Extra material: https://onlinelibrary.wiley.com/doi/10.1111/test.12321

Module 2A : Probability

Frequentist and Bayesian building blocks

Agenda:

- **Frequentist probability**
  - Venn Diagram & definition of Event
  - Addition Rule
    - Mutually exclusive, not mutually exclusive
  - Multiplication Rule
    - Independent, not independent
  - Conditional Probability
  - Important differentiation: $P(A \cup B)$, $P(A \cap B)$, $P(A|B)$

Probability:

Two major splits in how probability is defined:

Frequency Interpretation:

Frequency of a particular outcome (an event) across many random trials

Subjective (Bayesian) Interpretation:

Subjective belief or opinion of the chance that a particular outcome (an event) will be realized

Random Experiment:

The process of observing the outcome of a chance event

ex - one roll of a die

Sample State Space:

A list of all possible elementary outcomes of an experiment

ex - {1, 2, 3, 4, 5, 6}

Event:

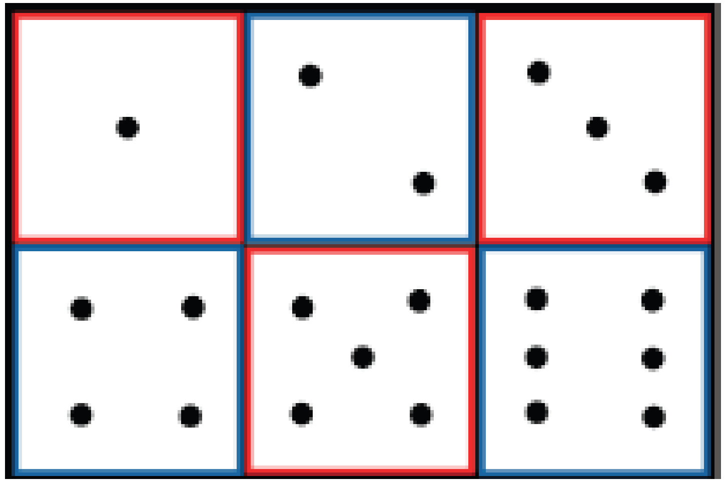
A set made up of elements from the sample space

ex - {1, 2, 3, 4, 5, 6}

ex - “an even number”

Sample Space examples:

