Rule # 1: In all cases,

$$0 \leq \Pr(E) \leq 1$$

Rule # 2: S is a set of events. $S = \{E_1, E_2, E_3, \ldots E_k\}$. If the events in S are mutually exclusive **and** S is exhaustive (every possible outcome is considered in the events in S), then:

$$\sum_{i=1}^{k} \Pr(E_i) = 1$$

Definition: Two events are **mutually exclusive** if they cannot both occur in the same trial:

$$\Pr(A \text{ and } B) = 0 \text{ if } A \text{ and } B \text{ are mutually exclusive}$$

Rule # 3: If A and B are mutually exclusive then

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$$

Rule # 4: The event "not A" means "any event other than A occurs"

$$\Pr(\text{not } A) = 1 - \Pr(A)$$

Rule # 5: The **general addition rule**:

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

Rule # 6: If A and B are independent events then:

$$\Pr(A \text{ and } B) = \Pr(A)\Pr(B)$$

Definition: If knowing that one event occurs tells you nothing about the probability of a second event, then the two events are **independent** of each other.

$$\Pr(A|B) = \Pr(A|\text{not } B) = \Pr(A) \text{ if } A \text{ and } B \text{ are independent}$$

Rule # 7: The **general multiplication rule**:

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

Rule # 8: Bayes' theorem:

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

Rule # 9: Law of total probability. If $S = \{E_1, E_2, E_3, \ldots E_k\}$, the events in S are mutually exclusive, and S is exhaustive:

$$\Pr(A) = \sum_{i=1}^{k} \Pr(A|E_i)\Pr(E_i)$$

General "Recipe" for solving probability problems.

You'll be given data from which you can extract probability statements.

You'll be asked to calculate a result (another probability statement, or a conclusion that relies on the calculation of a probability).

1. Write down the data and the quantity to be calculated in the correct notation:
   - What is the event that the probability statement applies to?
   - What is the sample space that the event is a part of?
   - Does the probability statement only apply to a special case of the entire realm of events? – In other words, is it a *conditional* probability statement (one that only applies in some cases)? If so, write it as a conditional probability statement.
2. Use the appropriate probability rules to solve for what you need to calculate.

Random trial: an experiment that could be repeated with differing outcomes.

Sample space: The set of all possible outcomes of a random trial.

Event: a possible result of a trial. Can be any subset of the sample space.

Probability of an event: The proportion of times that an event would occur if we could repeat a random trial many times.

Bayesian definition of probability of an event: A numerical representation of our certainty of an event ($0 \rightarrow$ impossible, and $1 \rightarrow$ certain).

Rule # 1:

$$0 \leq \Pr(E_1) \leq 1$$

$\Pr(E_1) = 0$ means that E_1 is impossible.

$\Pr(E_1) = 1$ means that E_1 is certain.

Definition: Two events are **mutually exclusive** if they cannot both occur in the same trial.

$\Pr(A \text{ and } B) = 0$ if A and B are mutually exclusive.

Rule # 2: If S is the sample space of **all** possible, **mutually exclusive** outcomes,

$S = \{E_1, E_2, E_3, \ldots E_k\}$, then:

$$\sum_{i=1}^{k} \Pr(E_i) = 1$$

"Something must happen" rule

If S is a mutually exclusive, and exhaustive description of sample space is:

$$S = \{E_1, E_2, E_3, \ldots, E_k\},$$

then we say the size of the sample space is $|S| = k$.

What is $\Pr(E_1)$?
We don't know, so let's set it to x:
Rule #2 says:

$$\begin{aligned} 1 &= \sum_{i=1}^{k} \Pr(E_i) \\ &= \sum_{i=1}^{k} x \\ &= x \sum_{i=1}^{k} 1 \\ &= kx \end{aligned}$$

$$x = \Pr(E_i) = 1/k$$

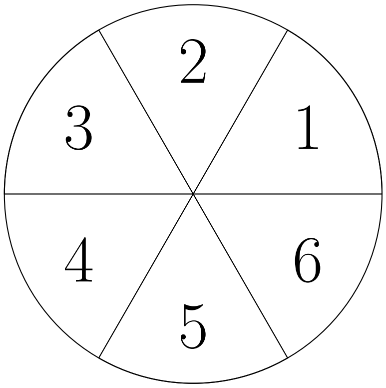
This is how rule #2 tells us the the probability of rolling a 3 on a six-sided dice is $1/6$.

Mutually exclusive addition rule.

Rule # 3: If A and B are mutually exclusive then

$$\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$$

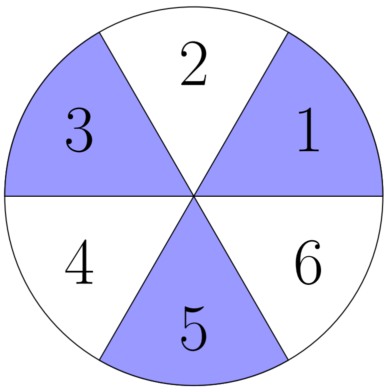
Example: What is the probability of rolling an odd number?



The sample space. $S = \{1, 2, 3, 4, 5, 6\}$

All events in S assumed to be equiprobable.

$\Pr(\text{roll} = 1) = 1/6$



Event $A = \text{roll is odd} = \{1, 3, 5\}$

$$
\begin{aligned}
\Pr(A) &= \Pr(1 \text{ or } 3 \text{ or } 5) \\
&= \Pr(1) + \Pr(3) + \Pr(5) \\
&= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}
\end{aligned}
$$

Complement rule.

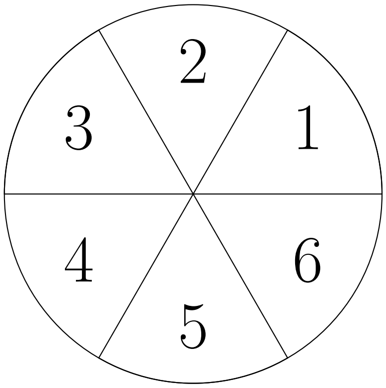
Events "A" and "not A" are mutually exclusive.

Events "A" or "not A" include the entire sample space.

Rule # 4:

$$\Pr(\text{not } A) = 1 - \Pr(A)$$

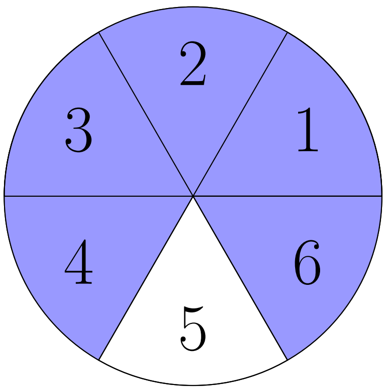
Example: What is the probability that the roll would be something other than 5?



The sample space. $S = \{1, 2, 3, 4, 5, 6\}$

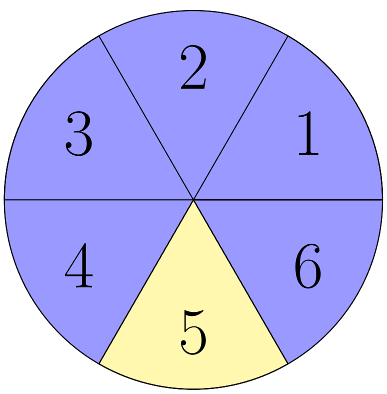
All events in S assumed to be equiprobable.

