Probability 2: Conditional Probability and Bayes' rule

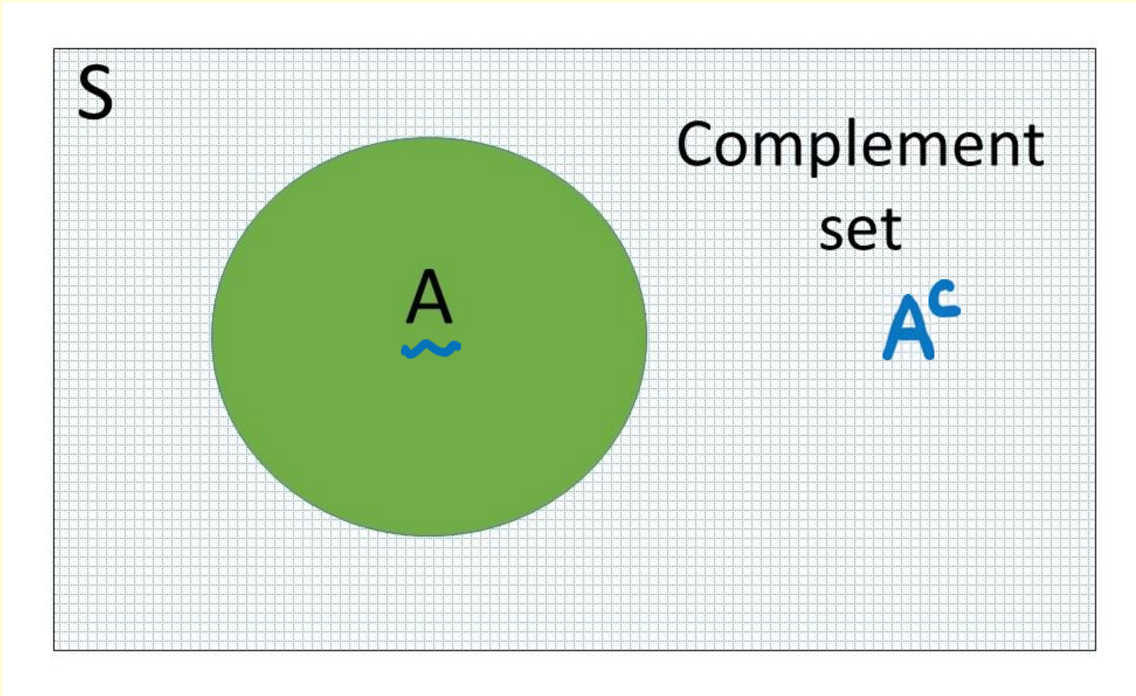
Module Outline

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

Review: Probability Basics

- *Sample space S:* set of <u>all</u> possible outcomes

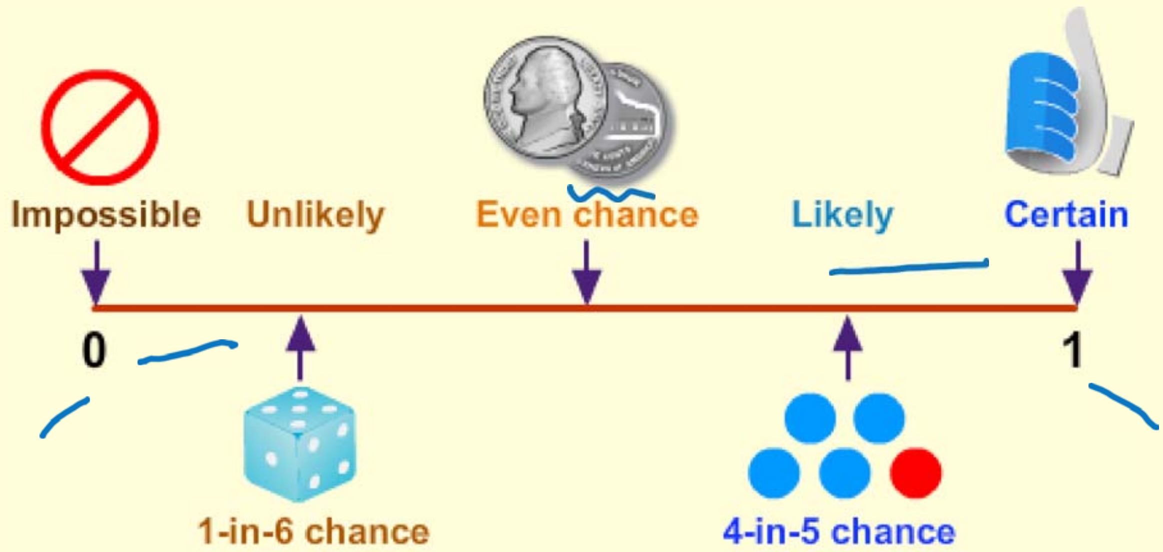
  Event: some subset of *S*



Review: Probability Basics

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

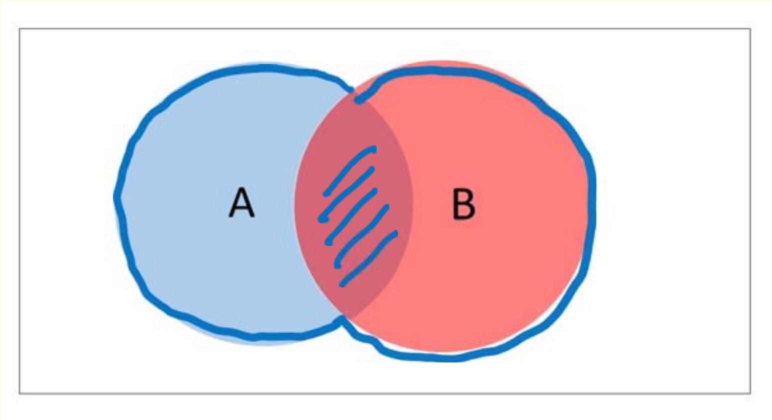
If A and B disjoint, $P(A \cup B) = P(A) + P(B)$

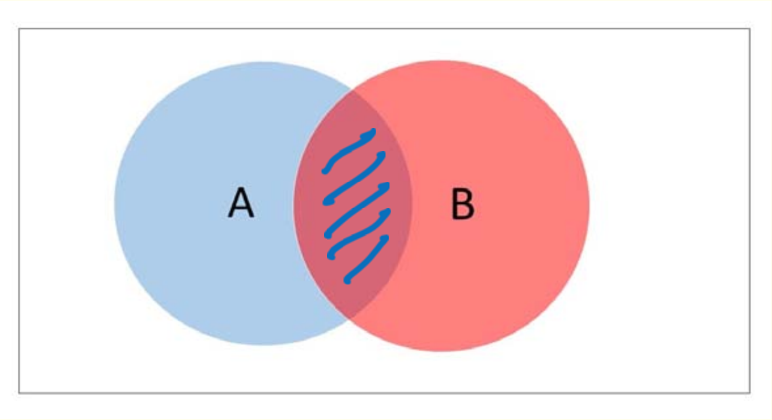


Review: Probability Basics

- $A \cup B$: 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 **<u>OR</u>** B (or both) occur

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

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

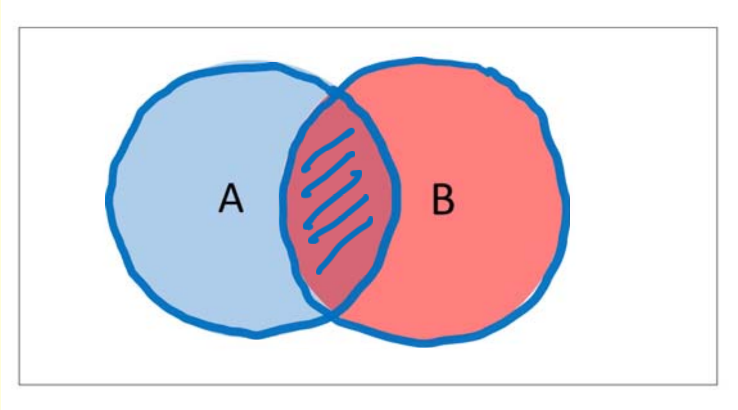
Review: Probability Basics

- $A \cup B$: 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]** $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

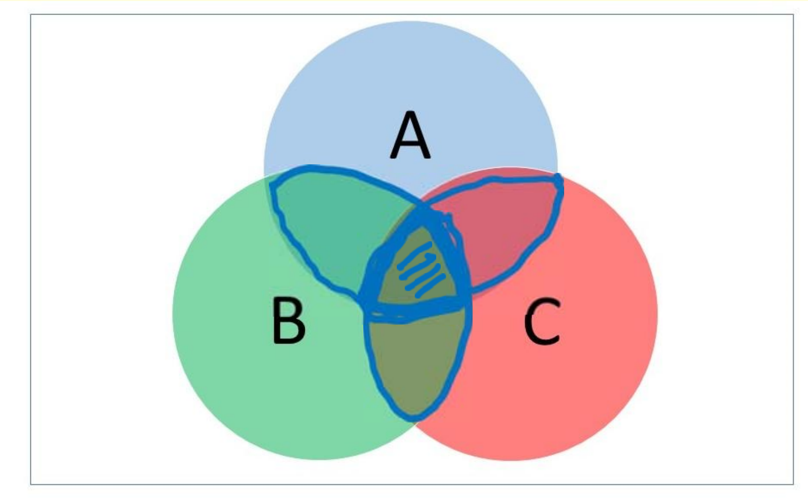


Review: Probability Basics

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

- **[Addition rule extension]**

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



Independent and Dependent events

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

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

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.

B given A

Example: Dependent events

- **Example 1**: Free-throws in basketball

  <u>1^{st} *throw*</u>: success-rate 73% (NBA: 2005-10):

