# Probability 2: Conditional Probability and Bayes' rule

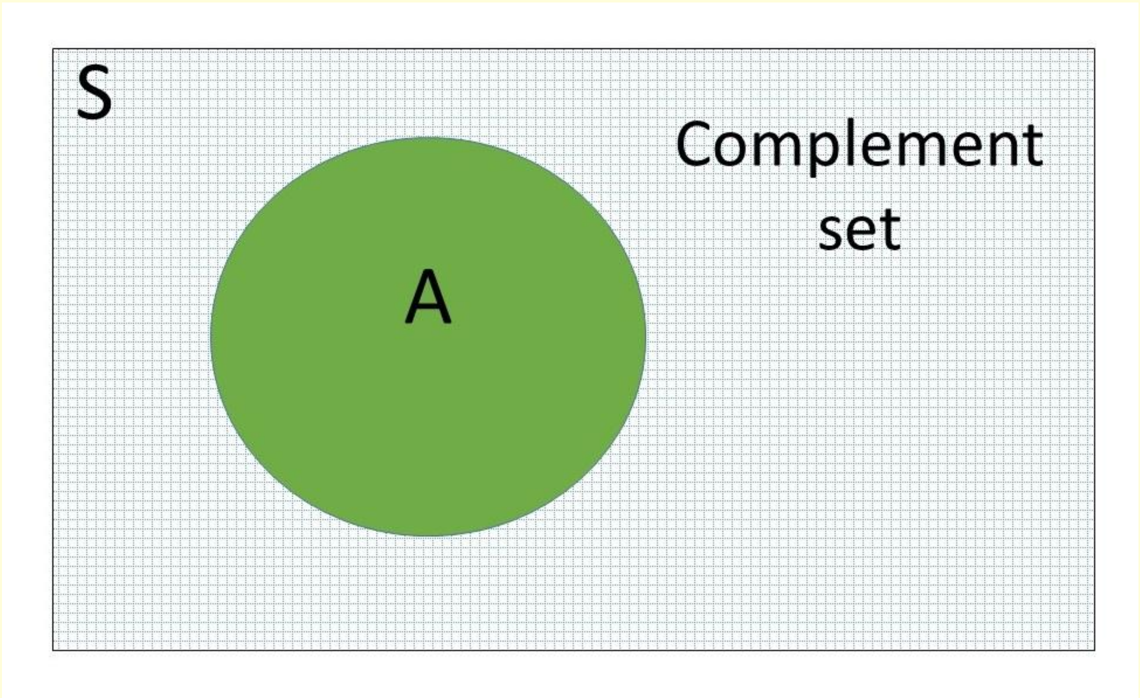
# Module Outline

- Review
- Dependence of events and conditional probability
- Probability trees
- Two-way tables
- Bayes' Rule

# Review: Probability Basics

- *Sample space S:* set of all possible outcomes

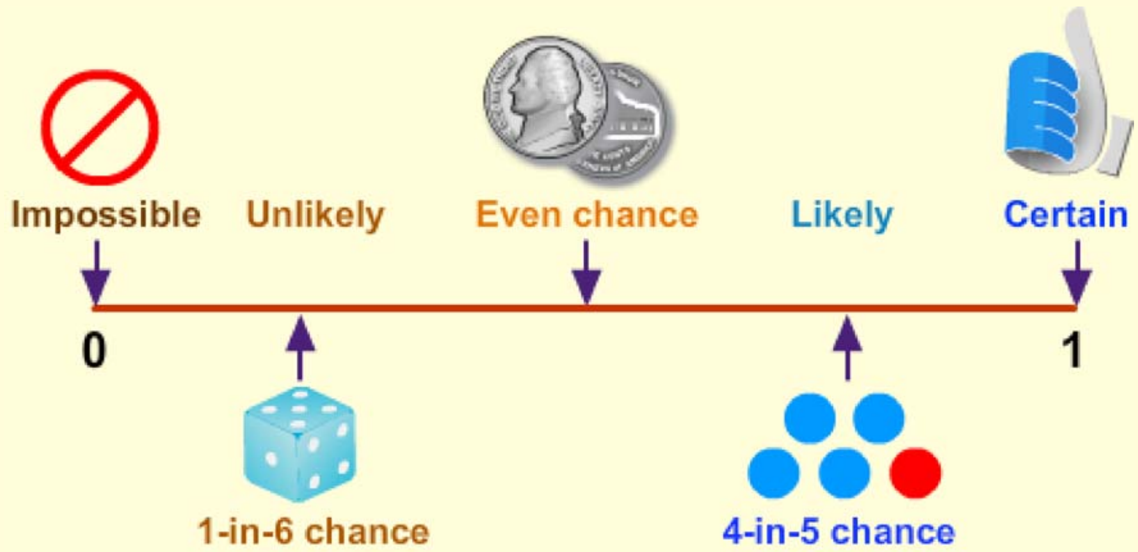
  *Event:* some subset of *S*



# Review: Probability Basics

- **Basic Axioms**: (i) $0 \leq P(A) \leq 1$ (ii) $P(S) = 1$

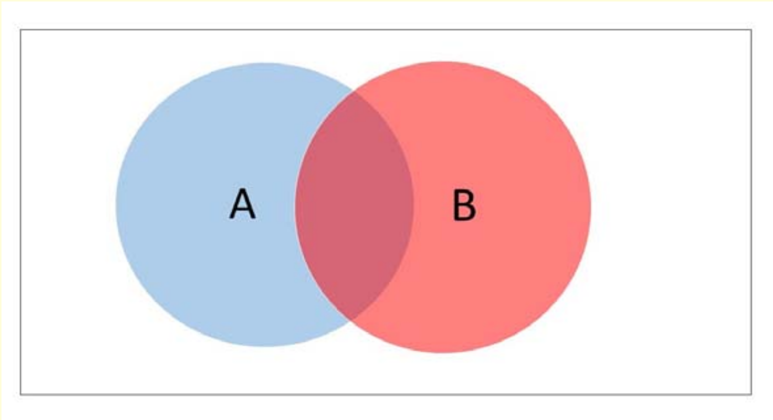
  If $A$ and $B$ disjoint, $P(A \cup B) = P(A) + P(B)$

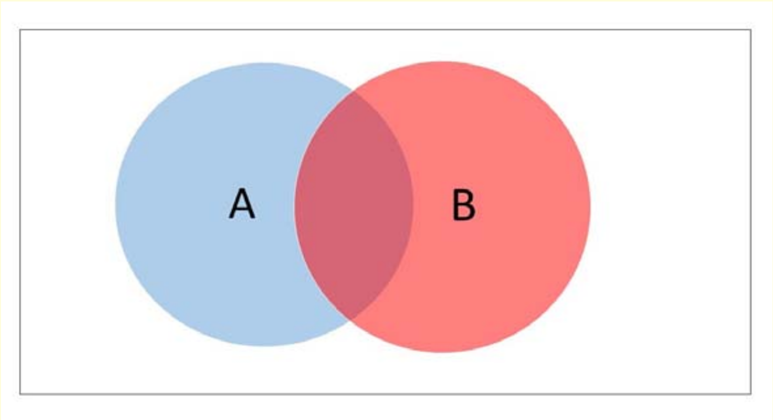


# Review: Probability Basics

- $A \cup B$ : event $A$ **OR** $B$ (or both) occur

  $A \cap B$ : event $A$ **AND** $B$ (i.e. both) occur



# Review: Probability Basics

- $A \cup B$ : event $A$ **OR** $B$ (or both) occur

  $A \cap B$ : event $A$ **AND** $B$ (i.e. both) occur

- **[Complement rule]** $P(A^c) = 1 - P(A)$

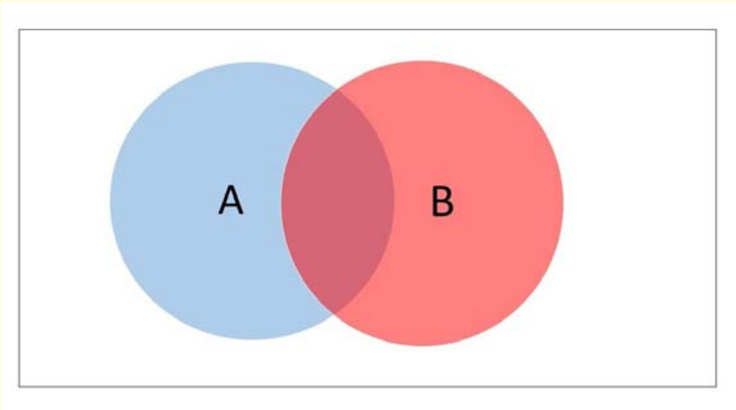
# Review: Probability Basics

- $A \cup B$ : event $A$ **OR** $B$ (or both) occur

  $A \cap B$ : event $A$ **AND** $B$ (i.e. both) occur

- **[Complement rule]** $P(A^c) = 1 - P(A)$

- **[Addition rule]** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

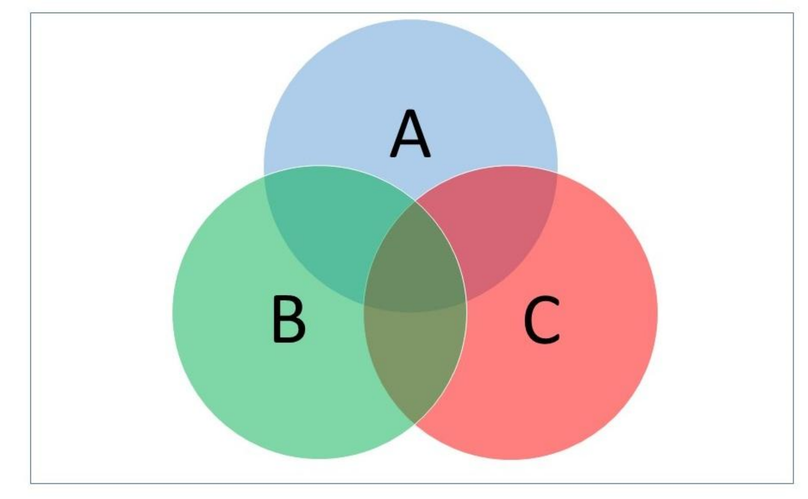


# Review: Probability Basics

- **[Addition rule]** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

- **[Addition rule extension]**

$$
\begin{aligned}
P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) \\
&\quad - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C)
\end{aligned}
$$



# Independent and Dependent events

- **Independent**: No relationship between the events:

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

# Independent and Dependent events

- **Independent**: No relationship between the events:

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

- **Dependent**: Some relationship between the events:

  $A$ may influence the chance of $B$ occurring

# Independent and Dependent events

- **Independent**: No relationship between the events:

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

- **Dependent**: Some relationship between the events:

  $A$ may influence the chance of $B$ occurring

- **Conditional probability** $P(B|A)$: The prob. of $B$ occurring, given that $A$ has occurred.

# Example: Dependent events

- **Example 1**: Free-throws in basketball (source: *Yaari and Eisenmann* (2011))

  $1^{st}$ *throw*: success-rate 73% (NBA: 2005-10):

# Example: Dependent events

- **Example 1**: Free-throws in basketball (source: *Yaari and Eisenmann* (2011))

  $1^{st}$ *throw*: success-rate 73% (NBA: 2005-10):

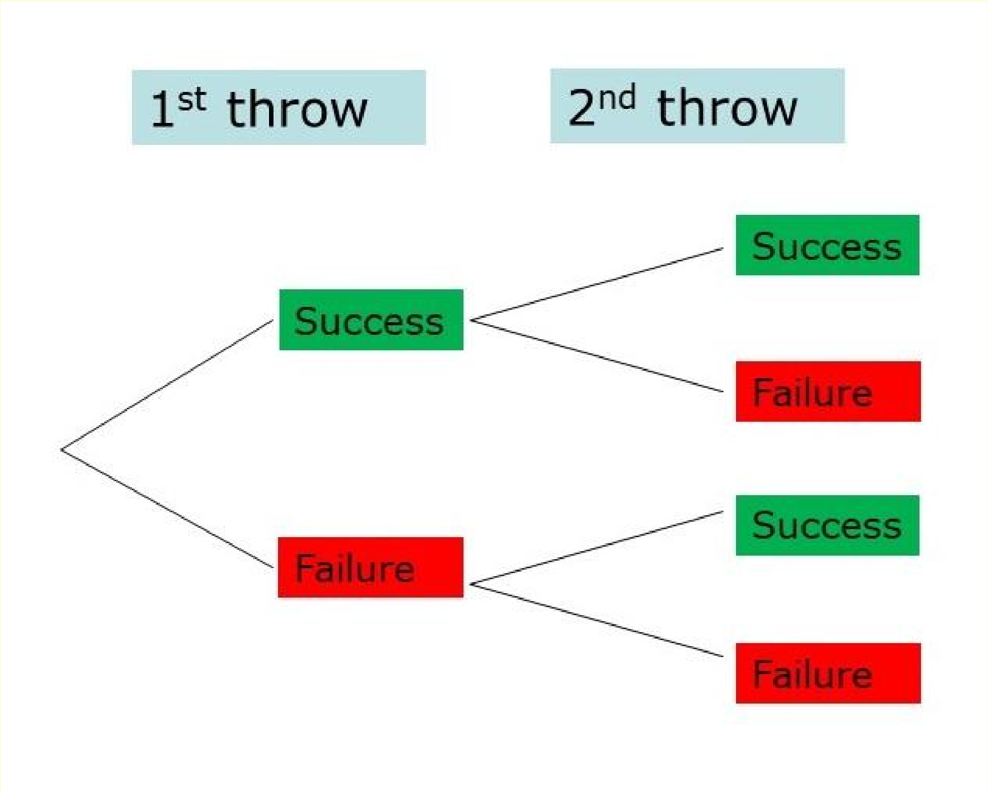
- $2^{nd}$ *throw*: success-rate depends on if $1^{st}$ was a success

