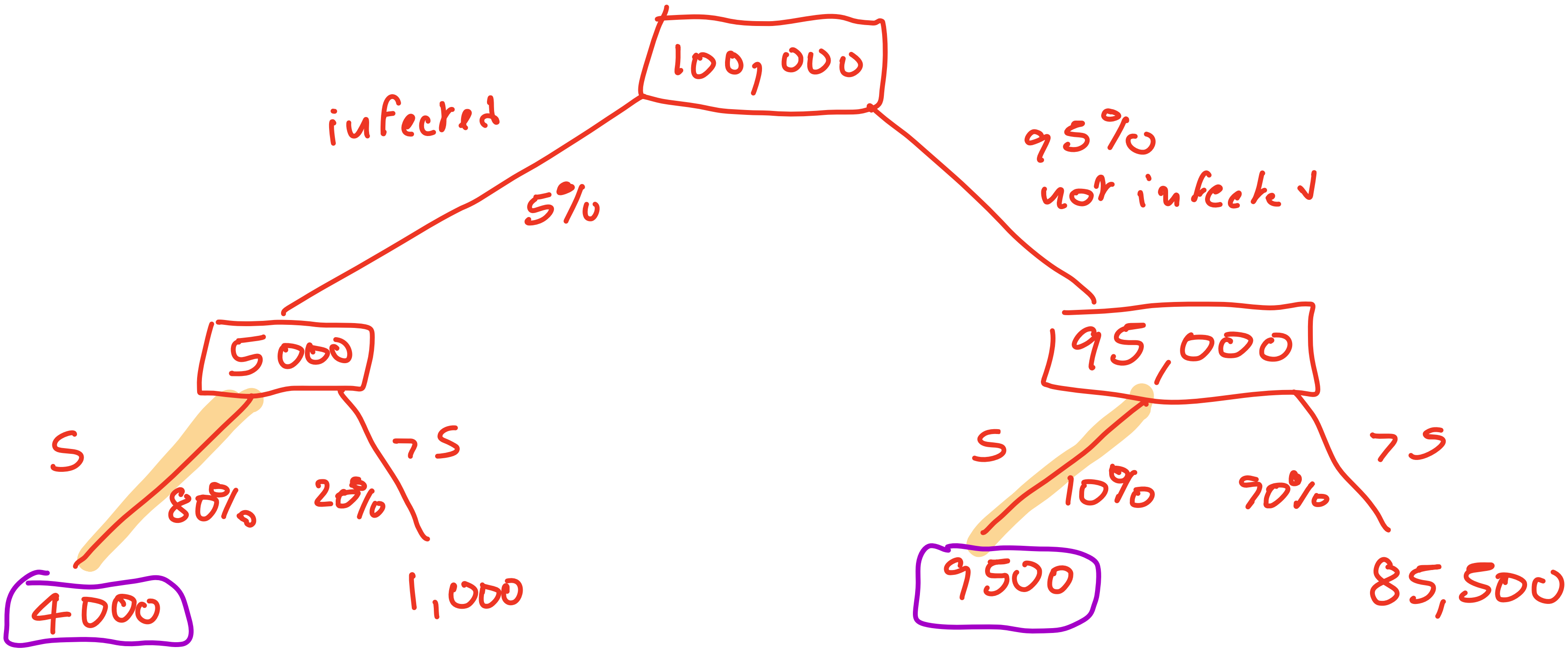
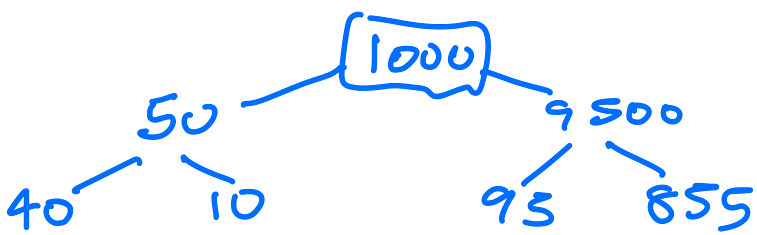
- Conditional reasoning: the frequency tree
- Conditional probability
- Updating probabilities
- Independence

