Probability rules

FW8051 Statistics for Ecologists

Department of Fisheries, Wildlife and Conservation Biology

Learning objectives

Understand and be able to work with basic rules of probability.

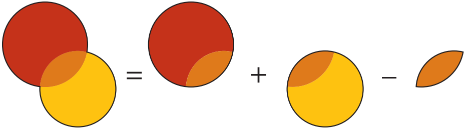
We will need these rules to understand Bayes Theorem, which is fundamental to Bayesian statistics.

Summary Probability Rules

1. $P(A \text{ and } B) = P(A)P(B \mid A) = P(B)P(A \mid B)$

Summary Probability Rules

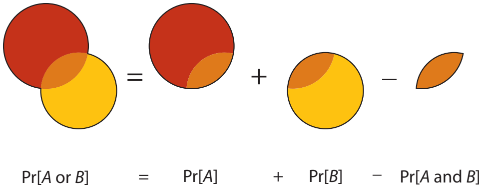
1. $P(A \text{ and } B) = P(A)P(B \mid A) = P(B)P(A \mid B)$
2. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$



Pr[*A* or *B*] = Pr[*A*] + Pr[*B*] − Pr[*A* and *B*]

Summary Probability Rules

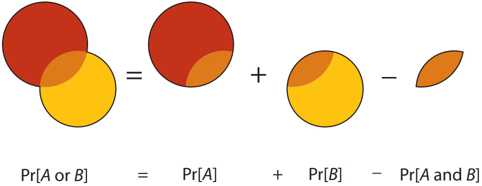
1. $P(A \text{ and } B) = P(A)P(B \mid A) = P(B)P(A \mid B)$
2. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$



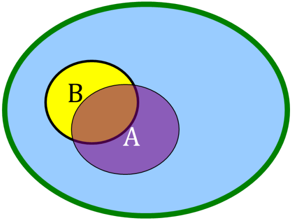
3. $P(\text{not } A) = 1 - P(A)$

Summary Probability Rules

1. $P(A \text{ and } B) = P(A)P(B \mid A) = P(B)P(A \mid B)$
2. $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$



3. $P(\text{not } A) = 1 - P(A)$
4. $P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$ (probability of "A given B")



[Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students who drink caffeine in the morning also drink caffeine in the afternoon?

[Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students who drink caffeine in the morning also drink caffeine in the afternoon?

Want to find: $P(A \mid M)$

[Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students who drink caffeine in the morning also drink caffeine in the afternoon?

Want to find: $P(A \mid M)$

$= P(A \text{ and } M)/P(M)$

[Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students who drink caffeine in the morning also drink caffeine in the afternoon?

Want to find: $P(A \mid M)$

$= P(A \text{ and } M)/P(M) = 0.37/0.50 = 0.74$

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

P(not(M or A))

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

P(not(M or A))

= 1 – P(M or A)

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

P(not(M or A))

= 1 – P(M or A)

= 1 – [P(M) + P(A) – P(M and A)]

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

P(not(M or A))

= 1 – P(M or A)

= 1 – [P(M) + P(A) – P(M and A)]

= 1 – [0.50 + 0.45 – 0.37]

Caffeine [Think-Pair-Share]

Based on recent survey data, 50% of students drink caffeine in the morning, 45% of students drink caffeine in the afternoon, and 37% drink caffeine in the morning and the afternoon. What percent of students do not drink caffeine in the morning or in the afternoon?

$P(\text{not}(M \text{ or } A))$

$= 1 - P(M \text{ or } A)$

$= 1 - [P(M) + P(A) - P(M \text{ and } A)]$

$= 1 - [0.50 + 0.45 - 0.37]$

$= 1 - 0.58 = 0.42$

[Think-Pair-Share]

A wildlife biologist surveys 100 different plots, looking for pheasants. Suppose:

- 30% of the plots contain pheasants.
- The biologist has a 60% chance of detecting pheasants when they are present.

On what percentage of the plots should we expect the wildlife biologist to see a pheasant.

[Think-Pair-Share]

A wildlife biologist surveys 100 different plots, looking for pheasants. Suppose:

- 30% of the plots contain pheasants.
- The biologist has a 60% chance of detecting pheasants when they are present.

On what percentage of the plots should we expect the wildlife biologist to see a pheasant.