  **if** $1^{st}$ **successful**: success-rate on $2^{nd}$ : 79%

  **if** $1^{st}$ **unsuccessful**: success-rate on $2^{nd}$ : 72%

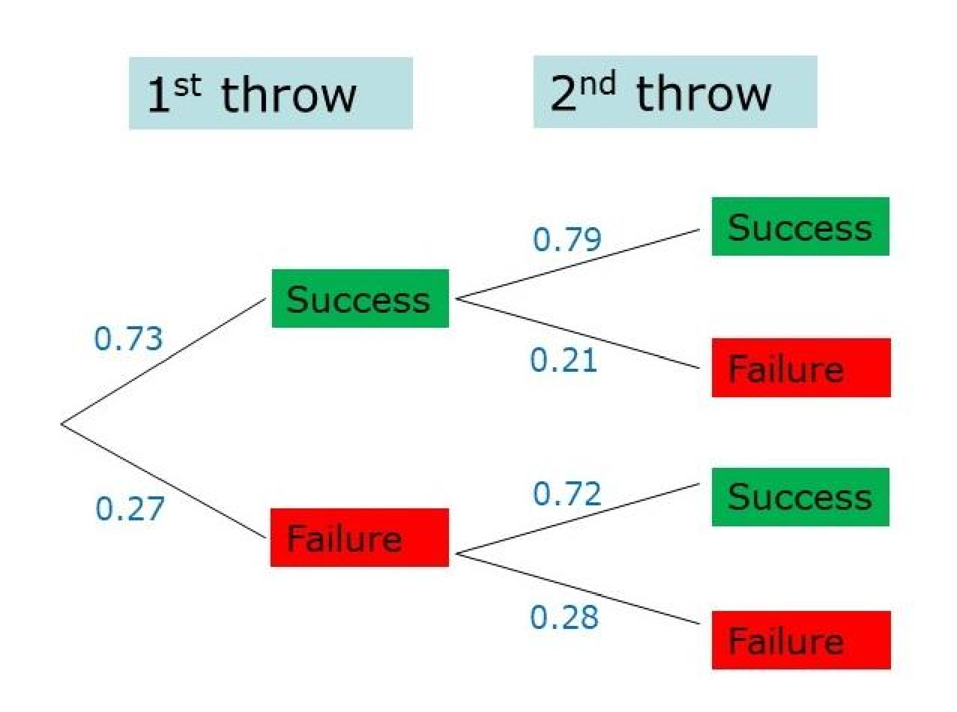
# Example 1

- Representation using a prob. tree



# Example 1

- Representation using a prob. tree
- **Q**: What is the prob. the player makes both free throws?



# General multiplication rule

- [**General multiplication Rule**]:

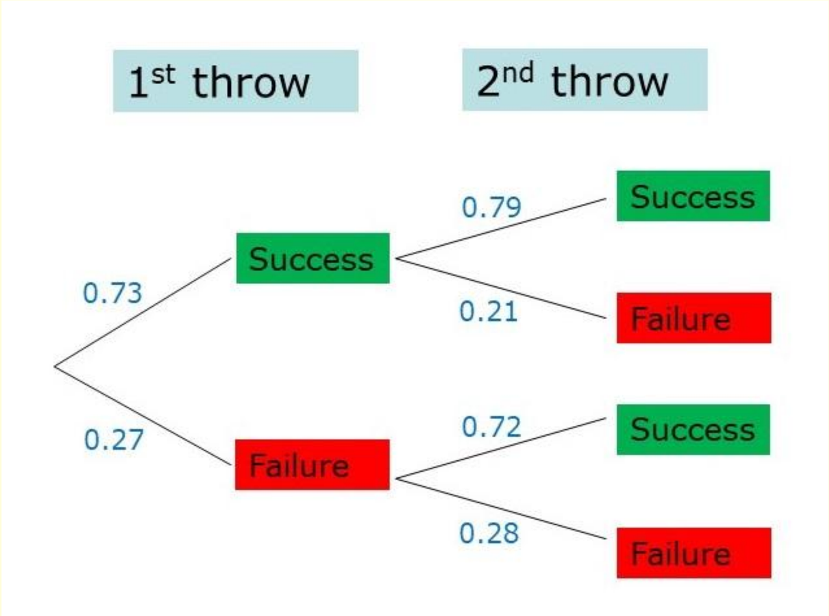
$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B|A)$$

# General multiplication rule

- **[General multiplication Rule]**:

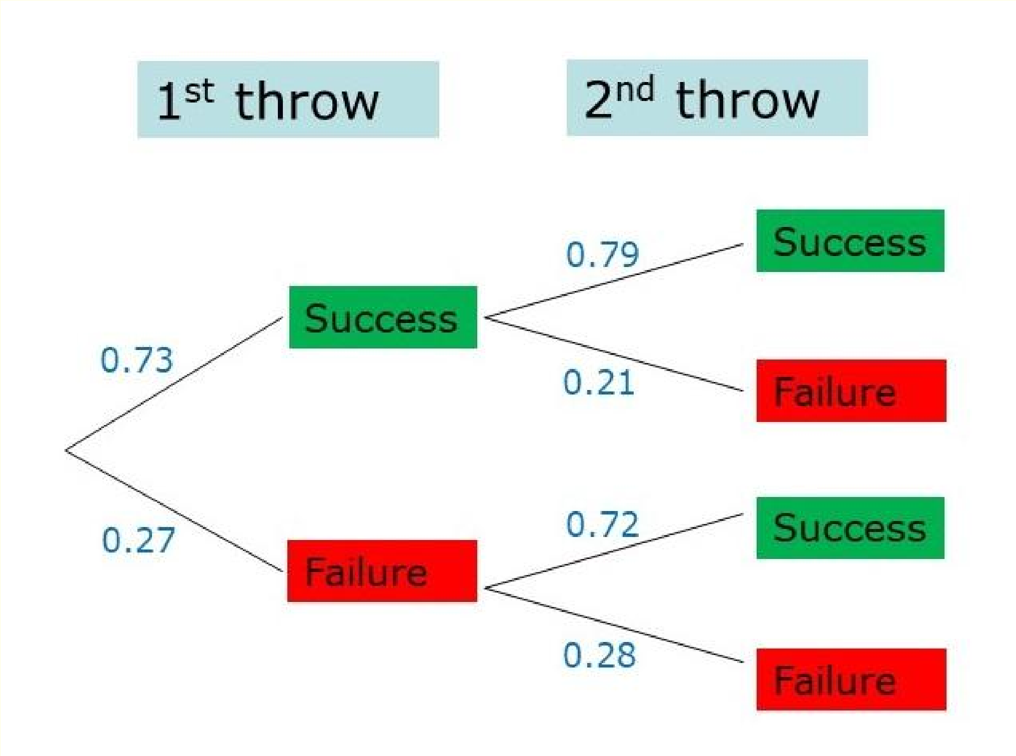
$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B|A)$$

- **Example 1**: Prob. the player makes both free throws?



# Example 1 contd.

- **Q**: Prob. the player makes only one?



# Dependent events: Dry summers and forest fires

- **Example 2**: A summer may be dry with prob. 0.4 : *affects the prob. of forest fires*

# Dependent events: Dry summers and forest fires

- **Example 2**: A summer may be dry with prob. 0.4 : *affects the prob. of forest fires*
-