$\Pr(\text{roll} = 1) = 1/6$



Event A = roll is anything other than 5

$$\Pr(A) = \Pr(1 \text{ or } 2 \text{ or } 3 \text{ or } 4 \text{ or } 6)$$
$$= \Pr(1) + \Pr(2) + \Pr(3) + \Pr(4) + \Pr(6)$$
$$= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{5}{6}$$



Event B = roll is 5

$\Pr(B) = \Pr(5) = \frac{1}{6}$

Events "B" and "not B" are:

mutually exclusive, and

exhaustive

$$\Pr(B) + \Pr(\text{not } B) = 1$$

Rule #4: $\Pr(\text{not } B) = 1 - \Pr(B)$

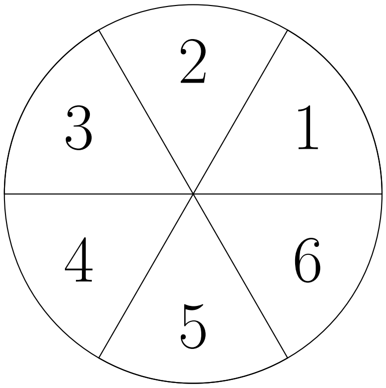
$$\Pr(\text{not } B) = 1 - \tfrac{1}{6} = \tfrac{5}{6}$$

General addition rule.

Rule # 5:

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

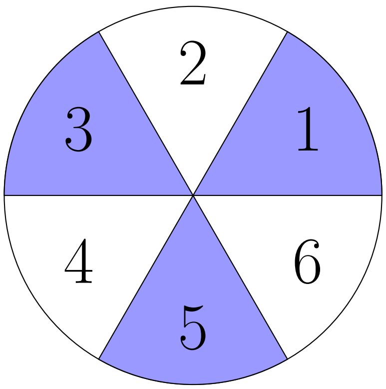
Example: What is the probability that a roll will be odd or less than 4 ?



The sample space. $S = \{1, 2, 3, 4, 5, 6\}$

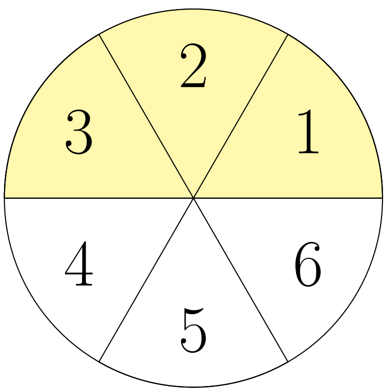
All events in S assumed to be equiprobable.

$\Pr(\text{roll} = 1) = 1/6$



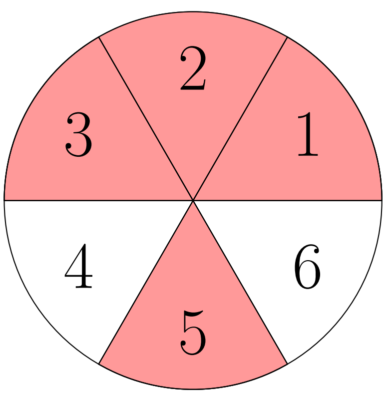
Event $A = \text{roll is odd} = \{1, 3, 5\}$

$$\begin{aligned}\Pr(A) &= \Pr(1 \text{ or } 3 \text{ or } 5)\\ &= \Pr(1) + \Pr(3) + \Pr(5)\\ &= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{2}\end{aligned}$$



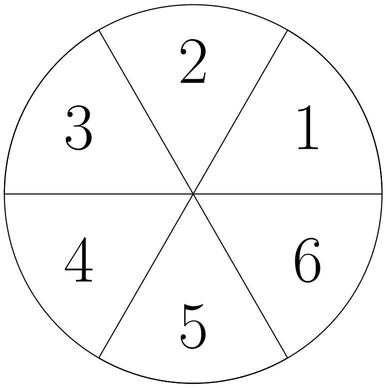
Event $B = \text{roll } < 4 = \{1, 2, 3\}$

$$\begin{aligned}\Pr(B) &= \Pr(1 \text{ or } 2 \text{ or } 3)\\ &= \Pr(1) + \Pr(2) + \Pr(3)\\ &= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{2}\end{aligned}$$



$$\begin{aligned}\text{Event } C &= \text{roll is odd or roll } < 4\\ &= A \cup B = \{1, 2, 3, 5\}\end{aligned}$$

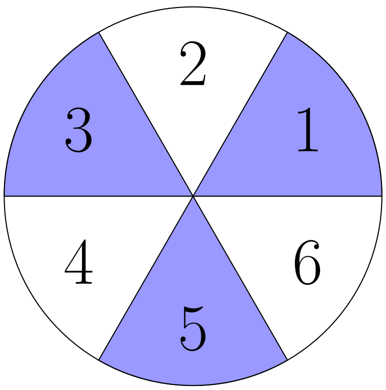
$$\begin{aligned}\Pr(C) &= \Pr(A \text{ or } B)\\ &= \Pr(1 \text{ or } 2 \text{ or } 3 \text{ or } 5)\\ &= \Pr(1) + \Pr(2) + \Pr(3) + \Pr(5)\\ &= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6}\\ &= \tfrac{2}{3}\end{aligned}$$



The sample space. $S = \{1, 2, 3, 4, 5, 6\}$

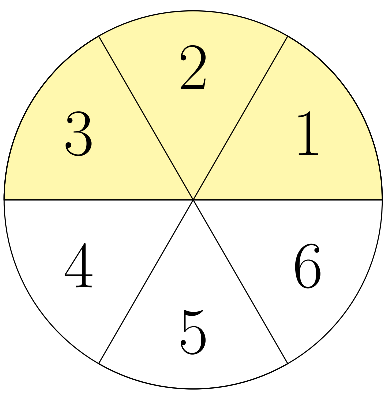
All events in S assumed to be equiprobable.

$\Pr(\text{roll} = 1) = 1/6$



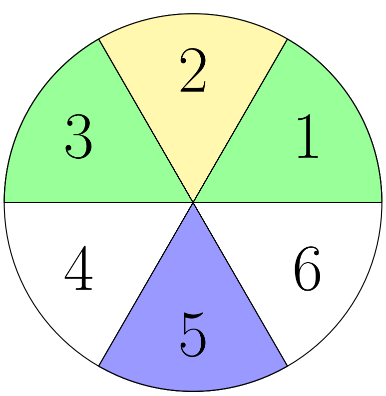
Event A = roll is odd

$$\Pr(A) = \Pr(1 \text{ or } 3 \text{ or } 5)$$
$$= \Pr(1) + \Pr(3) + \Pr(5)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$



Event B = roll < 4

$$\Pr(B) = \Pr(1 \text{ or } 2 \text{ or } 3)$$
$$= \Pr(1) + \Pr(2) + \Pr(3)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$



Event C = roll is odd or roll < 4

$$\Pr(C) = \Pr(A \text{ or } B)$$
$$= \Pr(A) + \Pr(B) - \Pr(A \text{ and } B)$$
$$= \frac{1}{2} + \frac{1}{2} - \frac{2}{6}$$
$$= \frac{4}{6} = \frac{2}{3}$$

Definition: If knowing that one event occurs tells you **nothing** about the probability of a second event, then the two events are **independent** of each other.

$$\Pr(A \mid B) = \Pr(A \mid \text{not } B) = \Pr(A)$$

Rule # 6: If A and B are **independent** events then:

$$\Pr(A \text{ and } B) = \Pr(A)\Pr(B)$$

Example: What is the probability that the next child born in the US will be female and will get an odd social security #?

Consider rolling two dice: What is the probability that they will both show 6?

The **general multiplication rule** (works whether or not A and B are independent).

Rule # 7:

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

Approximately 16.5% of males will develop prostate cancer. Assuming 50:50 gender ratio at birth, what is the probability that a randomly chosen infant will be male and will prostate cancer?

The purple petal allele (P) is dominant to white (p).
The round pea allele (R) is dominant to wrinkled (r).
The petal trait is inherited independently of the round/wrinkled trait.