P(present and seen) = P(Seen | present)P(present)

[Think-Pair-Share]

A wildlife biologist surveys 100 different plots, looking for pheasants. Suppose:

- 30% of the plots contain pheasants.
- The biologist has a 60% chance of detecting pheasants when they are present.

On what percentage of the plots should we expect the wildlife biologist to see a pheasant.

P(present and seen) = P(Seen | present)P(present)

= 0.3 x 0.6 = 0.18

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true: P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true: P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

What about:

- Mammal and lay eggs

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true: P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

What about:

- Mammal and lay eggs Not mutually exclusive, platypus + four species of echidna

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true: P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

What about:

- Mammal and lay eggs Not mutually exclusive, platypus + four species of echidna
- Teeth and feathers

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true: P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

What about:

- Mammal and lay eggs Not mutually exclusive, platypus + four species of echidna
- Teeth and feathers Yes, mutually exclusive

Mutually Exclusive Events

Two events are mutually exclusive if they cannot both be true:
P(A and B) = 0.

Sample Space

- Role of a die: {1,2,3,4,5,6}
- Location: {On campus, off campus} (disjoint areas)

What about:

- Mammal and lay eggs Not mutually exclusive, platypus + four species of echidna
- Teeth and feathers Yes, mutually exclusive

P(A or B) = P(A) + P(B) for mutually exclusive events.

Independence

Events A and B are independent if $P(A \mid B) = P(A)$

Independence

Events A and B are independent if $P(A \mid B) = P(A)$

Intuitively, knowing that event B happened does not change the probability that event A happened.

Independence

Events A and B are independent if $P(A \mid B) = P(A)$

Intuitively, knowing that event B happened does not change the probability that event A happened.

If A and B are independent then:

$P(A \text{ and } B) = P(A)P(B \mid A) = P(A)P(B)$

We have been using this rule to construct Likelihoods!

Summary of Special Cases

If events A and B are mutually exclusive:

- $P(A \text{ or } B) = P(A) + P(B)$
- $P(A \text{ and } B) = 0$

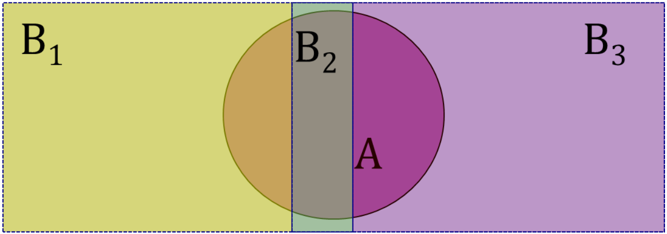
If events A and B are independent:

- $P(A \mid B) = P(A)$
- $P(A \text{ and } B) = P(A)P(B)$

Law of Total Probability

If events $B_1, B_2, \ldots, B_k$ are mutually exclusive and together make up all possibilities, then:

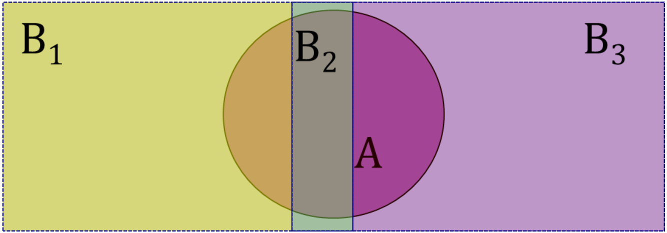
$P(A) = \sum_i P(A|B_i)P(B_i)$



Law of Total Probability

If events $B_1, B_2, \ldots, B_k$ are mutually exclusive and together make up all possibilities, then:

$P(A) = \sum_i P(A|B_i)P(B_i)$



Special Case: $P(A) = P(A \text{ and } B) + P(A \text{ and } (\text{not } B))$

Jewel wasp: from Whitlock and Schluter Example 5.8

Females can manipulate sex of the eggs they lay

- Previously parasitized hosts <- lay more male eggs
- Host that have not been previously parasitized <- lay more female eggs

Jewel Wasp [Exercise]

Suppose:

- When a wasp finds a host, there is a 0.20 probability another wasp has already laid eggs in it
- If the host is unparastized, the female lays a male egg with prob = 0.05 (and female egg with prob = 0.95)
- If the host already has eggs, the female lays a male egg with prob = 0.90 (and female egg with prob = 0.10)