| | Chance of forest fires |
|---|---|
| Dry summer | 0.6 |
| Wet summer | 0.2 |

# Clicker question 1

- A summer may be dry with prob. 0.4

| | Chance of forest fires |
|---|---|
| Dry summer | 0.6 |
| Wet summer | 0.2 |

**Q**: What is $P(\text{fire} \mid \text{dry summer})$?

**(a)** 0.2 **(b)** 0.4

**(c)** 0.6 **(d)** 1

# Solution to clicker question 1

- A summer may be dry with prob. 0.4

| | Chance of forest fires |
|---|---|
| Dry summer | 0.6 |
| Wet summer | 0.2 |

**Q**: What is $P(\text{fire} \mid \text{dry summer})$?

**(a)** 0.2

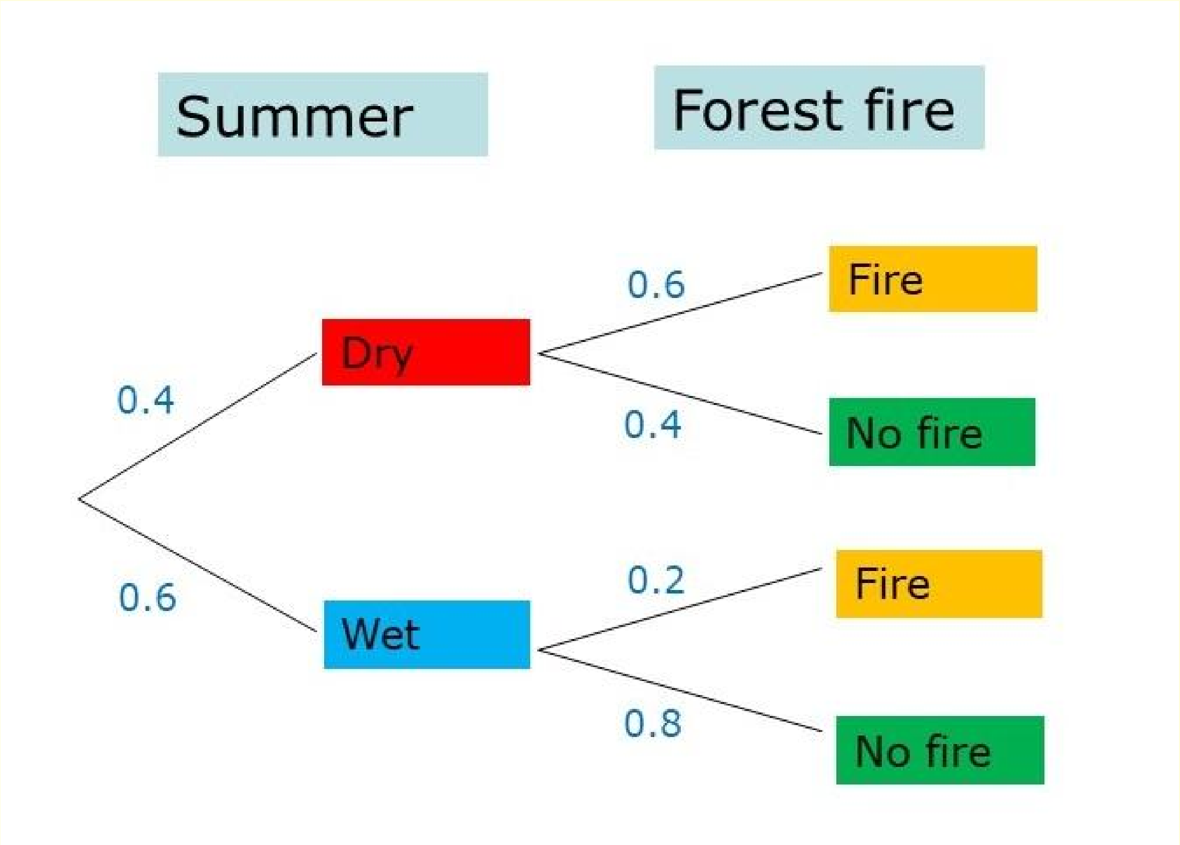
**(b)** 0.4

**(c)** 0.6

**(d)** 1

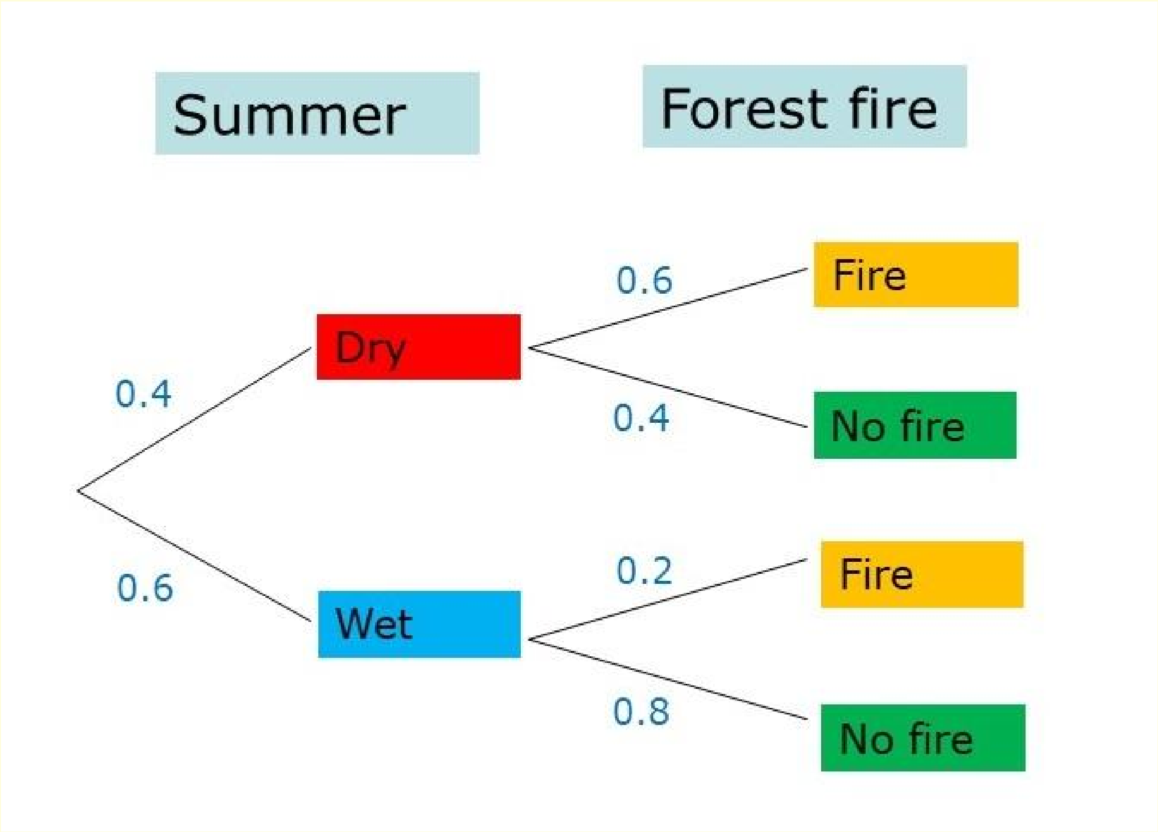
# Example 2

- Representation using a prob. tree



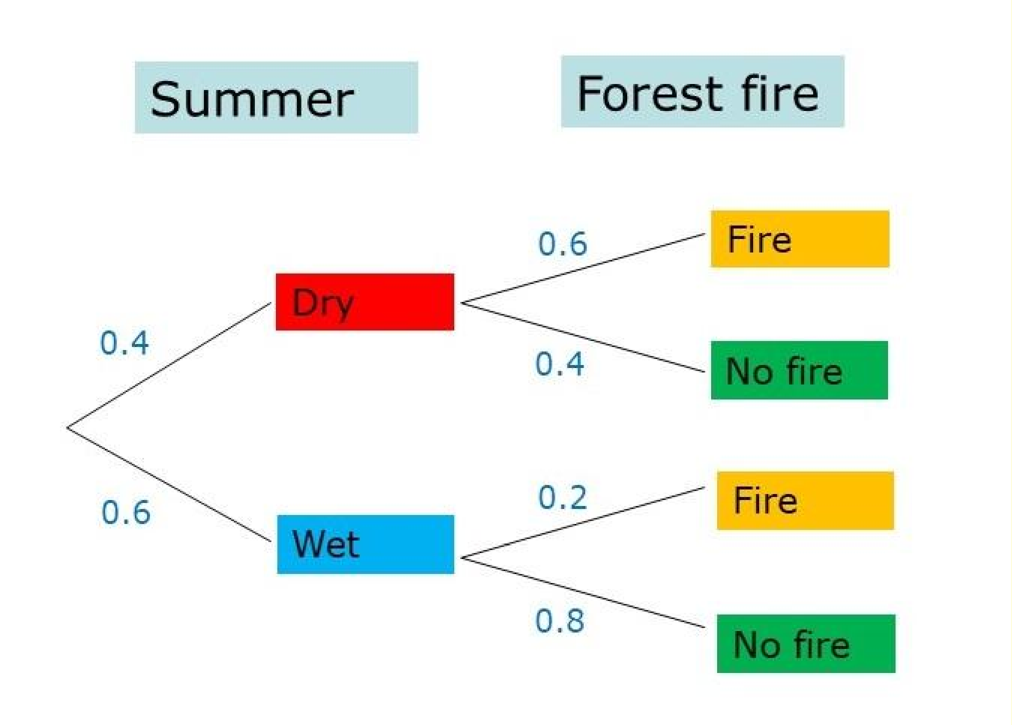
# Example 2

- Representation using a prob. tree
- $P(\text{fire} \mid \text{dry summer}) = 0.6$. What is $P(\text{fire} \cap \text{dry summer})$?



# Example 2

- Representation using a prob. tree
- $P(\text{fire} \mid \text{dry summer}) = 0.6$. What is $P(\text{fire} \cap \text{dry summer})$?
- **Q**: What is $P(\text{fire})$?



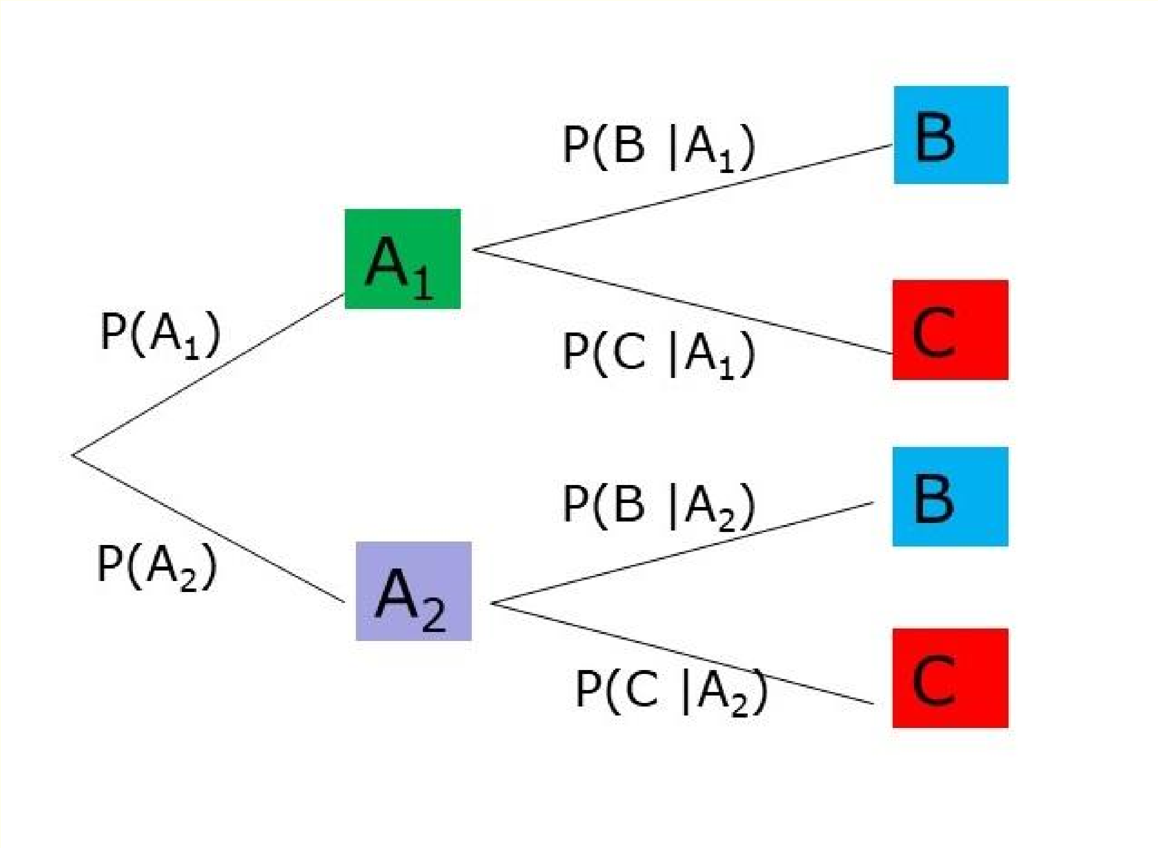
# Example 2

- Representation using a prob. tree
- $P(\text{fire} \mid \text{dry summer}) = 0.6$. What is $P(\text{fire} \cap \text{dry summer})$?
- **Q**: What is $P(\text{fire})$?
- General idea: $P(\mathit{fire}) =$

$$P(\mathit{dry})P(\mathit{fire}|\mathit{dry}) + P(\mathit{wet})P(\mathit{fire}|\mathit{wet})$$

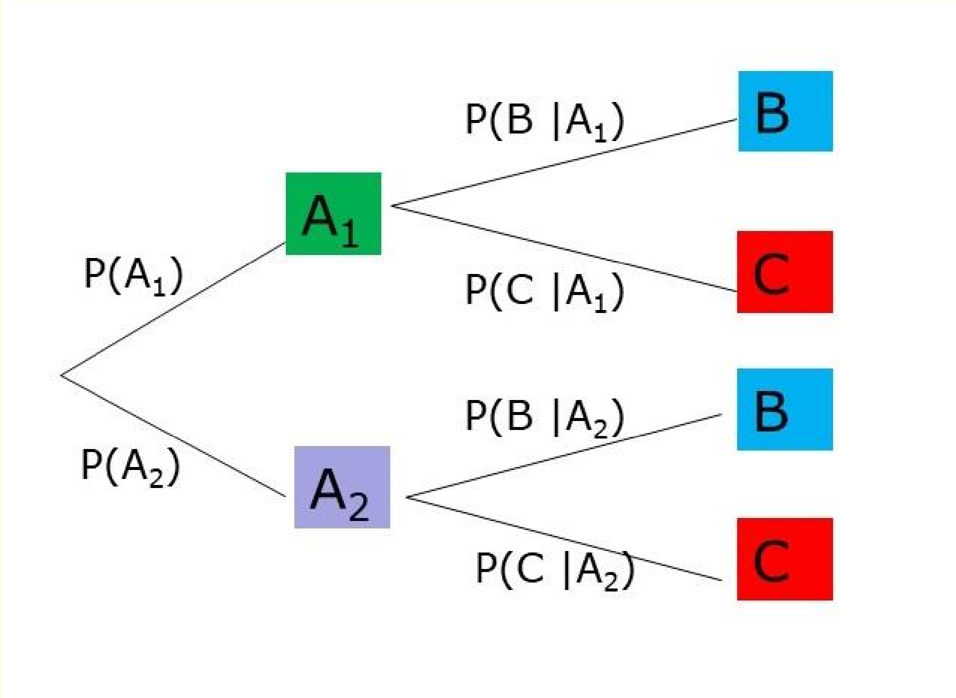
# Law of Total Probability

- Event $B$ can come about in many different ways.



# Law of Total Probability

- Event $B$ can come about in many different ways.
- $P(B) = P(A_1 \cap B) + P(A_2 \cap B)$
  $= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$



# Law of Total Probability

- Event $B$ can come about in many different ways.
- $P(B) = P(A_1 \cap B) + P(A_2 \cap B)$

  $= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$
- $P(B) = P(A \cap B) + P(A^c \cap B)$

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

# Law of Total Probability

- $P(B) = P(A_1 \cap B) + P(A_2 \cap B)$

  $= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$

# Law of Total Probability

- $P(B) = P(A_1 \cap B) + P(A_2 \cap B)$
  $= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$
- *More general:*

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) + .... + P(A_n \cap B) \\ &= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + ... + P(A_n)P(B|A_n) \\ &= \sum_{i=1}^{n} P(A_i)P(B|A_i) \end{aligned}$$

where $A_1, A_2, ......, A_n$ are disjoint

*AND* encompass all the possibilities.

# Example 2 extended

- **Example 2**: A summer may be *dry* (prob. 0.4), *normal* (prob. 0.3) or *wet* (prob. 0.3)

# Example 2 extended

- **Example 2**: A summer may be *dry* (prob. 0.4), *normal* (prob. 0.3) or *wet* (prob. 0.3)
-

| | Chance of forest fires |
|---|---|
| Dry summer | 0.6 |
| Normal summer | 0.4 |
| Wet summer | 0.2 |

# Example 2 extended

- Prob. tree

  **Q**: What is $P(\text{fire})$?

# Example 3: Picking from an unmarked platter

- **Example 3**: A party platter with 3 avocado and 7 cheese sandwiches.

# Example 3: Picking from an unmarked platter

- **Example 3**: A party platter with 3 avocado and 7 cheese sandwiches.
- *Problem*: sandwiches are unmarked!

# Example 3: Picking from an unmarked platter

- **Example 3**: A party platter with 3 avocado and 7 cheese sandwiches.
- *Problem*: sandwiches are unmarked!
- **Q**: *If you pick* $1^{st}$, prob. of picking an avocado sandwich?

