MATH 115
LECTURE 15

: Conditional Probability and Independence.

1| Conditional probability

Example: In a group of 30 athletes, 18 are women, 12 are swimmers and 10 are neither. A person is chosen at random. What is the probability that it is a female swimmer?

Solution: Events → $S = \{\text{being a swimmer}\}$, $W = \{\text{being a woman}\}$ - $P(W \cap S)$?

We know that

$$P(W) = \frac{18}{30}, \quad P(S) = \frac{12}{30}, \quad P((W \cup S)^c) = \frac{10}{30}.$$

Therefore,

$$P(W \cap S) = P(W) + P(S) - P(W \cup S) =$$

$$= \frac{18}{30} + \frac{12}{30} - \left(1 - \frac{10}{30}\right) = \frac{10}{30} = 1/3,$$

where we have used the inclusion-exclusion principle.

- Now, suppose that we choose a woman. Knowing this, what is the probability that she is a swimmer?

In this second case, the sample space has women. So the result is

$$P(S|W) = \frac{\# \text{ swimmers that are women}}{\# \text{ women}} = \frac{10}{18} = \frac{5}{9}.$$

This is read as "probability that the event S ocurrs knowing that W has ocurred",

or simply "probability of S conditioned to W".

• Def: The conditional probability of E given F, denoted $P(E|F)$, is defined by

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

↳ Particular case. If S is an equiprobable space, that is, if all the events have the same probability, then

$$P(E|F) = \frac{n(E\cap F)}{n(F)}$$ (this is what we did in the example of swimmers).

It follows from the general definition:

$$P(E\cap F) = \frac{n(E\cap F)}{n(S)},\ P(F) = \frac{n(F)}{n(S)} \Rightarrow P(E|F) = \frac{n(E\cap F)}{\cancel{n(S)}}\,\frac{\cancel{n(S)}}{n(F)} = \frac{n(E\cap F)}{n(F)}.$$

- A "trivial" consequence is the product rule (not so trivial in terms of its interpretation):

<u>Product rule:</u> $\boxed{P(E \cap F) = P(F)P(E|F)}$

<u>Ex</u>: One can check in the previous example that

$$\underbrace{P(W \cap S) = \frac{1}{3}}_{\text{Prob. of being women and swimmer}} = \underbrace{P(W)}_{\text{Prob. of being woman}} \times \underbrace{P(S|W)}_{\text{Prob. of being swimmer if the person is a woman.}} = \frac{18}{30} \cdot \frac{5}{9}.$$

- <u>Problem</u>: Two students are chosen, one after the other, from a group of 50 students, 20 of which are freshmen and 30 sophomores.

a) Probability that the 1^{st} is a freshman and the 2^{nd} a sophomore?

b) If three are chosen, what is the probability that the first is a sophomore and the next two freshmen?

<u>Solution</u>:

a) Let's do it as if we didn't know about conditional probability first.

Let's consider the event "1^{st} is a freshman and 2^{nd} a sophomore", and call this event A.

Then, $P(A) = \dfrac{\#\text{ of ways of choosing 1}^{st}\text{ freshman, 2}^{nd}\text{ sophomore}}{\#\text{ of ways of choosing two people}} = \dfrac{\# n(A)}{n(S)}$.

50 students $\begin{cases} 20 \text{ freshmen} \\ 30 \text{ soph.} \end{cases}$

$n(S) = 50 \cdot 49$, $n(A) = 20 \cdot 30 \Rightarrow P(A) = \dfrac{20 \cdot 30}{50 \cdot 49}$.

→ Let's do it now using the multiplication principle:

Denote $E = \{ 1^{st} \text{ is a freshman} \}$,

$F = \{ 2^{nd} \text{ is a sophomore} \}$.

Then,

$$P(E \cap F) = P(E) \cdot P(F | E), \text{ and clearly}$$

$$P(E) = \frac{20}{50},$$

$$P(F | E) = \frac{30}{49}.$$

· <u>Problem</u>: (For home)

A lot contains 12 items of which 4 are defective. Three items are drawn at random from the lot one after the other. Find the probability that all 3 are nondefective.

↳ Recommended: Try doing it using the multiplication principle.

2) Independence

• **Def**: Two events are independent if $P(E \cap F) = P(E)P(F)$.

Notice that this implies that $P(E|F) = P(E)$, that is, it doesn't matter what happens with F, it doesn't affect to E, they are "independent".

