Recitation Assignment 11 - 17 - 09

This assignment pertains to chapter 15 and focuses on:

- Rules of probability, conditional probability
- Addition rule in Venn diagram, $\bigcup_{or}$, $\bigcap_{and}$
- Appearance of rules in tables
- Appearance of rules in trees
- Independence vs dependence of events
- Appearance of independence in tables
- Appearance of independence in trees
- Trees as an aid to reliable calculation
- Bayes' formula (formula aside, it is just the rules in action)

You will also see some classical models in play.

1. An aircraft has two engines. The left engine has probability 0.0001 of failing over the course of a flight. The right engine has probability 0.0002 of failing over the course of a flight. Failure of both engines will be catastrophic. What is the probability that at least one engine will not fail?

This cannot be answered using the information given. The reason is that there are *four* regions of the Venn Diagram and we know only *three* things:

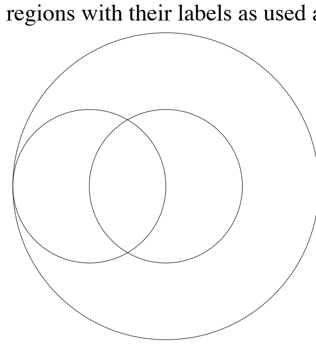
0.0001 = P(A) = P(left engine fails)
0.0002 = P(B) = P(right engine fails)
sum of probabilities of the four regions = 1

↦ we need the sum of these three ↤

$A\bigcap_{and}B$	$A\bigcap\overline{B}$	$\overline{A}\bigcap B$	$\overline{A}\bigcap\overline{B}$
both fail	only A fails	only B fails	neither fails

(we know only that)
↦ sum of these two is 0.0001 ↤
↖ sum of these two is **0.0002** ↗
↦ sum of all four probabilities is 1 ↤

a. Fill out a Venn Diagram, labeling A, B, and identify the four regions with their labels as used above.

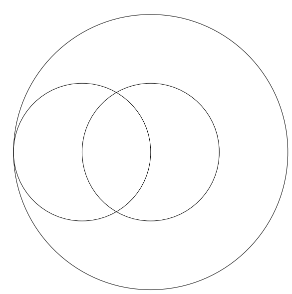


b. **What if A, B are independent events?** We say that events A, B are *statistically independent* if the conditional probability $P(B|_{if}A)$ (that is, the *conditional probability* the right engine fails if the left engine is known to have failed) remains at P(B). If engine failures are independent the probability P(B) of failure of the right engine is not changed if we know that the left engine has failed. **Suppose this is true in the present case.** We now have one more piece of information for a total of four.

Use the **multiplication rule $P(A\bigcap B) = P(A)\ P(B\ |\ A)$** to determine $P(A\bigcap B)$, the probability that both engines will fail. Place this value in the Venn diagram and solve for each of the three remaining probabilities of the regions of the diagram.

$P(A\bigcap\overline{B}) = P(B) - P(A\bigcap B) =$
$P(\overline{A}\bigcap B) =$
$P(\overline{A}\bigcap\overline{B}) = 1$ - sum of the other three probabilities

Fill out the diagram, labeling all four events and indicating their probabilities in the diagram.



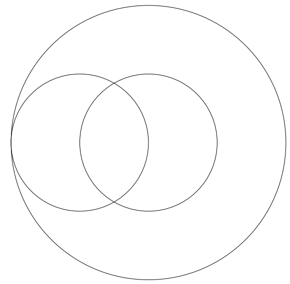
c. Once you have the probabilities of each of the four regions (from (b)) you can use the display above (a) to calculate P(at least one engine does not fail). Shade the relevant region in the diagram.

d. There is another way to determine the probability that at least one engine does not fail. From the display above (a) see that the probability at least one engine does not fail is just 1 - P(both fail) = $1 - P(A\bigcap B)$. Obtain by this means the probability that at least one engine does not fail.

2. An aircraft has two engines, continued. As in #1 we will assume P(A) = 0.0001 and P(B) = 0.0002. But this time we will not assume engine failures are independent events. Engine failures are often caused by flocks of birds at airports. If an engine fails it is likely due to this cause which means that failure of the left engine should raise the probability for failure of the right engine. Let us suppose that he probability of right engine failure rises from P(B) = 0.0002 to P(B | A) = 0.07 when A occurs.

a. Using the multiplication rule $P(A \cap B) = P(A)\ P(B \mid A)$ determine the probability both engines fail in this dependent setup.

b. Re-do the entire Venn Diagram (the probabilities will differ what they were in the independent case).



c. From (a) determine the probability that at least one engine will not fail in this dependent setup. Is it larger or smaller than when the engine failures were independent?

d. From the completed Venn Diagram determine
P(left engine fails | right engine fails)
= $P(A \mid B) = P(A \cap B)\ /\ P(B)$.

e. If one engine were to fail, which would concern you most, failure of the left engine or failure of the right engine? Consult P(B | A) versus P(A | B).

3. A box has colored balls [3R 4G 5Y]. We will sample three balls without replacement and with equal probability on those then remaining in the box.

a. What is P(R1)?

b. What are
$P(R2 \mid R1) =$ (from [2R 4G 5Y])
$P(R2 \mid \overline{R1}) =$ (from [3R 8other])

b. What is P(R2)? We may get it using the rules:
$P(R2) = P(R1\ R2) + P(\overline{R1}\ R2) - 0$ (addition rule)
$P(R1\ R2) = P(R1)\ P(R2 \mid R1)$ (multiplication rule)
$P(\overline{R1}\ R2) = P(\overline{R1})\ P(R2 \mid \overline{R1})$ (multiplication rule)
Make the calculations and see that you get P(R2) = P(R1) (establishing that "order of the deal does not matter) even though draws are *without replacement*.

c. Use the *definition of conditional probability* to find P(R1 | R2) = P(R1 R2) / P(R2) (its just the multiplication rule $P(A \cap B) = P(A)\ P(B \mid A)$ taking A = R2, B = R1).