What is the probability the offspring of a PpRr $\times$ PpRr cross will have white flowers and round peas?

$$
\begin{aligned}
\Pr(\text{round}) &= \Pr(RR \text{ or heterozygous } Rr) \\
&= \Pr(RR) + \Pr(\text{heterozygous } Rr) \\
\Pr(\text{white}) &= \Pr(pp)
\end{aligned}
$$

$$
\begin{aligned}
\Pr(RR) &= \Pr(\female = R \text{ and } \male = R) \\
\Pr(\text{heterozygous } Rr) &= \Pr([\female = R \text{ and } \male = r] \\
&\qquad \text{or } [\female = r \text{ and } \male = R])
\end{aligned}
$$

What is the probability of a woman becoming pregnant at least one time over a ten year time span?

For a single year:

$$\Pr(\text{pregnancy}|\text{no contraception}) = 0.85$$

Typical use:

$$\Pr(\text{pregnancy}|\text{typical use of condom}) = 0.15$$
$$\Pr(\text{pregnancy}|\text{typical use of pill}) = 0.08$$

$$\Pr(\text{pregnancy}|\text{perfect use of condom}) = 0.02$$
$$\Pr(\text{pregnancy}|\text{perfect use of pill}) = 0.003$$

Data from http://www.contraceptivetechnology.org/table.html

$$
\begin{aligned}
\Pr(\text{pregnancy in 1 year}|\text{no contraception}) &= 0.85 \\
\Pr(\text{no pregnancy in 1 year}|\text{no contraception}) &= 0.15
\end{aligned}
$$

$$
\begin{aligned}
\Pr(\text{P in 1}|\text{None}) &= 0.85 \\
\Pr(\text{not P in 1}|\text{None}) &= 0.15 \\
\Pr(\text{not P in 10}|\text{None}) &= \Pr(\text{not P in 1}|\text{None})^{10} \\
\Pr(\text{P in 10}|\text{None}) &= 1 - \Pr(\text{not P in 10}|\text{None})
\end{aligned}
$$

$$
\Pr(\text{at least 1 P in 10}|\text{None}) = 1 - 5.8 \times 10^{-09}
$$

$$
\begin{aligned}
\Pr(\text{No P in 10}|\text{Typical condom use}) &= 0.197 \\
\Pr(\text{at least 1 P in 10}|\text{Typical condom use}) &= 0.803 \\
\Pr(\text{No P in 10}|\text{Perfect condom use}) &= 0.817 \\
\Pr(\text{at least 1 P in 10}|\text{Perfect condom use}) &= 0.183
\end{aligned}
$$

Rule # 8: Bayes' theorem:

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

Rule # 9: Law of total probability. If $\{E_1, E_2, E_3, \ldots E_k\}$ covers the entire sample space and these events are all mutually exclusive:

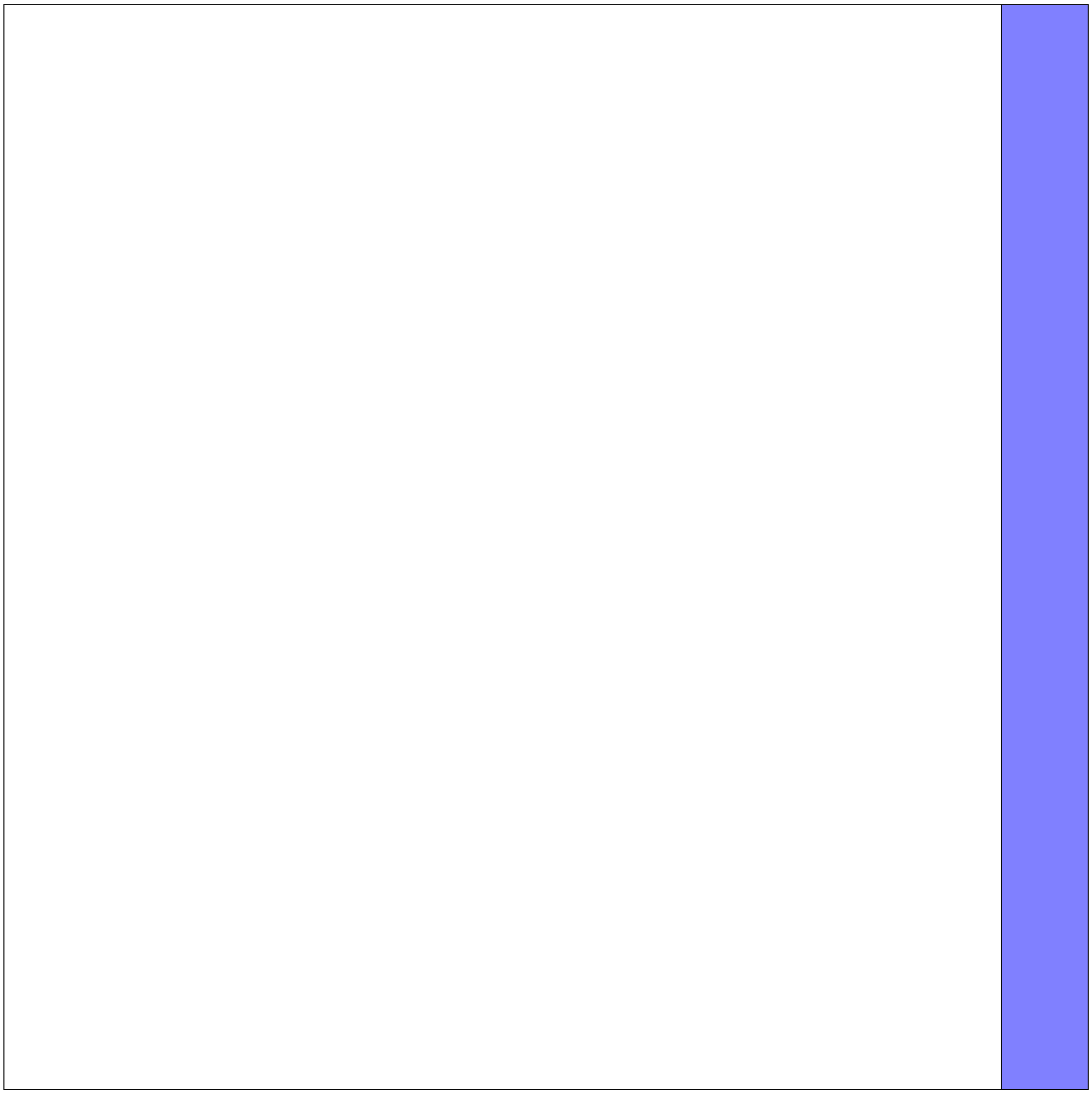
$$\Pr(A) = \sum_{i=1}^{k} \Pr(A|E_i)\Pr(E_i)$$

A disease is found in 8% of the population. There is a diagnostic test. The test has a false positive rate of 10% (so 10% of healthy people will have positive test result), and a false negative rate of 5% (someone with the disease will have a negative result for the test 5% of the time). If someone has a positive result, what is the chance that they have the disease?

$$\begin{aligned} \Pr(D) &= 0.08 \\ \Pr(T_+|\text{not } D) &= 0.1 \\ \Pr(\text{not } T_+|D) &= 0.05 \end{aligned}$$