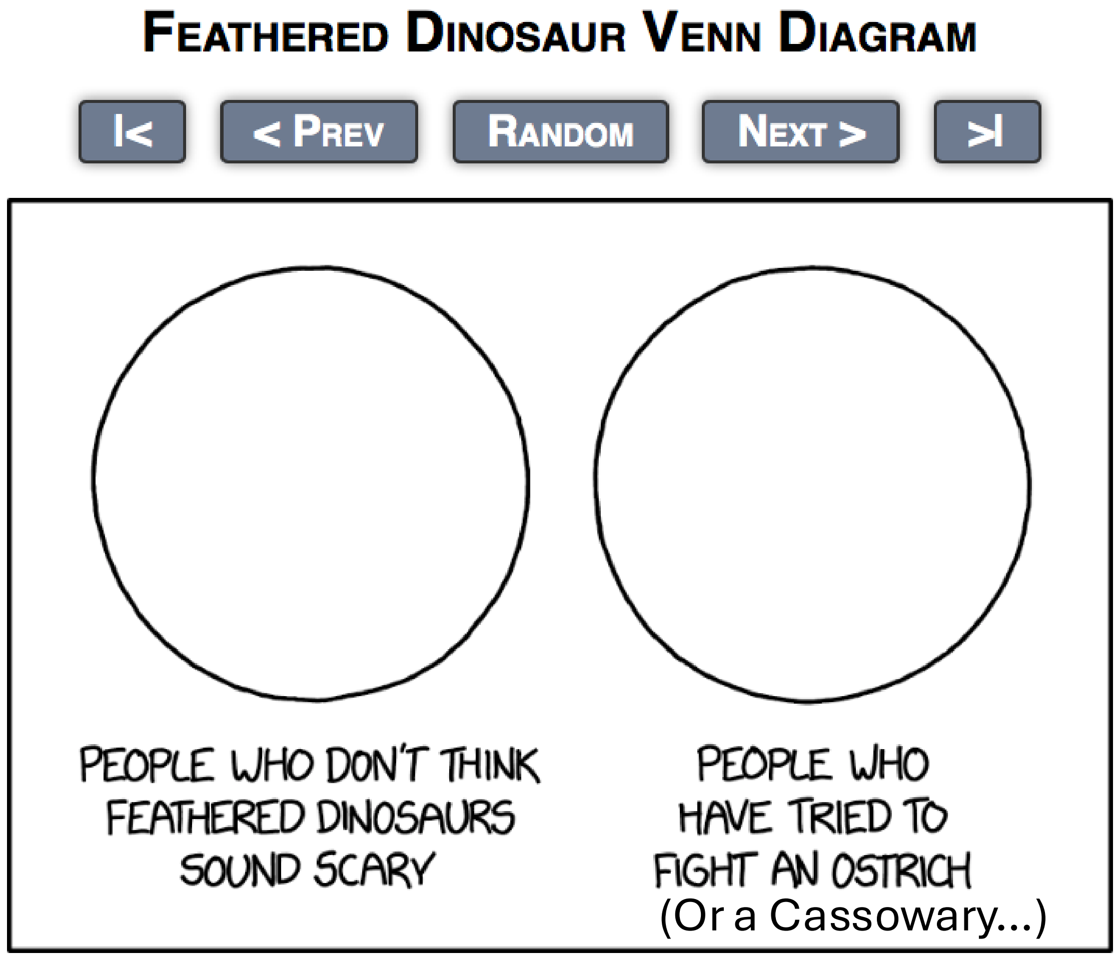
One Die Toss



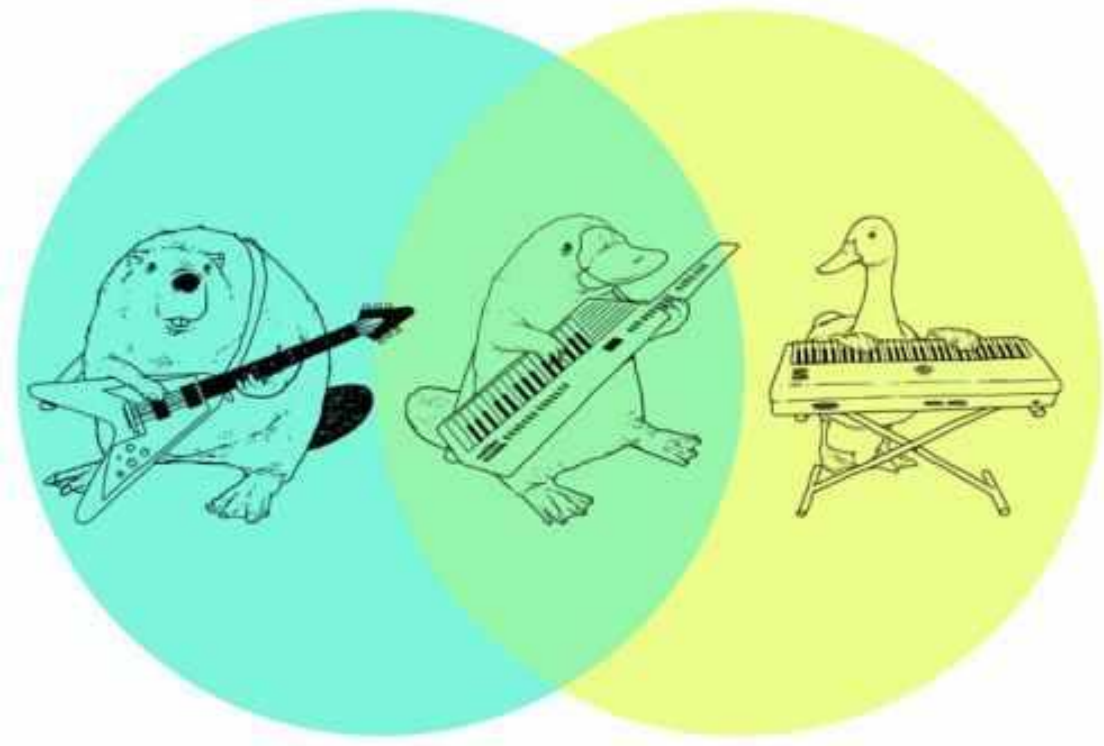
{1,2,3,4,5,6}

One Coin Toss

{heads, tails}



The BEST Venn diagram

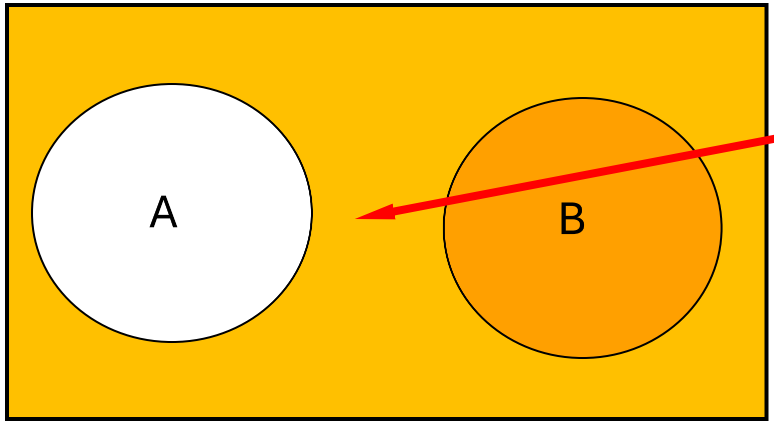


https://xkcd.com/2090/

http://i.imgur.com/CsaCFLc.jpg

NOT A

The event does not occur (complement); everything except A

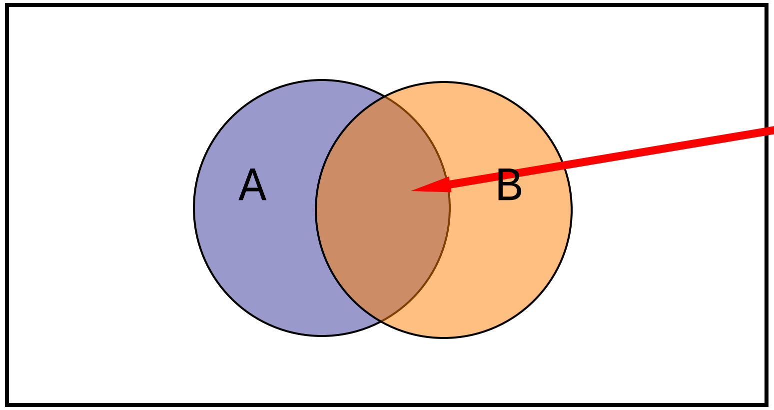


Mutually exclusive

$$P(A \cap B) = 0$$

AND

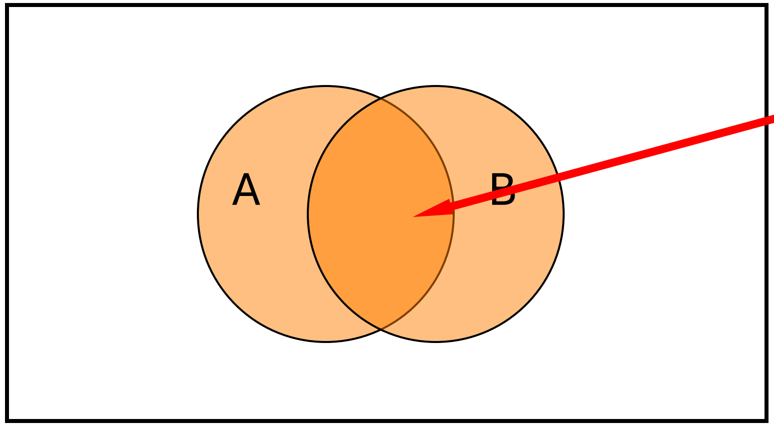
Both events occur (intersection)



$$P(A \cap B) \neq 0$$

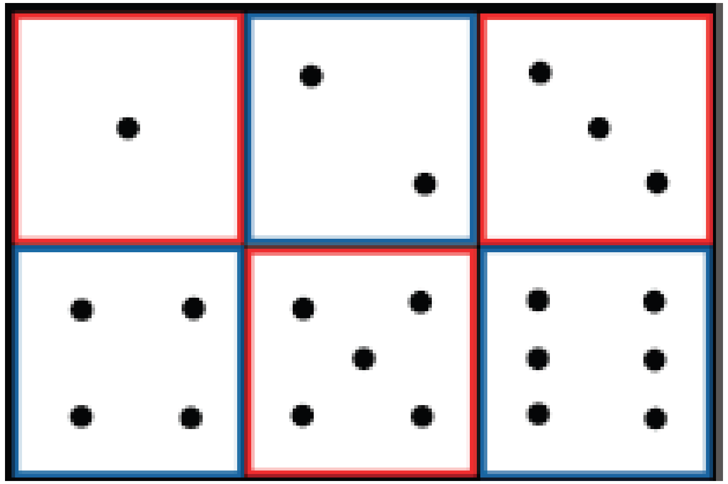
OR

Either events or both events occur (union)

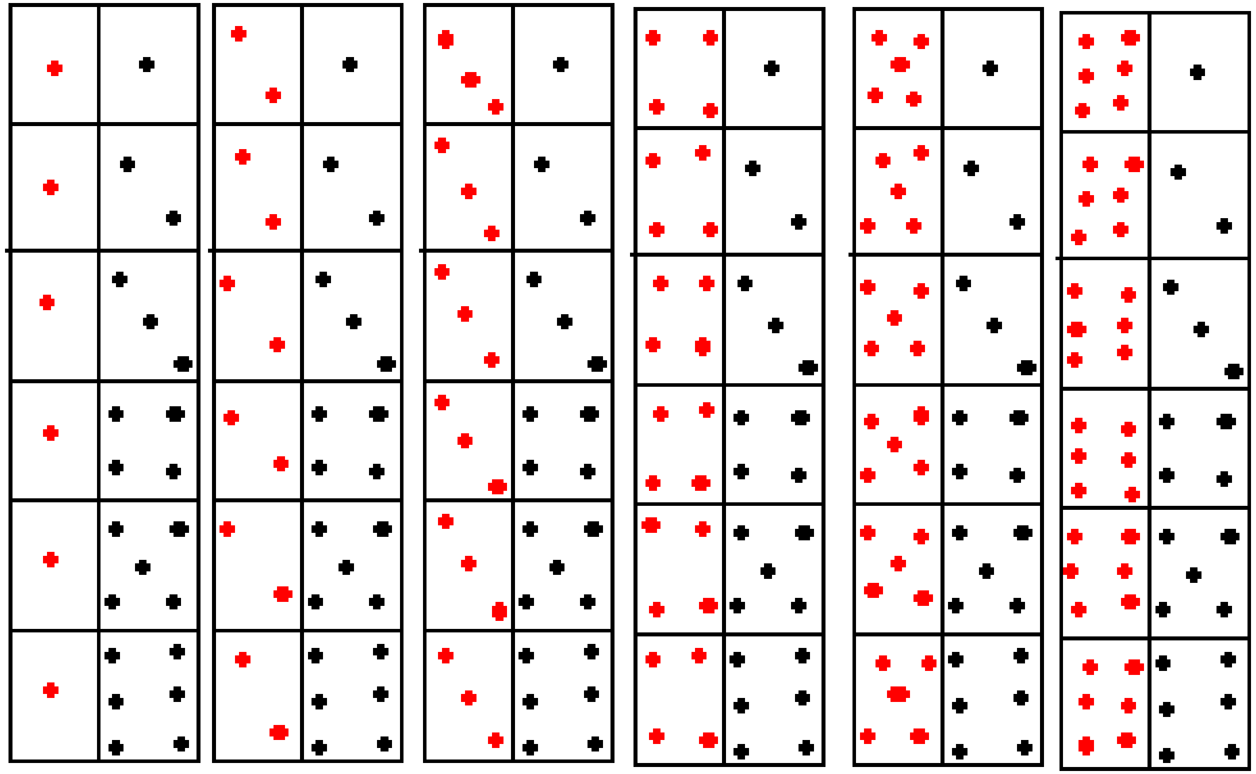


$$P(A \cup B) \neq 0$$

Mutually Exclusive outcomes:



