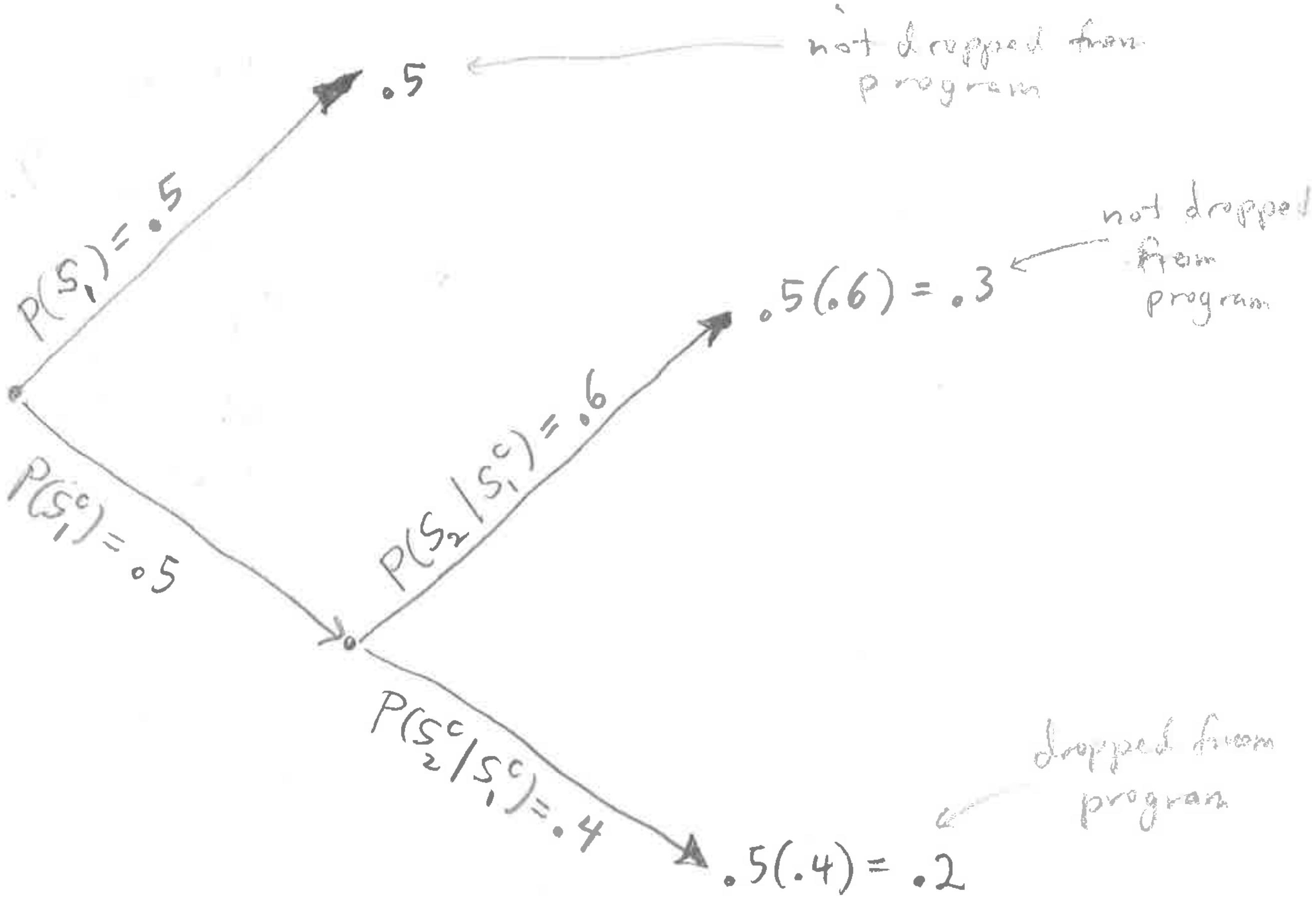
Example 50% of 1st-year Ph.D pass a qualifying exam on their first try. Of those that don't pass on the first try, 60% pass on their second try. If students don't pass by their second try, they are dropped from the program.