- 5% of the population are infected
- 80% of those infected have the symptoms
- 10% of those **not** infected have the symptoms



$$P(I \mid S) = \frac{4{,}000}{4{,}000 + 9{,}500} = 29.63\%$$



$$P(I \mid S) = \frac{40}{40+95} = 29.63\%$$

A test is now available. The probability of testing positive is independent of whether or not you have symptoms:

- If you are infected, the probability of testing positive is 80% (whether or not you have the symptoms)

true positive

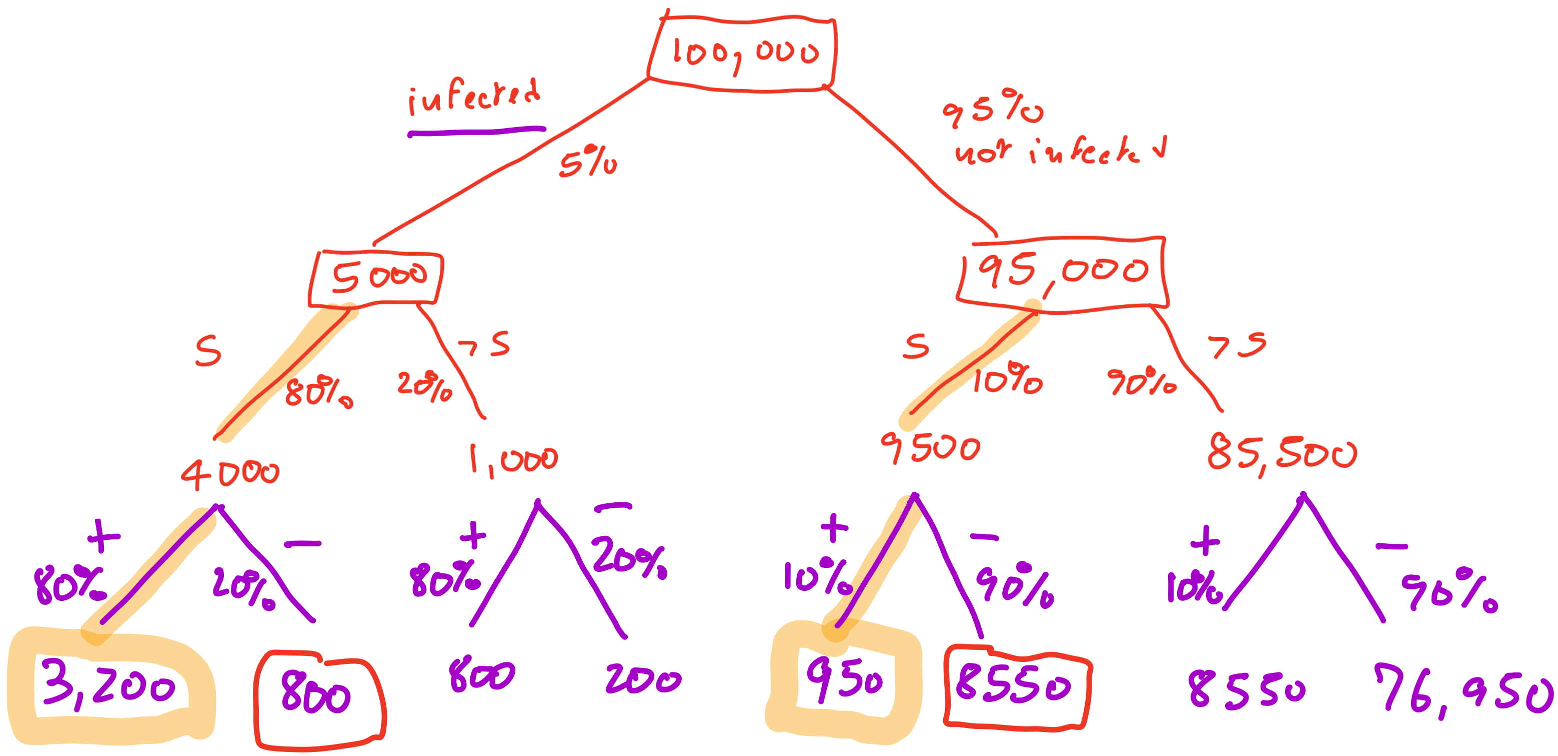
- If you are **not** infected, the probability of testing positive is 10% (whether or not you have the symptoms)

false positive

Since you woke up with symptoms, you decided to get tested and the result was positive. How likely is it that you are infected?