P(rolling 1) =P(rolling 2) = ...= P(rolling 6) = 1/6

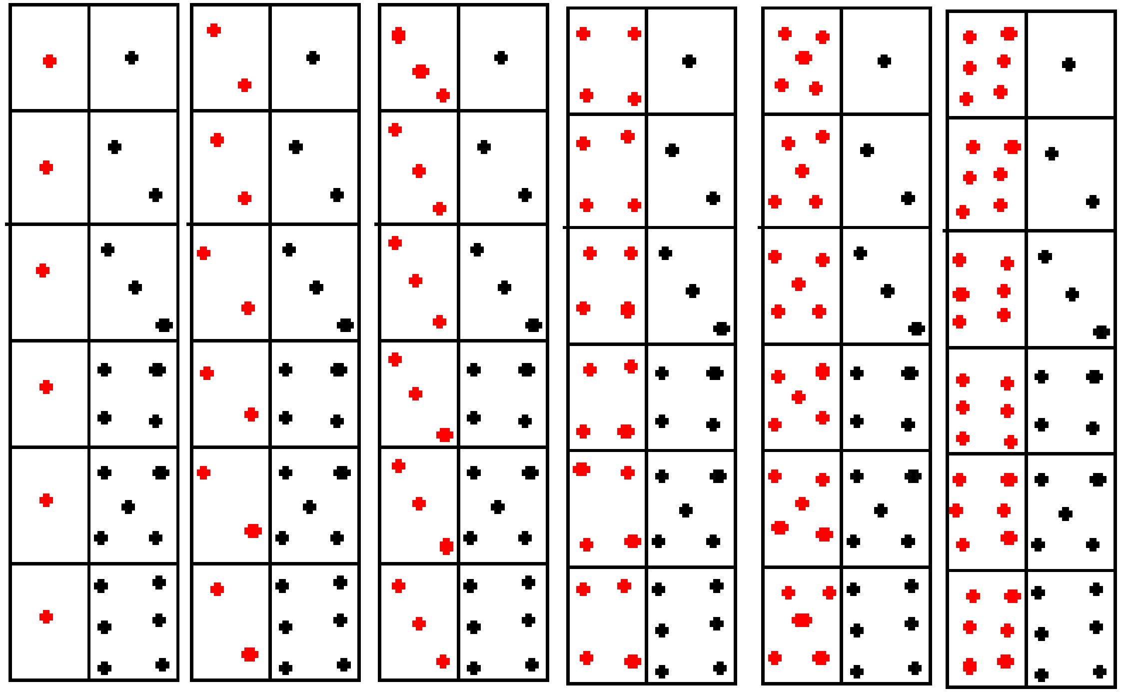


P(rolling 1∩ rolling 1) =P(rolling 1∩ rolling 2) = ...= P(rolling 1∩ rolling 6) = 1/36

ex. Two dice are rolled. The event of interest is the two dice faces add up to 3. What is the probability of this event?

The probability of this event is the sum of the probabilities of the elementary outcomes in this subset.

<u>The full set of elementary outcomes of rolling two dice</u>



P(showing 3 dots)=2 ways= 1/36+1/36 = 2/36

Two dice are rolled and the red die shows 1, what is the probability of rolling a total of 3?

P(total 3|Die 1=1) = 1/6

A <u>fair</u> coin is tossed six times, and the results are recorded in the order that they appear. Which of the following outcomes is most likely to occur.

H=heads; T=tails

1. HHHTTT: _ _ _ _ _ _
2. HTHTHT
3. HTTHTT
4. 1 and 2 are equally likely
5. 1,2 and 3 are all equally likely

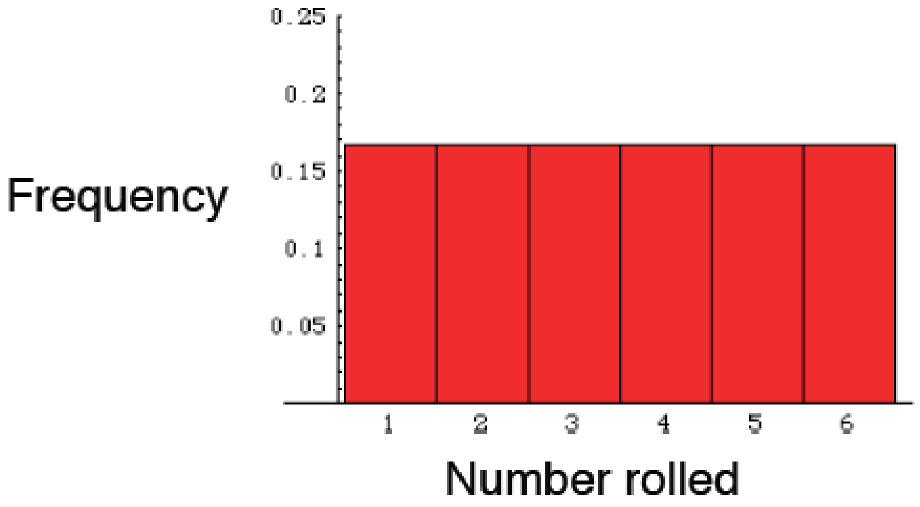
For independent events, you can think of them as: _ _ _ _ _ _

Discussion:

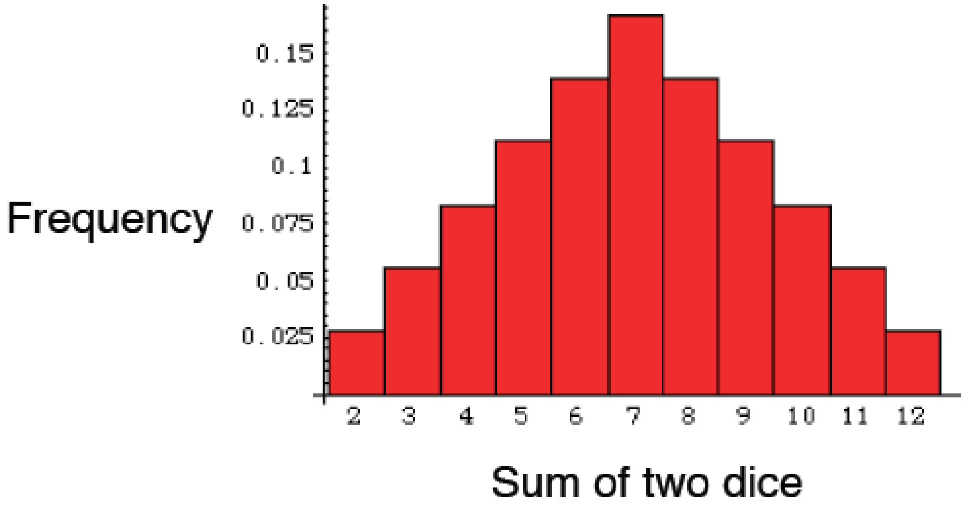
What if a coin is flipped five times and comes up heads each time. Is a tail "due" and therefore more likely than not to occur on the next flip?