Compute the probability that a 1st-year Ph.D student will not be dropped from the program.

Solution. Let S_i be the event a 1st-year Ph.D student passes on their i^{th} try. $(i=1,2)$



$P(\text{not dropped from program}) = .5+.3=.8$

Also $= 1-P(S_1^c \cap S_2^c) = 1-.2 = .8$

The Method of randomized response (Warner, 1965)

A researcher wants to estimate the proportion of a population with a certain property, say, for example, would (honestly) answer Yes to a (possibly) sensitive question.

There are two implementations of this method (typically).

Implementation #1

Give each person in a representative sample of the population a balanced coin. Ask them to flip the coin (privately). Then tell them to do the following:

1. If they tossed heads, answer the question honestly.
2. If they tossed tails, just respond Yes (regardless of their actual answer).

What the experimenter will observe is a sequence of Yes, No responses. If the experimenter hears a Yes, then they don't know if this is an honest answer to the question or if they just flipped a tail.

On the otherhand, this implementation has the feature that when the experimenter hears a No, then they know that this is an honest answer to the Question.

The probability idea works like this:

Let Y be the event that the experimenter hears a Yes. Then

$$Y = (Y \cap \{\text{Head tossed}\}) \cup (Y \cap \{\text{tail tossed}\})$$

Consequently

$$P(Y) = P(Y \cap \{\text{Head tossed}\}) + P(Y \cap \{\text{tail tossed}\})$$

$$= P(\text{Head tossed}) P(Y|\text{Head tossed}) + P(\text{tail tossed}) P(Y|\text{tail tossed})$$

this is what the experimenter wants to know

$= 1$ by the design

$$P(Y) = \frac{1}{2} P(Y|\text{Head}) + \frac{1}{2} \cdot 1$$

Now we estimated $P(Y)$ by the proportion of our sample that answered Yes. Call this quantity p.

$$p \approx \frac{1}{2} P(Y|\text{Head}) + \frac{1}{2} \implies P(Y|\text{Head}) \approx 2p-1.$$

Implementation #2

This implementation requires that randomization to be a bit biased (so not 50-50) Say, with prob. p they are asked to answer the sensitive question and, therefore, with probability $1-p$ that are asked to answer the "complementary" question

For example if the sensitive question is

Q_s : "Have you ever cheated on an exam?"

the complementary question is

Q_s^c : "Have you NEVER cheated on an exam?"

The point is: if you answer YES to one of these then this is the same as answering No to the other.

Now,

$$P(Y) = P(Y \cap Q_s) + P(Y \cap Q_s^c)$$

$$= P(Q_s) P(Y|Q_s) + P(Q_s^c) P(Y|Q_s^c)$$

$$= p\, P(Y|Q_s) + (1-p) P(N|Q_s)$$

$$P(Y) = p\, P(Y|Q_s) + (1-p)(1 - P(Y|Q_s))$$

Solve for $P(Y|Q_s)$:

$$P(Y|Q_s) = \frac{P(Y) - (1-p)}{2p-1} \longleftarrow \text{requires } p \neq \tfrac{1}{2}.$$

We saw in a few examples where it helped to decompose an event into disjoint events. For example

$$(*) \qquad B_2 = (B_2 \cap B_1) \cup (B_2 \cap B_1^c).$$

helped us compute the probability of getting a blue marble on the second draw.

This is just a special case of a more general result called the <u>Law of total probability</u>.

called a partition

Suppose we can decompose a sample space Ω into disjoint events $B_1, B_2, \ldots, B_m$ each having positive probability. Then if $A \subseteq \Omega$ is any event, then

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_m)$$

$$= P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \cdots + P(B_m)P(A|B_m)$$

We've already used this formula. But, let's do some more examples to demonstrate its usefulness.