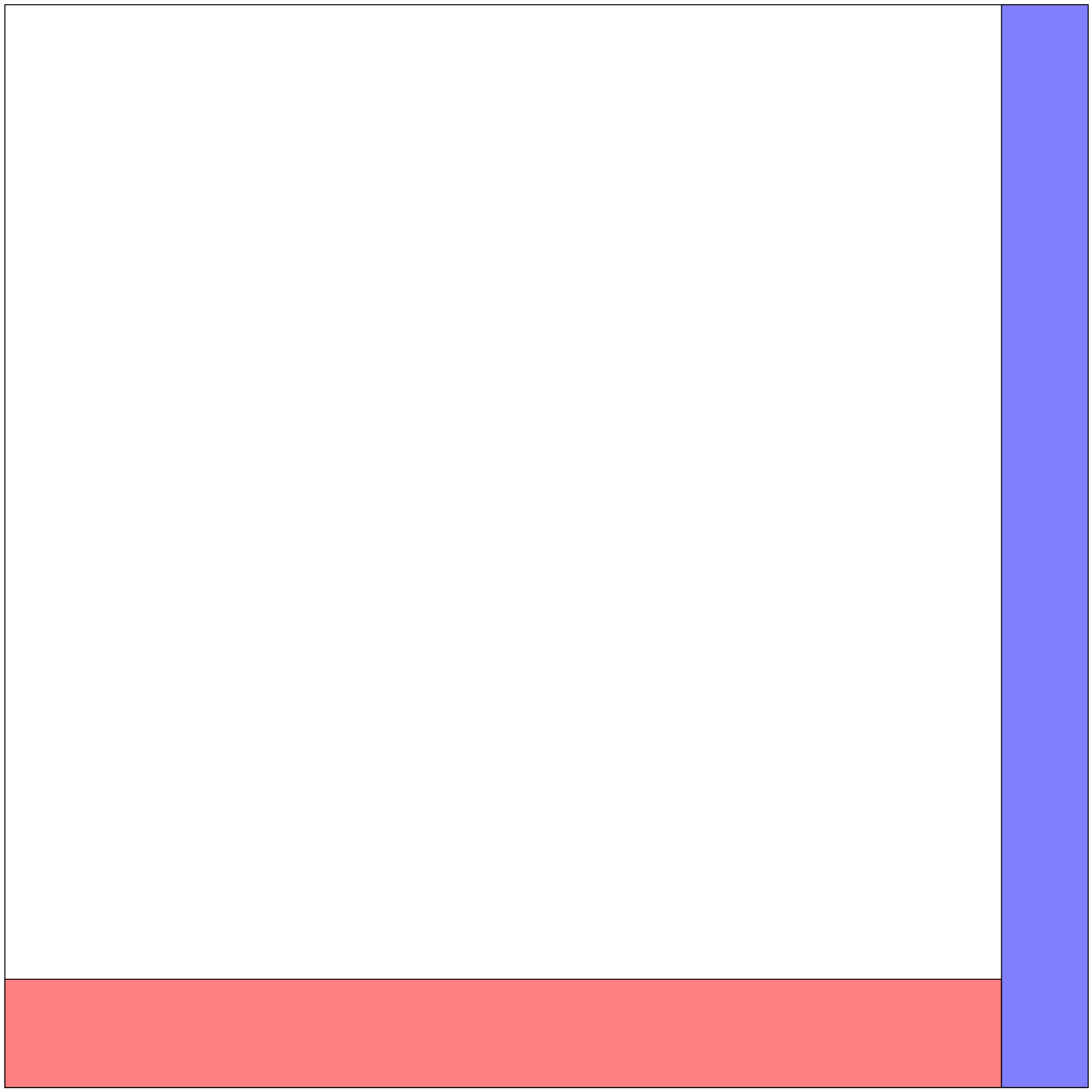
$$\begin{aligned}
\Pr(\text{not } D) &= 0.92 \\
\Pr(T_+|D) &= 0.95 \\
\Pr(\text{not } T_+|\text{not } D) &= 0.90
\end{aligned}$$

$$\begin{aligned}
\Pr(D|T_+) &= \Pr(T_+|D)\Pr(D)/\Pr(T_+) \\
\Pr(T_+|D)\Pr(D) &= 0.95 \times 0.08 = 0.076 \\
\Pr(T_+) &= \Pr(T_+|D)\Pr(D) + \Pr(T_+|\text{not } D)\Pr(\text{not } D) \\
&= 0.95 \times 0.08 + 0.1 \times 0.92 \\
&= 0.168 \\
\Pr(D|T_+) &= \frac{\Pr(T_+|D)\Pr(D)}{\Pr(T_+)} \\
&= \frac{.076}{0.168} = 0.452
\end{aligned}$$