$$P(I \mid S \text{ and } +) = ?$$

- If you are infected, the probability of testing positive is 80% (whether or not you have the symptoms)
- If you are **not** infected, the probability of testing positive is 10% (whether or not you have the symptoms)



$$P(I \mid S \text{ and } +) = \frac{3,200}{3,200 + 950} = 77.11\%$$

$$P(I \mid S \text{ and } -) = \frac{800}{800 + 8550} =$$

One more example

Base rate of a disease: percentage of the population that has the disease

Sensitivity of a test: percentage of those who have the disease that tests positive true positive

Specificity of a test: percentage of those who **do not** have the disease that tests negative true negative

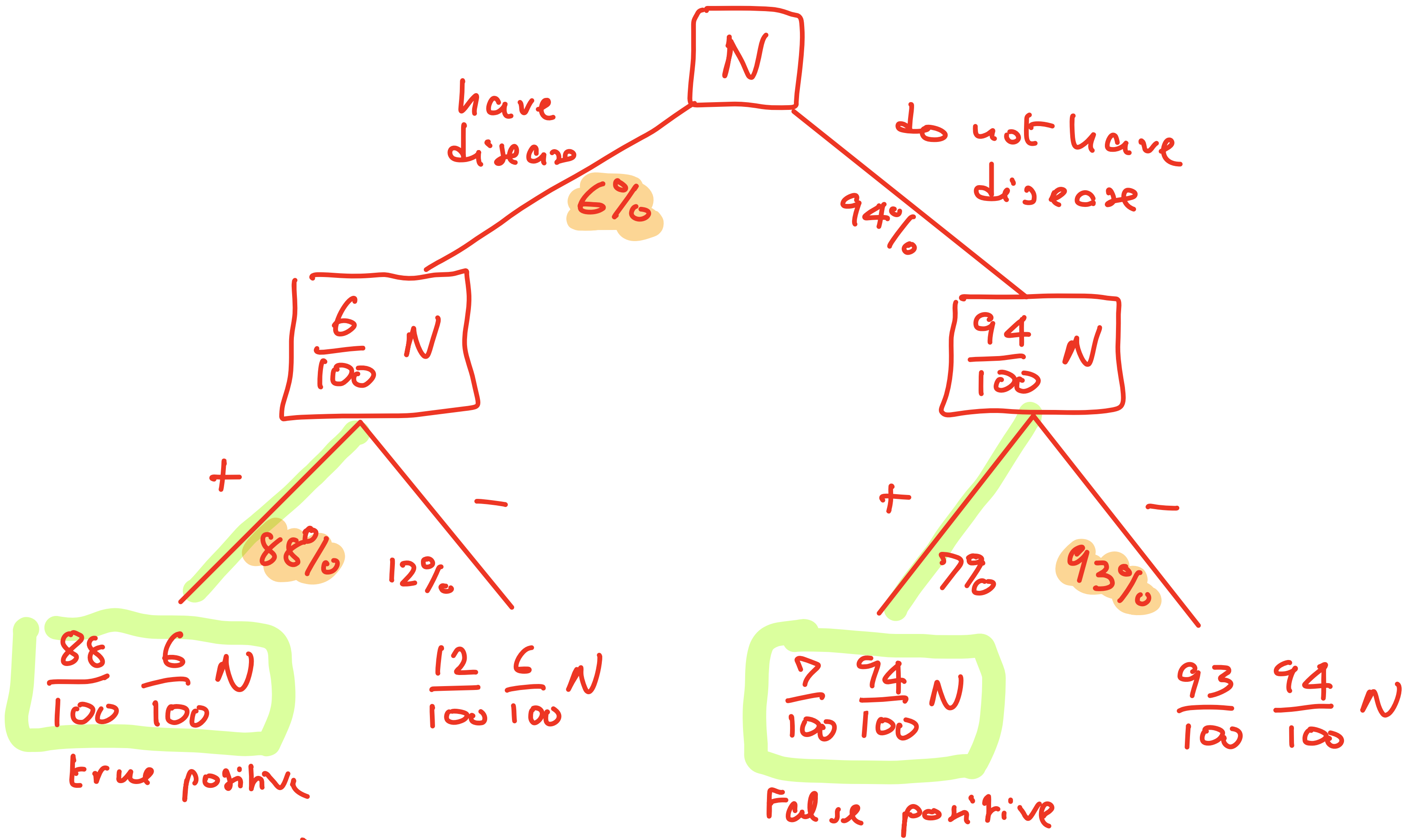
Suppose:

Base rate = 6%

Sensitivity = 88%

Specificity = 93%

Suppose you test positive. What is the probability that you have the disease?