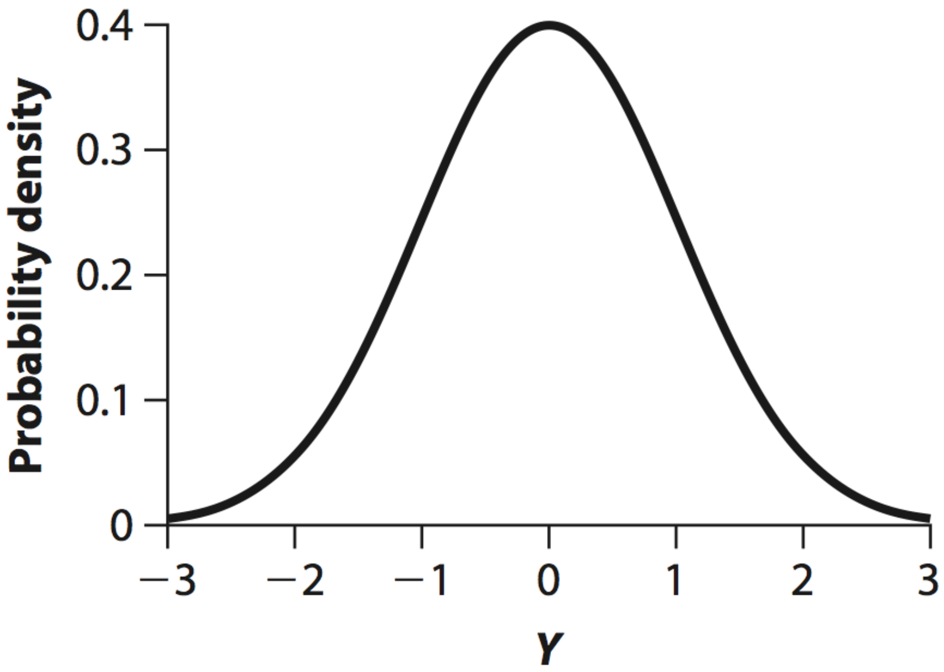
http://onlinestatbook.com/2/probability/gambler.html

Probability distribution for the outcome of a roll of one die:



Probability distribution for the sum of a roll of two dice:

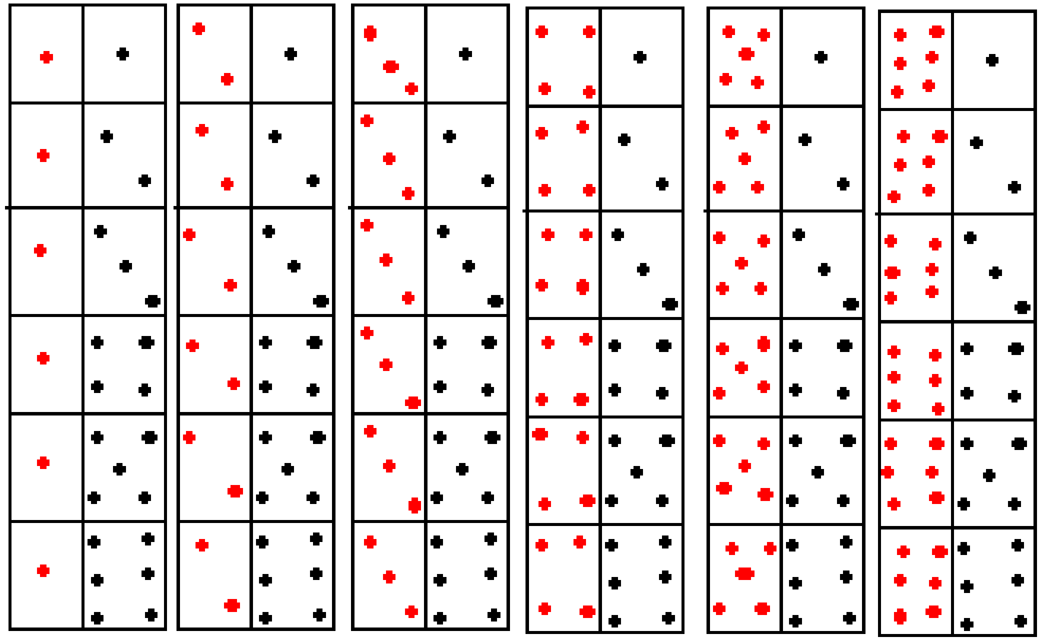




Ex. Event A = Black die is 1, Event B = Red die is 1

*P(A **OR** B) = ?*

a. 6/36
b. 12/36
c. 11/36
d. 1/36



*P(A **AND** B) =?*

a. 6/36
b. 12/36
c. 11/36
d. 1/36

Ex. Event A = Black die is 1, Event B = Red die is 1

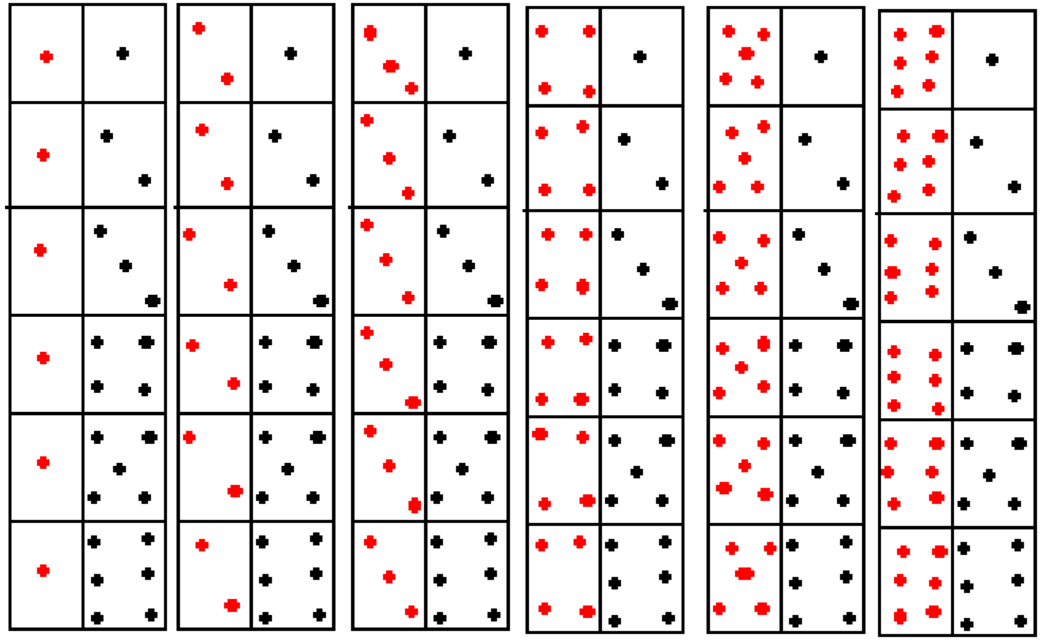
*P(**NOT** A) = ?*

a. 1/36

b. 35/36

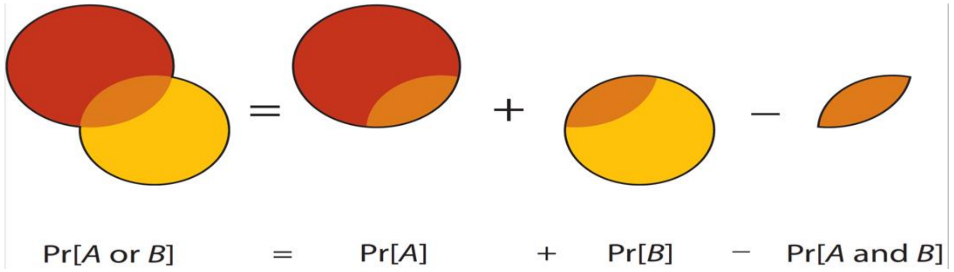
c. 6/36

d. 30/36



Addition Rule:

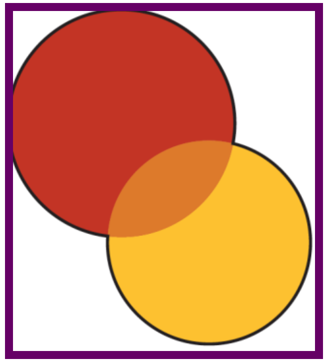
$$P[A \cup B] = P[A] + P[B] - P[A \cap B]$$