blue: $\Pr(D) = 0.08$

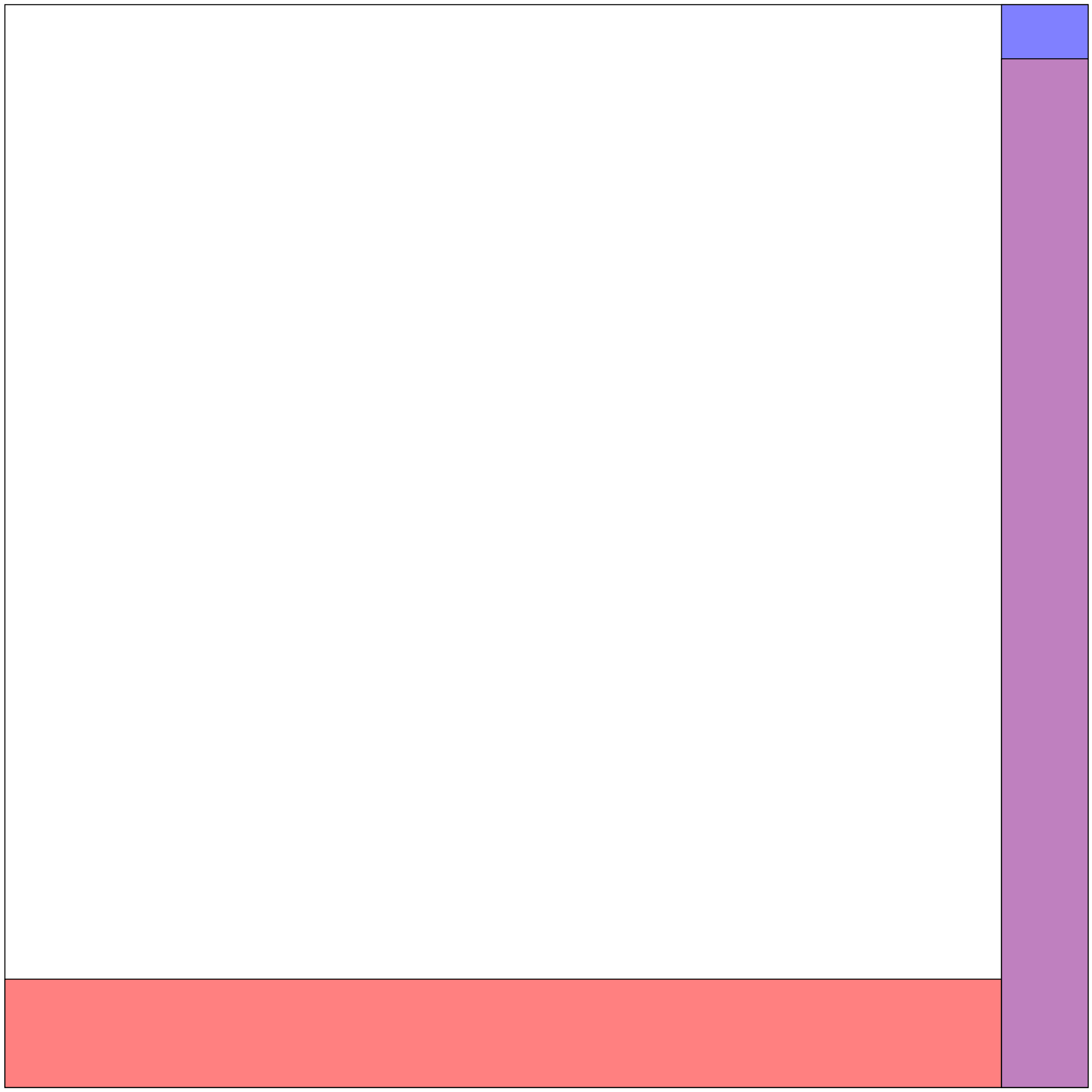
white: $\Pr(\text{not } D) = 0.92$



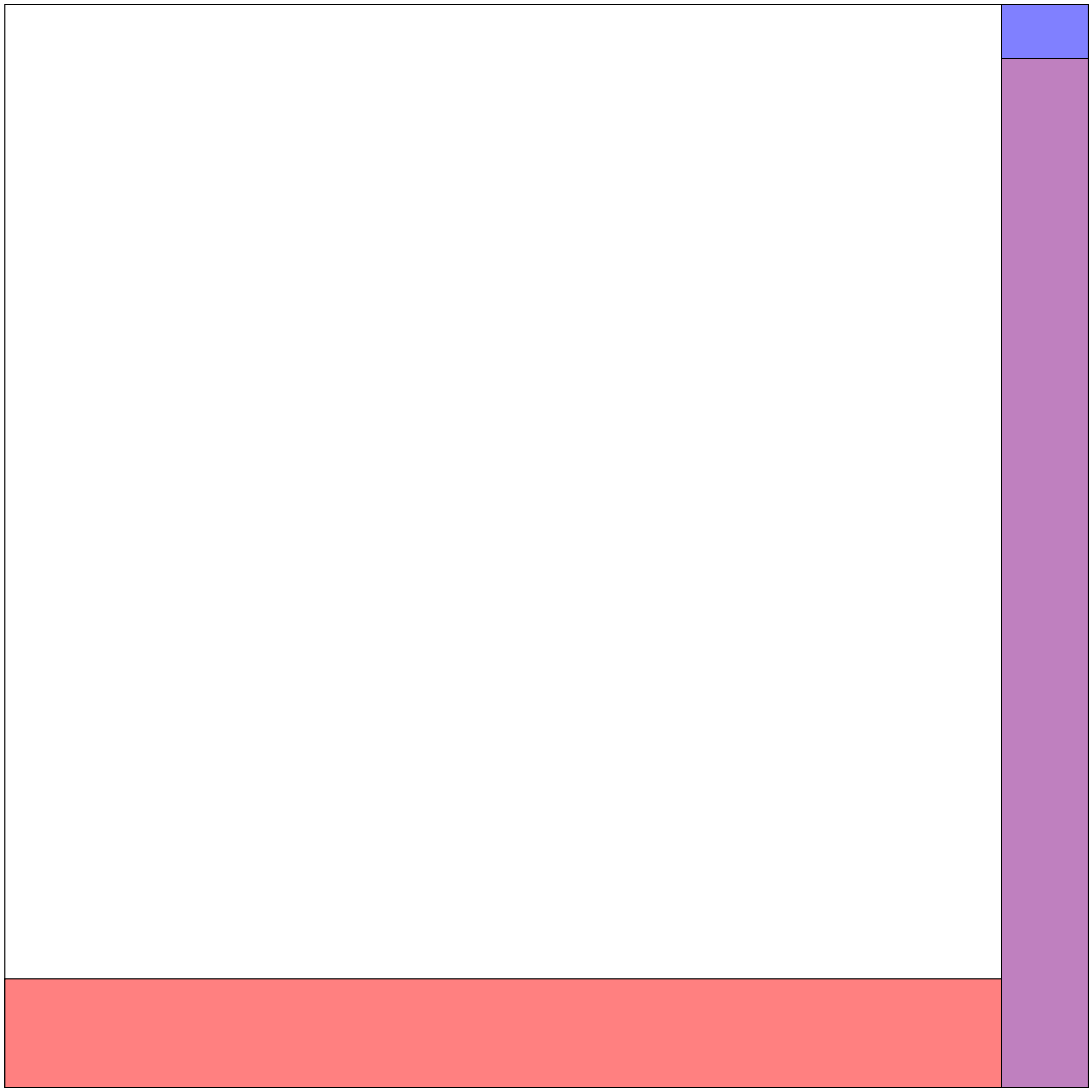
$$\begin{aligned}
\text{blue: } \Pr(D) &= 0.08 \\
\text{red: } \Pr(T_+ \text{ and not } D) &= \Pr(\text{not } D)\,\Pr(T_+|\text{not } D) \\
&= (0.92)(0.10) \\
&= 0.092
\end{aligned}$$

$$\begin{aligned}
\Pr(\text{not } D) &= \Pr(T_+ \text{ and not } D) + \Pr([\text{not } T_+] \text{ and not } D) \\
\Pr([\text{not } T_+] \text{ and not } D) &= \Pr(\text{not } D) - \Pr(T_+ \text{ and not } D)
\end{aligned}$$

$$\begin{aligned}
\text{white: } \Pr([\text{not } T_+] \text{ and not } D) &= 0.92 - 0.092 \\
&= 0.828
\end{aligned}$$



$$
\begin{aligned}
\text{purple: } \Pr(T_+ \text{ and } D) &= \Pr(D)\Pr(T_+|D) \\
&= 0.076 \\
\text{blue: } \Pr([\text{not } T_+] \text{ and } D) &= \Pr(D) - \Pr(T_+ \text{ and } D) \\
&= 0.004 \\
\text{red: } \Pr(T_+ \text{ and not } D) &= 0.092 \\
\text{white: } \Pr([\text{not } T_+] \text{ and not } D) &= 0.828
\end{aligned}
$$



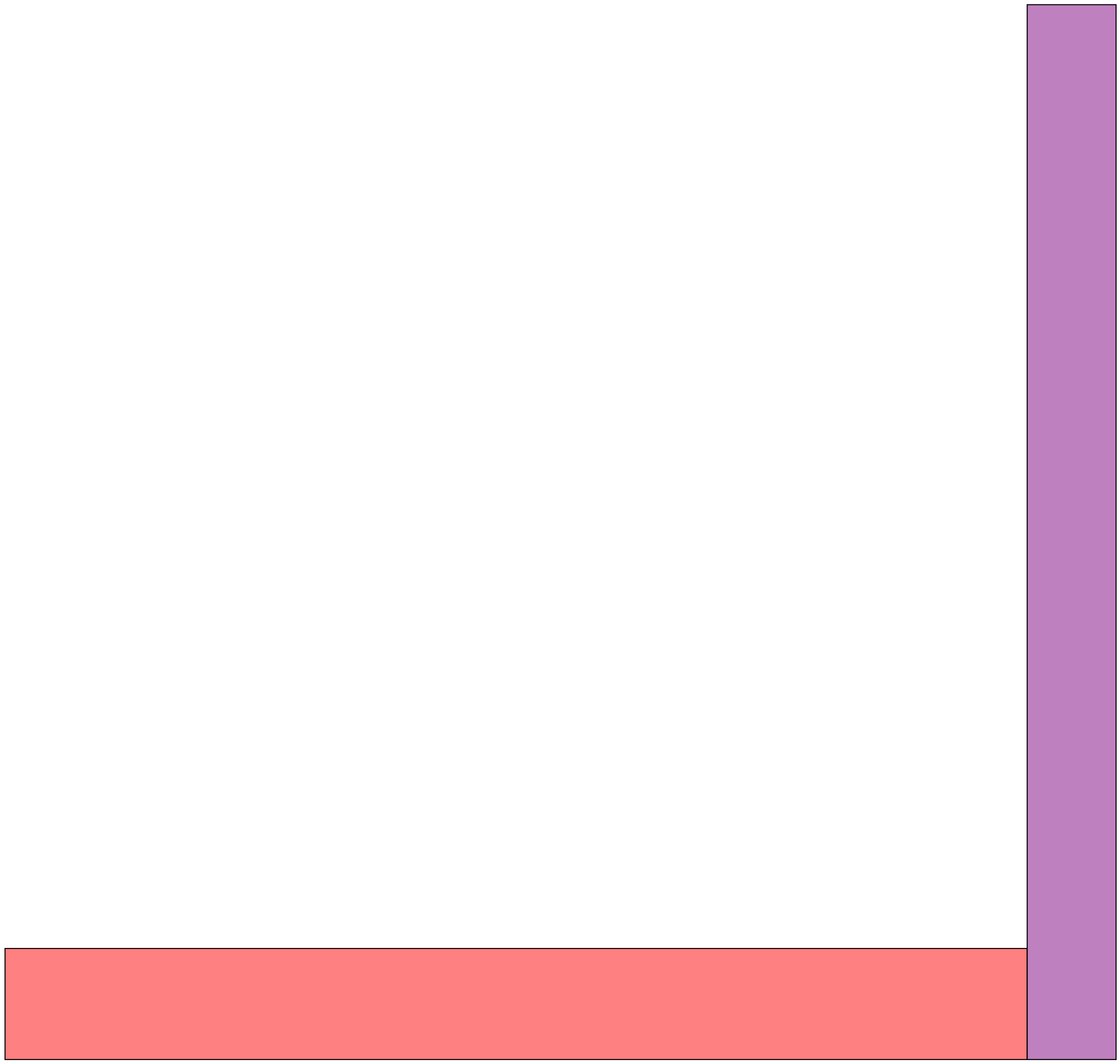
$T_{-} = \text{not } T_{+} \qquad H = \text{not } D$