$P(D \mid +)$

prob. of having disease given that you test positive

$$= \frac{\frac{88}{\cancel{100}}\frac{6}{\cancel{100}}\cancel{N}}{\frac{88}{\cancel{100}}\frac{6}{\cancel{100}}\cancel{N} + \frac{7}{\cancel{100}}\frac{94}{\cancel{100}}\cancel{N}} = \frac{88 \cdot 6}{88 \cdot 6 + 7 \cdot 94}$$

$$= 44.52\%$$

MORE THAN TWO CATEGORIES

Enrollment in a class

ECN	*ARE*	*PSY*	*Other*
38%	20%	12%	30%

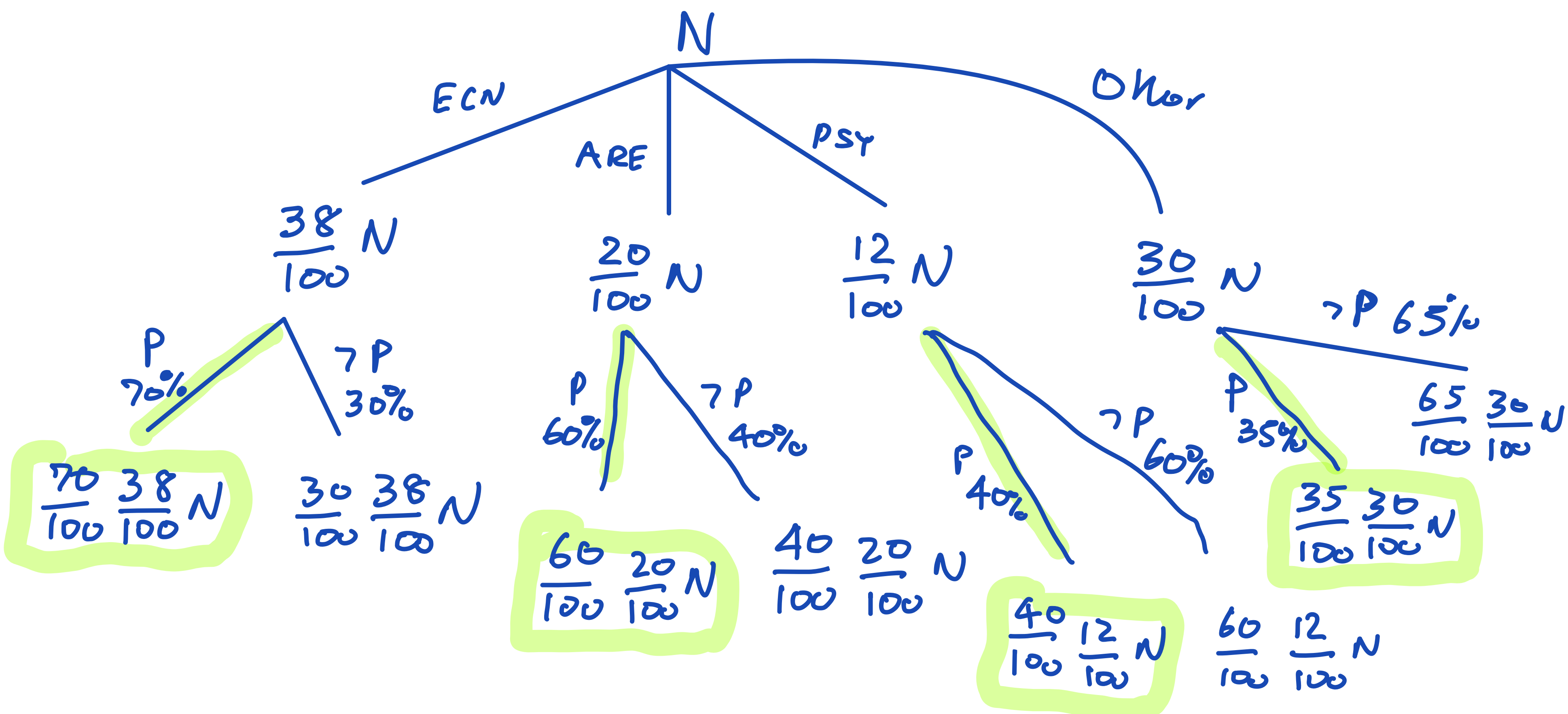
Percentages of those who passed:

major	*ECN*	*ARE*	*PSY*	*Other*
percentage who passed	70%	60%	40%	35%

You learn that Ann passed the class. How likely is it that Ann is a PSY major?