Example: Dependent events

- **Example 1**: Free-throws in basketball

  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%

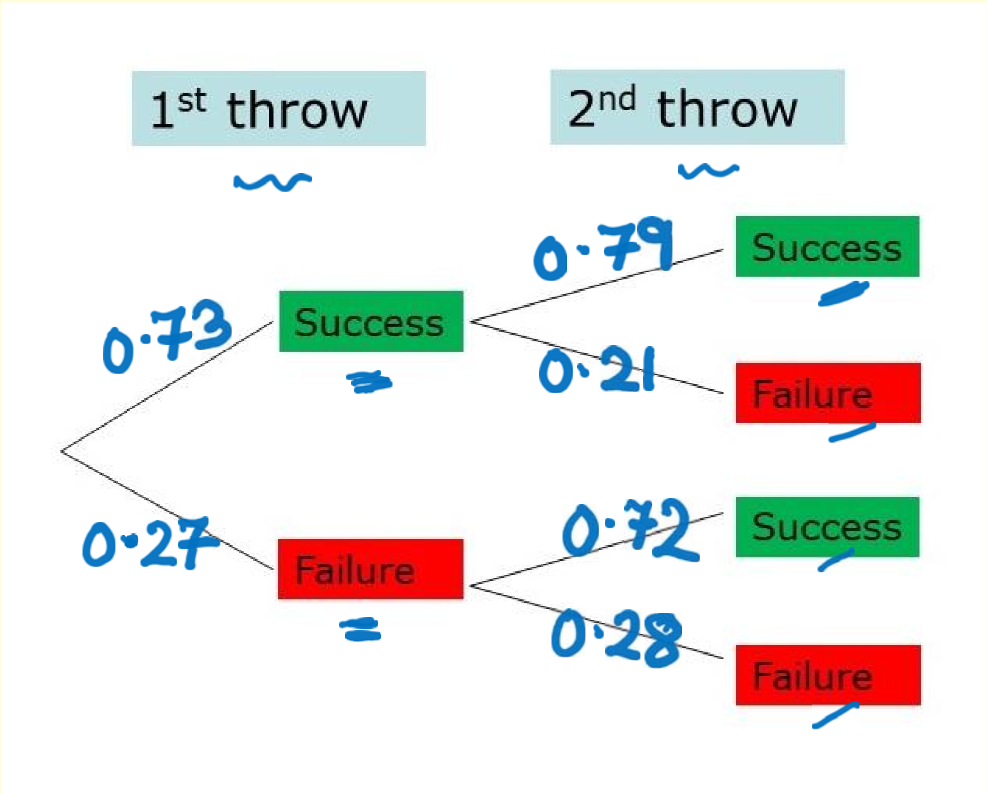
  if 1^{st} **unsuccessful**: success-rate on 2^{nd} : 72%

$$P(2^{nd} S \mid 1^{st} S) = 0.79$$

$$P(2^{nd} S \mid 1^{st} F) = 0.72$$

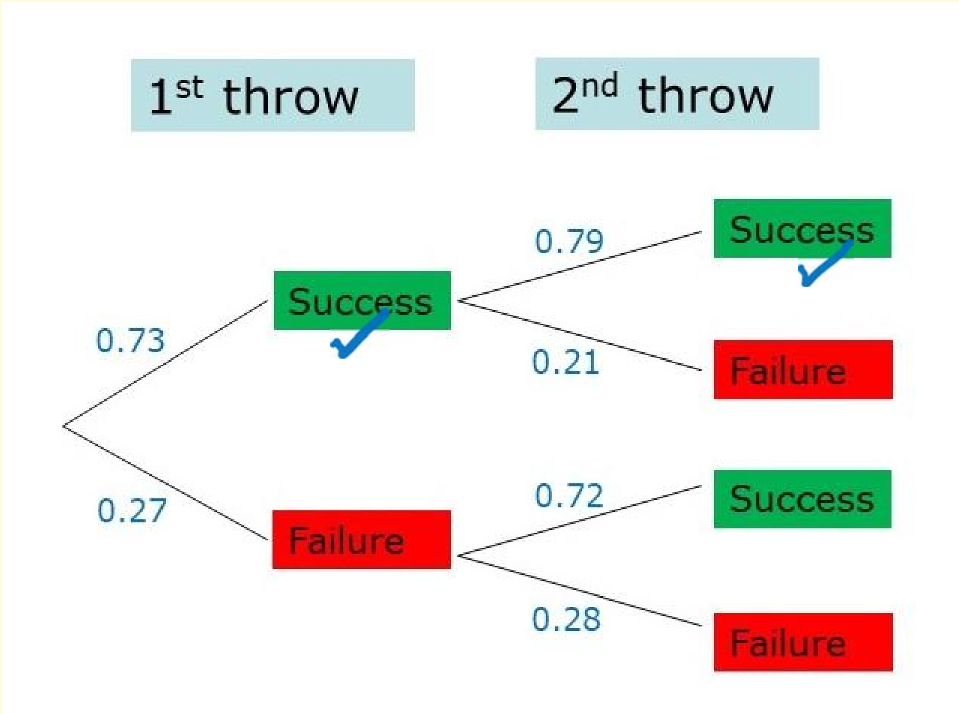
Example 1

- Representation using a prob. tree



Example 1

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



$Pr(\underline{1^{st}S} \cap \underline{2^{nd}S})$

$= ?$

General multiplication rule

- **[General multiplication Rule]**:

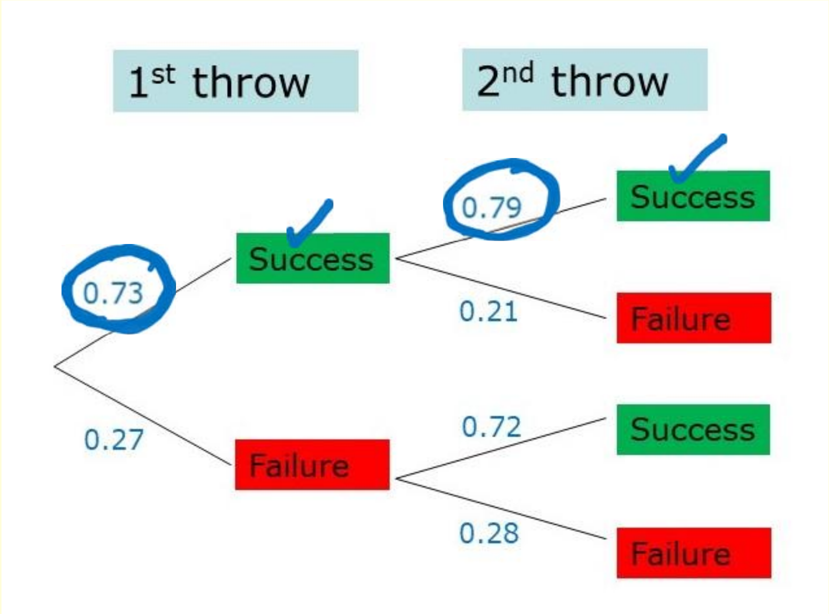
$$P(A \text{ and } B) = P(A \cap B) = P(A)P(B|A)$$

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?



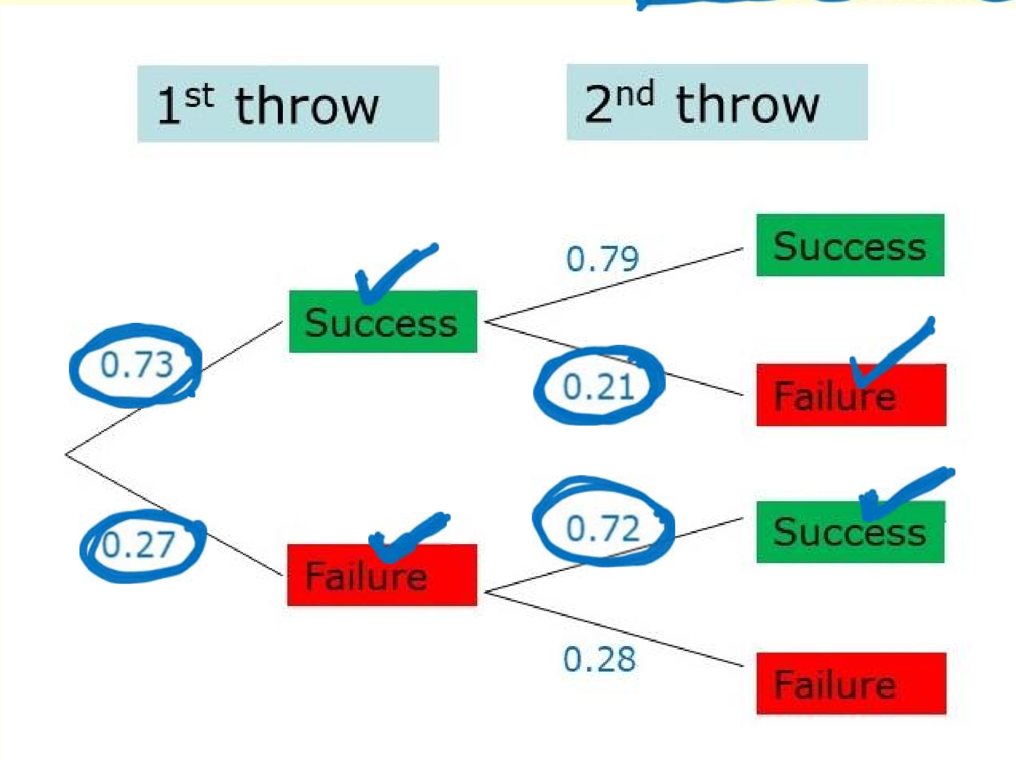
$= P(1^{st}\, S) \times P(2^{nd}\, S \,|\, 1^{st}\, S)$

$= 0.73 \times 0.79$

$= 0.58$

Example 1 contd.

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



$0.73 \times 0.21 = 0.153$

$+ 0.27 \times 0.72 = 0.194$

$= \boxed{0.347}$

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

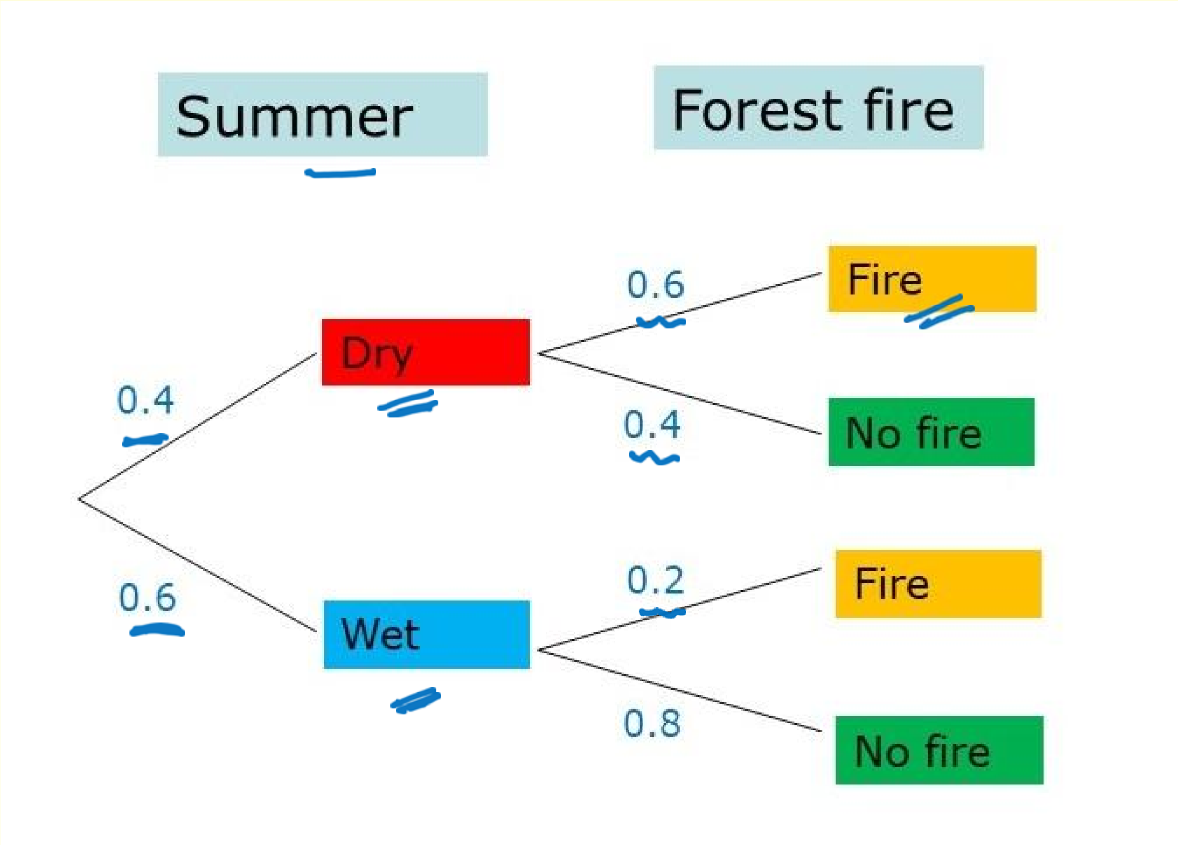
Q: What is $P(\text{fire} \mid \text{dry summer})$?