purple: $\Pr(T_{+} \text{ and } D) = 0.076$

blue: $\Pr(T_{-} \text{ and } D) = 0.004$

red: $\Pr(T_{+} \text{ and } H) = 0.092$

white: $\Pr(T_{-} \text{ and } H) = 0.828$



$$
\begin{aligned}
\text{purple: } \Pr(+ \text{ and } D) &= 0.076 \\
\text{red: } \Pr(+ \text{ and } H) &= 0.092 \\
\text{red or purple: } \Pr(+) &= \Pr(+ \text{ and } D) + \Pr(+ \text{ and } H) \\
&= 0.076 + 0.092 \\
&= 0.168
\end{aligned}
$$

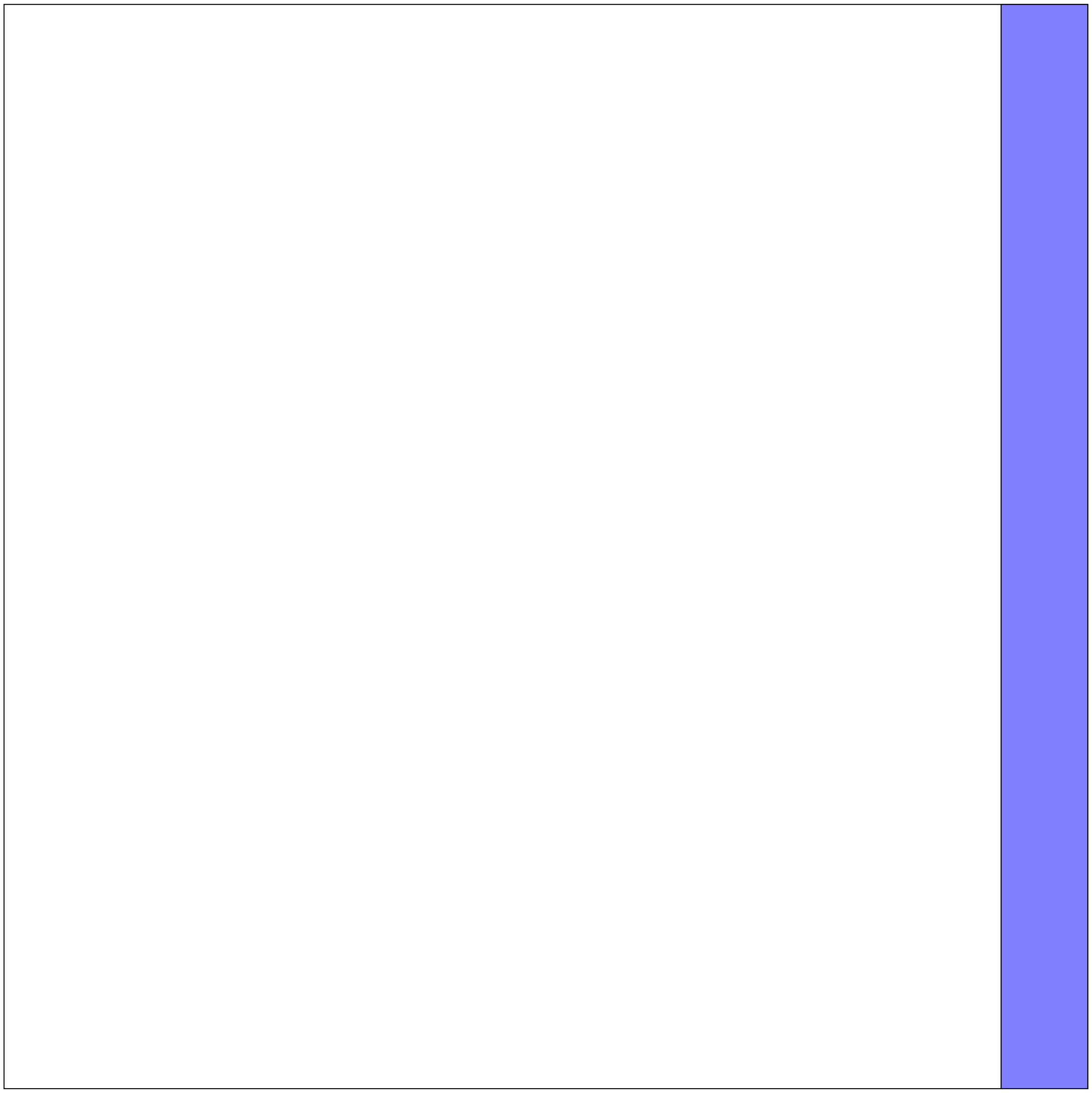
$$
\begin{aligned}
\Pr(D|+) &= \frac{\Pr(+|D)\Pr(D)}{\Pr(+)} = \frac{\Pr(+ \text{ and } D)}{\Pr(+)} \\
&= \frac{0.076}{0.168} \\
&= 0.452
\end{aligned}
$$

A disease is found in 8% of the population. There is a diagnostic test for the disease, but the test has a false positive rate of 10% (in other words 10% of healthy people will have positive result for the test), and a false negative rate of 5% (someone with the disease will not have a positive result for the test 5% of the time). If someone has a positive result, what is the chance that they have the disease?

$$\begin{aligned} \Pr(D) &= 0.08 \\ \Pr(+|\text{not } D) &= 0.1 \\ \Pr(\text{not } +|D) &= 0.05 \\ \Pr(D|+) &= ? \end{aligned}$$