# Example 3: Picking from an unmarked platter

- **Example 3**: A party platter with 3 avocado and 7 cheese sandwiches.

  **Q**: You pick $2^{nd}$, prob. of picking an avocado sandwich?

# Example 3: Picking from an unmarked platter

- **Example 3**: A party platter with 3 avocado and 7 cheese sandwiches.

  **Q**: You pick $2^{nd}$, prob. of picking an avocado sandwich?

- *If $1^{st}$ has got avocado*: $P(2^{nd}\ A | 1^{st}\ A) =$

# Clicker question 2

- A party platter with 3 avocado and 7 cheese sandwiches.

  If the $1^{st}$ person has got a cheese sandwich, what is the prob. that the $2^{nd}$ will get an avocado sandwich? i.e. What is $P(2^{nd}\ A|1^{st}\ C)$?

  **(a)** $\frac{3}{10}$ **(b)** $\frac{7}{10}$

  **(c)** $\frac{3}{9}$ **(d)** $\frac{6}{9}$

# Solution to Clicker question 2

- A party platter with 3 avocado and 7 cheese sandwiches.

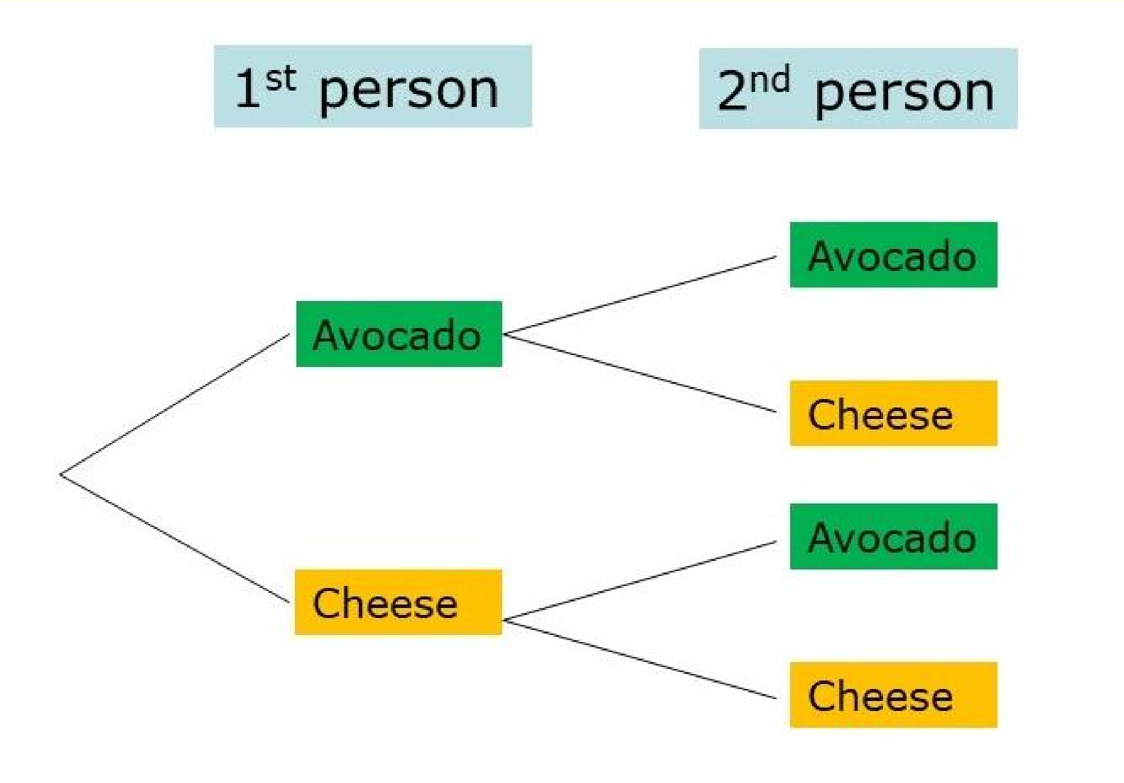
  If the $1^{st}$ person has got a cheese sandwich, what is the prob. that the $2^{nd}$ will get an avocado sandwich? i.e. What is $P(2^{nd}\ A|1^{st}\ C)$?

  **(a)** $\frac{3}{10}$ **(b)** $\frac{7}{10}$

  **(c)** $\frac{3}{9}$ **(d)** $\frac{6}{9}$

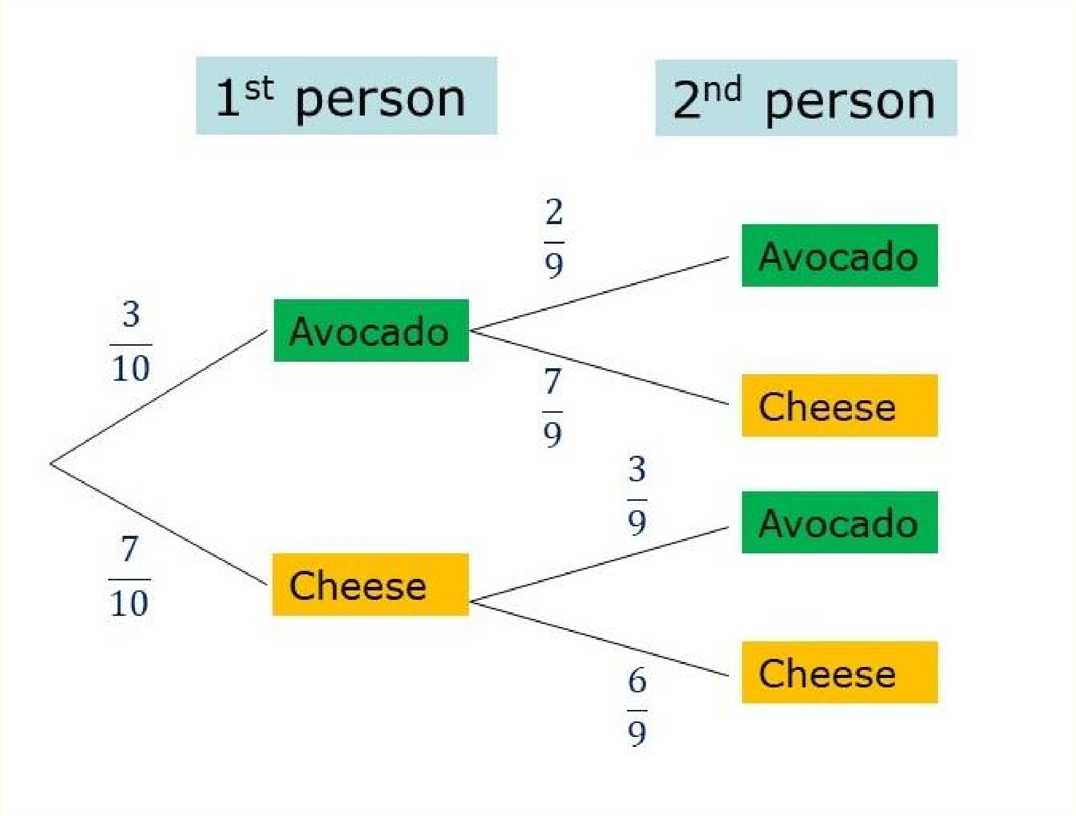
# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $2^{nd}$, prob. of picking an avocado sandwich?



# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $2^{nd}$, prob. of picking an avocado sandwich?
- Prob. tree
- Overall prob. of picking avocado, when you pick $2^{nd}$ :



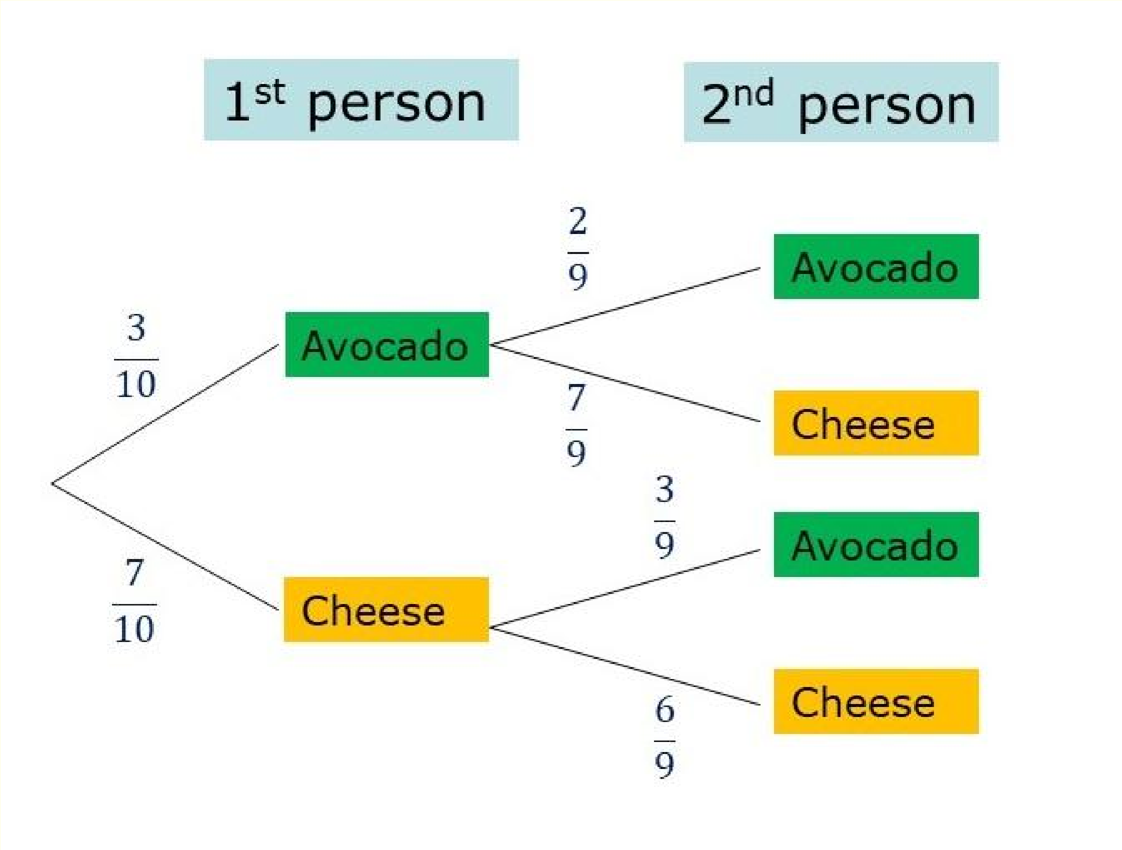
# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $2^{nd}$, prob. of picking an avocado sandwich?
- Prob. tree
- Overall prob. of picking avocado, when you pick $2^{nd}$ :
- General idea: $P(2^{nd}\ A)$

$$P(1^{st} A)P(2^{nd} A|1^{st} A) + P(1^{st} C)P(2^{nd} A|1^{st} C)$$

# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $3^{rd}$, prob. of picking an avocado sandwich?



# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $3^{rd}$, prob. of picking an avocado sandwich?
- First two outcomes: $AA$, $AC$, $CA$, $CC$

# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $3^{rd}$, prob. of picking an avocado sandwich?
- First two outcomes: $AA$, $AC$, $CA$, $CC$
- $P(AA) = \quad , P(AC) = \quad , P(CA) = \quad , P(CC) = \quad ,$
  $P(3^{rd}A|AA) = \quad , P(3^{rd}A|AC) = \quad ,$
  $P(3^{rd}A|CA) = \quad , P(3^{rd}A|CC) =$

# Example 3: Platter with 3 avocado and 7 cheese

- **Q**: You pick $3^{rd}$, prob. of picking an avocado sandwich?
- First two outcomes: $AA$, $AC$, $CA$, $CC$
- $P(AA) = \quad , P(AC) = \quad , P(CA) = \quad , P(CC) = \quad ,$
  $P(3^{rd} A|AA) = \quad , P(3^{rd} A|AC) = \quad ,$
  $P(3^{rd} A|CA) = \quad , P(3^{rd} A|CC) =$
- Overall prob. of $3^{rd}$ person picking avocado:

# Independent events

- Event $A$ has no influence on occurrence of $B$ :
  $P(B|A) = P(B) \quad \Rightarrow P(A \cap B) = P(A)P(B)$

# Independent events

- Event $A$ has no influence on occurrence of $B$ :
  $P(B|A) = P(B) \quad \Rightarrow P(A \cap B) = P(A)P(B)$

