Topic 3 – Conditional Probability

Montey Hall problem

- See Numberphile video and 21 video from website first. (*)
- Suppose you always start by picking door #1. Then Montey Hall reveals a goat behind either door 2 or door 3. Then asks if you want to switch or stay on door 1. What do you do?

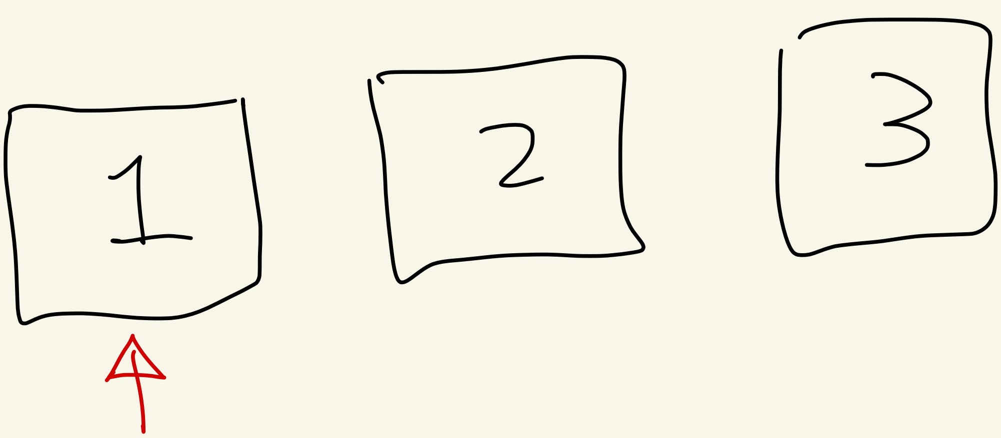


Table of possibilities

door 1	door 2	door 3	Stay w/ door 1 strategy	switch from door 1 strategy
car	goat	goat	WIN	LOSE
goat	car	goat	LOSE	WIN
goat	goat	car	LOSE	WIN

always starting with door 1 as first choice

Staying you win 1/3 of the time

Switching you win 2/3 of the time

You should always switch!

Ex: Suppose we roll two 6-sided dice, a green die and a red die. Suppose the green die stops rolling and lands on a 3, but the red die keeps rolling.

What's the probability that the sum of the dice is 8?

$\boxed{3}$ $\boxed{?}$

starting sample space

S

(1,1)
(1,2)
(1,3)
(1,4)
⋮

(3,1)
(3,2)
(3,3)
(3,4)
(3,5)
(3,6)
⋮

(6,5)
(6,6)

36 outcomes

(green, red)

new sample space

S'

(3,1)
(3,2)
(3,3)
(3,4)
(3,5)
(3,6)

6 outcomes

only one has sum of dice being 8.

So, the probability is $1/6$.

Let's make a formula for this without having to shrink the sample space S and also a method that generalizes even to spaces where the outcomes are not equally likely.

Let E = the event in S where the sum of the dice is 8.

Let $F = S'$ = the event in S where the green die is 3.

We want to know the "conditional probability" of the event E occuring given that F has "already occured".

$F = S'$

(3,1)
(3,2)
(3,3)
(3,4)
(3,5)
(3,6)

E

(6,2) (5,3) (4,4) (3,5) (2,6)

$E \cap F$

(3,5)

S

(1,1) (2,1) (4,1) (5,1) (6,1)
(1,2) (2,2) (4,2) (5,2) (6,3)
(1,3) (2,3) (4,3) (5,4) (6,4)
(1,4) (2,4) (4,5) (5,5) (6,5)
(1,5) (2,5) (4,6) (5,6) (6,6)
(1,6)

$$\underbrace{\frac{|E\cap F|}{|F|}}_{\text{we did this to get } 1/6} = \frac{|E\cap F| / |S|}{|F| / |S|} = \underbrace{\frac{P(E\cap F)}{P(F)}}_{\text{probability in } S} \qquad \frac{\left(1/36\right)}{\left(6/36\right)}$$

ok since all outcomes are equally likely

$\boxed{1/6}$

Def: Let (S, Ω, P) be a probability space. Let E and F be two events. Suppose $P(F) > 0$. Define the <u>conditional probability</u> that E occurs given that F occured to be

$$\underbrace{P(E|F)}_{\text{notation}} = \underbrace{\frac{P(E \cap F)}{P(F)}}_{\text{these probabilities are calculated in } S}$$

Ex: (HW 3 #3 modified)