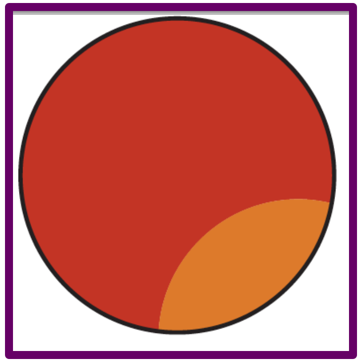


if mutually exclusive: $P[A \cup B] = P[A] + P[B]$

Multiplication Rule:

$$\boxed{P[A \cap B] = P[A \mid B]P[B] = P[B \mid A]P[A] \xrightarrow{IF F independent} P[A]P[B]}$$





- *Two events are independent if the occurrence of one gives no information about whether the second will occur*
- *Two events are dependent if the probability or outcome of one event <u>changes</u> because of the outcome of a second event*

Example: In a forest, imagine that 1% of the trees are infected by a fungal rot and 0.1% have owl nests. What is the probability that a tree has both fungal rot and an owl nest if:

The two events are independent?

a. P(Owls)*P(rot) = 0.01*0.001

b. 0.01+0.001

c. 0.01+0.001-0.01*0.001

The two events are mutually exclusive?

a. 0

b. 0.01*0.001

c. 0.01+0.001

Suppose you take out two cards from a standard pack of cards (52) one after another, without replacing the first card. What is probability that the first card is the ace of spades, and the second card is a heart?

So far, we have learned a great many things about probability:

1. Sample space is made up of elementary outcomes
2. Events can be elementary outcomes or groupings of elementary outcomes
3. Logic operators on probabilities: AND, NOT, OR
4. General Addition rule: $P(A \bigcup B) = P(A) + P(B) - P(A \bigcap B)$
5. IF events A and B are mutually exclusive, then general addition rule collapses into special addition rule: $P(A \bigcup B) = P(A) + P(B)$
6. General Multiplication rule: P[A and B]=P[A|B]xP[B] = P[B|A]XP[A]
7. If events A and B are independent, general multiplication rule collapses into special multiplication rule
   - allows one to test whether or not two events are independent

8. What about if they are not independent?

The Bruce Effect

A female house mouse mates and enters very early pregnancy. In her territory, **unfamiliar males** sometimes intrude. Exposure to the **scent of an unfamiliar male** (urine/pheromones) can trigger the **Bruce effect**—pregnancy block via implantation failure. This functions as an adaptive strategy in environments where infanticide risk from non-sire males is high: rather than invest in a likely doomed litter, the female aborts early and re-mates.

- If **no unfamiliar male is present**, pregnancy usually proceeds.
- If an **unfamiliar male is present**, pregnancy block is much more likely.

Thus, the **state of the environment** the female encounters (unfamiliar male present vs not) and the **outcome** (pregnancy continues vs blocks) are **dependent variables**

The Bruce Effect

State:

U(nfamiliar) =1 (present), =0 (absent)

Outcome:

B(lock) = 1, =0 (no block, pregnancy proceeds)

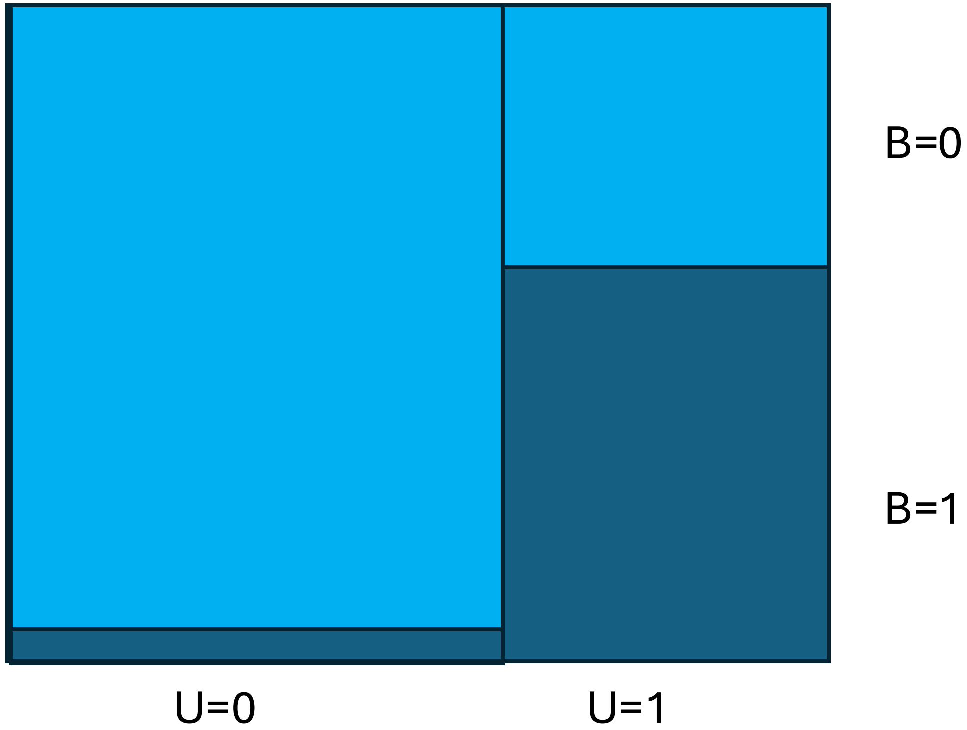
Prior:

P(U=1) = 0.40; P(U=0)=0.60

Conditional:

P(B=1|U=1) = 0.6

P(B=1|U=0)=0.05



These are made up numbers to make the calculations easy, but
This paper characterized the Bruce effect!
https://www.nature.com/articles/184105a0

The Bruce Effect

Q1. What proportion of pregnancies block overall?

Q2. If you observe a pregnancy block, what’s the posterior probability an unfamiliar male was present?

Q3. If pregnancy continues, what’s the probability an unfamiliar male was present?

Another Example: *Nasonia vitripennis*, a parasitoid wasp, lays eggs in fly pupae; larval wasps then hatch inside, feed on host, and emerge as adults; the males and females then mate on the spot.

***Nasonia* females manipulate sex of their offspring depending on if host fly pupa previously parasitized**.

- If host not yet parasitized, then *Nasonia* lays mainly female eggs and produces only a few males (one male can fertilize multiple females).
- If host already parasitized, then *Nasonia* lays mostly male eggs.

The state of the host encountered by a female and the sex of an egg laid are **dependent variables** (Werren, 1980)

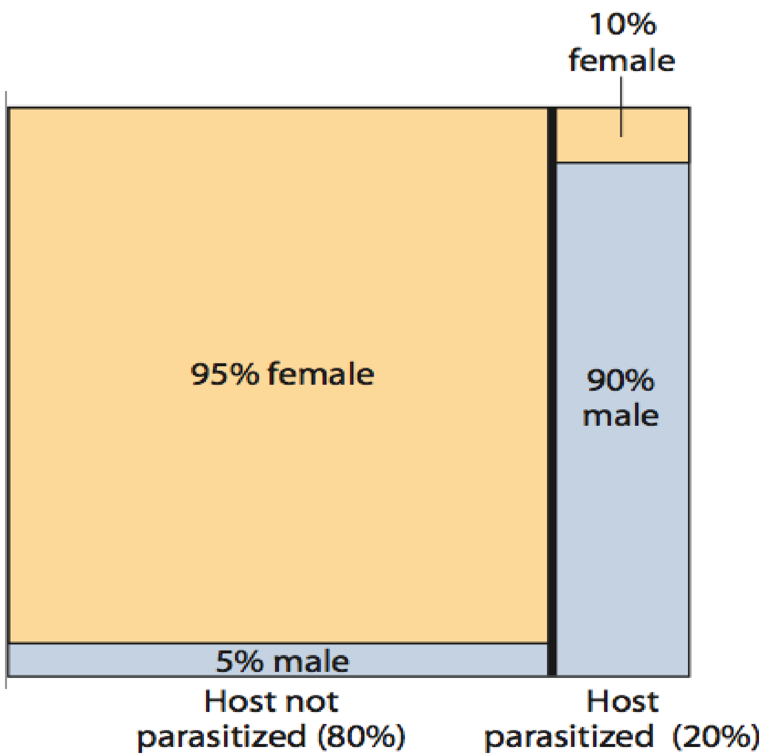
* If host not yet parasitized, then *Nasonia* lays mainly female eggs and produces only a few males (one male can fertilize multiple females).