major	*ECN*	*ARE*	*PSY*	*Other*
enrollment	38%	20%	12%	30%
percentage who passed	70%	60%	40%	35%

Ann passed the class. How likely is it that she is a PSY major?



P(Ann is PSY | Passed)

$$= \frac{\frac{40}{\cancel{100}}\,\frac{12}{\cancel{100}}\,\cancel{N}}{\frac{70}{\cancel{100}}\,\frac{38}{\cancel{100}}\,\cancel{N} + \frac{60}{\cancel{100}}\,\frac{20}{\cancel{100}}\,\cancel{N} + \frac{40}{\cancel{100}}\,\frac{12}{\cancel{100}}\,\cancel{N} + \frac{35}{\cancel{100}}\,\frac{30}{\cancel{100}}\,\cancel{N}}$$

$$= \frac{40 \cdot 12}{70 \cdot 38 + 60 \cdot 20 + 40 \cdot 12 + 35 \cdot 30} = 8.9\%$$

Probability and conditional probability

Finite set of ***states*** $S = \{s_1, s_2, \ldots, s_n\}$. Subsets of S are called ***events***.

Probability distribution over S:

$0 \le p_1 \le 1$, $0 \le p_2 \le 1$, ...

$$\begin{array}{cccc} s_1 & s_2 & \ldots & s_n \\ p_1 & p_2 & \ldots & p_n \end{array}$$

$p_1 + p_2 + \ldots + p_n = 1$

Denote the probability of state s by $p(s)$.

$p(s_2) = p_2$

Given an event $E \subseteq S$, the probability of E is:

$$P(E) = \begin{cases} 0 & \text{if } E = \emptyset \\ \sum_{s \in E} p(s) & \text{if } E \neq \emptyset \end{cases}$$

Denote by $\neg E$ the complement of $E \subseteq S$.

$\overline{E}$

Example

$S = \{a, b, c, d, e, f, g\}$ $A = \{\dot{a}, \dot{c}, \dot{d}, \dot{e}\}$ $B = \{a, e, g\}$

$\neg A = \{b, f, g\}$ $\neg B = \{b, c, d, f\}$

Given

a	b	c	d	e	f	g
$\frac{1}{14}$	$\frac{2}{14}$	0	$\frac{1}{14}$	$\frac{6}{14}$	$\frac{1}{14}$	$\frac{3}{14}$

$P(A) = \frac{1}{14} + 0 + \frac{1}{14} + \frac{6}{14} = \frac{8}{14}$ $P(B) = \frac{1}{14} + \frac{6}{14} + \frac{3}{14} = \frac{10}{14}$

$A \cap B = \{a, e\}$ $P(A \cap B) = \frac{1}{14} + \frac{6}{14} = \frac{7}{14}$

$A \cup B = \{a, c, d, e, g\}$ $P(A \cup B) = \frac{1}{14} + 0 + \frac{1}{14} + \frac{6}{14} + \frac{3}{14} = \frac{11}{14}$

Note: for every two events E and F:

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

We denote by $P(E \mid F)$ the probability of E **conditional on** F and define it as:

$$\boxed{P(E \mid F) = \frac{P(E \cap F)}{P(F)}}$$

ASSUMING THAT $P(F) > 0$

Continuing the example above where

a	b	c	d	e	f	g
$\frac{1}{14}$	$\frac{2}{14}$	0	$\frac{1}{14}$	$\frac{6}{14}$	$\frac{1}{14}$	$\frac{3}{14}$