- **Example 4**: A stock's value goes up or down by \$1 each day

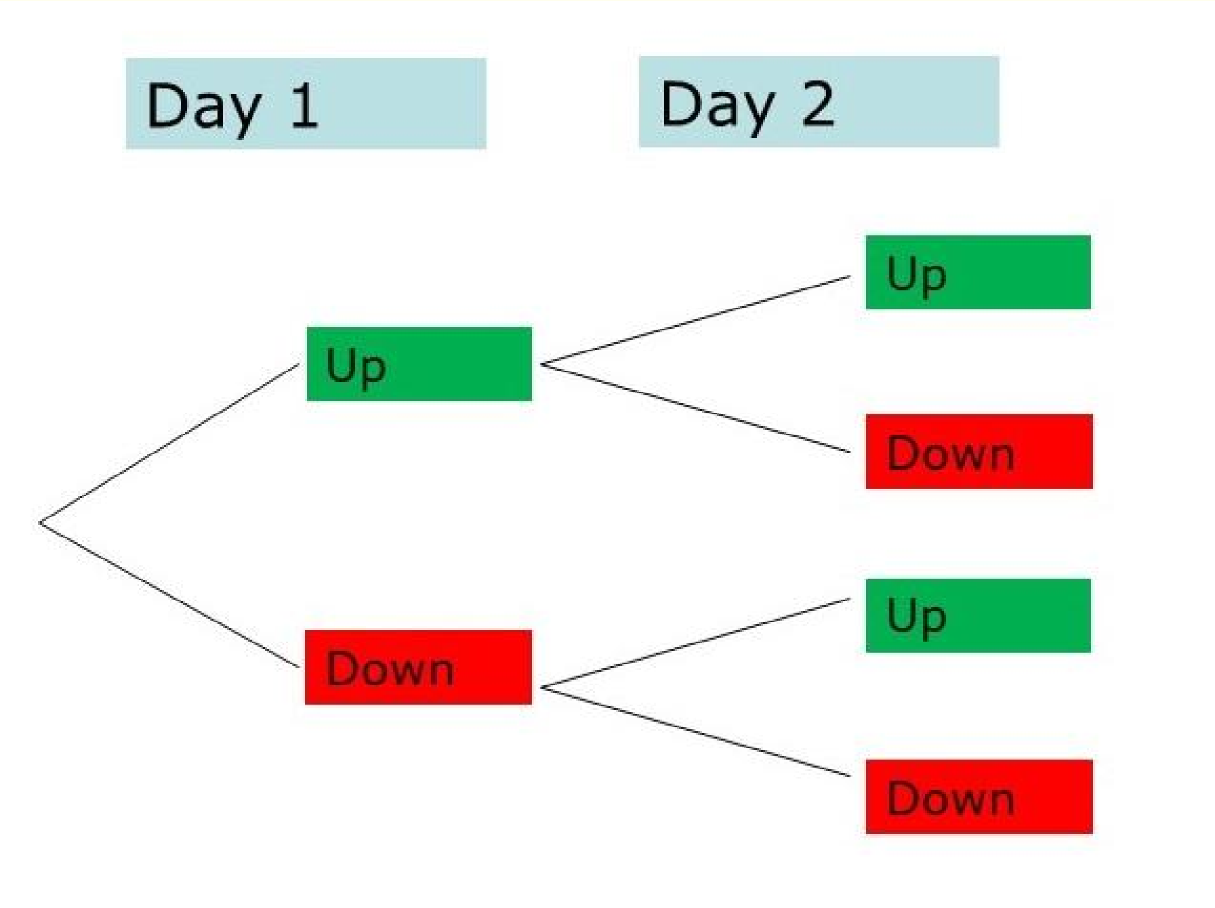
# Independent events

- Event $A$ has no influence on occurrence of $B$ :
$P(B|A) = P(B) \quad \Rightarrow P(A \cap B) = P(A)P(B)$
- **Example 4**: A stock's value goes up or down by \$1 each day
- The outcome DOES NOT depend on what happened the previous day:

| | Up tomorrow | Down tomorrow |
|---|---|---|
| Up today | 0.8 | 0.2 |
| Down today | 0.8 | 0.2 |

# Prob. tree with independence

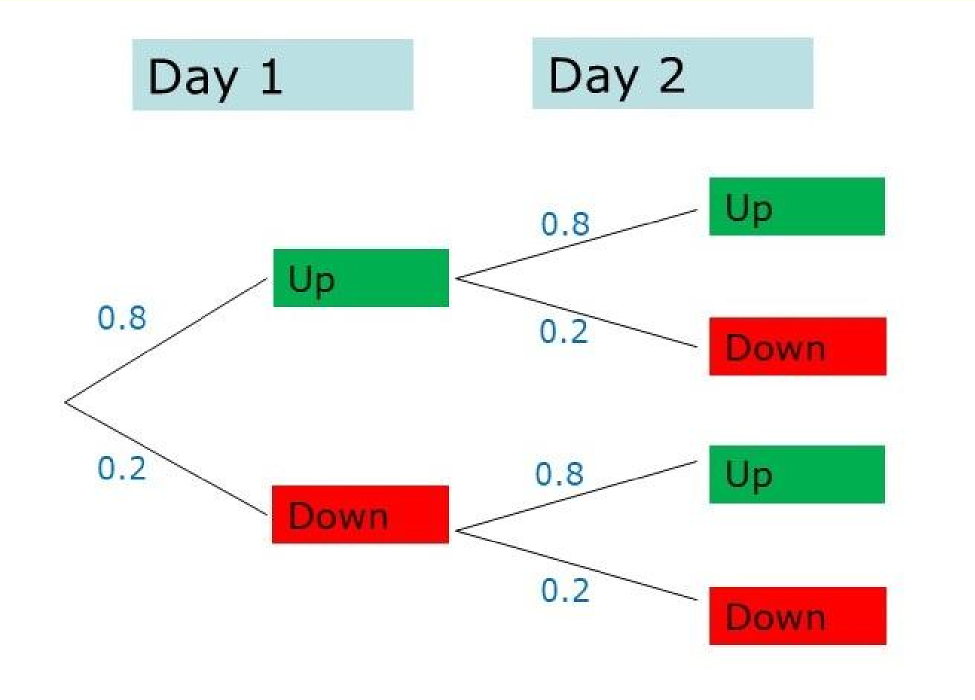
- Prob. tree



# Prob. tree with independence

- Prob. tree
- If the stock is up today, what is the prob. that

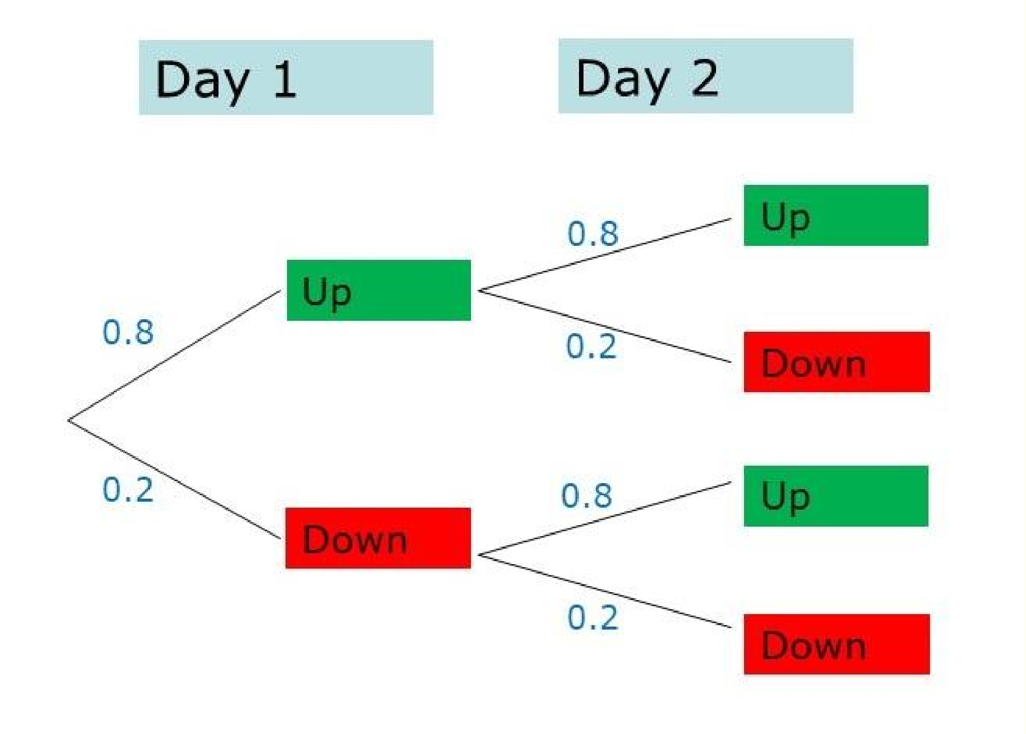
  (a) it is up the next 2 days?



# Example 4: Stock value

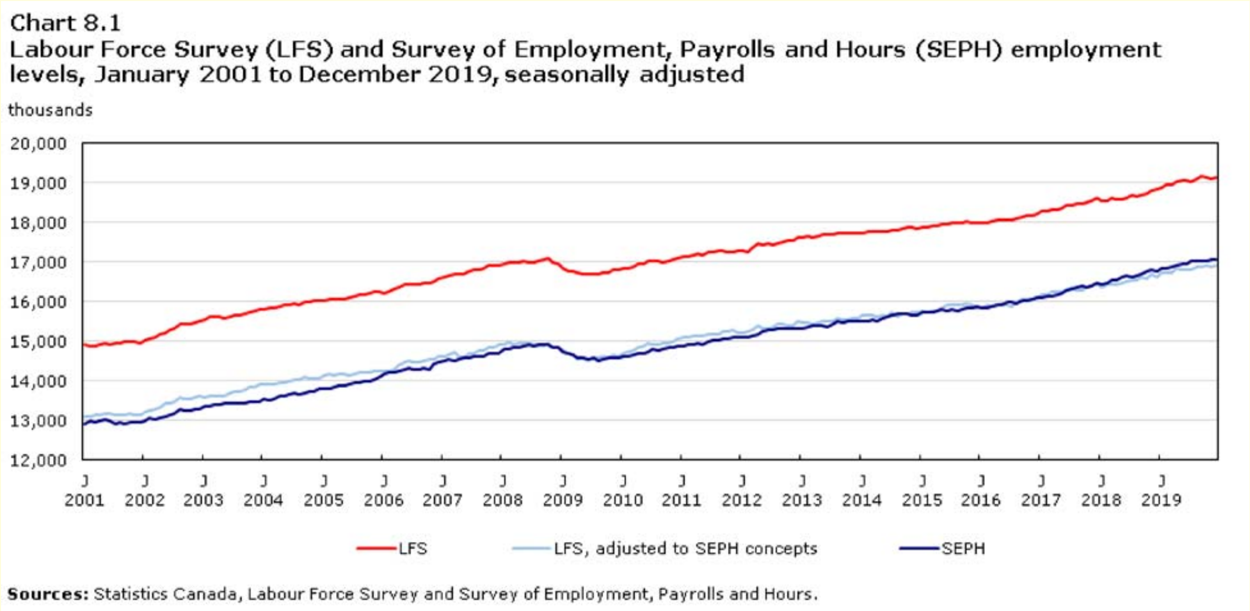
- If the stock is up today, what is the prob. that

  (b) it returns to today's value after 2 days?



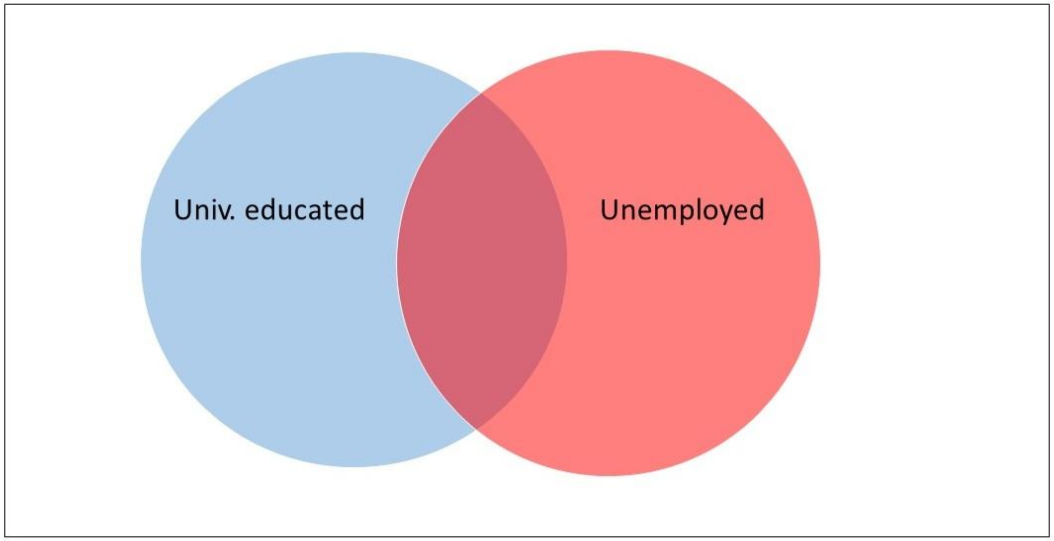
# Finding Conditional Probability

- **Example 5**: *Labour Force Survey* of 1000 people



# Finding Conditional Probability

- **Example 5**: *Labour Force Survey* of 1000 people
- *Unemployed*: 110, *Univ. educated*: 300,
  *Unemployed and Univ. educated*: 24