↳ **Remark**: It also implies that $P(F|E) = P(F)$.

→ The definition is important because it isn't always clear when two events are independent:

Examples:

1) ("a clear one") Roll a die twice. Let E be "get a 1 on 1st roll", F be "get a 3 on 2nd roll".

Are E, F independent?

Sol: The intuition says they are independent. Indeed,

$$P(E) = \frac{1}{6}, \quad P(F) = \frac{1}{6}, \quad P(E \cap F) = \frac{1}{36}$$

here $S = \{1, 2, 3, 4, 5, 6\}$
$n(S) = 6$

here $S = \{(1,1), (1,2), \boxed{(1,3)}, \ldots$
$(2,1), (2,2), (2,3), \ldots$
$\vdots$
$(6,1), (6,2), (6,3), \ldots \}$, $n(S) = 36$.

2) A card is to be drawn from a full deck. Let the events $E =$ "the card is a 4", $F =$ "the card is a spade".

2.1) Are E, F independent?

<u>Sol</u>: Yes, $P(E) = \frac{4}{52}$, $P(F) = \frac{13}{52}$. $P(E \cap F) = \frac{1}{52}$.

So $P(E)P(F) = \frac{4 \cdot 13}{52^2} = \frac{\cancel{52}}{52^{\cancel{2}}} = \frac{1}{52} = P(E \cap F)$ //.

2.1) Are E, F independent if the original deck was missing the 7 of clubs?

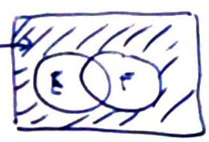
<u>Sol</u>: Let's see. $P(E) = \frac{4}{51}$, $P(F) = \frac{13}{51}$, $P(E \cap F) = \frac{1}{51}$.

But $P(E)P(F) = \frac{4 \cdot 13}{51^2} = \frac{52}{51}\frac{1}{51} \neq \frac{1}{51} = P(E \cap F) \Rightarrow$

$\Rightarrow E, F$ are not independent in this case.

• <u>Exercise</u>: Show that if E, F are independent, then so are E^c and F^c. Also E and F^c.

$E^c \cap F^c$



<u>Sol</u>: $P(E^c \cap F^c) = 1 - P((E^c \cap F^c)^c) = 1 - P(E \cup F) =$

$= 1 - P(E) - P(F) + P(E \cap F) = 1 - P(E) - P(F) + P(E)P(F) =$

$= 1 - P(E)(1 - P(F)) - P(F) = (1 - P(F))(1 - P(E)) = P(F^c)P(E^c)$ //.

$P(E \cap F^c) = \ldots$ (home).

• Def: A collection of events are independent if for any subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ we have

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \ldots P(A_{i_k}).$$

For example, E, F, G are independent if and only if

$$\left.\begin{array}{l} P(E \cap F) = P(E)P(F) \\ P(E \cap G) = P(E)P(G) \\ P(F \cap G) = P(F)P(G) \end{array}\right\} \text{These three alone are not enough.}$$

$$P(E \cap F \cap G) = P(E)P(F)P(G)$$

• Exercise: E, F, G are independent events with $P(E) = 5/10$, $P(F) = 4/10$, $P(G) = 3/10$. Find:

a) $P(E \cap F \cap G)$, b) $P(E \cap G^c)$, c) $P(E \cap (F \cup G)^c)$, d) $P(E \cup (F \cap G)^c)$

(For home)

3) Independent repeated trials.

We have already seen some examples of this: toss a coin 10 times.

It is convenient to define the following:

• **Def:** Let S be a finite probability space. The probability space of n independent trials, S_n, consists of ordered n-tuples of elements of S, with probability

$$P((s_1, s_2, \dots, s_n)) = P(s_1)\,P(s_2)\dots P(s_n).$$

• **Example:** (Important)

A machine produces defective items with probability p.

a) If 10 items are chosen at random, what is the probability that exactly 3 are defective?

b) What is the probability of finding at least one defective item in the 10 chosen?

c) If we observe the items one at a time as they come off the line, what is the probability that the 3rd defective item is the 10th item observed?

Solution:

a) Consider all the possible ways of obtaining exactly 3 defective items out of 10. Let's denote d = "defective", n = "non-defective".

$P(d) = p$, $P(n) = 1-p$.

Now, $P((d,d,d,n,n,n,n,n,n,n)) = p^3(1-p)^7$.