$A = \{a, c, d, e\}$ $\quad B = \{a, e, g\}$

$$P(A) = \tfrac{1}{14} + 0 + \tfrac{1}{14} + \tfrac{6}{14} = \tfrac{8}{14}, \quad P(B) = \tfrac{1}{14} + \tfrac{6}{14} + \tfrac{3}{14} = \tfrac{10}{14}, \quad P(A \cap B) = \tfrac{1}{14} + \tfrac{6}{14} = \tfrac{7}{14}$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{7}{14}}{\frac{10}{14}} = \frac{7}{10}$$

$A \cap B = B \cap A$ always

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{\frac{7}{14}}{\frac{8}{14}} = \frac{7}{8}$$

The conditional probability formula can also be applied to individual states:

$E = \{s\}$

$P(E \mid F) = P(\{s\} \mid F)$

$= P(s \mid F)$

$$p(s \mid F) = \begin{cases} 0 & \text{if } s \notin F \\ \dfrac{P(s)}{P(F)} & \text{if } s \in F \end{cases}$$

$\{s\} \cap F = \{s\}$

assuming $P(F) > 0$

$P(A) = \frac{8}{14}$ $\quad$ $P(A|B) = \frac{7}{8}$

We can think of $p(\cdot|E)$ as a probability distribution on the entire set S. Continuing the example above where $S = \{a,b,c,d,e,f,g\}$, $\boxed{A = \{a,c,d,e\}}$ and

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{1}{14} & \frac{2}{14} & 0 & \frac{1}{14} & \frac{6}{14} & \frac{1}{14} & \frac{3}{14} \end{array}$$

(so that $P(A) = \frac{8}{14}$)

$$P_{new}: \begin{array}{ccccccc} \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{6}{8} & 0 & 0 \\ a & b & c & d & e & f & g \end{array}$$

$$p(\cdot|A): \begin{array}{ccccccc} \frac{1/14}{8/14} & 0 & \frac{0}{8/14} & \frac{1/14}{8/14} & \frac{6/14}{8/14} & 0 & 0 \end{array}$$

for any event E

$P(E|A) = P_{new}(A)$

Shortcut to obtain the revised or updated probabilities:

Initial or prior probabilities. Note that here they all have the **same denominator**.	$\begin{array}{cccc} a & b & c & d \end{array}$ $\begin{pmatrix} \frac{15}{100} & \frac{70}{100} & \frac{5}{100} & \frac{10}{100} \end{pmatrix}$
Information or conditioning event: $F = \{a,b,d\}$	
STEP 1. Set the probability of every state which is not in F to zero:	$\begin{array}{cccc} a & b & c & d \end{array}$ $\begin{pmatrix} & & 0 & \end{pmatrix}$
STEP 2. For the other states write new fractions with the same numerators as before:	$\begin{array}{cccc} a & b & c & d \end{array}$ $\begin{pmatrix} \frac{15}{\dots} & \frac{70}{\dots} & 0 & \frac{10}{\dots} \end{pmatrix}$
STEP 3. In every denominator put the sum of the numerators: 15+70+10=95. Thus the updated probabilities are:	$\begin{array}{cccc} a & b & c & d \end{array}$ $\begin{pmatrix} \frac{15}{95} & \frac{70}{95} & 0 & \frac{10}{95} \end{pmatrix}$

as denominator put $15+70+10 = 95$

initially, $P(\{a,d\} \mid \underbrace{\{a,b,d\}}_{F}) = P_{new}(a) + P_{new}(d) = \frac{15}{95} + \frac{10}{95} = \frac{25}{95}$

↑ verify

$$= \frac{P(\{a,d\} \cap \{a,d,b\})}{P(\{a,d,b\}} = \frac{\frac{25}{100}}{\frac{95}{100}} = \frac{25}{95}$$

$$\{a,d\} \cap \{a,b,d\} = \{a,d\} \qquad P(\{a,d\}) = \frac{15}{100} + \frac{10}{100}$$

$$P(\{a,b,d\}) = \frac{15}{100} + \frac{70}{100} + \frac{10}{100} = \frac{95}{100}$$

In the above example, where $\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{1}{14} & \frac{2}{14} & 0 & \frac{1}{14} & \frac{6}{14} & \frac{1}{14} & \frac{3}{14} \end{array}$ and $A = \{a, c, d, e\}$, to compute $p(\bullet | A)$

Step 1: assign zero probability to states in $\neg A$:

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \end{array}$$

Step 2: keep the same numerators for the states in A:

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \\ & 0 & & & & 0 & 0 \end{array}$$

Step 3: since the sum of the numerators is 8, put 8 as the denominator:

$$\begin{array}{ccccccc} a & b & c & d & e & f & g \\ \frac{1}{8} & 0 & 0 & \frac{1}{8} & \frac{6}{8} & 0 & 0 \end{array}$$

$\frac{3}{20} \quad \frac{6}{20} \quad \frac{1}{20} \quad 0 \quad \frac{8}{20} \quad \frac{2}{20}$

EXAMPLE 2. Sample space or set of states: $\{a,b,c,d,e,f\}$.