# Finding Conditional Probability

- **Example 5**: *Labour Force Survey* of 1000 people
- *Unemployed*: 110, *Univ. educated*: 300,

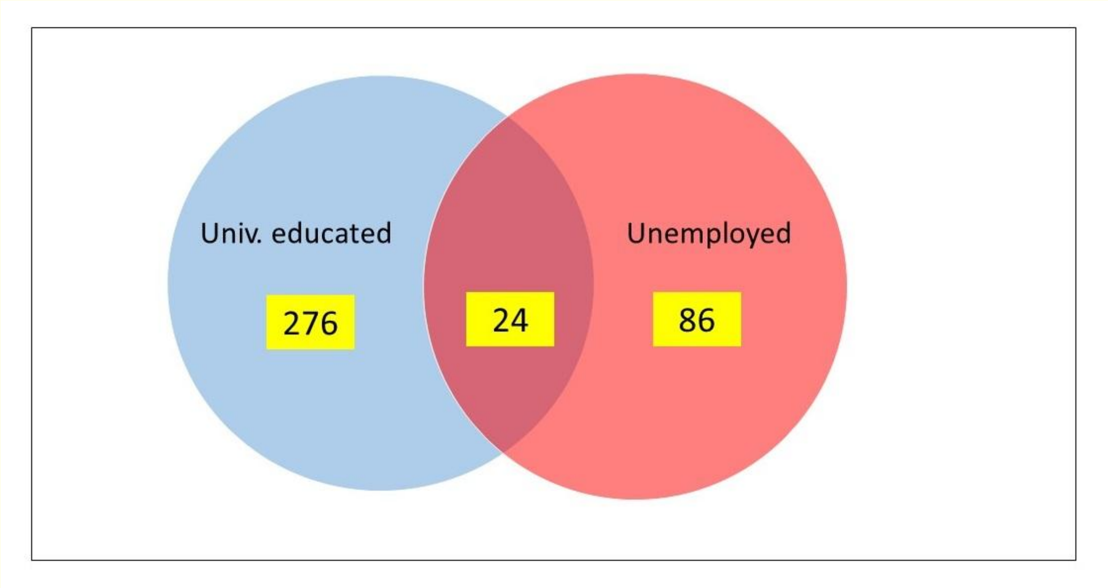
  *Unemployed and Univ. educated*: 24
- Venn diagram:

  **Q**: Unemployment rate?

  **Q**: Unemp. rate among univ. educated?

# Example 5: Labour Force Survey

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) =$

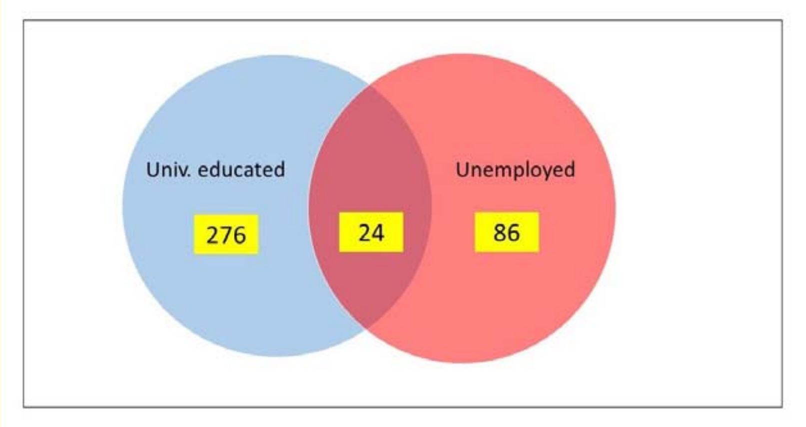


# Example 5: Labour Force Survey

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) =$
- Unemp. rate among univ.: $P(unemp|univ) =$

# Example 5: Labour Force Survey

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) =$
- Unemp. rate among univ.: $P(unemp|univ) =$
- Unemp. rate among non-univ.: $P(unemp|no\ univ) =$



# Finding Conditional Probability

- $P(A \cap B) = P(A)P(B|A) \Rightarrow$

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

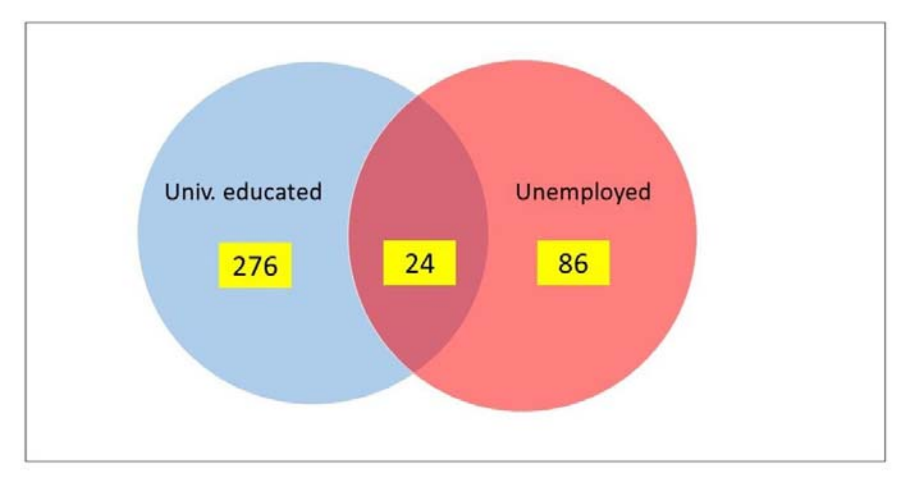
# Finding Conditional Probability

- $P(A \cap B) = P(A)P(B|A) \Rightarrow$

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

- Unemp. rate among univ.:

$$P(unemp \mid univ) = \frac{P(unemp \,\&\, univ)}{P(univ)} =$$



# Finding Conditional Probability

- $P(A \cap B) = P(A)P(B|A) \Rightarrow$

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

- Unemp. rate among univ.:

  $P(unemp \mid univ) = \frac{P(unemp \ \& \ univ)}{P(univ)} =$

- Unemp. rate among non-univ.:

  $P(unemp \mid no \ univ) = \frac{P(unemp \ \& \ no \ univ)}{P(no \ univ)} =$

# Example 5: Marginal probability

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Unemp. rates: Overall: $P(unemp) = 0.11$
- Among univ and non-univ.:

  $P(unemp|univ) = 0.08$, $P(unemp|no\ univ) = 0.123$

- **Marginal prob..** $P(unemp) =$

  $$P(univ)P(unemp \mid univ) + P(no\ univ)P(unemp \mid no\ univ)$$

# Conditional Probability: another example

- **Example 6**: 400 Stanford bicyclists at lunch time (source: *Stanford review* Sep. 2021 (author: Maxwell Meyer))

# Conditional Probability: another example

- **Example 6**: 400 Stanford bicyclists at lunch time (source: *Stanford review* Sep. 2021 (author: Maxwell Meyer))
- *Masked*: 163, *Helmeted*: 71,

  *Masked and Helmeted*: 29

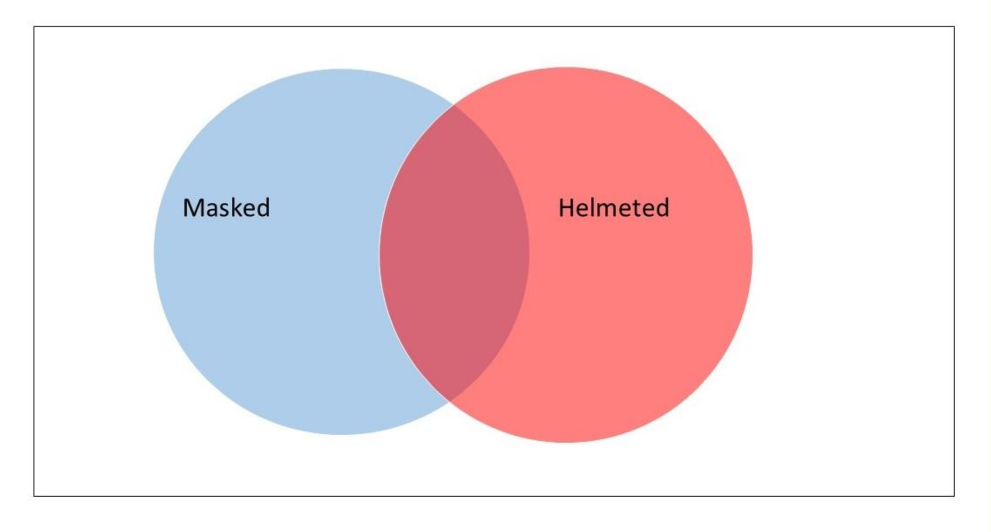
# Conditional Probability: another example

- **Example 6**: 400 Stanford bicyclists at lunch time (<u>source</u>: *Stanford review* Sep. 2021 (author: Maxwell Meyer))
- *Masked*: 163, *Helmeted*: 71,

  *Masked and Helmeted*: 29
- Venn diagram:

  **Q**: Masking rate

  **Q**: Helmeting rate?

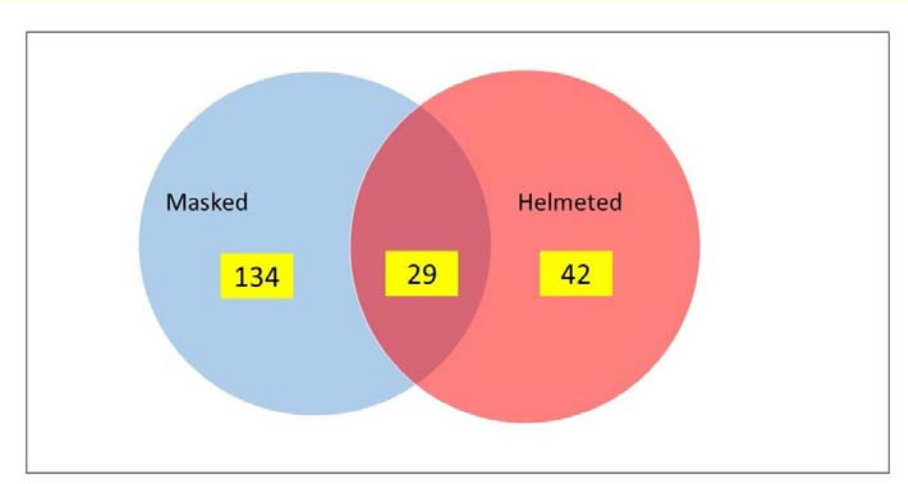


# Example 6: Stanford students

- **Q**: Is there a difference in masking between helmeted and non-helmeted students?

# Example 6: Stanford students

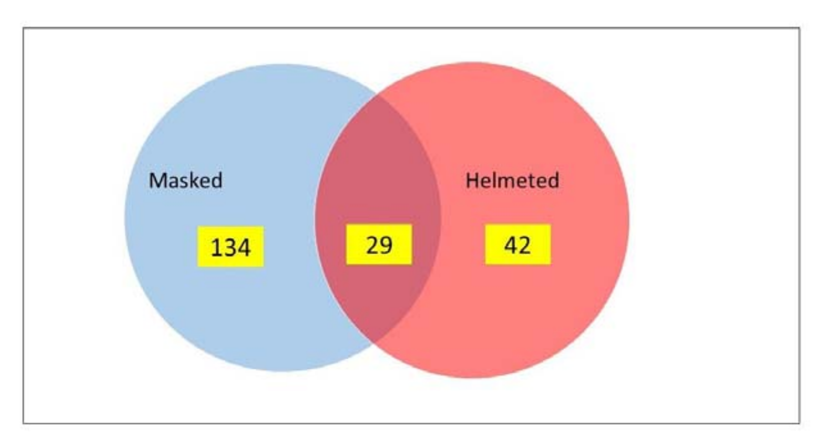
- **Q**: Is there a difference in masking between helmeted and non-helmeted students?
- Masking. rate among helmeted: $P(mask|helmet) =$



# Clicker question 3

- The following represents mask and helmet usage among Stanford cyclists:

  **Q**: What is $P(mask|helmet)$?



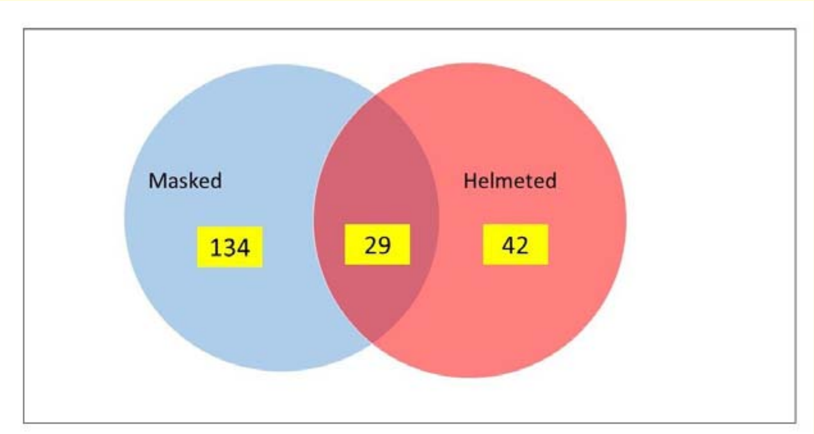

**(a)** $\frac{163}{400}$ **(b)** $\frac{134}{400}$

**(c)** $\frac{42}{71}$ **(d)** $\frac{29}{71}$

# Solution to Clicker question 3

- The following represents mask and helmet usage among Stanford cyclists:

  **Q**: What is $P(mask|helmet)$?




**(a)** $\frac{163}{400}$ **(b)** $\frac{134}{400}$

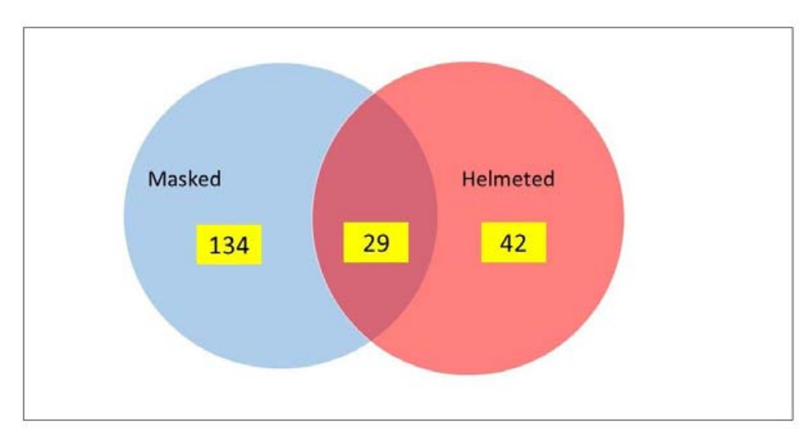
**(c)** $\frac{42}{71}$ **(d)** $\frac{29}{71}$

# Example 6: Stanford students

- **Q**: Is there a difference in masking between helmeted and non-helmeted students?

- Masking. rate among helmeted: $P(mask|helmet) = \frac{29}{71} = 0.4085$

- $P(mask|helmet) = \frac{P(mask \ \& \ helmet)}{P(helmet)} =$



# Example 6: Stanford students

- **Q**: Is there a difference in masking between helmeted and non-helmeted students?
- Masking. rate among helmeted: $P(mask|helmet) = 0.4085$
- Masking. rate among non-helmeted: $P(mask|no\ helmet) =$

# Example 6 contd.: Marginal probability

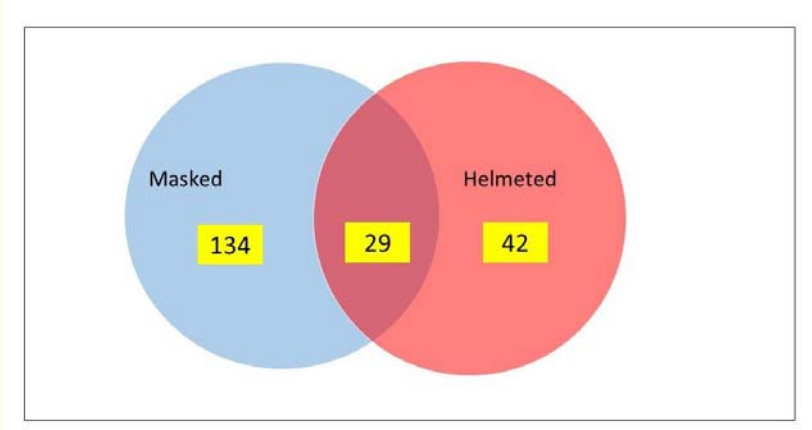
- *Overall masking rate*: $P(mask) = 0.4075$

  *Masking rate among helmeted*: $P(mask|helmet) = 0.408$

  *Among non-helmeted*: $P(mask|no\ helmet) = 0.407$

- *Marginal prob.*. $P(mask) =$

  $P(helmet)P(mask \mid helmet) + P(no\ helmet)P(mask \mid no\ helmet)$



# Prob. from two way frequency tables

- Useful way of presenting data with two characteristics

# Prob. from two way frequency tables

- Useful way of presenting data with two characteristics
- **Example 6**: Mask and helmet usage among Stanford cyclists:

| | Helmet | No helmet | *Total* |
|---|---|---|---|
| Mask | 29 | | 163 |
| No mask | | | |
| *Total* | 71 | | 400 |



# Prob. from two way frequency tables: Example 6

- 

| | Helmet | No helmet | *Total* |
|---|---|---|---|
| Mask | $\mathrm{P}(H \cap M) =$ | $\mathrm{P}(NH \cap M) =$ | $\mathrm{P}(M) =$ |
| No mask | $\mathrm{P}(H \cap NM) =$ | $\mathrm{P}(NH \cap NM) =$ | $\mathrm{P}(NM) =$ |
| *Total* | $\mathrm{P}(H) =$ | $\mathrm{P}(NH) =$ | 1 |

- $P(M|H) = \frac{P(M \cap H)}{P(H)} =$
- $P(M|NH) = \frac{P(M \cap NH)}{P(NH)} =$

# Clicker question 4

- The following represents mask and helmet usage among Stanford cyclists:

| | Helmet | No helmet | *Total* |
|---|---|---|---|
| Msk | 0.0725 | 0.335 | $\mathrm{P}(M) = 0.4075$ |
| No msk | 0.105 | 0.4875 | $\mathrm{P}(NM) = 0.4925$ |
| *Total* | $\mathrm{P}(H) = 0.1775$ | $\mathrm{P}(NH) = 0.8225$ | 1 |

**Q**: What is $P(H|M)$?

**(a)** $\frac{335}{4075}$ **(b)** $\frac{105}{1775}$

**(c)** $\frac{725}{1775}$ **(d)** $\frac{725}{4075}$

# Solution to Clicker question 4

- The following represents mask and helmet usage among Stanford cyclists:

| | Helmet | No helmet | *Total* |
|---|---|---|---|
| Msk | 0.0725 | 0.335 | $\text{P}(M) = 0.4075$ |
| No msk | 0.105 | 0.4875 | $\text{P}(NM) = 0.4925$ |
| *Total* | $\text{P}(H) = 0.1775$ | $\text{P}(NH) = 0.8225$ | 1 |

**Q**: What is $P(H|M)$?

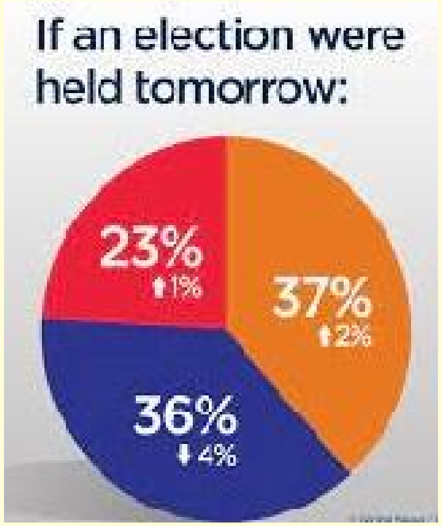
**(a)** $\frac{335}{4075}$ **(b)** $\frac{105}{1775}$

**(c)** $\frac{725}{1775}$ **(d)** $\frac{725}{4075}$

# Prob. from two way frequency tables: Example 7

- **Example 7**: Survey of 1000 voters about support for 3 parties:

| | Under 30 yrs | 30-64 yrs | 65+ years | *Total* |
|---|---|---|---|---|
| Party *A* | 180 | 110 | 80 | |
| Party *B* | 120 | 120 | 50 | |
| Party *C* | 100 | 170 | 70 | |
| *Total* | | | | 1000 |



# Prob. from two way frequency tables: Example 7

- **Example 7**: Survey of 1000 voters about support for 3 parties:

| | Under 30 yrs | 30-64 yrs | 65+ years | *Total* |
|---|---|---|---|---|
| Party $A$ | 180 | 110 | 80 | |
| Party $B$ | 120 | 120 | 50 | |
| Party $C$ | 100 | 170 | 70 | |
| *Total* | | | | 1000 |

- Support among under 30s:

$P(A|young) = \quad , P(B|yng) = \quad , P(C|yng) = \quad ,$

# Prob. from two way frequency tables: Example 7

- **Example 7**: Survey of 1000 voters about support for 3 parties:

| | Under 30 yrs | 30-64 yrs | 65+ years | *Total* |
|---|---|---|---|---|
| Party $A$ | 180 | 110 | 80 | |
| Party $B$ | 120 | 120 | 50 | |
| Party $C$ | 100 | 170 | 70 | |
| *Total* | | | | 1000 |

- Support among under 30s:

  $P(A|young) =$ , $P(B|yng) =$ , $P(C|yng) =$ ,

- Support among 30-64s:

  $P(A|middle) =$ , $P(B|mid) =$ , $P(C|mid) =$ ,

# Prob. from two way frequency tables: Example 7

- **Example 7**: Survey of 1000 voters about support for 3 parties:

| | Under 30 yrs | 30-64 yrs | 65+ years | *Total* |
|---|---|---|---|---|
| Party $A$ | 180 | 110 | 80 | |
| Party $B$ | 120 | 120 | 50 | |
| Party $C$ | 100 | 170 | 70 | |
| *Total* | | | | 1000 |

- Support among under 30s:

$P(A|young) = \quad , P(B|yng) = \quad , P(C|yng) = \quad ,$

- Support among 30-64s:

$P(A|middle) = \quad , P(B|mid) = \quad , P(C|mid) = \quad ,$

- Support among 65+s:

$P(A|older) = \quad , P(B|older) = \quad , P(C|older) = \quad ,$

# Example 7: Support for 3 parties

- $P(A|young) = 0.45$, $P(B|young) = 0.3$, $P(C|young) = 0.25$,
  $P(A|mid) = 0.275$, $P(B|mid) = 0.3$, $P(C|mid) = 0.425$,
  $P(A|older) = 0.4$, $P(B|older) = 0.25$, $P(C|older) = 0.35$,

# Example 7: Support for 3 parties

- $P(A|young) = 0.45, P(B|young) = 0.3, P(C|young) = 0.25,$

  $P(A|mid) = 0.275, P(B|mid) = 0.3, P(C|mid) = 0.425,$

  $P(A|older) = 0.4, P(B|older) = 0.25, P(C|older) = 0.35,$

- If each group constitutes $\frac{1}{3}$ of the electorate, which party is likely to win? *Marginal probs.*:

# Example 7: Support for 3 parties

- $P(A|young) = 0.45$, $P(B|young) = 0.3$, $P(C|young) = 0.25$,
  $P(A|mid) = 0.275$, $P(B|mid) = 0.3$, $P(C|mid) = 0.425$,
  $P(A|older) = 0.4$, $P(B|older) = 0.25$, $P(C|older) = 0.35$,

- If young voters constitute 25% of the electorate and middle-aged voters constitute 50%, which party is likely to win?

# Reverse Conditioning: Bayes' rule

- You know $P(A|B)$. **Q: How to find** $P(B|A)$?

# Reverse Conditioning: Bayes' rule

- You know $P(A|B)$. **Q: How to find** $P(B|A)$?
- *Example:* Driving while drunk increases chance of accidents.

  **Q**: Prob. someone who had crashed was drunk?

# Reverse Conditioning: Bayes' rule

- You know $P(A|B)$. **Q: How to find** $P(B|A)$?
- *Example:* Driving while drunk increases chance of accidents.

  **Q**: Prob. someone who had crashed was drunk?
- From research studies, you know $P(accident|drunk)$

  **Q**: How to find $P(drunk|accident)$?

# Reverse Conditioning: Bayes' rule

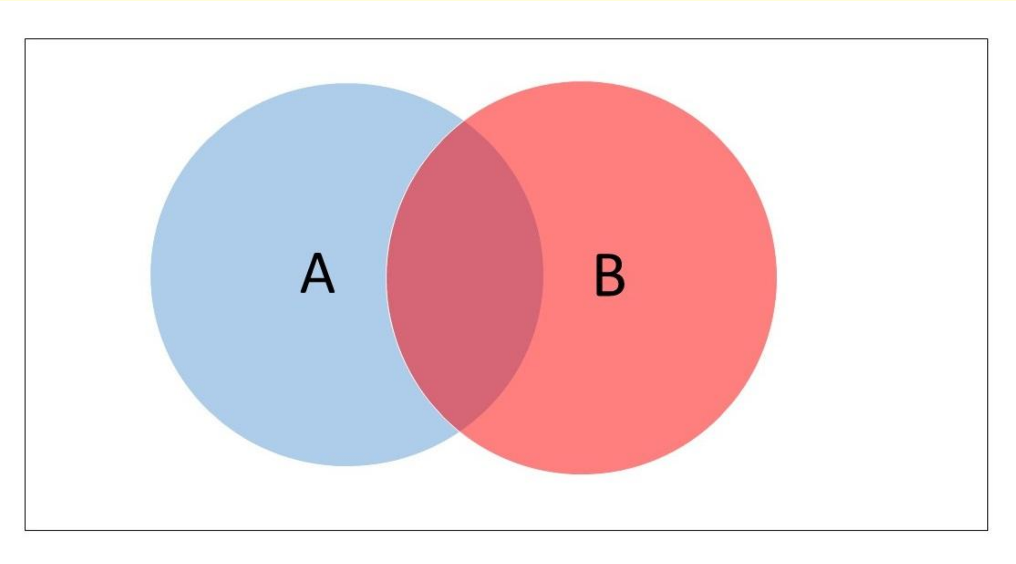
- You know $P(A|B)$. **Q: How to find** $P(B|A)$?
- *Example:* Driving while drunk increases chance of accidents.