Example Roll a balanced 6-sided die twice (36 possible outcomes). Then it is easy to see that the event F = sum of upfaces is 5 has

$$P(F) = \frac{4}{36} = \frac{1}{9}.$$

This is because it is easy to enumerate the subset of Ω whose sum is 5:

$$F = \{(1,4), (2,3), (3,2), (4,1)\}$$

So $P(S) = \frac{|F|}{|\Omega|} = \frac{4}{36}$ ✓

Here's another way using the Law of total probability. Let B_i be the event that the first roll shows i (for $i = 1,2,3,4,5,6$). Then $B_1, B_2, B_3, B_4, B_5, B_6$ forms a partition and $P(B_i) = \frac{1}{6}$ for all i.

With this choice

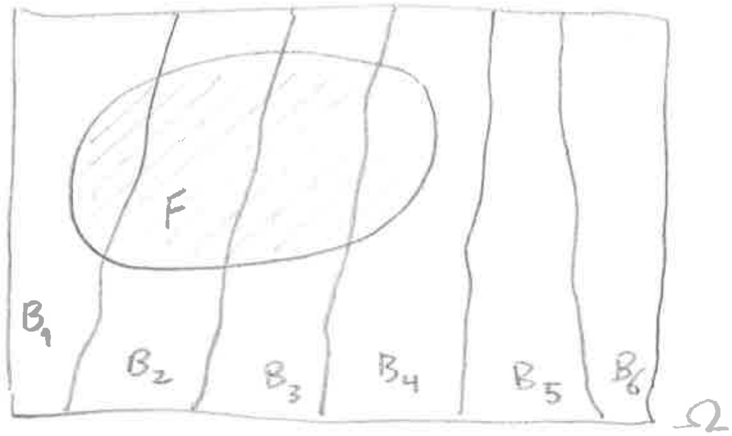
$$P(F|B_1) = \frac{1}{6} \quad P(F|B_2) = \frac{1}{6} \quad P(F|B_3) = \frac{1}{6} \quad P(F|B_4) = \frac{1}{6}$$

$$P(F|B_5) = 0 \quad P(F|B_6) = 0.$$

So all together

$$P(F) = \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot\frac{1}{6} + \frac{1}{6}\cdot 0 + \frac{1}{6}\cdot 0 = \frac{4}{36}$$ ✓

A Venn diagram picture of the previous example



The event F is shaded, $B_1, B_2, \ldots, B_6$ forms a partition of Ω.

Notice $F \cap B_5 = \emptyset$ and $F \cap B_6 = \emptyset$ in this example.

Example Two people play the following game – call them Alan and Betty. They each toss a balanced coin. Alan tosses 3 times, but Betty only twice.

Declare Alan the winner if Alan tosses more heads than Betty. If not, then declare Betty the winner.

Compute the probability that Alan wins.

Solution

Let A be the event that Alan wins. (A^c is the event that Betty wins.)

4 possible outcomes for Betty TT, TH, HT, HH (all equally-likely)

8 possible outcomes for Alan: TTT, TTH, THT, THH, HTT, HTH, HHT, HHH

also all equally-likely.

Obviously the event that Alan wins depends on what Betty tosses. But more importantly if we knew what Betty tossed, it would be easy to compute the probability that Alan wins.

To this end, let

B_0 = event Betty tosses 0 heads

B_1 = " " 1 head

B_2 = " " 2 heads

Then this forms a partition of Ω.

Moreover, $P(B_0) = \frac{1}{4}$ $P(B_1) = \frac{1}{2}$ and $P(B_2) = \frac{1}{4}$.

Now

$$P(A) = P(B_0)P(A|B_0) + P(B_1)P(A|B_1) + P(B_2)P(A|B_2)$$

$$= \frac{1}{4} \cdot \frac{7}{8} + \frac{1}{2} \cdot \frac{4}{8} + \frac{1}{4} \cdot \frac{1}{8}$$

$$= \frac{1}{2}.$$

$P(A^c) = 1 - \frac{1}{2} = \frac{1}{2}$.

Alan and Betty have the same chance of winning!

$\blacksquare$

Recap of the Law of total probability.