Initial or prior probabilities:	$\begin{array}{c} \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{pmatrix} \frac{3}{20} & \frac{3}{10} & \frac{1}{20} & 0 & \frac{2}{5} & \frac{1}{10} \end{pmatrix} \end{array}$
Information:	$F = \{a,b,d,e\}$
STEP 0. Rewrite all the probabilities **with the same denominator:**	$\begin{array}{c} \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{pmatrix} & & 0 & & & 0 \end{pmatrix} \end{array}$
STEP 1. Change the probability of every state which is not in F to zero:	$\begin{array}{c} \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{pmatrix} & & 0 & & & 0 \end{pmatrix} \end{array}$
STEP 2. Write new fractions which have the same numerators as before:	$\begin{array}{c} \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{pmatrix} & & 0 & & & 0 \end{pmatrix} \end{array}$
STEP 3. In every denominator put the sum of the numerators: 3+6+8=17.	$\begin{array}{c} \begin{matrix} a & b & c & d & e & f \end{matrix} \\ \begin{pmatrix} \frac{3}{17} & \frac{6}{17} & 0 & 0 & \frac{8}{17} & 0 \end{pmatrix} \end{array}$

INDEPENDENT EVENTS.

We say that two events A and B are independent if

$$P(A \cap B) = P(A)\,P(B) \qquad (*)$$

It follows from this and the definition of conditional probability that if A and B are independent then

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

Alternatively, one can take one of the two equalities in (**) as definition of independence and derive both the other and (*). Thus (*) and (**) are equivalent.

$$P(A \mid B) = \frac{\overbrace{P(A \cap B)}^{\text{by independence } = P(A) \cdot P(B)}}{P(B)} = \frac{P(A) \cdot \cancel{P(B)}}{\cancel{P(B)}} = P(A)$$

$$P(B \mid A) = P(B)$$

Are A and B independent? Is it true that $P(A \cap B) = P(A) \cdot P(B)$?

Going back to our example where $S = \{a,b,c,d,e,f,g\}$, $A = \{a,c,d,e\}$, $B = \{a,e,g\}$, $A \cap B = \{a,e\}$ and

a	b	c	d	e	f	g
$\frac{1}{14}$	$\frac{2}{14}$	0	$\frac{1}{14}$	$\frac{6}{14}$	$\frac{1}{14}$	$\frac{3}{14}$

$P(A) = \frac{8}{14}$, $P(B) = \frac{10}{14}$, $P(A \cap B) = \frac{7}{14}$ $\quad P(A)P(B) = \frac{8}{14} \cdot \frac{10}{14} = \frac{80}{14^2}$

$\neq$ so A and B are not independent

On the other hand, if $S = \{a,b,c,d,e,f,g,h,i\}$ and

a	b	c	d	e	f	g	h	i
$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	0	$\frac{2}{9}$	0	$\frac{1}{9}$	$\frac{1}{9}$

Then $E = \{a,b,c,e\}$ and $F = \{c,d,e,g\}$ are independent. In fact, $P(E) = \underbrace{\frac{1}{9} + \frac{1}{9} + \frac{1}{9}}_{\frac{3}{9}} + 0$, $P(F) = \frac{1}{9} + \frac{2}{9} + 0 + 0 = \frac{3}{9}$

$\underbrace{E \cap F = \{c,e\}}$, $P(E \cap F) = \frac{1}{9} + 0 = \frac{1}{9}$ and thus $P(E \cap F) = P(E)P(F)$.

$P(E \cap F) = \frac{1}{9}$ $\qquad P(E) \cdot P(F) = \frac{3}{9} \cdot \frac{3}{9} = \frac{9}{81} = \frac{1}{9}$

Same so E and F are independent