* If host already parasitized, then *Nasonia* lays mostly male eggs.

State of host (parasitized, not parasitized)

<u>Possibly Dependent variable based on mosaic plot</u>

Sex of egg (male, female)



Example: Offspring of two carriers (Nn x Nn):

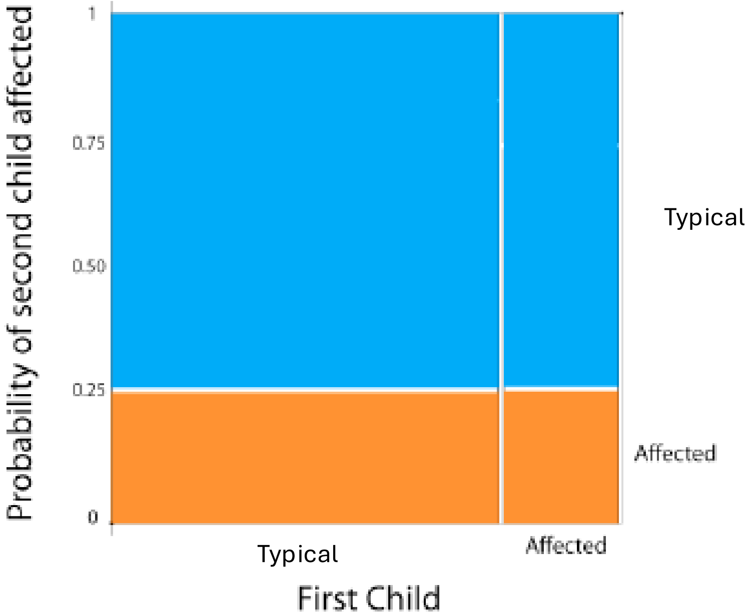
P[night blindness]=0.25

	N	**n**
N	NN	Nn
n	nN	nn

What is the probability that two kids from this family both have night blindness?

P[(1st child night blindness) AND (2nd child night blindness)]
= 0.25 x 0.25 = 0.0625

Possibly Independent variable based on mosaic plot:



Probability trees provide a straightforward method to determine independence or dependence between variables

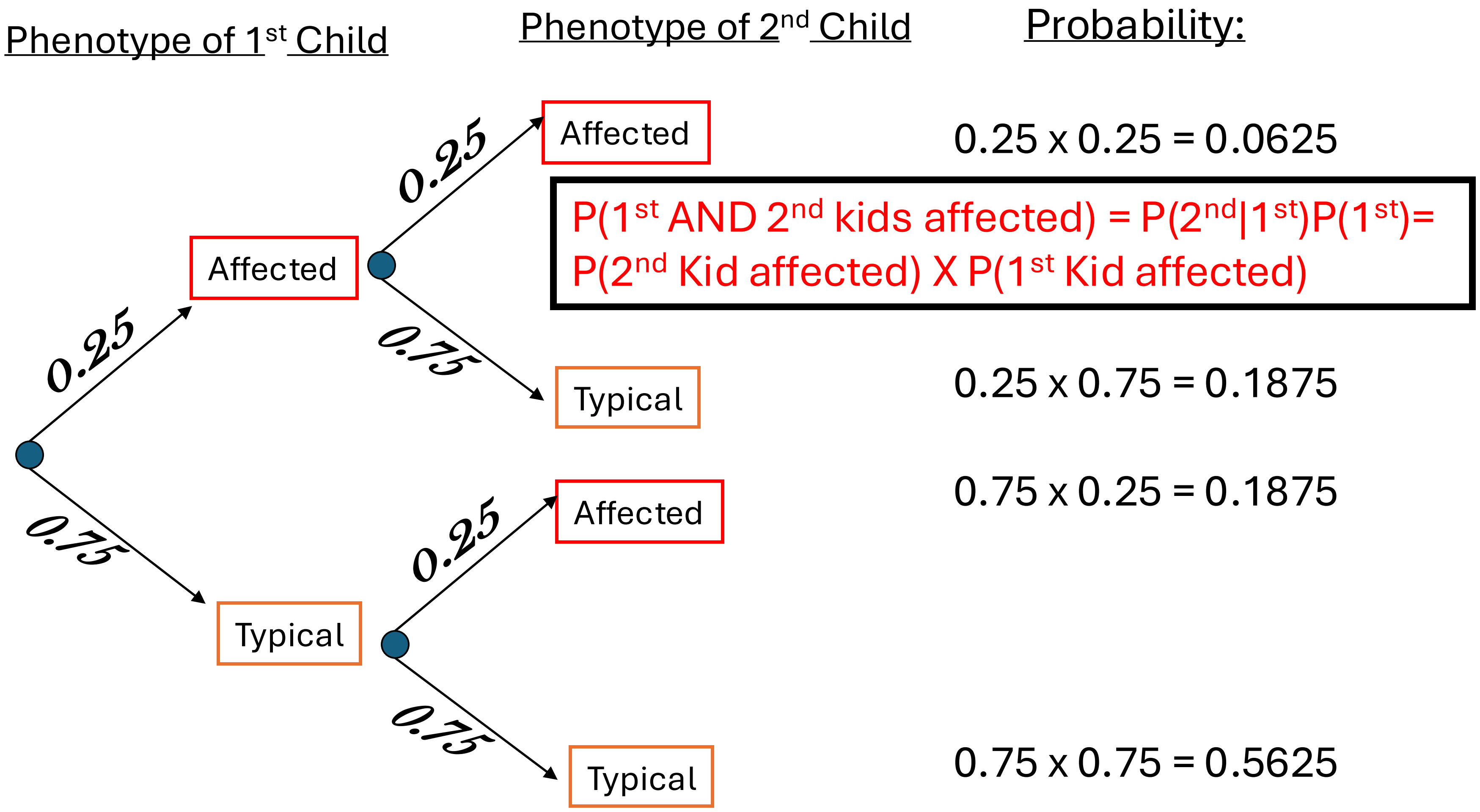
- Map out probabilities of all mutually exclusive outcomes of variables

<u>Additional Benefits:</u>

- easy to calculate the probability of any possible outcome sequence for the variables under consideration

- easy to double check that all possibilities have been enumerated

Probability Trees: two events, meiosis is independent



So far, we have learned a great many things about probability:

1. Sample space is made up of elementary outcomes
2. Events can be elementary outcomes or groupings of elementary outcomes
3. Logic operators on probabilities: AND, NOT, OR
4. General Addition rule: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. IF events A and B are mutually exclusive, then general addition rule collapses into special addition rule: $P(A \cup B) = P(A) + P(B)$
6. General Multiplication rule: P[A and B]=P[A|B]xP[B]
7. If events A and B are independent, general multiplication rule collapses into special multiplication rule
   - allows one to test whether or not two events are independent

8. What about if they are not independent?

A large population of giant pandas has five alleles at one gene labeled: $\mathbf{A_1, A_2, A_3, A_4, A_5}$. They have corresponding frequencies in the population population 0.1, 0.15,0.6, 0.05, 0.1. In this randomly mating population, the two alleles present in any individual are independently sampled from the population as a whole.

a. What is the probability that a single allele chosen at random from this population either $\mathbf{A_1}$ or $\mathbf{A_4}$ **= $P(A_1)+P(A_4)$ = 0.1 +0.05 = 0.15**

b. What is the probability that the individual has **two A_1** alleles? $\underline{P(A_1)}*\underline{P(A_1)}$ = 0.1*0.1 = 0.01

c. What is the probability that an individual is ***not*** $\mathbf{A_1A_1}$? 1-P(A_1 A_1) = 1-0.01 = 0.99

d. What is the probability, if you drew two individuals at random from this population that neither of them would have an $\mathbf{A_1A_1}$ genotype?