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

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

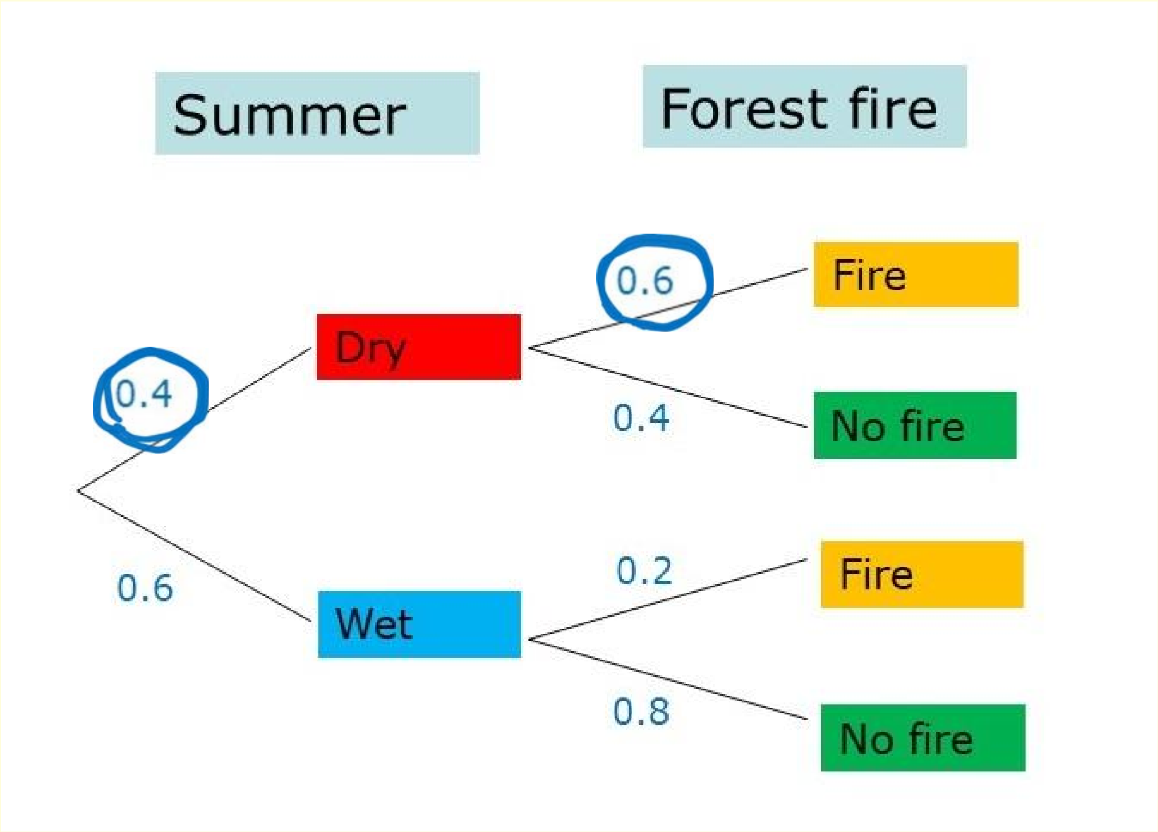
Example 2

- Representation using a prob. tree



Example 2

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



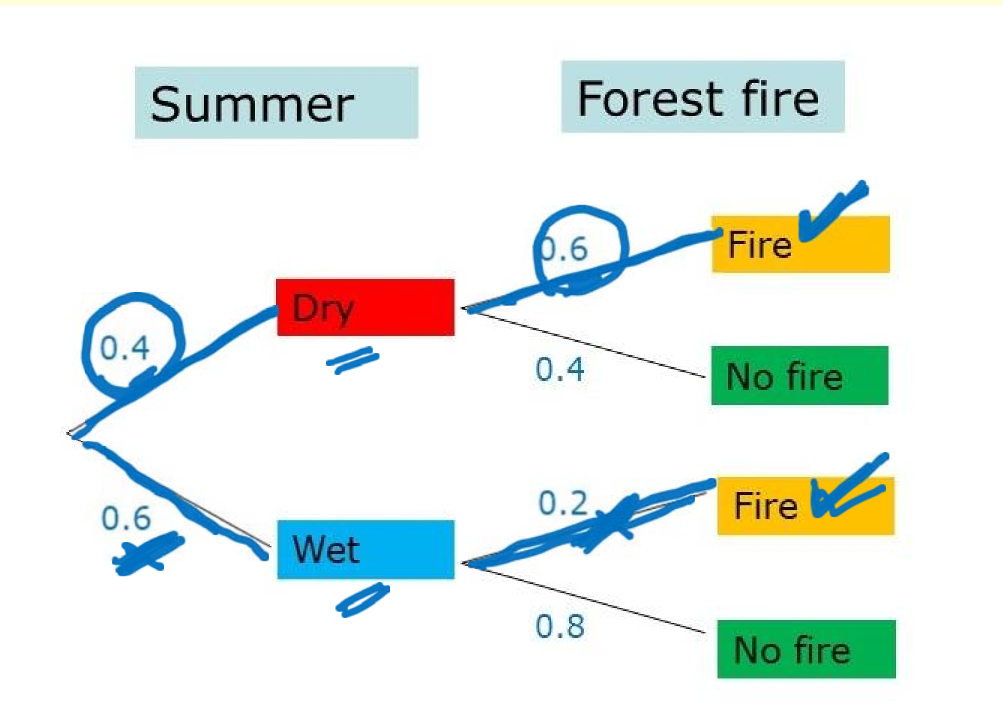
$= Pr(dry)\, Pr(fire \mid dry)$

$= 0.4 \times 0.6$

$= 0.24$

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 $P(\text{fire})$?



$= 0.4 \times 0.6$

$+ 0.6 \times 0.2$

$= 0.24 + 0.12 = 0.36$

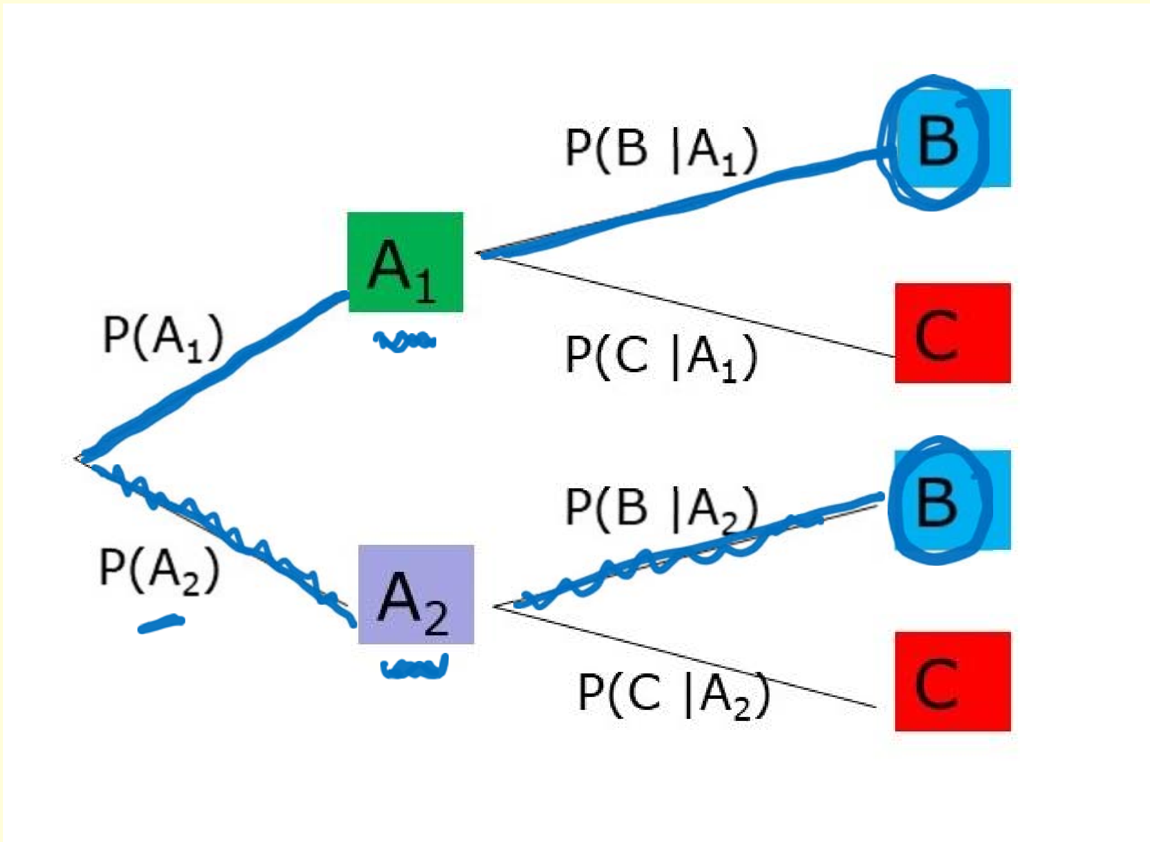
Example 2

- 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: $P(fire) =$

$$P(dry)P(fire|dry) + P(wet)P(fire|wet)$$

Law of Total Probability

- Event B can come about in many different ways.