Notice that this is the same for all the cases with exactly 3 defective items (like (d,n,d,d,n,n,n,n,n,n)).

Now, the number of ways of obtaining exactly 3 def. items is

$$\frac{10!}{3!\,7!}$$

Therefore,

$$P(\text{"exactly 3 def. out of 10"}) = \frac{10!}{3!\,7!} \cdot p^3(1-p)^3 \equiv \binom{10}{3} p^3 (1-p)^7$$

Remark: This will appear again in Chapter 6 when we study the Binomial distribution.

(the last term should remind you of the binomial theorem).

b) $P(\text{"at least 1 def. in 10"}) = 1 - P(\text{"0 def. in 10"}) =$

$$= 1 - (1-p)^{10}$$

c) Here we want to find the probability of events like

$$\underbrace{\underline{d}\ \underline{d}\ \underline{n}\ \underline{n}\ \underline{n}\ \underline{n}\ \underline{n}\ \underline{n}\ \underline{n}}_{\text{here exactly two defect.}}\ \underline{d}$$

this has to be defective.

probability of these events is $p^3(1-p)^7$.

there are $\dfrac{9!}{2!\,7!}$ ways of this happening

Therefore,

$$P(\text{"3rd defect. is the 10th observed"}) = \frac{9!}{2!\,7!}\, p^3 (1-p)^7 .$$

notice that we obtained the same using the product rule:

$E =$ "10th item is defective"

$F =$ "exactly 2 defect. items in the first 9"

$$P(E \cap F) = P(E)\,P(F) = p \cdot \underbrace{\frac{9}{2!\,7!}\, p^2 (1-p)^7}_{\text{this is completely analogous to part a).}}$$

E, F are of course independent

4.) Finite stochastic processes and tree diagrams.

• Def: A finite stochastic process is a finite sequence of experiments where each experiment has a finite number of outcomes with given probabilities.

We use tree diagrams to describe them

• Example: A city of 100000 people is broken into 4 precincts of unequal size (P_1, P_2, P_3, P_4). Their populations are:

$P_1 \to 10000$, $P_2 \to 20000$, $P_3 \to 30000$, $P_4 \to 40000$.

A review of crimes recorded shows that:

- 20% of records in P_1 contain errors.
- 5% " " " P_2 " "
- 10% " " " P_3 " "
- 5% " " " P_4 " " .

a) Draw a tree diagram describing the results.

b) Probability that a record has an error and is in P_3?

c) Probability that a record has an error?

d) Probability that a record is in P_3 given that it has an error?

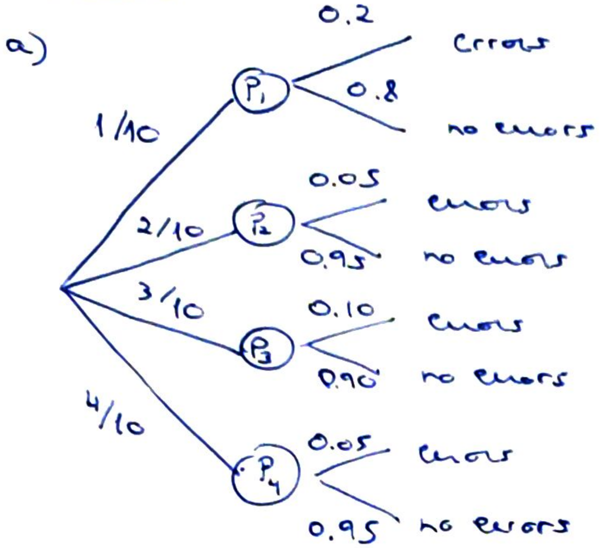
Solution:

These are conditional probabilities.

$P(E|P_1), P(NE|P_1), \ldots$

error — no errors

a)



b) We have to follow the path $P_3 \to$ errors:

$$P(P_3 \cap \text{error}) = P(P_3)P(E|P_3) = \frac{3}{10}\cdot 0.10 = \frac{3}{100} = 0.03$$

c) We have to add all the paths (branches) that lead to errors:

$$P(\text{errors}) = \frac{1}{10}\cdot(0.2) + \frac{2}{10}\cdot(0.05) + \frac{3}{10}\cdot(0.10) + \frac{4}{10}\cdot(0.05) = \frac{8}{100} = \frac{2}{25}$$

$\|$ Strictly, we are computing the following:

$$P(E) = P\big((P_1\cap E)\cup(P_2\cap E)\cup(P_3\cap E)\cup(P_4\cap E)\big)$$

Since P_1, P_2, P_3, P_4 are mutually disjoint, then $P_1\cap E, P_2\cap E, P_3\cap E, P_4\cap E$ are also disjoint, and thus the inclusion-exclusion principle gives that

$P(E) = P(P_1\cap E) + P(P_2\cap E) + \ldots + P(P_4\cap E)$, and for each one we can use the multiplication principle: $P(P_i \cap E) = P(P_i)P(E|P_i), \ldots$ $\blacksquare$

d) We want to compute $P(P_3|E)$.