If A is an event and $B_1, B_2, \dots, B_m$ is any partition of Ω, then we can compute $P(A)$ as

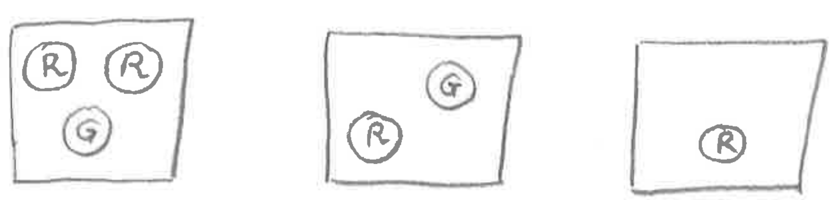
$$(*) \qquad P(A) = \sum_{j=1}^{m} P(A \cap B_j) = \sum_{j=1}^{m} P(B_j) P(A \mid B_j).$$

The formula $(*)$ is called the Law of total probability and is especially useful to compute probabilities of events A from experiments that can be thought of as sequential (i.e., experiments performed in stages)

However, this is not the only situation it is useful.

Consider the following example

Three boxes of Marbles in the following configurations:



Box 1	Box 2	Box 3
↑	↑	↑
2 red, 1 green	one red, one green	only one red

The experiment is to pick one of the 3 Boxes at random (equally-likely — i.e., pick Box i with prob. $= 1/3$ for each $i = 1, 2, 3$). Then pick one marble (equally-likely) at random from that box.

(Notice this experiment can be thought of a sequential:

$$\underbrace{\text{Pick a Box}}_{\text{Stage 1}} \xrightarrow{\text{then}} \underbrace{\text{Pick a marble}}_{\text{Stage 2}}$$
)

Let's define the event R to be the event that a red marble is drawn. Compute $P(R)$.

Let B_i be the event that Box i was selected $(i = 1, 2, 3)$

Then B_1, B_2, B_3 is a partition.

By the law of total probability ....

$$P(R) = P(B_1)P(R|B_1) + P(B_2)P(R|B_2) + P(B_3)P(R|B_3)$$

$$= \frac{1}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1$$

$$= \frac{2}{9} + \frac{1}{6} + \frac{1}{3} = \frac{4}{18} + \frac{3}{18} + \frac{6}{18} = \frac{13}{18}.$$

Remark In this example, notice that $P(R) \neq \frac{4}{6} = \frac{2}{3}$

Naïvely one might think $P(R) = \frac{4}{6}$ since there are a total of 6 marbles (in all 3 boxes combined) of which 4 are red.

However, a probability of $\frac{4}{6}$ in not correct for our situation. since if all 6 marbles we combined into one Box and one is selected at random, then although the Prob. of getting a red is $\frac{4}{6}$ there is a better than $\frac{1}{3}$ chance that the marble you select comes from Box 1 because Box 1 has 3 marbles (half of the 6 total).

From the last example we showed

$$P(R) = P(R \cap B_1) + P(R \cap B_2) + P(R \cap B_3)$$

$$= \frac{4}{18} + \frac{3}{18} + \frac{6}{18}$$

Now, a new question: If we are told a Red marble is drawn, can we compute the probability that it came from a particular Box ?

For example, let's compute $P(B_1|R)$.

$$P(B_1|R) = \frac{P(B_1 \cap R)}{P(R)} = \frac{P(B_1 \cap R)}{P(B_1 \cap R) + P(B_2 \cap R) + P(B_3 \cap R)}$$

$$= \frac{\frac{4}{18}}{\frac{4}{18} + \frac{3}{18} + \frac{6}{18}} = \frac{4}{13}.$$

Similarly,

$$P(B_2|R) = \frac{\frac{3}{18}}{\frac{4}{18} + \frac{3}{18} + \frac{6}{18}} = \frac{3}{13}$$

$$P(B_3|R) = \frac{\frac{6}{18}}{\frac{4}{18} + \frac{3}{18} + \frac{6}{18}} = \frac{6}{13}.$$

So that given knowledge of the color drawn, we can update the probabilities on B_1, B_2, B_3.

This demonstrates an application of the Bayes' rule.