$P(B) =$

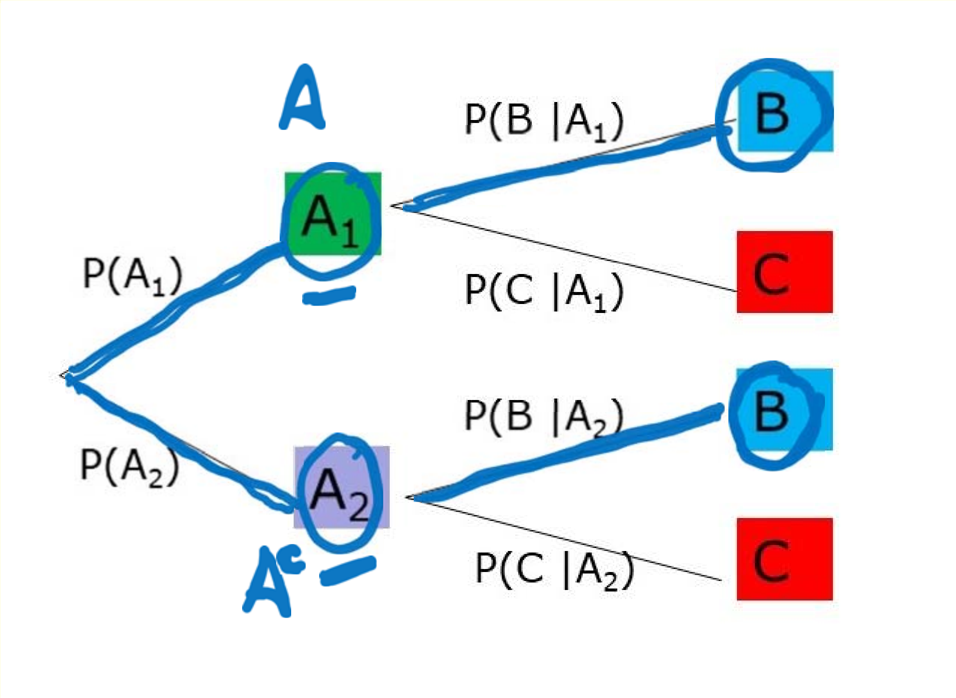
$P(A_1)P(B|A_1)$

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

$=$

Law of Total Probability

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



Law of Total Probability

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

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

  $= P(A)P(B|A) + P(A^c)P(B|A^c)$ ← imp. for Bayes' rule

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

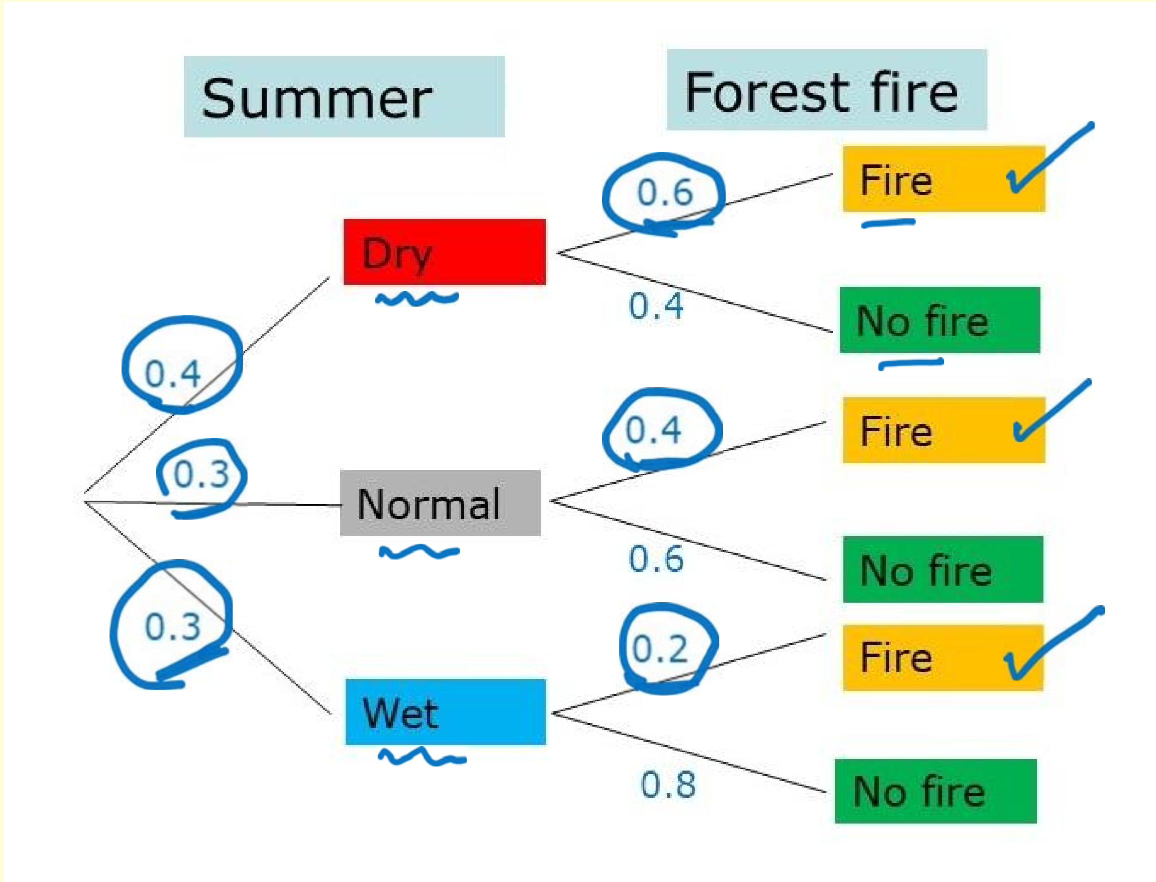
- **Example 2**: A summer may be *dry* (prob. 0.4), *normal* (prob. 0.3) or *wet* (prob. 0.3)

	Chance of forest fires
Dry summer	0.6 ✓
Normal summer	0.4 ✓
Wet summer	0.2 ✓

Example 2 extended

- Prob. tree

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



$$= P(\text{dry})P(\text{fire}|\text{dry}) + P(\text{normal})P(\text{fire}|\text{normal}) + P(\text{wet})P(\text{fire}|\text{wet})$$

$$= 0.4 \times 0.6 + 0.3 \times 0.4 + 0.3 \times 0.2$$

$$= 0.24 + 0.12 + 0.6 = 0.42$$

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?

$$\frac{3}{10} = 0.3$$

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*: $P(2^{nd}\ A|1^{st}\ A) = \frac{2}{9}$

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

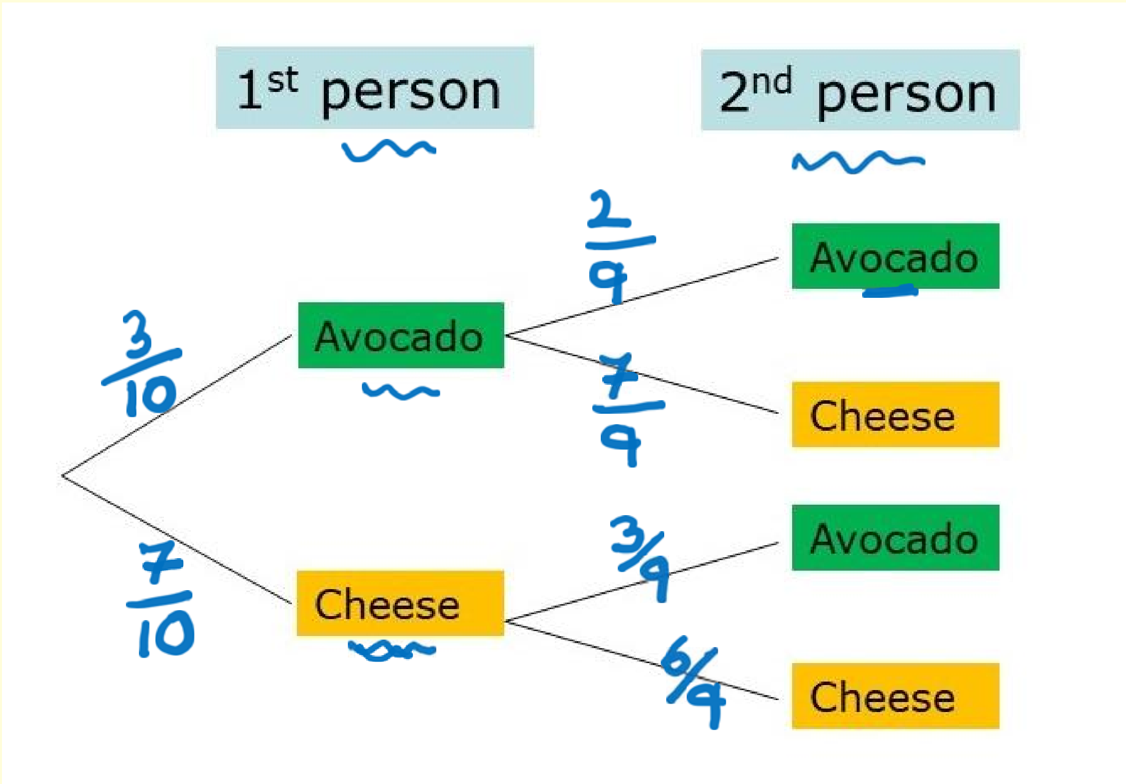
  $\frac{3}{9}$

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

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

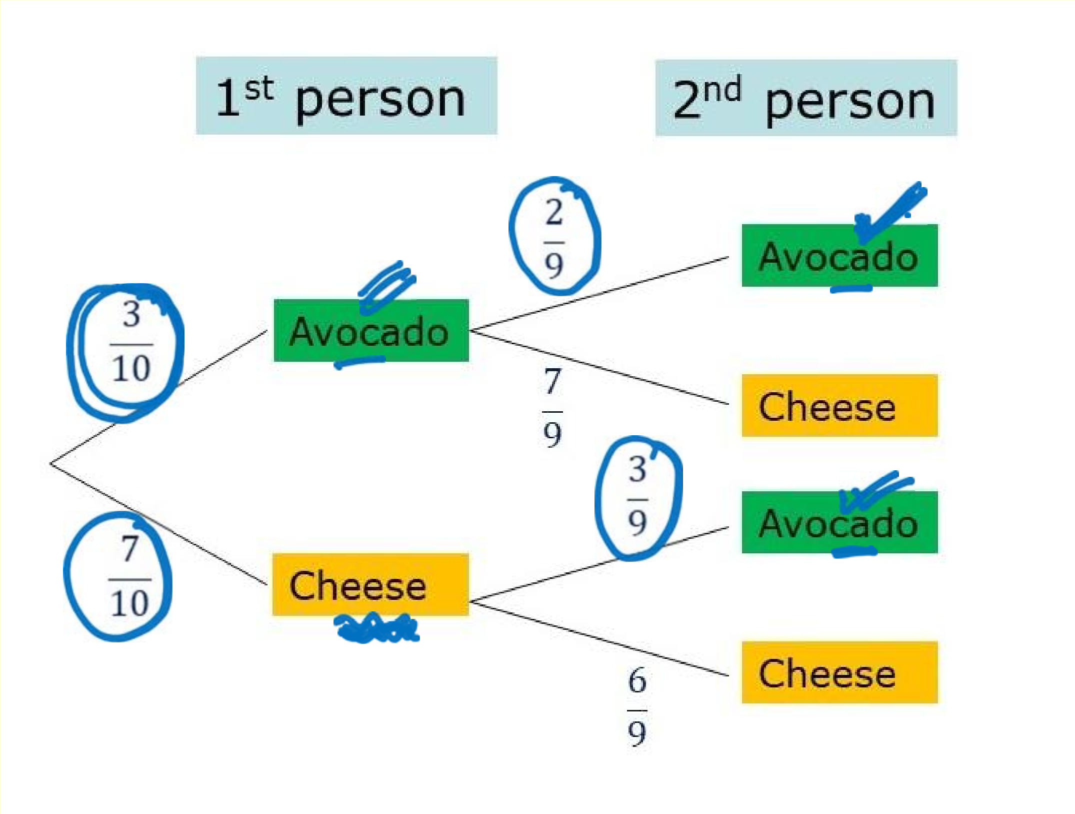
Example 3: Platter with 3 avocado and 7 cheese

- **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} :