Calculation (c) above is an example of *Bayes' Formula* except that we've not even had to see the formula. What has happened is that we've determined the probability of the "earlier" draw having been red based on our observation that the "later" draw is red. We've reasoned from observation back to an underlying cause. Bayes (see chapter 15) wrote out the calculations in detail and sensed that he was onto something. It is greatly important that you too realize the power of his seemingly innocuous calculation for (although you cannot be expected to see it) this idea establishes a type of "logical" engine that can place "informed by experts" probabilities on things not known and update these probabilities in the light of new evidence. This is called the "Bayesian Method." I have seen it used to restore images from the faulty Hubble Telescope (UNC 1990) (the things not known being the true star fields as they should be imaged) and to determine how many extremely expensive older rockets to test (to destruction) in order to establish reliability of the fleet, the things not known being which rockets in the fleet are faulty).*

*Experts are consulted "informing us" as to the probabilities they might place on different forms of true star fields being the actual ones, or different patterns of fault that might be present among the fleet of rockets. Massive math models and calculations are then brought to bear which calculate the "updated" probabilities of different images being the correct ones conditional on what Hubble reports. We then pick the conditionally most probable true star field and print it out. It is involved but hey, we have computers. Guess what? They fixed the Hubble and those early restorations (from when it was faulty) were really good.

4. Weather. A forecast says there is 80% chance of rain Sat and 60% chance of rain Sun.

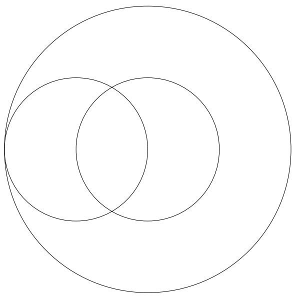
a. Do you believe these events are *independent* (assume East Lansing is the locale)?

b. If they *were* independent you could find

P(rain both days) (decimal) =

P(no rain on the weekend) =

What the heck, do the whole Venn:



c. If the events are not independent but instead

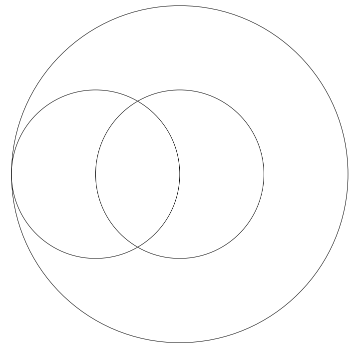
P(rain Sun | rain Sat) = **70%**

find

P(rain both days) (decimal) =

P(no rain on the weekend) =

What the heck, re-do the whole Venn:



d. I arrive from out of town to find rain Sunday. What is the conditional probability P(rain Sat | rain Sun)?

5. The Tree and Bayes' Method. A potential oilfield is thought to have probability .3 of having oil in commercially viable amount. Call this event OIL. A test can be performed (involves seismic readings etc.) and the following claims are made for the response of this test to OIL, or "no OIL" (i.e. $\overline{\text{OIL}}$, sometimes denoted OIL^c):

P(- | OIL) = .17 (17% false negatives)

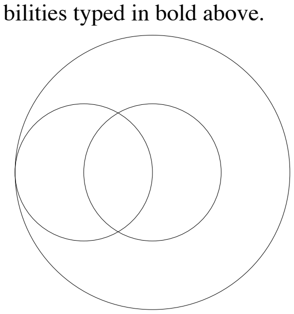
P(+ | no OIL) = .08 (8% **false positives**)

a. Lay out a tree diagram for this information (see chapter 15) with branching

P(OIL) = .3
- P(+ | OIL) = .83 **P(OIL+) = .3 .83**
- P(- | OIL) = .17 **P(OIL-) =**

P(no OIL) = .7
- P(+ | no OIL) = .08 **P(noOIL+) =**
- P(- | no OIL) = .92 **P(noOIL-) =**

b. Fill out the complete Venn Diagram using the endpoint probabilities typed in bold above.



c. Read off the value P(+) = P(OIL+) + P((noOIL) +).

d. Use the *definition* P(OIL | +) = P(OIL+) / P(+) to find the *conditional probability* that OIL is present *if the test yields a positive result*. This is Bayes' idea.*

*You could use the formula in chapter 15 but can just as easily get it from the tree as we just did. Bayesians (people who specialize in the method) are much assisted by the formula in the many derivations and manipulations they have to make, but we don't do that here.

e. Find P(OIL | -). Does it make sense that it should be less than P(OIL | +)?

f. Contemplate the fact that we might have to consult expert opinion in order to come up with numbers for P(OIL), P(- | OIL), P(+ | no OIL). That does not make them correct but it can be a very promising approach to making the decision as to whether or not we drill the field. It separates assumptions from logical deductions flowing from them, focusing attention on getting the best judgment possible on each input to the calculation. Repeated in many drilling operations this can be a more perfectible process than "seat of the pants."

```
In[1]:= .3 × .17 / (.3 × .17 + .7 × .92)

Out[1]= 0.07338129496402877`
```