$\boxtimes$

Bayes' rule

Let A be any event and $B_1, B_2, \ldots, B_m$ a partition. Then for each $j = 1, 2, \ldots, m$.

$$P(B_j \mid A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(B_j)P(A \mid B_j)}{\sum_{i=1}^{m} P(B_i)P(A \mid B_i)}$$

Example A company buys resistors in bulk from two sources: S_1 and S_2. 5% of resistors from S_1 are defective, while only 2% of resistors from S_2 are defective. Due to price and competitiveness the company buys 70% of their resistors from source S_1 and 30% from source S_2.

(a) If a resistor is randomly selected from their supply what is the probability it is defective?

(b) If a resistor is found to be defective, what's the probability it came from source S_1?

Told

$P(D|S_1) = .05 \quad P(D|S_2) = .02$

$P(S_1) = .70 \quad P(S_2) = .30$

(a) $P(D) = P(S_1)P(D|S_1) + P(S_2)P(D|S_2)$

$= .70(.05) + .30(.02)$

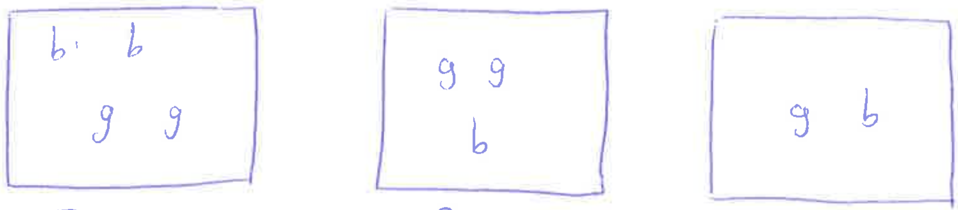
$= .035 + .006$

$= .041$

(b) $P(S_1|D) = \dfrac{P(S_1)P(D|S_1)}{P(D)} = \dfrac{.70(.05)}{.041} = \dfrac{.035}{.041}$

$\cong .853659\ldots$

__Example__ Three boxes of marbles:



Box 1 has 2 blue, 2 green

Box 2 has 1 blue, 2 green

Box 3 has 1 blue, 1 green

The experiment is to randomly (equally-likely) select a marble from box 1, transfer this marble to box 2; then randomly select a marble from box 2, transfer the marble to box 3; then finally randomly draw a marble from box 3.

What is the probability that a blue marble is selected in the final draw?

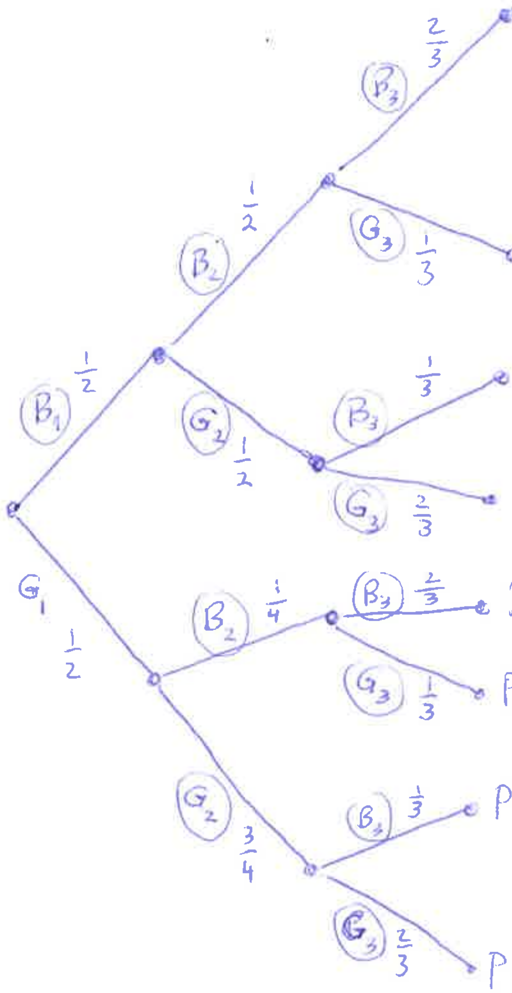
To fix notation, let's set B_i to be the event that a blue marble is drawn in the i-th stage $(i = 1, 2, 3)$.