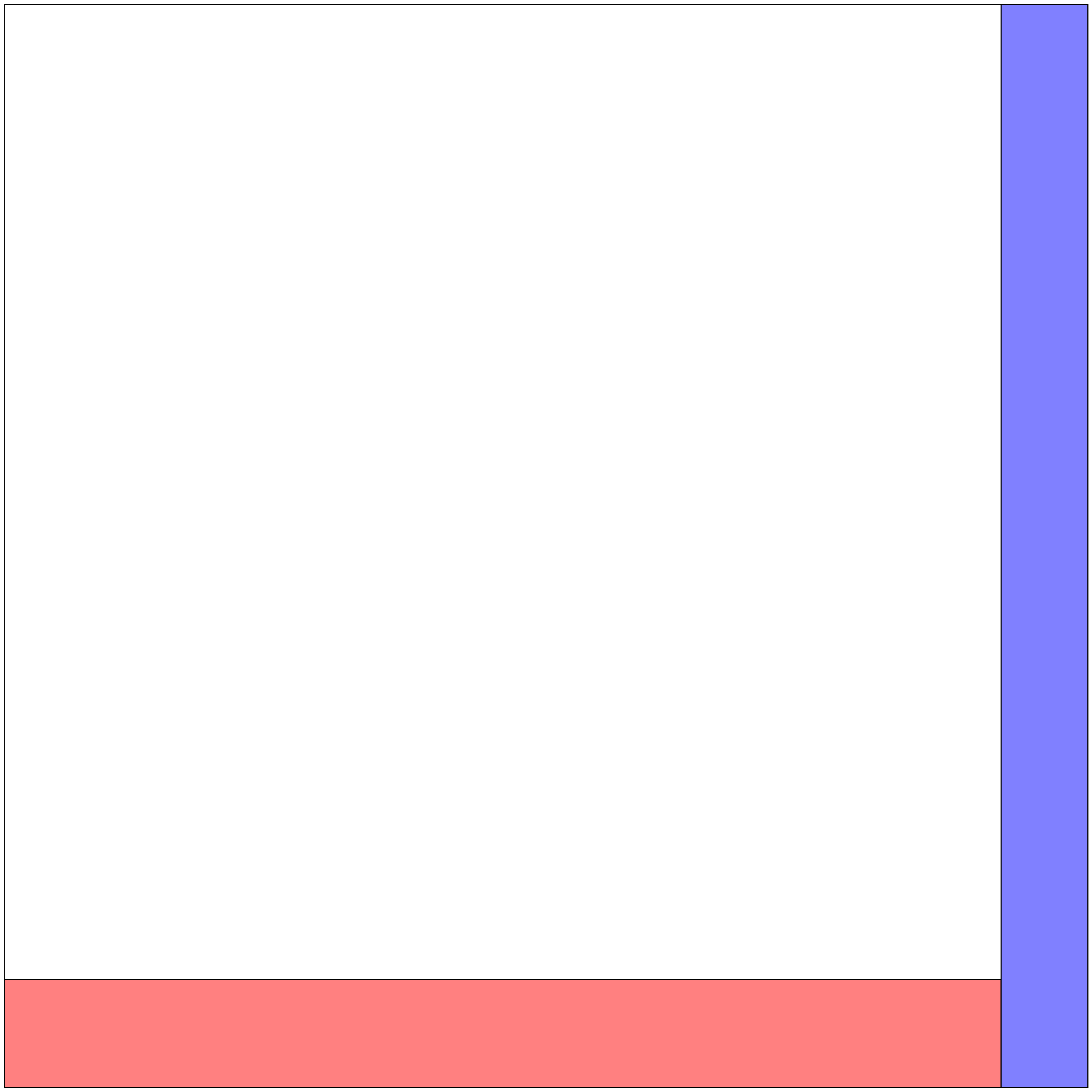
$$\begin{aligned} \Pr(\text{not } D) &= 0.92 \\ \Pr(+|D) &= 0.95 \\ \Pr(\text{not } +|\text{not } D) &= 0.90 \end{aligned}$$

$$\begin{aligned} \Pr(D|+) &= \Pr(+|D)\Pr(D)/\Pr(+) \\ \Pr(+|D)\Pr(D) &= 0.95 \times 0.08 = 0.076 \\ \Pr(+) &= \Pr(+|D)\Pr(D) + \Pr(+|\text{not } D)\Pr(\text{not } D) \\ &= 0.95 \times 0.08 + 0.1 \times 0.92 \\ &= 0.168 \\ \Pr(D|+) &= \frac{\Pr(+|D)\Pr(D)}{\Pr(+)} \\ &= \frac{.076}{0.168} = 0.452 \end{aligned}$$