Law of total prob. =

$$\frac{3}{10}\times\frac{2}{9}+\frac{7}{10}\times\frac{3}{9}$$

$$=\frac{6}{90}+\frac{21}{90}=\frac{\cancel{27}^{\,3}}{\cancel{90}_{\,10}}$$

$$=0.3$$

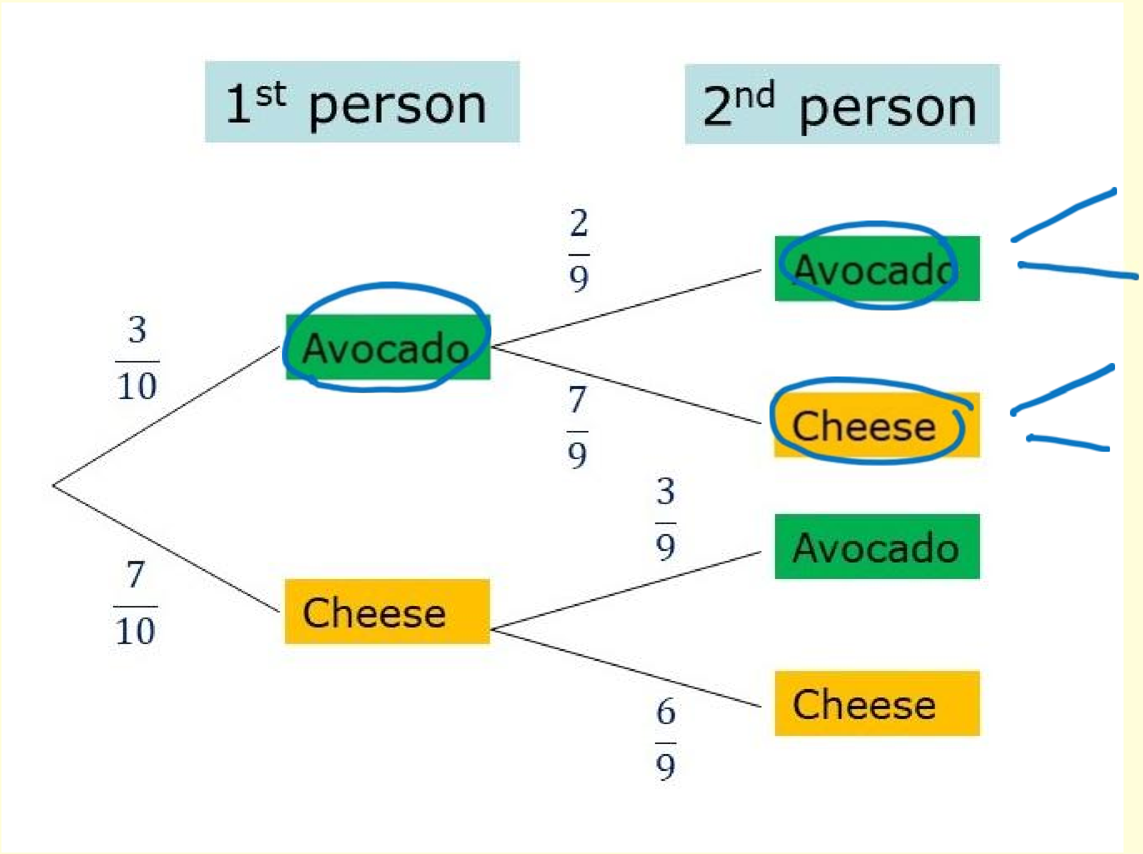
Example 3: Platter with 3 avocado and 7 cheese

- **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
- $P(AA) = \frac{3}{10}\cdot\frac{2}{9}$, $P(AC) = \frac{3}{10}\cdot\frac{7}{9}$, $P(CA) = \frac{7}{10}\cdot\frac{3}{9}$, $P(CC) = \frac{7}{10}\cdot\frac{6}{9}$,
  $P(3^{rd}A|AA) = \frac{1}{8}$, $P(3^{rd}A|AC) = \frac{2}{8}$,
  $P(3^{rd}A|CA) = \frac{2}{8}$, $P(3^{rd}A|CC) = \frac{3}{8}$

$$P(AA)P(3^{rd}A|AA) + \ldots\ldots$$
$$= \frac{6}{90}\cdot\frac{1}{8} + \frac{21}{90}\cdot\frac{2}{8} + \frac{21}{90}\cdot\frac{2}{8} + \frac{42}{90}\cdot\frac{3}{8}$$
$$= \frac{216}{720} = 3/10 = 0.3$$

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:
$P(B|A) = P(B) \quad \Rightarrow P(A \cap B) = P(A)P(B)$

$$\left.\begin{array}{c} P(B|A)\,P(A) \\ \parallel \\ P(B)\,P(A) \end{array}\right\} \text{indep. events}$$

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

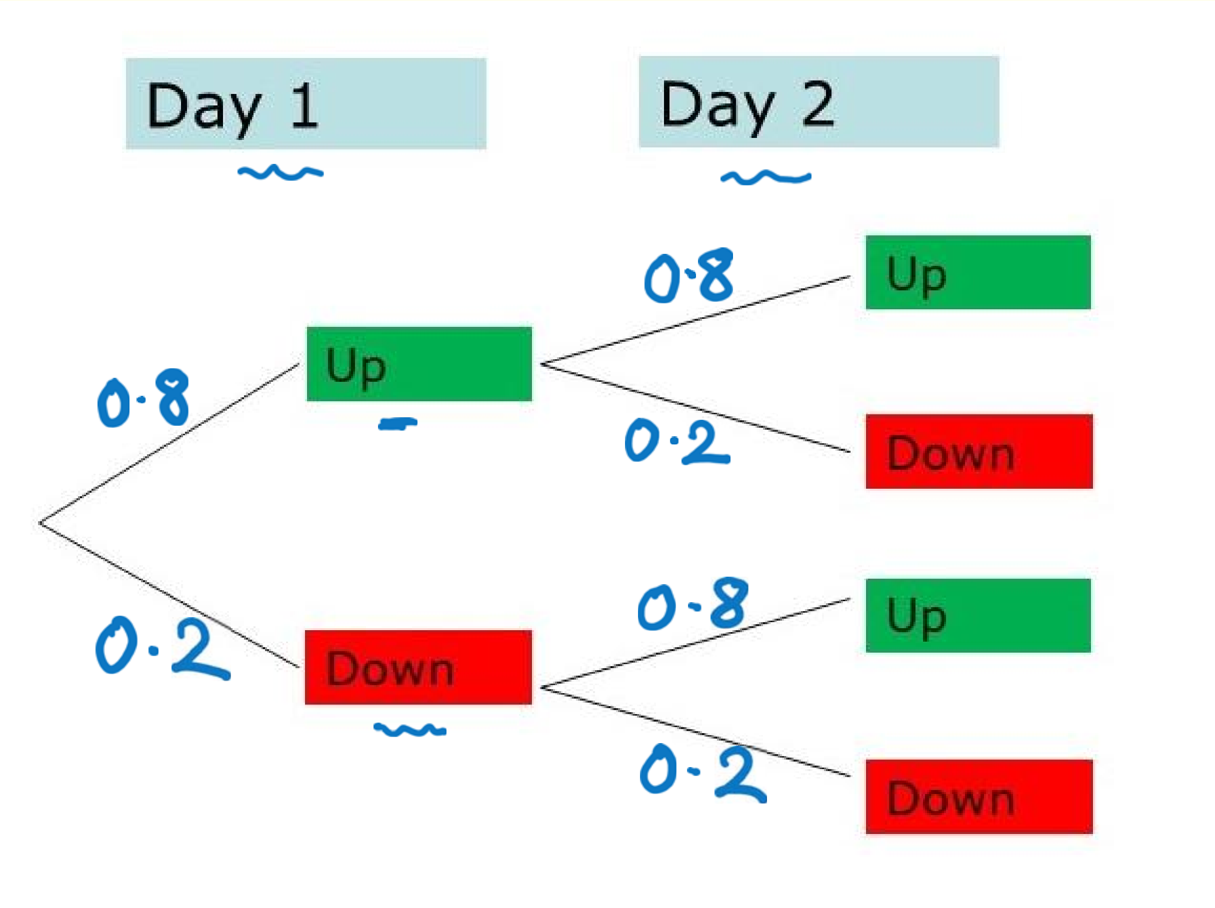
- The outcome DOES NOT depend on what happened the previous day:

	Up tomorrow	Down tomorrow
Up today	0.8 ↑	0.2 ↓
Down today	0.8	0.2

P(up tomorrow | up today)
= P(up tomorrow | Down today)

Prob. tree with independence

- Prob. tree

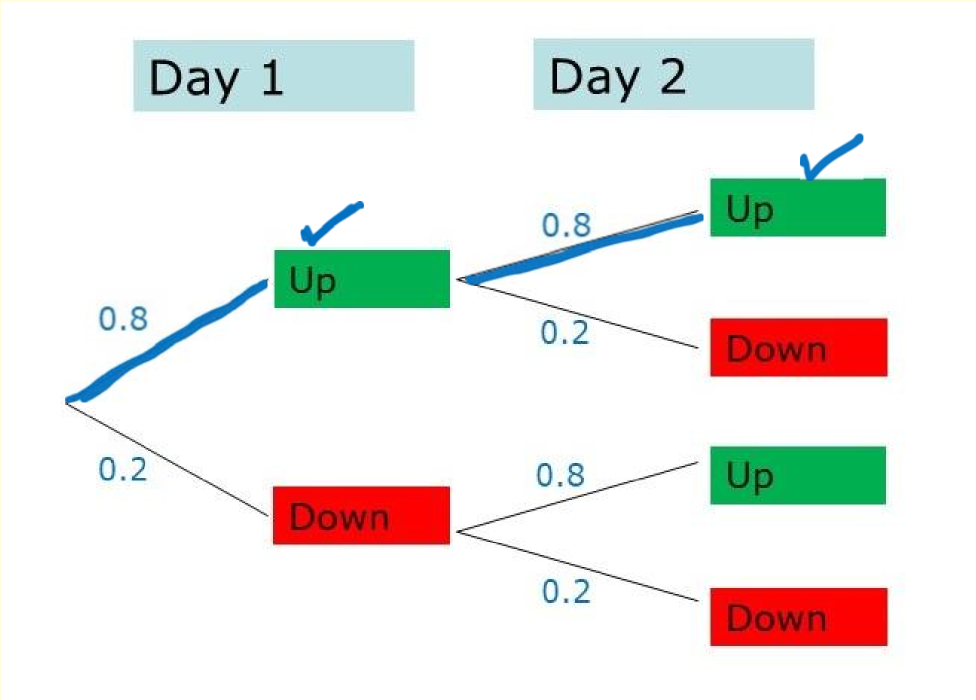


Prob. tree with independence

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

  (a) it is up the next 2 days?