We can use its definition:

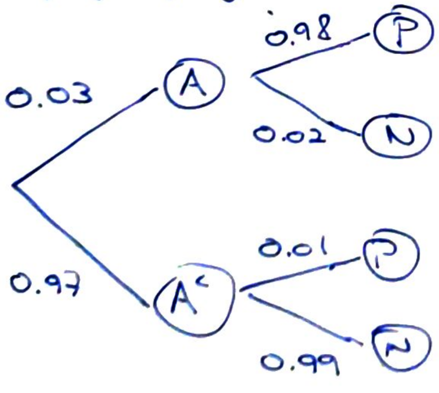
$$P(P_3|E) = \frac{P(P_3 \cap E)}{P(E)} = \frac{3/100}{2/25} = \frac{3}{8}.$$

• <u>Problem</u>: A test for a certain allergy tests positive 98% of the time if the person has that allergy, while it only tests positive 1% of the time if the person doesn't have it (false positives). Given that only 3% of people have this allergy, what is the probability that a patient has the allergy if it tests positive?

<u>Solution</u>: $A = \{\text{person is allergic}\}$, $A^c = \{\text{person is not allergic}\}$

$P = \{\text{tests positive}\}$, $N = \{\text{tests negative}\}$

$P(A|P)$?



Using the tree diagram we see that

$$P(A|P) = \frac{P(A \cap P)}{P(P)} = \frac{(0.03)\cdot(0.98)}{(0.03)\cdot(0.98) + (0.97)(0.01)}$$

• Problem: A crate of apples contains 3 bad apples and 7 good ones. Apples are chosen until we pick a good one. What is the probability that it takes at least 3 picks to get a good one?

Solution:

$$P(\text{"it takes at least 3 picks."}) = 1 - P(\text{takes 1 or 2 picks})$$

Let $A = \{\text{It takes 1 pick to get a good one}\}$

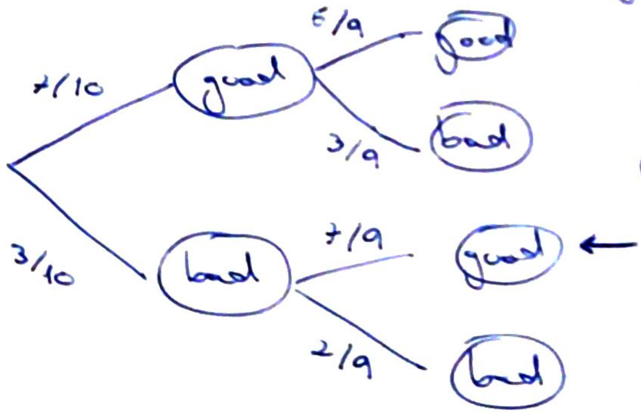
$B = \{\text{It takes 2 picks to get a good one}\}$.

Clearly these two event are disjoint: $P(A \cap B) = 0$ (both cannot happen at the same time).

Therefore,

$$P(\text{"it takes at least 3.."}) = 1 - P(A \cup B) = 1 - P(A) - P(B).$$

Now we can use a tree diagram:



Clearly, $P(A) = \frac{7}{10}$, while for B we have

$$P(B) = \frac{3}{10} \cdot \frac{7}{9}$$

[Remark: The tree diagram is not necessary, but it might help you]

5.1 Bayes' Theorem

The example of the previous section showed a fact that is true in general:

- **Theorem:** If $A_1, A_2, \dots, A_n$ are $\overbrace{\text{mutually exclusive}}^{P(A_i \cap A_j) = 0 \ (\text{for } i \neq j)}$ events whose union is the whole sample space, then for any event B we have that

$$P(A_i \mid B) = \frac{P(A_i)\, P(B \mid A_i)}{P(B)} = \frac{P(A_i)\, P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \dots + P(A_n)P(B \cap A_n)}$$

this is called Law of Total Probability.