blue: $\Pr(D) = 0.08$

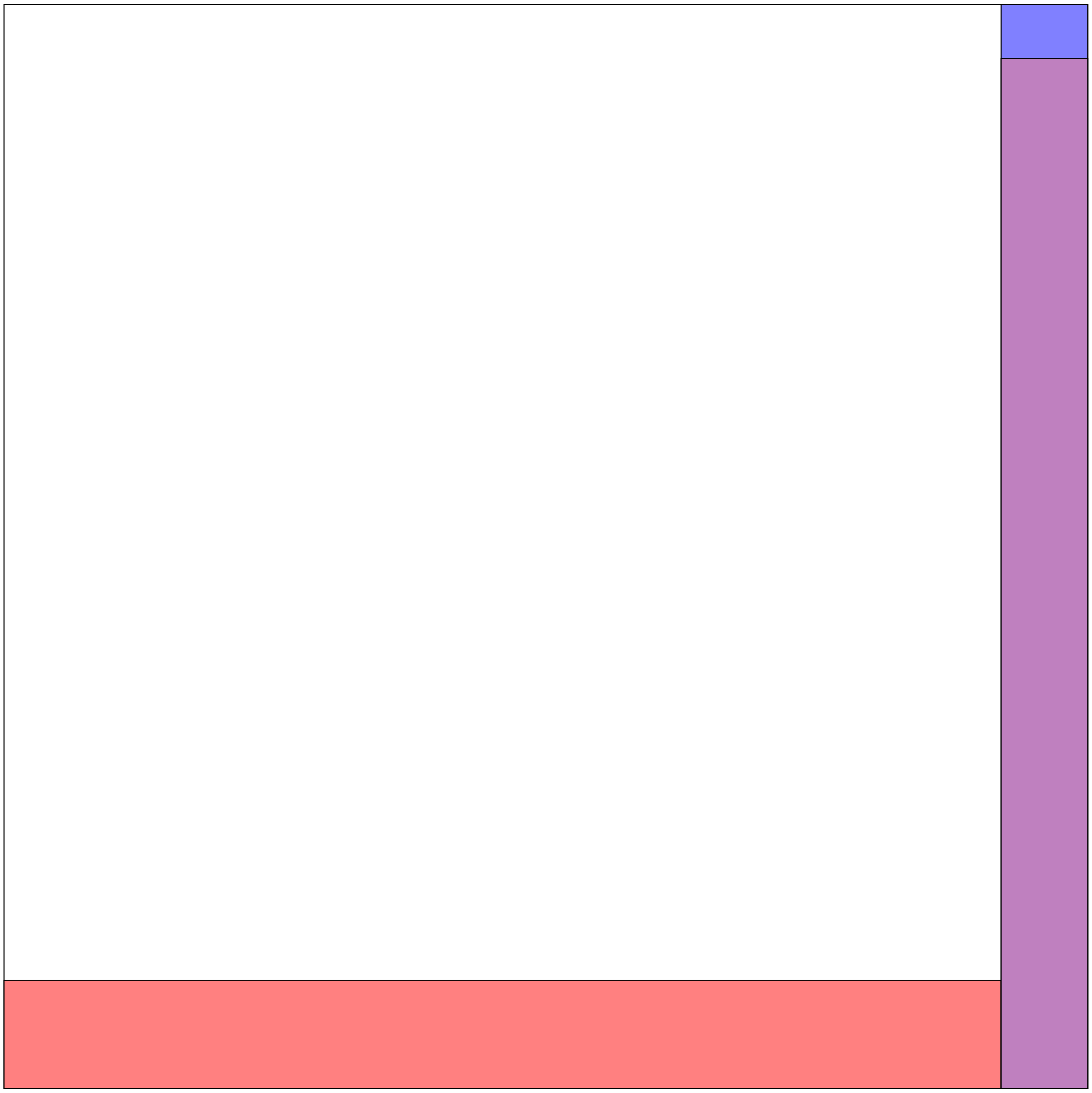
white: $\Pr(\text{not } D) = 0.92$



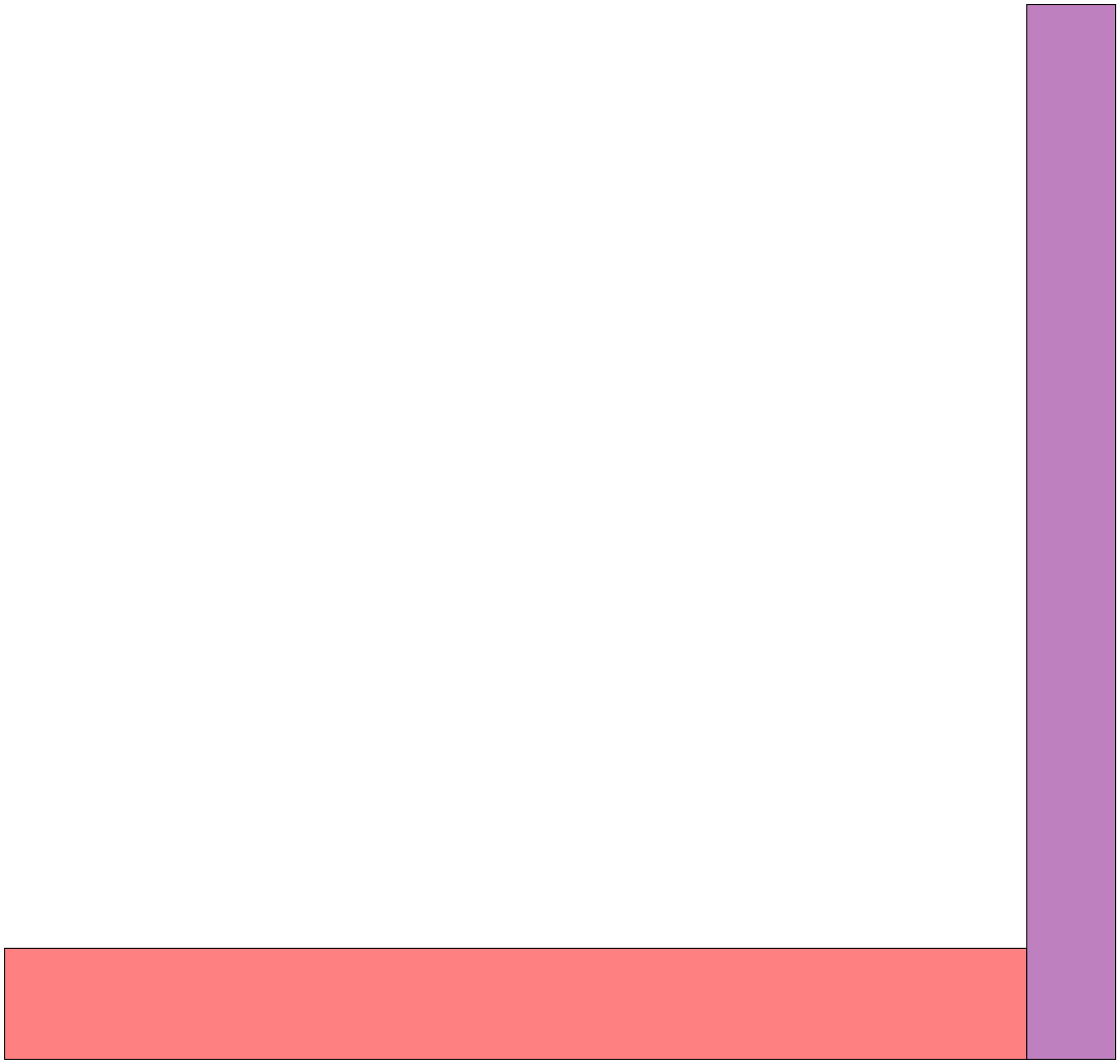
blue: $\Pr(D) = 0.08$

red: $\Pr(+ \text{ and not } D) = 0.092 = \Pr(\text{not } D)\Pr(+|\text{not } D)$

white: $\Pr(\text{not } + \text{ and not } D) = 0.828$



$$\begin{aligned}
\text{purple:}\quad & \Pr(+ \text{ and } D) = 0.076 = \Pr(D)\Pr(+|D) \\
\text{blue:}\quad & \Pr(D) = 0.004 \\
\text{red:}\quad & \Pr(+ \text{ and not } D) = 0.092 \\
\text{white:}\quad & \Pr(\text{not } + \text{ and not } D) = 0.828
\end{aligned}$$



$$\begin{aligned}
\text{purple:} \quad & \Pr(+ \text{ and } D) = 0.076 \\
\text{red:} \quad & \Pr(+ \text{ and not } D) = 0.092 \\
& \Pr(D|+) = \frac{\Pr(+ \text{ and } D)}{\Pr(+ \text{ and } D) + \Pr(+ \text{ and not } D)} \\
& = \frac{0.076}{0.076 + 0.092} \\
\Pr(D|+) \quad & = 0.452
\end{aligned}$$

Ethologists determine that when a jay attacks a skink, 10% of the time the skink loses its tail and the rest of the time the skink loses its life. If we exhaustively 1-year old skinks from a field site and find 996 skinks survived with tails, and 4 without tails, then how many were eaten by jays? Assume that tail-loss only occurs from jay attack, that jays are the only source of mortality for the skinks, and that we can ignore the chance of a tail-less skink being attacked another time.

$$\begin{aligned}
\Pr(TL|A) &= 0.1 \\
\Pr(T|\text{not } D) &= 0.996 \\
\Pr(TL|\text{not } D) &= 0.004 \\
\Pr(TL|\text{not } A) &= 0 \\
\Pr(D|\text{not } A) &= 0 \\
\Pr(T|\text{not } A) &= 1 \\
\Pr(D) &= ?
\end{aligned}$$

What is the chance of rolling a 5 on the first roll but not on the second?

A bag has 5 black marbles, 4 white marbles, and 1 green marble. If we reach into the bag and draw out 2 marbles, what is the chance that have the same color as each other?

A locus has two alleles Q and q. If allele q has a frequency of 0.1 in the population, and we assume random mating (maternal and paternal alleles are randomly drawn from a gene pool), What is the probability that an individual will be heterozygous?

The ABO blood type frequencies in Japan are $\Pr(A) = .279$, $\Pr(B) = 0.172$, and $\Pr(O) = 0.549$. Recall that the O-type is recessive (the genotype is OO). We randomly select a two-child family and assay the first child. If that child is type O, what is the probability that the second child in the family is type O?

A pollinator requires the nectar tube of a flower to be $\leq$ 1.2cm. Consider two populations of flowers. One has a median nectar tube length of 1.2cm. In the other population nectar tube lengths follow the normal distribution with the mean length of 1.4cm and a standard deviation of 0.1cm. If we draw one individual from each population, what is the probability that the will both meet the pollinator's requirments?

Ability to differentiate red and green is conferred by the dominant R allele. The trait is sex-linked (carried on the X-chromosome).

- the sex-ratio at birth for humans is roughly equal,
- In the population at large the frequency the r allele is 10%,
- Male children get their X chromosome from their mother,

A woman has genotype is Rr for red-green color blindness. What is the chance that the child will be color blind?

#20 from your book. In Vancouver, the probability of rain during a winter day is 0.58, for a spring day is 0.38, for a summer day is 0.25, and for a fall day is 0.53. Each of the season lasts one quarter of the year.

a. What is the probability of rain on a randomly chosen day?

b. If you were told it was raining on a day, what is the probability that the day would be winter?

Winning a match in tennis means winning two out of three sets of games. A study of tennis indicates that "momentum" appears to affect the outcome; even when two evenly matched players compete, the winner of the previous set in that match has a 60% chance of winning the next set. Player A and player B are evenly matched. What is the probability of a come from behind victory in a match between these players?

Lecture #7
probability densities
probability trees

Start of Chapter #6: Hypothesis testing