$0.8 \times 0.8 = 0.64$



Example 4: Stock value

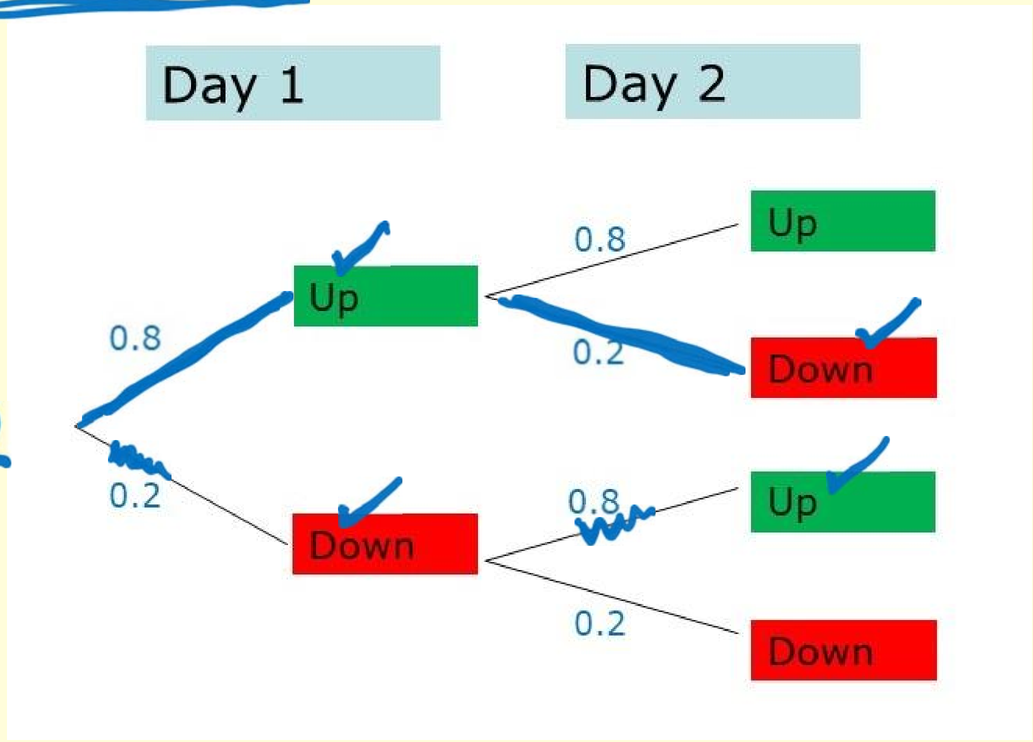
- If the stock is up today, what is the prob. that

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

0.8×0.2

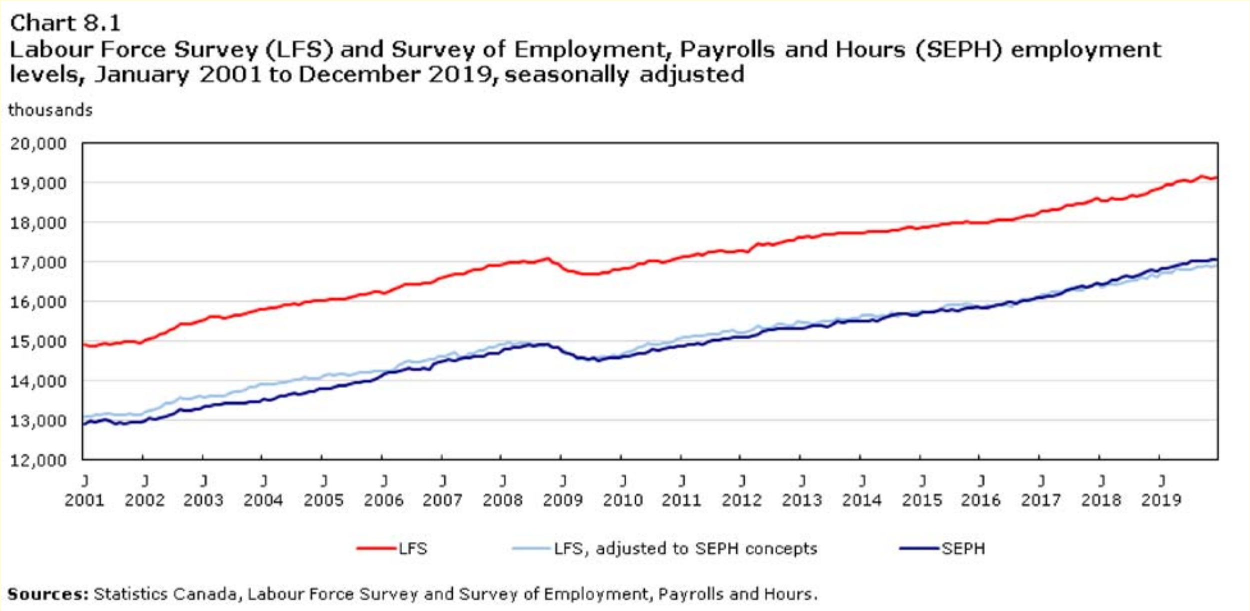
$+ 0.2 \times 0.8$

$= 0.16 + 0.16 = 0.32$



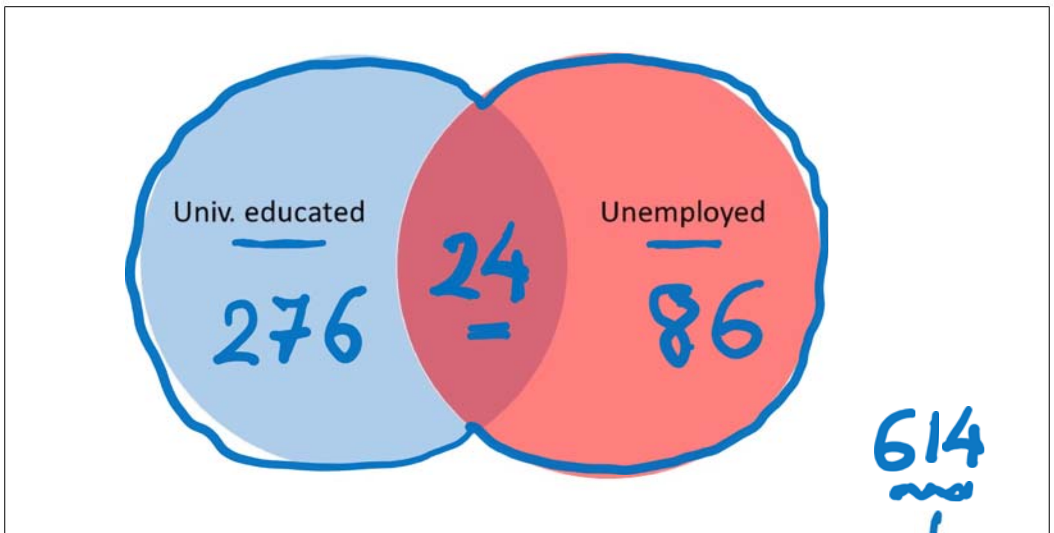
Finding Conditional Probability

- **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



$$\begin{array}{r} 300 \\ -24 \\ \hline 276 \end{array}$$

$276 + 24 + 86 = 386$

614

employed & no univ. educn

$$\begin{array}{r} 110 \\ -24 \\ \hline 86 \end{array}$$

Finding Conditional Probability

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

  Unemployed and Univ. educated: 24
- Venn diagram:

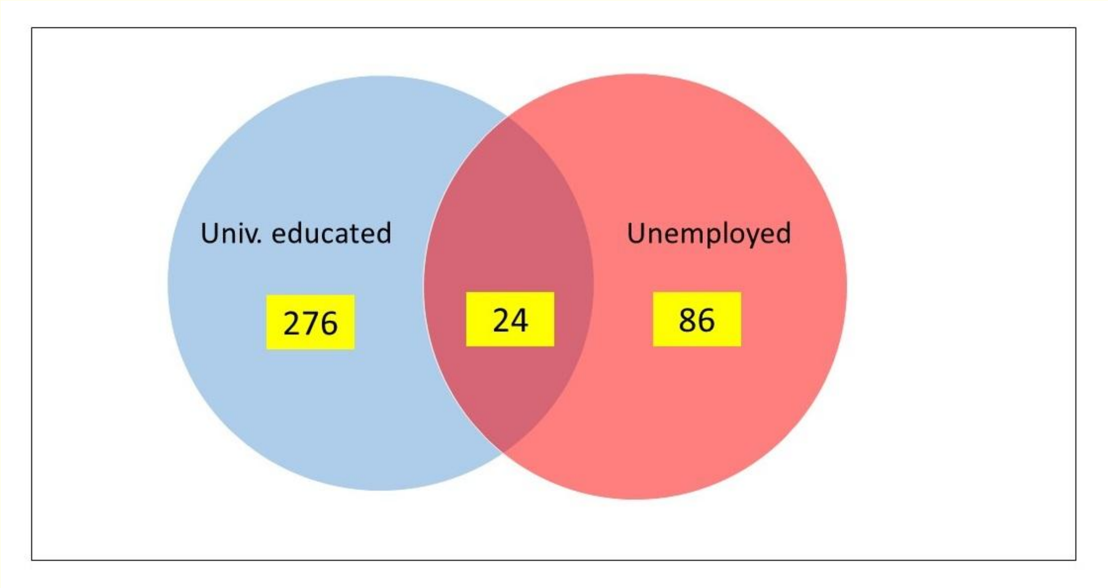
  Q: Unemployment rate? $\frac{110}{1000} = 0{\cdot}11$

  Q: Unemp. rate among univ. educated? P(unemp | univ. educated)

Example 5: Labour Force Survey

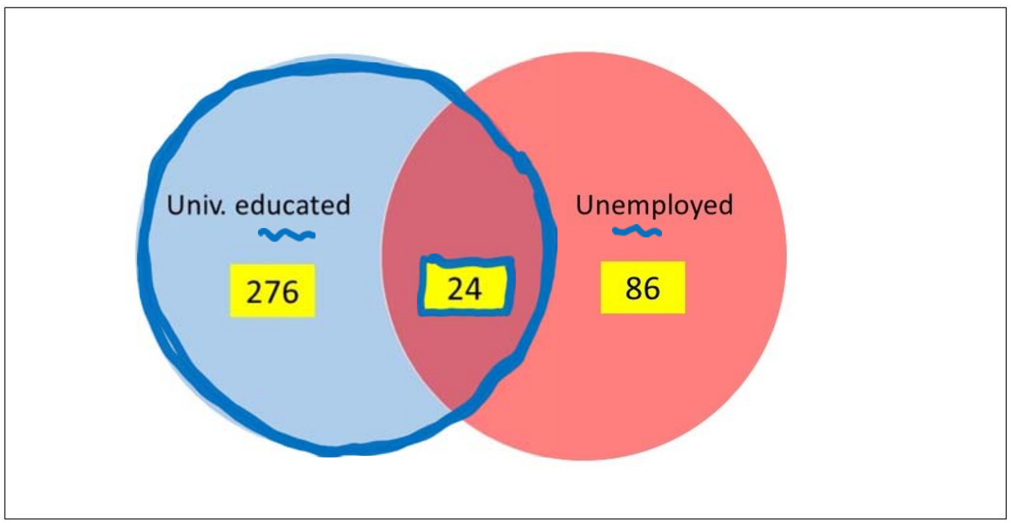
1000

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) = \frac{110}{1000} = 0.11$



Example 5: Labour Force Survey

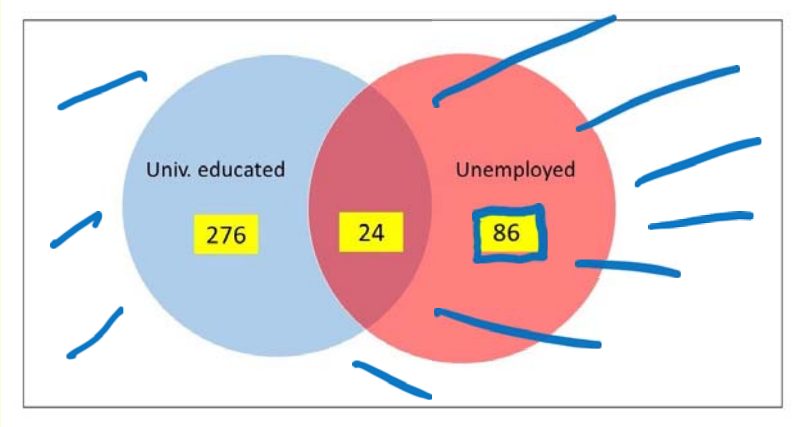
- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) =$
- Unemp. rate among univ.: $P(unemp | univ) = \dfrac{24}{300} = 0.08$