Similarly, $G_i = B_i^c$ means we drew a green marble in i-th stage.

Then with this notation we are trying to compute

$$P(B_3).$$

Here's a graphical depiction of this problem:



$$P(B_1 \cap B_2 \cap B_3) = P(B_1)P(B_2 | B_1)P(B_2 | B_1 \cap B_2) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{6}$$

$$P(B_1 \cap B_2 \cap G_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12}$$

$$P(B_1 \cap G_2 \cap B_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{12}$$

$$P(B_1 \cap G_2 \cap G_3) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{6}$$

$$P(G_1 \cap B_2 \cap B_3) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{2}{3} = \frac{1}{12}$$

$$P(G_1 \cap B_2 \cap G_3) = \frac{1}{2} \cdot \frac{1}{4} \cdot \frac{1}{3} = \frac{1}{24}$$

$$P(G_1 \cap G_2 \cap B_3) = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{3} = \frac{1}{8}$$

$$P(G_1 \cap G_2 \cap G_3) = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{4}$$

To compute $P(B_3)$, for instance, we just need to sum all the probabilities in the above graph that lead to the occurrence of B_3:

$$\frac{1}{6} + \frac{1}{12} + \frac{1}{12} + \frac{1}{8} = \frac{11}{24} = P(B_3)$$

In fact, the graph above can be used to compute all probabilities of events that can be observed from this experiment ....

For example, $P(G_2)$ can be computed by summing all the probabilities in the leaves of the tree that contain G_2:

$$P(G_2) = \frac{1}{12} + \frac{1}{6} + \frac{1}{8} + \frac{1}{4} = \frac{15}{24}$$

By the way, we would immediately have that $P(B_2) = 1 - \frac{15}{24} = \frac{11}{24}$.

(This is coincidence here.)

How about $P(G_1 \cap B_3)$?

We would sum all the probabilities in the leaves of the tree that contain $G_1 \cap B_3$:

$$P(G_1 \cap B_3) = \frac{1}{12} + \frac{1}{8} = \frac{5}{24},$$

Etc.

Now, a Bayes' rule type question? If a Blue marble was drawn from Box 3 (in this experiment), what is the probability that a Green marble was drawn from Box 1?

Namely, compute $P(G_1 \mid B_3)$.

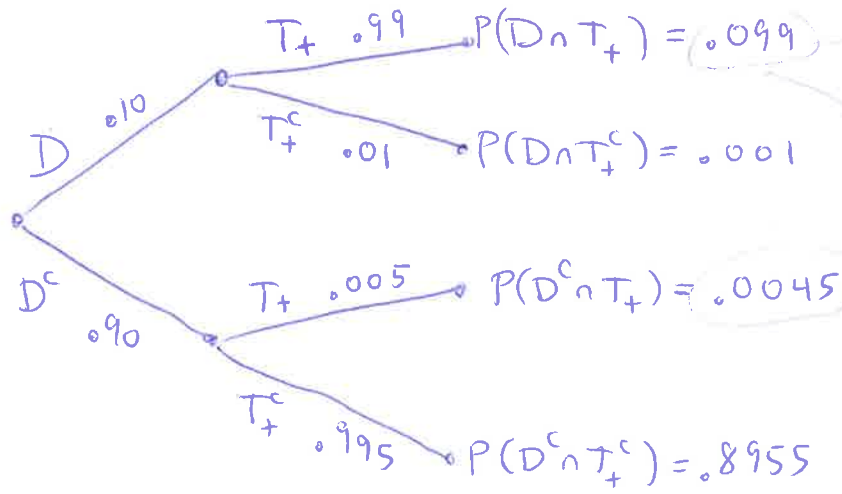
Now, A less involved example of the graphical approach....

<u>Example</u> Imagine that in a certain population, 10% are diseased. A company markets a kit to test for this disease. The company quotes that if a person is diseased, then the test will return a positive for the disease with probability 99%. On the other hand if a person is non-diseased, then the kit will test positive for the disease with probability .005 (.5%). The question is: If you randomly select a person and they test positive for the disease, what is the probability they are diseased?