- **Problem:** We have two coins. Coin 1 is fair while Coin 2 has two heads. We select a coin randomly and toss it. Say a head comes up.

a) What is the probability that is Coin 1?

b) Flip the coin again and say a head comes up again. What is the probability that it is Coin 1?

Solution: Let $A_1 = \{\text{Coin is Coin 1}\}$

$A_2 = \{\text{Coin is Coin 2}\}$

$H_1 = \{\text{a head comes up at } 1^{st} \text{ flip}\}$

$H_2 = \{\text{a head comes up at } 2^{nd} \text{ flip}\}$

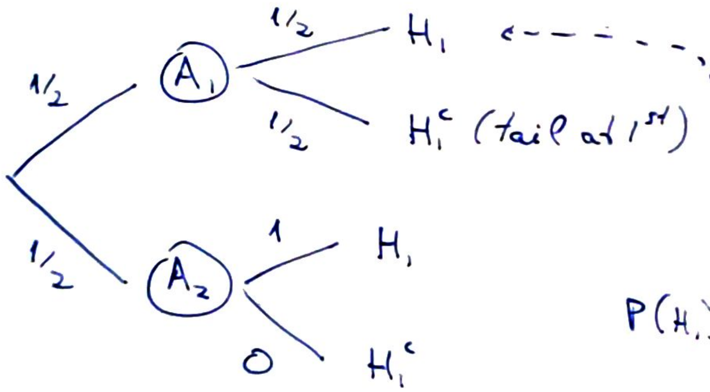
a) $P(A_1 \mid H_1)$?

(this is Bayes' formula).

$$P(A_1 \mid H_1) = \frac{P(A_1 \cap H_1)}{P(H_1)} = \frac{P(A_1)P(H_1 \mid A_1)}{P(A_1)P(H_1 \mid A_1) + P(A_2)P(H_1 \mid A_2)} =$$

$$= \frac{(0.5)(0.5)}{(0.5)(0.5) + (0.5)\cdot 1} = \frac{1}{3}.$$

Remark: We could have done this "without" memorizing Baye's. Draw a tree diagram



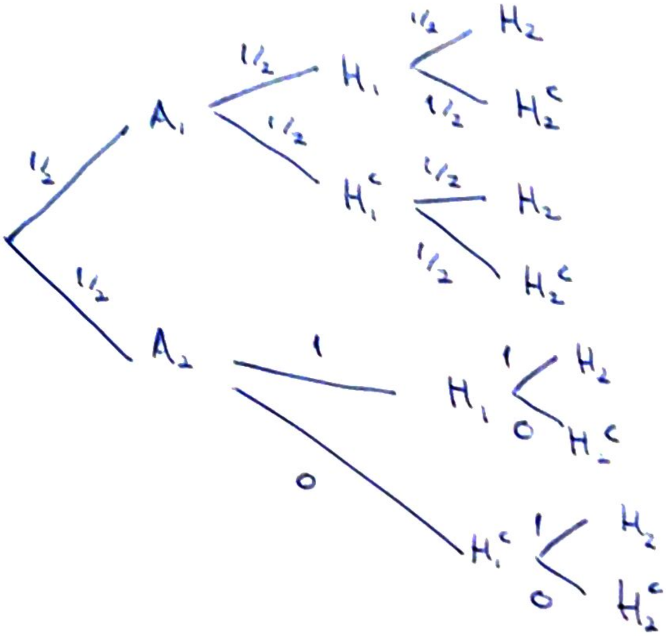
$$P(A_1 \mid H_1) = \frac{P(A_1 \cap H_1)}{P(H_1)}, \text{ with}$$

$$P(A_1 \cap H_1) = \frac{1}{2} \cdot \frac{1}{2}$$

$$P(H_1) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot 1$$ (all branches that lead to H_1).

b) $P(A_1 | H_1 \cap H_2)$?

we can use Bayes' "or" a tree diagram (they are indeed the same)



$$P(A_1 | H_1 \cap H_2) = \frac{P(A_1 \cap H_1 \cap H_2)}{P(H_1 \cap H_2)}$$

- $P(A_1 \cap H_1 \cap H_2) = \frac{1}{2}\frac{1}{2}\frac{1}{2}$ (one branch)
- $P(H_1 \cap H_2) = \frac{1}{2}\frac{1}{2}\frac{1}{2} + \frac{1}{2} \cdot 1 \cdot 1$ (two branches).