Example 5: Labour Force Survey

- *Unemp*: 110, *Univ.*: 300, *Unemp & Univ.*: 24
- Overall unemp. rate: $P(unemp) =$
- Unemp. rate among univ.: $P(unemp|univ) =$ 0.08
- Unemp. rate among non-univ.: $P(unemp|no\ univ) =$ $\frac{86}{700}$

$= 0.123$

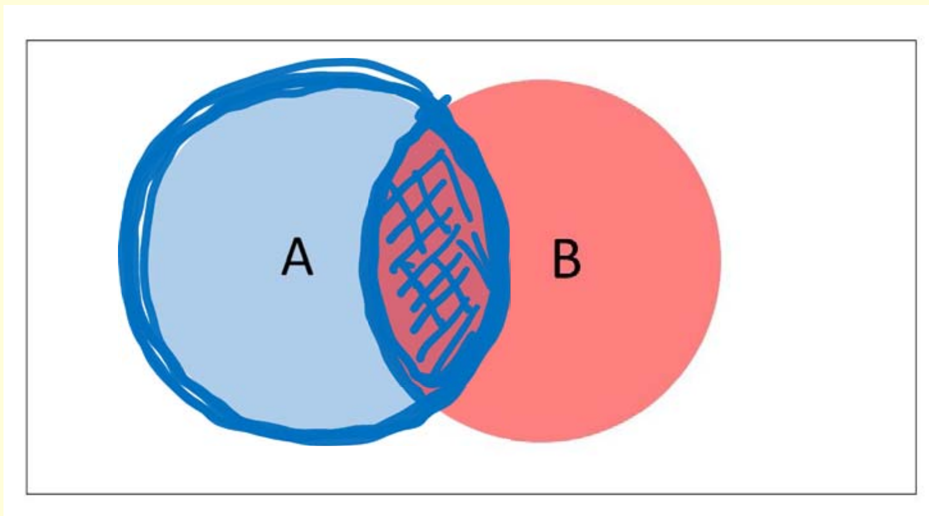


Finding Conditional Probability

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

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

$P(B|A)$



Finding Conditional Probability

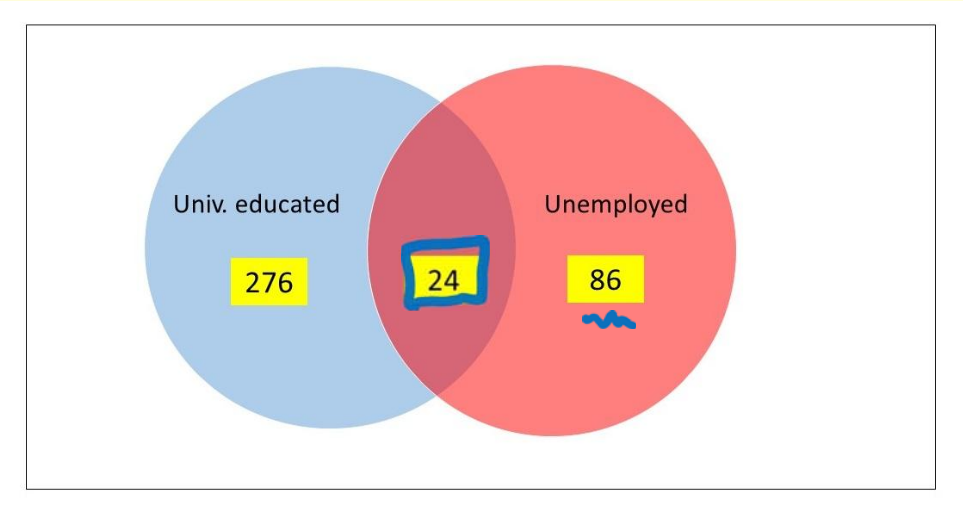
- $P(A \cap B) = P(A)P(B|A) \Rightarrow$

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

- Unemp. rate among univ.:

$$P(unemp \mid univ) = \frac{P(unemp \;\&\; univ)}{P(univ)} = \frac{24/1000}{300/1000}$$

$$= \frac{24}{300} = 0.08$$



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.:

$$P(unemp \mid no\ univ) = \frac{P(unemp \ \& \ no\ univ)}{P(no\ univ)} = \frac{86/1000}{700/1000}$$

$$= \frac{86}{700} = 0.123$$

Example 5: Marginal probability

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

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

total prob.

- $\boxed{\textbf{Marginal prob.}}$ $P(unemp) =$

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

$$\frac{300}{1000} = 0.3 \times 0.08 + \frac{700}{1000} = 0.7 \times 0.123$$

$$= 0.110$$

Conditional Probability: another example

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

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

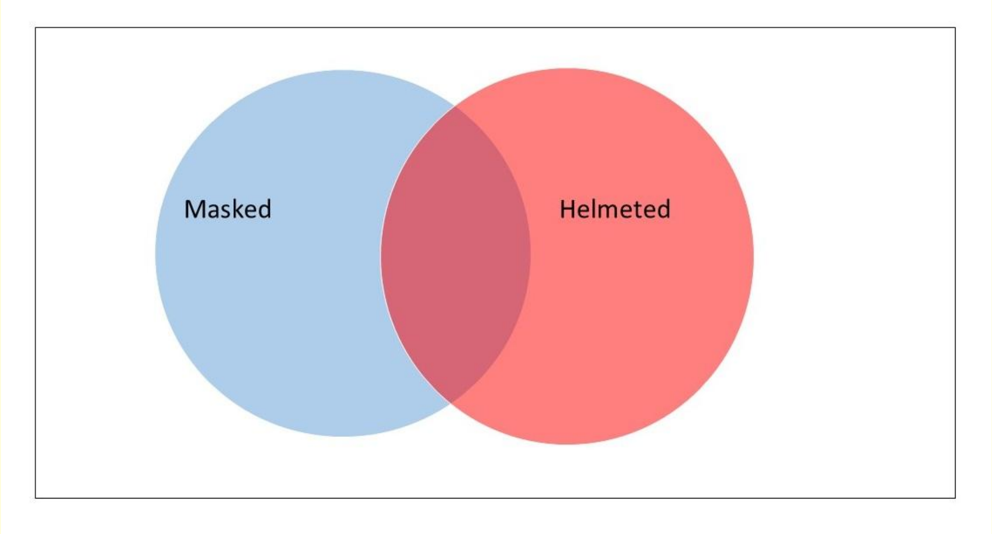
Conditional Probability: another example

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

  Masked and Helmeted: 29
- 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?

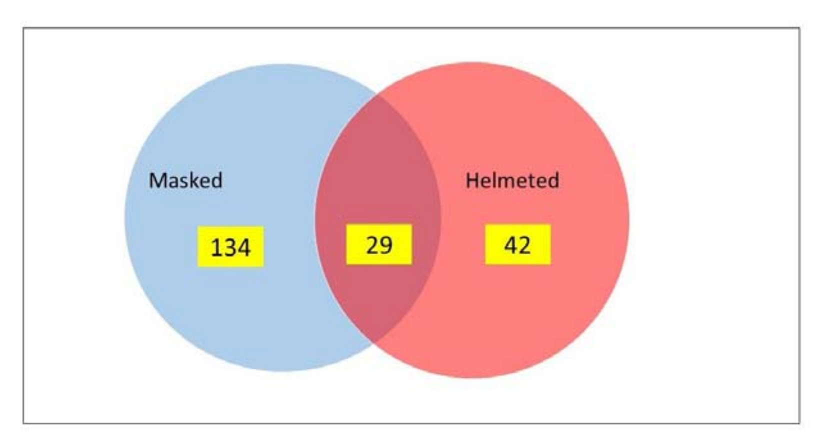
- Masking. rate among helmeted: $P(\underset{\sim}{mask}|helmet) =$

$=$

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)** $\frac{29}{71}$

$$\frac{29}{71}$$

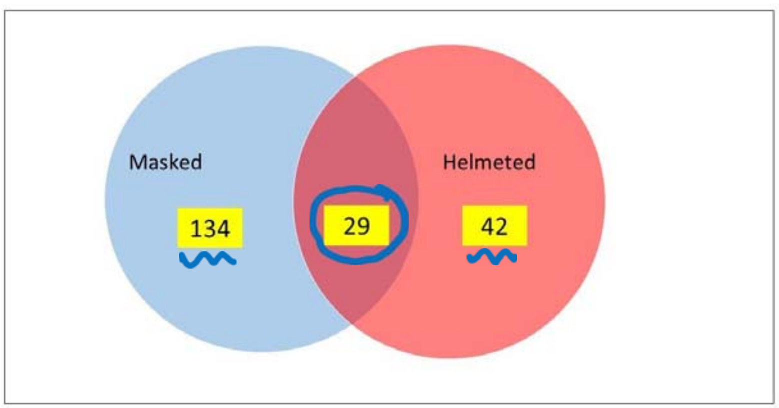
Example 6: Stanford students

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

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

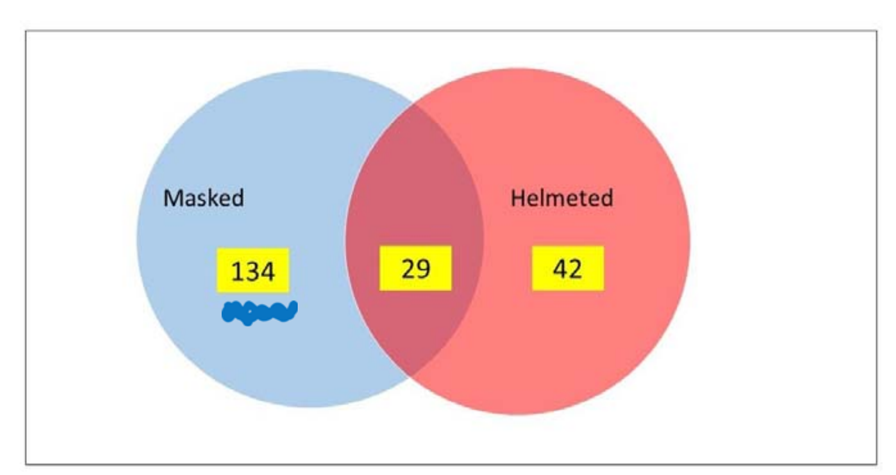
- $P(mask|helmet) = \frac{P(mask \ \& \ helmet)}{P(helmet)} =$

$$\frac{29/400}{71/400} = \frac{29}{71}$$



Example 6: Stanford students

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



$$\begin{array}{r} 400 \\ -\ 71 \\ \hline 329 \end{array}$$

$$\frac{134}{329} = 0.4073$$

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	134	163
No mask	42	195	237
Total	71	329	400

Prob. from two way frequency tables: Example 6

-

	Helmet	No helmet	*Total*
Mask	$\mathrm{P}(H \cap M) = \frac{29}{400} = 0.0725$	$\mathrm{P}(NH \cap M) = \frac{134}{400} = 0.335$	$\mathrm{P}(M) = \frac{163}{400}$
No mask	$\mathrm{P}(H \cap NM) =$	$\mathrm{P}(NH \cap NM) =$	$\mathrm{P}(NM) =$
Total	$\boxed{\mathrm{P}(H) =} \frac{71}{400} = 0.1775$	$\mathrm{P}(NH) =$	1