  **Q**: Prob. someone who had crashed was drunk?
- From research studies, you know $P(accident|drunk)$

  **Q**: How to find $P(drunk|accident)$?
- **Answer**: Bayes' rule

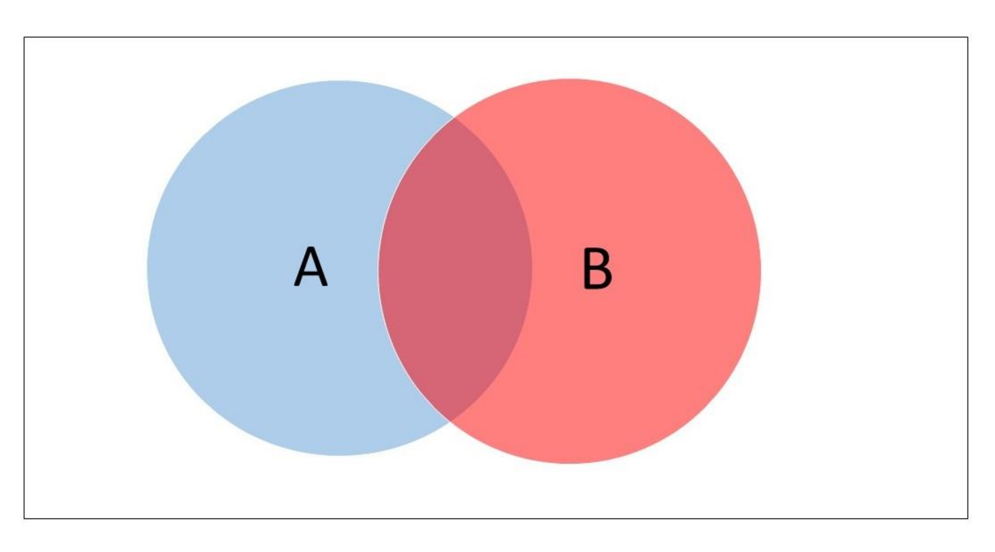
# Reverse Conditioning: Bayes' rule

- Basic idea: $P(B|A) = \frac{P(A \cap B)}{P(A)}$ e.g. $B \equiv$booze, $A \equiv$accident



# Reverse Conditioning: Bayes' rule

- Basic idea: $P(B|A) = \frac{P(A \cap B)}{P(A)}$ e.g. $B \equiv$booze, $A \equiv$accident
- *Also:* $P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Rightarrow P(A \cap B) = P(A|B)P(B)$



# Reverse Conditioning: Bayes' rule

- Basic idea: $P(B|A) = \frac{P(A \cap B)}{P(A)}$ e.g. $B \equiv$booze, $A \equiv$accident
- *Also:* $P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \Rightarrow P(A \cap B) = P(A|B)P(B)$
- **Bayes' rule (almost):** $P(B|A) = \dfrac{P(A|B)P(B)}{P(A)}$

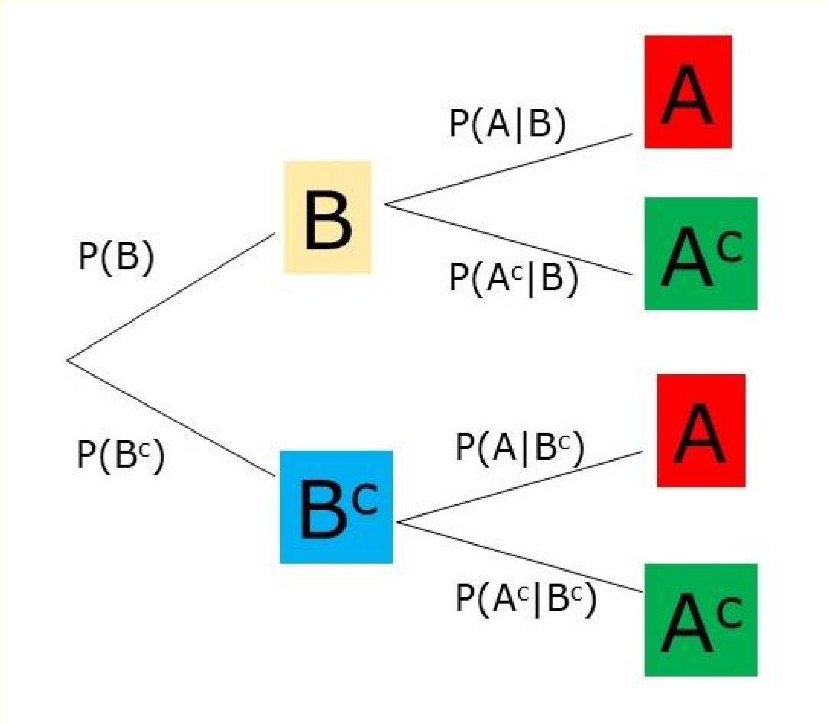
# Bayes' rule

- *Bayes' rule (almost):* $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$

# Bayes' rule

- *Bayes' rule (almost):* $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$
- *One more step:* How can $A$ occur?

  Marginal prob.: $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$



# Bayes' rule

- *Bayes' rule (almost):* $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$
- *One more step:* How can $A$ occur?

  Marginal prob.: $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$
- Finally **Bayes' rule**:

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

# Bayes' rule: Application

- **Example 8**: $P(Accident|Binge) = 0.5$,
  $P(Accident|Sober) = 0.1$, $P(Binge) = 0.2$

# Bayes' rule: Application

- **Example 8**: $P(Accident|Binge) = 0.5$, $P(Accident|Sober) = 0.1, \quad P(Binge) = 0.2$
- *An accident occurs:* What chance the driver was drunk?

# Bayes' rule: Application

- **Example 8**: $P(Accident|Binge) = 0.5$, $P(Accident|Sober) = 0.1, \quad P(Binge) = 0.2$
- *An accident occurs:* What chance the driver was drunk?
- **Bayes' rule**:

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

# Bayes' rule: Terminology

- **Bayes' rule**:

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

# Bayes' rule: Terminology

- **Bayes' rule**:

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

- Start with $P(B) \equiv$ **prior**,

  *Updating due to new information*: $P(B|A) \equiv$ **posterior**

# Bayes' rule: Terminology

- **Bayes' rule**:

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

- Start with $P(B) \equiv$ **prior**,

  *Updating due to new information*: $P(B|A) \equiv$ **posterior**

- **Example 9**: Stock $S$ maybe hot or cold!

  Prior prob. $P(hot) = 0.4$

# Bayes' rule: Terminology

- **Bayes' rule**:

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

- Start with $P(B) \equiv$ **prior**,

  *Updating due to new information*: $P(B|A) \equiv$ **posterior**

- **Example 9**: Stock $S$ maybe hot or cold!

  Prior prob. $P(hot) = 0.4$

- *Dr. Stock* gives it a "Buy" recommendation.

  Updated $P(hot|\text{"Buy"}) =?$

# Bayes' rule: Example 9

- **Example 9**: Prior prob. $P(hot) = 0.4$

  *Dr. Stock* gives it a "Buy" recommendation.

  Problem: *Dr. Stock* is sometimes wrong!

# Bayes' rule: Example 9

- **Example 9**: Prior prob. $P(hot) = 0.4$

  *Dr. Stock* gives it a "Buy" recommendation.

  Problem: *Dr. Stock* is sometimes wrong!

- *Dr. Stock*: $P(\text{"Buy"} \mid \text{hot}) = 0.8$, $P(\text{"Buy"} \mid \text{cold}) = 0.4$

# Bayes' rule: Example 9

- **Example 9**: Prior prob. $P(hot) = 0.4$

  *Dr. Stock* gives it a "Buy" recommendation.

  Problem: *Dr. Stock* is sometimes wrong!

- *Dr. Stock*: $P(\text{"Buy"} \mid \text{hot}) = 0.8$, $P(\text{"Buy"} \mid \text{cold}) = 0.4$

- Bayes' rule:

$$P(hot|''Buy'') = \frac{P(''Buy''|hot)P(hot)}{P(''Buy'')}$$

# Bayes' rule: Example 9

- **Example 9**: Prior prob. $P(hot) = 0.4$

  *Dr. Stock*: $P(\text{"Buy"} \mid \text{hot}) = 0.8$, $P(\text{"Buy"} \mid \text{cold}) = 0.4$

- Bayes' rule: $P(hot|"Buy") = \frac{P("Buy"|hot)P(hot)}{P("Buy")}$

$$= \frac{P("Buy"|hot)P(hot)}{P("Buy"|hot)P(hot) + P("Buy"|cold)P(cold)}$$

# Bayes' rule: Example 9 contd.

- **Example 9 contd**: Prior prob. $P(hot) = 0.4$

  What if *Dr. Stock* could predict perfectly?

  *Dr. Stock*: $P(\text{"Buy"} \mid \text{hot}) = 1$, $P(\text{"Buy"} \mid \text{cold}) = 0$

# Bayes' rule: Example 9 contd.

- **Example 9 contd**: Prior prob. $P(hot) = 0.4$

  What if *Dr. Stock* could predict perfectly?

  *Dr. Stock*: $P(\text{"Buy"} \mid \text{hot}) = 1$, $P(\text{"Buy"} \mid \text{cold}) = 0$

- Bayes' rule: $P(hot|"Buy") =$

$$\frac{P("Buy"|hot)P(hot)}{P("Buy"|hot)P(hot) + P("Buy"|cold)P(cold)}$$

# Clicker question 5

- Given his symptoms, Jim assesses that his probability of being infected is 0.3. So he decides to take a test.

  The test shows either $+ive$ or $-ive$, but it's not fully accurate. It has the following probabilities of detecting infection:

  $P(+ive \mid \text{infected}) = 0.7$, $P(+ive \mid \text{not infected}) = 0.1$.

  **Q**: What is Jim's prior here?

  **(a)** 0.1 **(b)** 0.2

  **(c)** 0.3 **(d)** 0.7

# Solution to Clicker question 5

- Given his symptoms, Jim assesses that his probability of

  being infected is 0.3. So he decides to take a test.

  The test shows either $+ive$ or $-ive$, but it's not fully accurate. It has the following probabilities of detecting infection:

  $P(+ive \mid \text{infected}) = 0.7$, $P(+ive \mid \text{not infected}) = 0.1$.

  **Q**: What is Jim's prior here?

  **(a)** 0.1 **(b)** 0.2

  **(c)** 0.3 **(d)** 0.7

# Bayes' Rule application: medical testing

- Prob. (infection) $= 0.3$.

  $P(+ive \mid \text{infected}) = 0.7$, $P(+ive \mid \text{not infected}) = 0.1$.

- False positives and false negatives

# Bayes' Rule application: medical testing

- Prob. (infection) $= 0.3$.

  $P(+ive \mid \text{infected}) = 0.7$, $P(+ive \mid \text{not infected}) = 0.1$.

- False positives and false negatives

- **Ex**: Use Bayes' rule to find $P(\text{infected} \mid +ive) =$

# Bayes' rule: Many applications

- Update belief about public figure/celebrity based on news revelations. Sometimes news is wrong!

# Bayes' rule: Many applications

- Update belief about public figure/celebrity based on news revelations. Sometimes news is wrong!
- *AI*: Spam filter, ads,.....

# Bayes' rule: Many applications

- Update belief about public figure/celebrity based on news revelations. Sometimes news is wrong!
- *AI*: Spam filter, ads,.....
- Update belief about effectiveness of drugs/vaccines based on test results. Update beliefs about theories based on facts/experimental results.

# Bayes' rule: Many applications

- Update belief about public figure/celebrity based on news revelations. Sometimes news is wrong!
- *AI*: Spam filter, ads,.....
- Update belief about effectiveness of drugs/vaccines based on test results. Update beliefs about theories based on facts/experimental results.
- Update beliefs about criminality/innocence of defendant based on testimonies/evidence.

# Recap

- Conditional probability
- General multiplication rule: $P(A \cap B) = P(B|A)P(A)$
- Total/Marginal probability:
  $P(B) = P(A)P(B|A) + P(A^c)P(B|A^c)$
  $= P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + ... + P(A_n)P(B|A_n)$
- Probability trees, two-way tables
- Bayes' rule: Introduction, main formula and applications.