Use the total law of probability to determine the probability(sex of new egg is male). Hint: let A = {male}, B = {previously parasitized, not previously parasitized}

Jewel Wasp

Probability(sex of new egg is male)

= P(male & previously parasitized) + P(male & not previously parasitized)

Jewel Wasp

Probability(sex of new egg is male)

= P(male & previously parasitized) + P(male & not previously parasitized)

= P(male | previously parasitized)P(previously parasitized) + P(male | not previously parasitized)*P(not previously parasitized)

Jewel Wasp

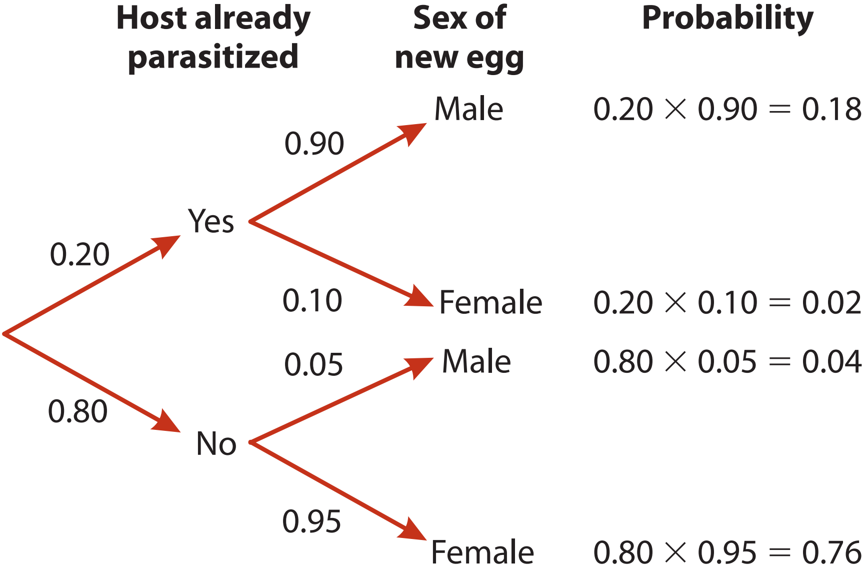
Probability(sex of new egg is male)

= P(male & previously parasitized) + P(male & not previously parasitized)

= P(male | previously parasitized)P(previously parasitized) + P(male | not previously parasitized)*P(not previously parasitized)

= 0.2 x 0.9 + 0.05 x 0.8 = 0.22

Tree Diagram



Bayes Theorem

Let $\bar{A}$ = not(A)

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$= \frac{P(B|A)P(A)}{P(B \text{ and } A)+P(B \text{ and } \bar{A})}$$

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

The last two expressions can be extended to more than 2 groups using the total law of probability

Example: Trisomy 21 or Down Syndrome

Caused by an extra copy of chromosome 21.

- 1 in 800 children have Down Syndrome, i.e., $P(D) = 1/800 = 0.00125$
- A multiple-marker screening test can be performed in the second trimester of pregnancy
- False Positive: $P(+|\bar{D}) = 0.05$
- False Negative: $P(-|D) = 0.19$

Given that one tests positive, what is the probability that the fetus has Down Syndrome? $P(D|+)$

Example: Trisomy 21 or Down Syndrome

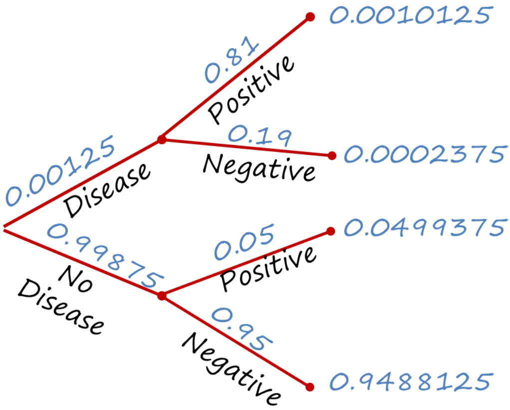
Caused by an extra copy of chromosome 21.