→ Marginal prob.

marginal prob. of M

- $P(M|H) = \frac{P(M \cap H)}{P(H)} = \frac{0.0725}{0.1775} = 0.408$
- $P(M|NH) = \frac{P(M \cap NH)}{P(NH)} =$

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	$\mathrm{P}(M) = 0.4075$
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 $P(H|M)$? $= \dfrac{P(H \cap M)}{P(M)} = \dfrac{0.0725}{0.4075} = \dfrac{725}{4075}$

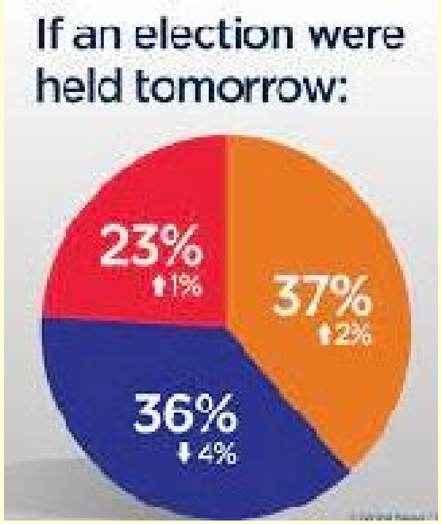
(a) $\frac{335}{4075}$ **(b)** $\frac{105}{1775}$

(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*	$\boxed{180}$	110	80	370
Party *B*	120	120	50	290
Party *C*	100	170	70	340
Total	400	400	200	1000

$P(A \cap Y) = \frac{180}{1000} = 0.18$

- Support among under 30s:

$P(A|young) = \frac{180}{400} = 0.45$, $P(B|yng) = \frac{120}{400} = 0.3$, $P(C|yng) = \frac{100}{400} = 0.25$,

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:

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

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		400	200	1000

- Support among under 30s:

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

- Support among 30-64s:

  $P(A|middle) = \frac{110}{400}$, $P(B|mid) = \frac{120}{400}$, $P(C|mid) = \frac{170}{400}$,

- Support among 65+s:

  $P(A|older) = \frac{80}{200}$, $P(B|older) =$, $P(C|older) =$,

  $= 0.4$

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

$$P(A) = \frac{1}{3} \times 0.45 + \frac{1}{3} \times 0.275 + \frac{1}{3} \times 0.4 = 0.375$$

$$P(B) = \frac{1}{3} \times 0.3 + \frac{1}{3} \times 0.3 + \frac{1}{3} \times 0.25 = 0.283$$

$$P(C) = \frac{1}{3} \times 0.25 + \frac{1}{3} \times 0.425 + \frac{1}{3} \times 0.35 = 0.342$$

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?

older $\rightarrow$ 25%

$$P(A) = \frac{1}{4} \times 0.45 + \frac{1}{2} \times 0.275 + \frac{1}{4} \times 0.4 = 0.35$$

$$P(B) = \frac{1}{4} \times 0.3 + \frac{1}{2} \times 0.3 + \frac{1}{4} \times 0.25 = 0.2875$$

$$P(C) = \frac{1}{4} \times 0.25 + \frac{1}{2} \times 0.425 + \frac{1}{4} \times 0.35 = 0.3625$$

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

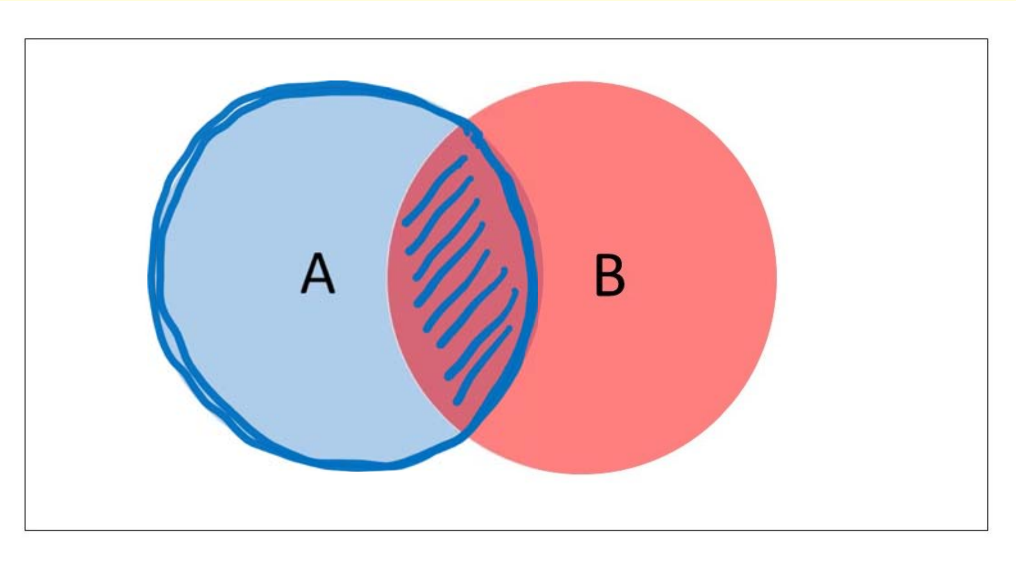
- You know $P(A|B)$. **Q: How to find** $P(B|A)$?
- *Example:* Driving while drunk increases chance of accidents.

  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

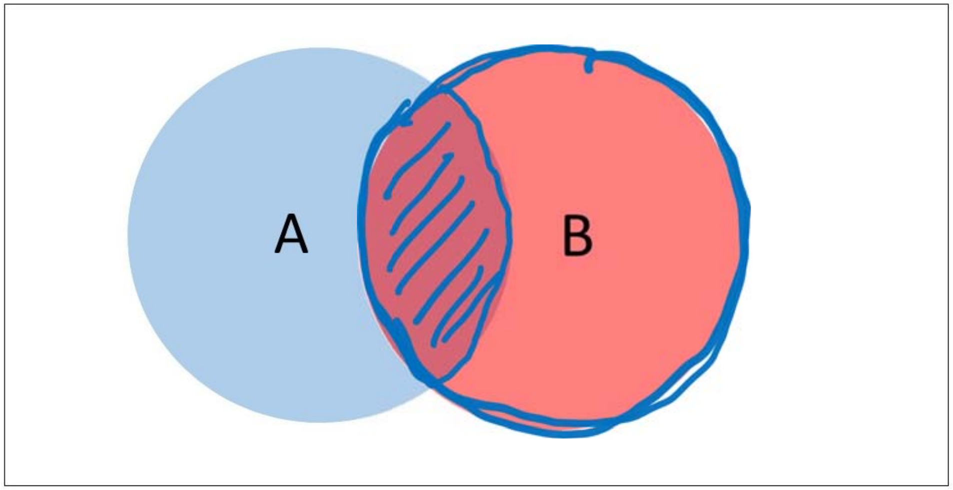
- Basic idea: $P(B|A) = \frac{P(A \cap B)}{P(A)}$ e.g. $B \equiv$booze, $A \equiv$accident



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:* $\underline{P(A|B)} = \frac{P(A \cap B)}{\underline{P(B)}} \quad \Rightarrow \underline{P(A \cap B)} = \underline{P(A|B)}\,\underline{P(B)}$



?

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

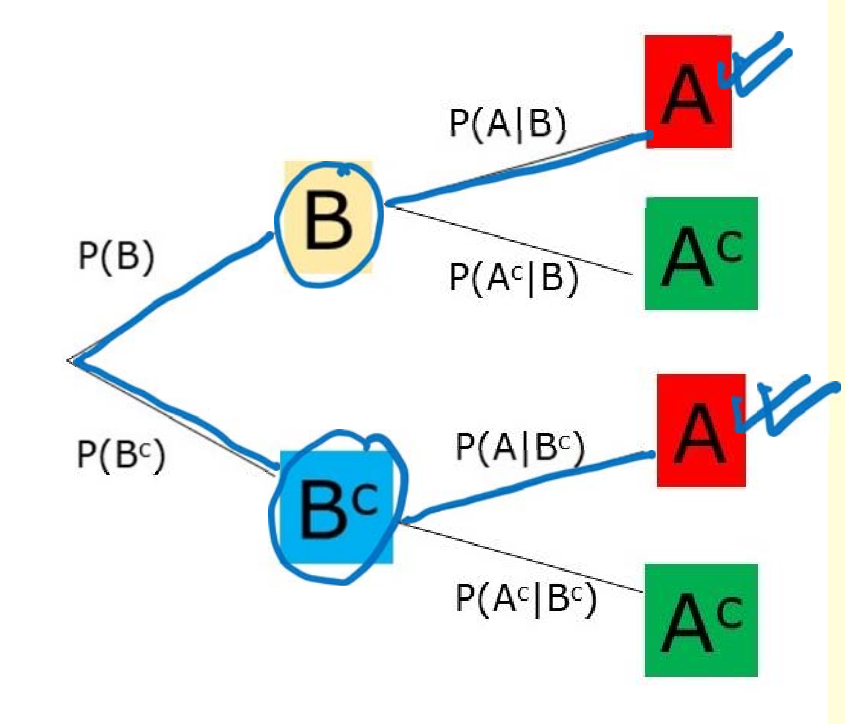
Bayes' rule

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

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.*:* $P(A) = \underline{P(A|B)P(B) + P(A|B^c)P(B^c)}$
- Finally **Bayes' rule**:

?

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

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**:

Binge

$$\underbrace{P(B|A)}_{0.56} = \frac{\overset{0.5}{P(A|B)}\overset{0.2}{P(B)}}{\underset{0.5}{P(A|B)}\underset{0.2}{P(B)} + \underset{0.1}{P(A|B^c)}\underset{0.8}{P(B^c)}}$$

Acc.

$$= \frac{0.5 \times 0.2}{0.5 \times 0.2 + 0.1 \times 0.8} = \frac{0.1}{0.1 + 0.08} = \frac{0.1}{0.18} = 0.56$$

Bayes' rule: Terminology

- **Bayes' rule**:

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

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

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:

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

Bayes' rule: Example 9

- **Example 9**: Prior prob. $P(hot) = 0.4$ $P(cold) = 0.6$

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

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

$$= \frac{\overset{0.8 \times 0.4}{P(''Buy''|hot)P(hot)}}{\underset{0.8 \times 0.4 + 0.4 \times 0.6}{P(''Buy''|hot)P(hot) + P(''Buy''|cold)P(cold)}}$$

$$= \frac{0.32}{0.32 + 0.24} = \frac{32}{56} = 0.57$$

posterior

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

$$\frac{\overset{1}{P(\text{"}Buy\text{"}|hot)}\,\overset{0.4}{P(hot)}}{\underset{1}{P(\text{"}Buy\text{"}|hot)}\,\underset{0.4}{P(hot)} + \underset{0}{P(\text{"}Buy\text{"}|cold)}\,\underset{0.6}{P(cold)}} = \frac{0.4}{0.4} = 1$$

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) $= \underset{\text{prior}}{0.3}$.

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

  $P(-ive \mid \text{infected}) = 0.3$

- False positives and false negatives

  $\underbrace{\text{False positives}}_{0.1}$ and $\underbrace{\text{false negatives}}_{0.3}$

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 $P(\text{infected} \mid +ive) =$ $\dfrac{\bigcirc}{\bigcirc + \bigcirc}$

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 $P(B|A)$ indep : $P(B|A) = P(B)$
- 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.