P(1st diploid isn't A1A1)? 0.99; P(2nd diploid isn't A1A1) = 0.99

0.99*0.99 = 0.9801

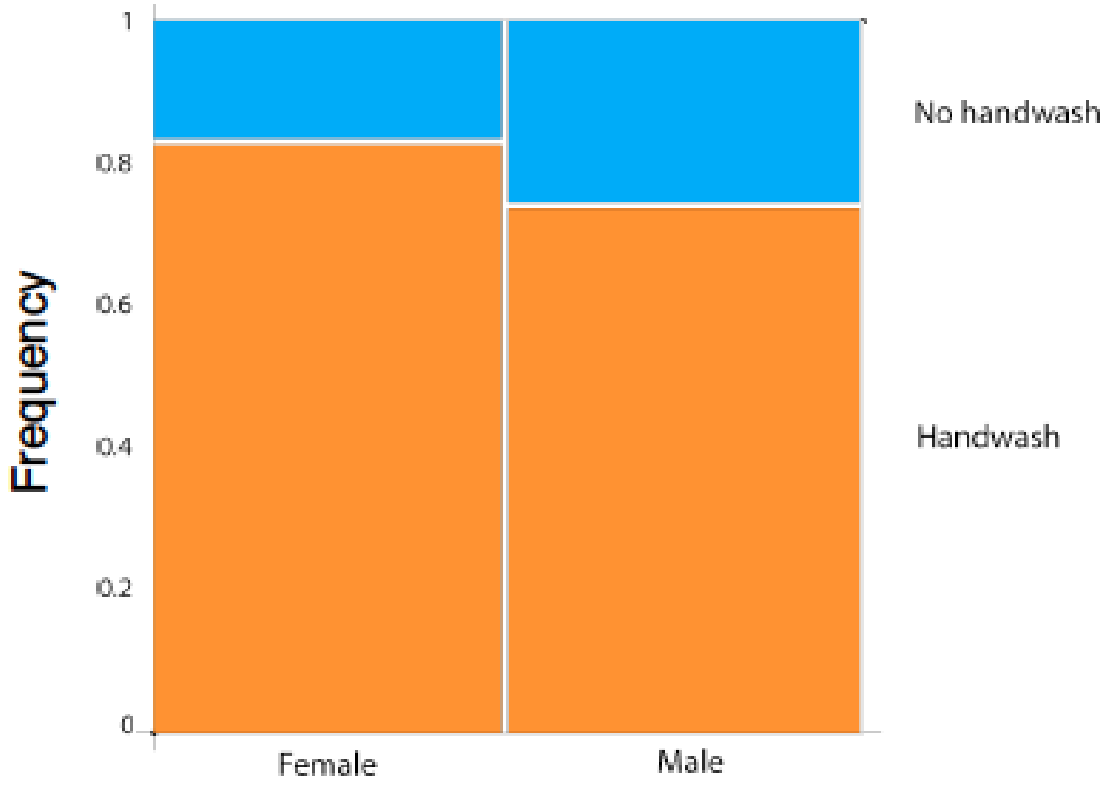
For your certificate of completion:

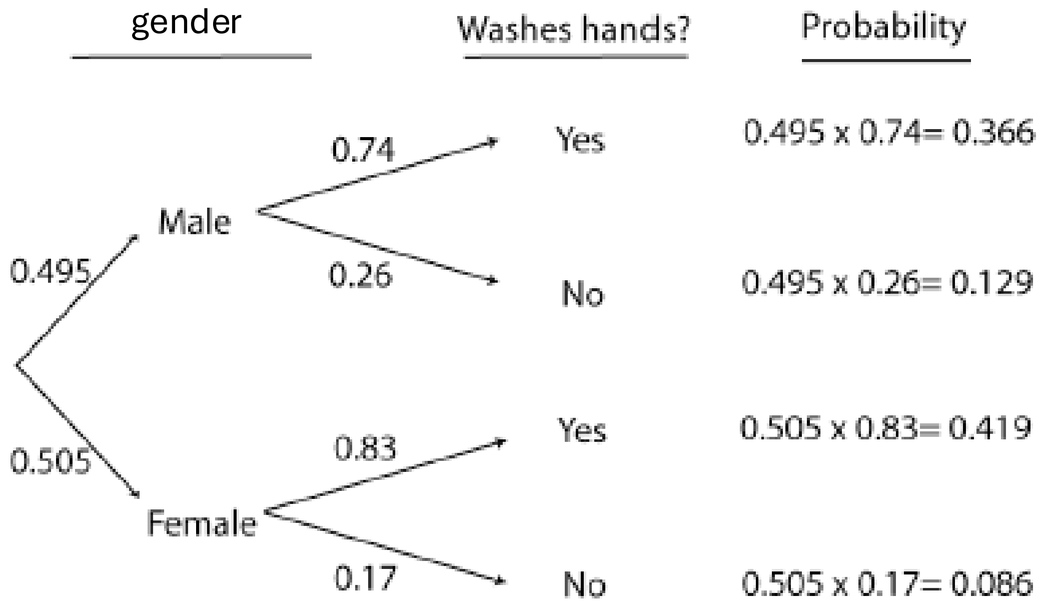
e. What is the probability, if you drew two individuals at random from this population that at least one of them would have an $\mathbf{A_1A_1}$ genotype?

f. What is the probability that three randomly chosen individuals would have ***no*** $\mathbf{A_2}$ or $\mathbf{A_3}$ alleles?

Example: Is washing your hands after using the washroom dependent on gender?

- P[male] = 0.495
- P[male washes his hands]=0.74
- P[female washes her hands]=0.83





Conditional Probability:

The probability that an event occurs given that a condition is met

P[X|Y] = P[X and Y]/P[Y]

This is read as “the probability of X given Y”

It means: the probability of X if Y is true

Fancier way of writing the total probability of an event:

$$\boxed{P[X] = \sum_{Y} P[X \mid Y] P[Y]}$$

Conditional Probability: **P[X|Y] = P[X and Y]/P[Y]**

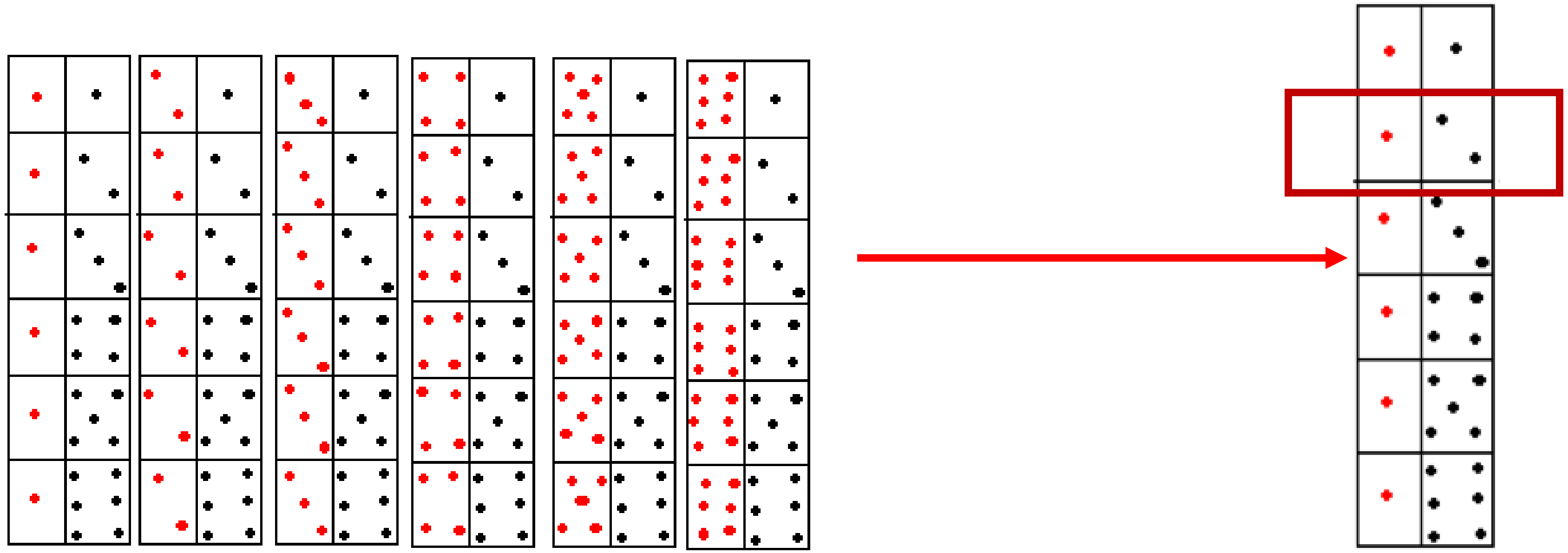
Example:

What is the probability that two dice will sum to three?