D = event diseased $\quad T_+$ = test positive

D^c = not diseased $\quad T_+^c$ = test negative

$$P(D|T_+) = \frac{P(D \cap T_+)}{P(T_+)}$$



$$= \frac{.099}{.099 + .0045}$$

$$\approx .9565$$

Check that $P(D^c|T_+^c) \approx .998885...$

See next Page for this and other calculations.

$$P(D^c \mid T_+) = \frac{P(D^c)P(T_+ \mid D^c)}{P(D^c)P(T_+ \mid D^c) + P(D)P(T_+ \mid D)}$$

$$= \frac{.9(.005)}{.9(.005) + .1(.99)} \approx .043478 \quad \left(\text{false positive rate}\right)$$

$$P(D \mid T_+^c) = \frac{P(D)P(T_+^c \mid D)}{P(D)P(T_+^c \mid D) + P(D^c)P(T_+^c \mid D^c)}$$

$$= \frac{.1(.01)}{.1(.01) + .9(.995)} \approx .001115 \quad \left(\text{false negative rate}\right)$$

$$P(D^c \mid T_+^c) = \frac{.9(.995)}{.9(.995) + .1(.01)} \approx .998885$$

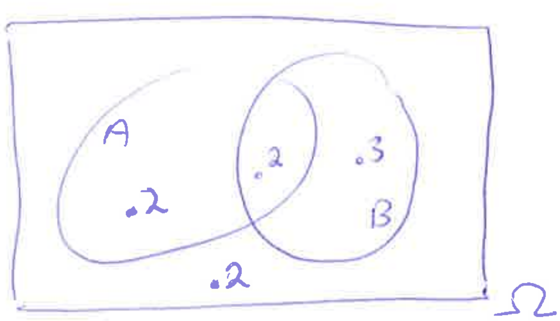
Independence (of two events)

We define two events A, B in a sample space Ω to be Independent to mean

$$(*) \quad \boxed{P(A \cap B) = P(A) \cdot P(B)}$$

Furthermore, if A and B are events that satisfy condition $(*)$ above, then we say A and B are (statistically) independent.

Consider the Venn diagrams below.



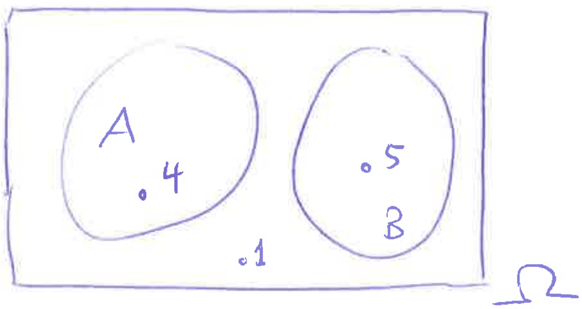
In this picture the numbers appearing in each region are the corresponding probabilities of those regions. For instance $P(A^c \cap B) = .3$ and $P(A \cap B) = .2$

Notice that from the Venn diagram we can read off that

$P(A) = .2 + .2 = .4$ and $P(B) = .2 + .3 = .5$. But, also, $P(A \cap B) = .2$ and in this example A, B are statistically independent:

$$.2 = P(A \cap B) = P(A)P(B) = .4(.5) = .2 .$$

Compare the above situation to this one (where A, B are disjoint)



Here, we see plainly that $P(A) = .4$ and $P(B) = .5$

But $P(A \cap B) = 0$ since A and B are disjoint.

Therefore,

$$0 = P(A \cap B) \neq P(A)P(B) = .4(.5) = .2.$$

and thus A and B are <u>NOT</u> independent.

The reason events A and B satisfying (*) are called independent follows from the result that A, B would satisfy

$$(\star\star) \qquad P(A|B) = P(A) \quad \text{if } P(B) > 0.$$

since $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ if $P(B>0)$ and if

$\dfrac{P(A \cap B)}{P(B)} = P(A)$ we would then have condition (*).

Condition ($\star\star$) says that if A, B are independent, then Given B occurs the probability ~~doesn't~~ of A doesn't change.