text{otherwise} \end{cases}$.

Are D_1 and S independent?

E.g., 2 fair independent 6-sided die D_1, D_2

Let $S := D_1 + D_2$ and $T := \begin{cases} 1 & \text{if } S = 7 \\ 0 & \text{otherwise} \end{cases}$.

Are D_1 and S independent? No!

E.g., 2 fair independent 6-sided die D_1, D_2

Let $S := D_1 + D_2$ and $T := \begin{cases} 1 & \text{if } S = 7 \\ 0 & \text{otherwise} \end{cases}$.

Are D_1 and S independent? No!

$$\Pr(S=12, D_1=1) = 0 \neq \Pr(S=12) \cdot \Pr(D_1=1)$$