*-this is really asking **P[X and Y]** where $\boldsymbol{X}_{red\ die}$ =1 or 2 and $\boldsymbol{Y}_{black\ die}$ = 1 or 2*

Now: **what if we already have rolled the first die and know that we have a one? Event X_{red}=1**

Reduced state space, from 36 to 6:



P[Sum to three] =2/36

P[Sum to three]=1/6

- <u>**Very**</u> important to understand conditional probability before we tackle Bayes'

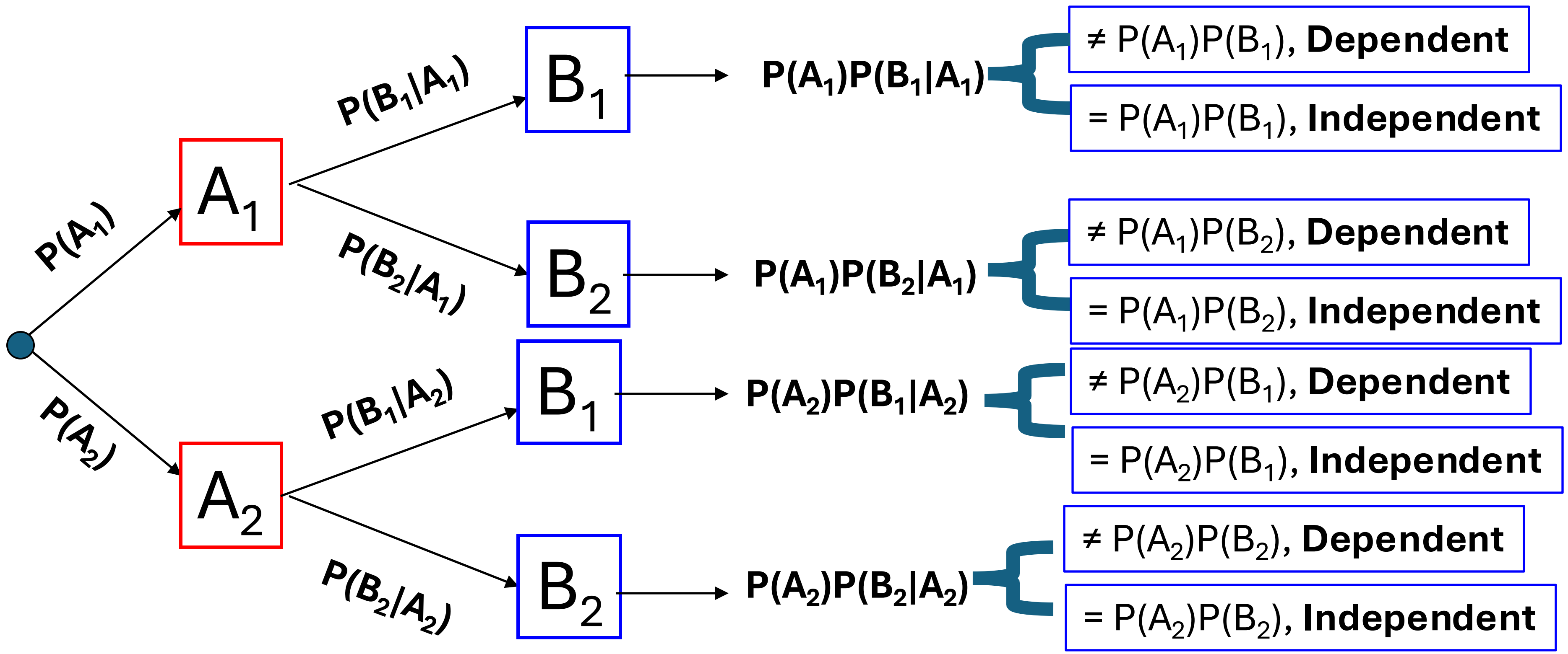
- Conditional probability can be a little confusing; sometimes using a Venn diagram with **3** events instead of 2 makes it clearer (example 5 and 6):

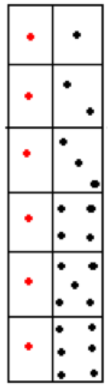
  https://www.nagwa.com/en/explainers/403141497934/

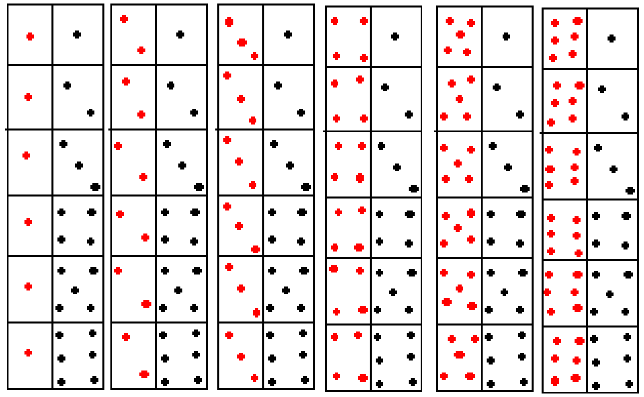
- Some learners have struggled with the difference between $P(A \cap B)$ and P(A|B)

- **Remember:** $\boxed{P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)}$

1st Variable | 2nd Variable | $P(A \cap B)$ | **Dependent or Independent?**







Module 2A Questions:

Upload to Canvas as part of your Certification of Completion

1. Finish up the blue questions in the allele question (Slide 24) that we began in class.

2. A man went to Monte Carlo to try and make his fortune. Whilst he was there, he had an opportunity to bet on the outcome of rolling dice. He was offered the same odds for each of the following outcomes (in each case the dice are rolled simultaneously):

   i. At least 1 six with 6 dice.

   ii. At least 2 sixes with 12 dice.

   iii. At least 3 sixes with 18 dice.

The man decided that these were all the same so he could choose one at random and bet on that. Was he right? Give your answer and explanation (in your own words). Note: to answer iii, you will need to use the binomial coefficient (we will cover this in 3A). I included this so you can see how probability connects to what we will focus on next and because you have been given the answer (you only must indicate by explaining the solution yourself that you truly understand the answer).

3. Here is the question that you can imagine represents migration between two originally isolated populations which, as you will recall from your biology courses, is one of the forces that shapes genetic variation (stolen from https://nrich.maths.org/1970)

   1. There are 5 yellow balls in bag One. One ball is transferred to bag Two which contains an unknown number of green balls.
   2. Bag Two is then shaken and a ball is selected at random and transferred to bag One without seeing its colour.
   3. Bag One is then shaken and a ball is selected at random and transferred to bag Two without seeing its colour.
   4. Finally, a ball is chosen at random from bag Two.

If I tell you before you carry out this process that there is a 3/5 chance you will end up with a green ball at the end, can you work out how many green balls are in bag Two at the start? **Draw a probability diagram (showing your solution) and explain the logic of your answer in your own words.**

Hint: You are following the state space of each of the two bags in this question. They start out as the following representation: (5,0)(0,n), as there are 5 yellow balls in bag 1 initially, and 0 yellow balls in bag 2, as well as an unknown number of green balls in bag 2. Note: There will be a bit of algebra involved in the solution, but we will discuss it in next class – just do your best! You can find the answer to this question online, too, but you will need to use your own words to explain your solution. If you can successfully answer this question, you have a solid understanding of conditional probability!