- 1 in 800 children have Down Syndrome, i.e., $P(D) = 1/800 = 0.00125$
- A multiple-marker screening test can be performed in the second trimester of pregnancy
- False Positive: $P(+|\bar{D}) = 0.05$
- False Negative: $P(-|D) = 0.19$

Given that one tests positive, what is the probability that the fetus has Down Syndrome? $P(D|+)$

Use Bayes rule: $= P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A)+P(B|\bar{A})P(\bar{A})}$. It might also help to draw a probability tree.

Probability Tree



$P(D \mid +) = P(D \text{ and } +)/P(+)$

= 0.0010125/[0.0010125 + 0.0499375] = 0.02

Lets Make A Deal

1. 3 doors (2 goats and 1 car)
2. Monte knows where the car is, but you don't
3. You pick a door and Monte opens one of the remaining doors holding a goat.
4. Should you switch doors?

Lets Make A Deal

1. 3 doors (2 goats and 1 car)
2. Monte knows where the car is, but you don't
3. You pick a door and Monte opens one of the remaining doors holding a goat.
4. Should you switch doors?

Answer

4 options, determined by 2 decisions

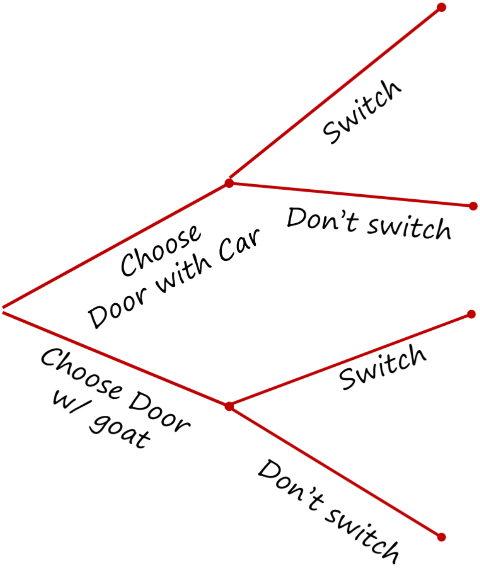
Step 1:

- You choose the door with the car behind it.
- You choose the door without the car behind it.

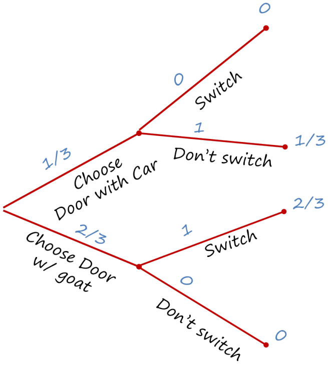
Step 2:

- You switch your choice
- You do not switch your choice.

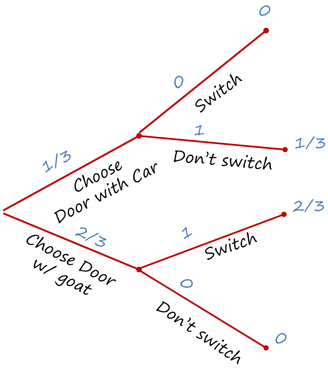
Sample Space



Probability Tree

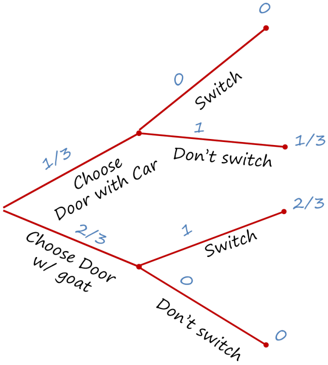


Probability Tree



P(Win | Switch) = 0 + 2/3

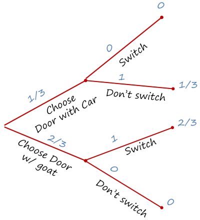
Probability Tree



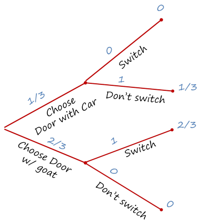
P(Win | Switch) = 0 + 2/3

P(Win | do not Switch) = 1/3 + 0 = 1/3

Probability Tree



Probability Tree



P(win | switch) = P(win & switch | car first)P(car first) + P(win & switch | goat first)P(goat first) = 0 + 2/3

P(win | stay put) = P(win & stay put | car first)P(car first) + P(win & stay put | goat first)P(goat first) = 1/3 + 0

For some interesting comments on the problem, see: Link