Suppose you roll two 8-sided dice. You can't see the outcome, but your friend can. They tell you that the sum of the dice is divisible by 5. What is the probability that both dice have landed on 5?

$$S = \{(a,b) \mid a,b = 1,2,\dots,8\}$$

$$|S| = 8^2 = 64$$

$$F = \{(a,b) \mid a+b \text{ is divisible by } 5\}$$

$$\boxed{E = \{(5,5)\}}$$

Want: $P(E|F) = \dfrac{P(E \cap F)}{P(F)}$

We have

$$\boxed{F = \{(1,4), (2,3), (2,8), (3,2), (3,7), (4,1), (4,6), (5,5), (6,4), (7,3), (7,8), (8,2), (8,7)\}}$$

$$\boxed{E \cap F = \{(5,5)\}}$$

$$P(E \mid F) = \frac{P(E \cap F)}{P(F)} = \frac{(1/64)}{(13/64)}$$

$$= 1/13$$

$$\approx 0.7692\ldots$$

$$\approx 7.7\%$$

Theorem: Let (S, Ω, P) be a probability space.

① Let A and B be events and $P(A) > 0$. Then

$$P(A \cap B) = P(A) \cdot P(B|A)$$

② Let $A_1, A_2, \ldots, A_n$ be events with $P(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$.

Then,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) =$$

$$P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot P(A_4|A_1 \cap A_2 \cap A_3) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

③ (Law of total probability)

Suppose $S = E_1 \cup E_2 \cup \cdots \cup E_n$ where each $E_i \neq \phi$, and $E_i \cap E_j = \phi$ if $i \neq j$,

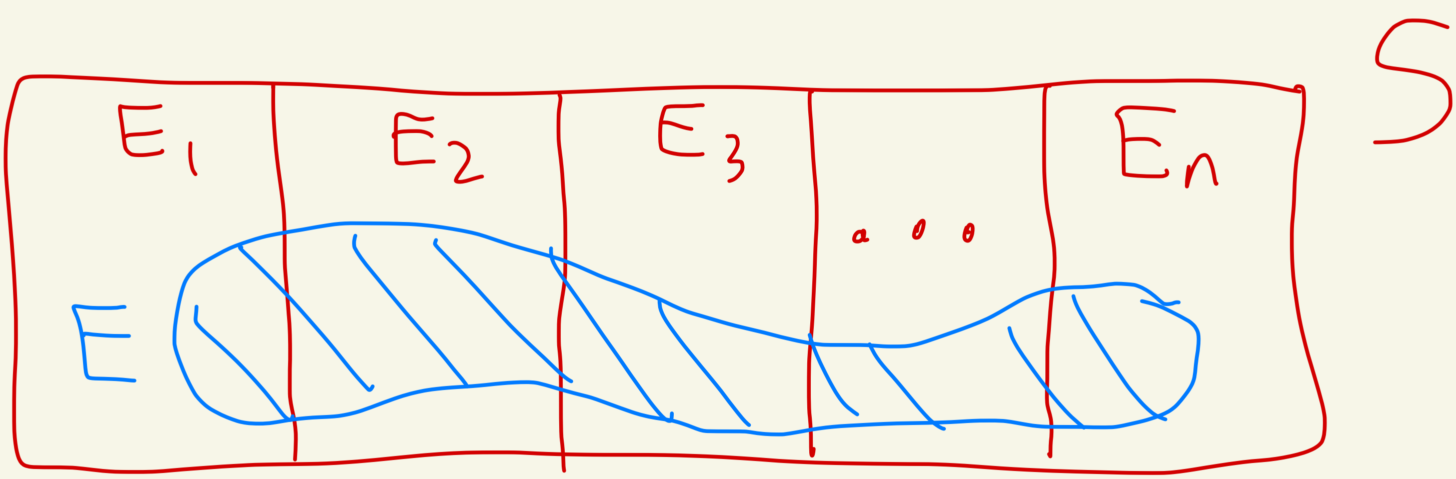
S is broken into n disjoint events

and $P(E_i) \neq 0$ for each i.

Then for event E we have

$$P(E) = P(E|E_1) \cdot P(E_1) + P(E|E_2) \cdot P(E_2) + \cdots + P(E|E_n) \cdot P(E_n)$$

$P(E \cap E_1)$
$+$
$P(E \cap E_2)$
$+$
$\vdots$
$+$
$P(E \cap E_n)$



proof:

① This follows from the definition

$$P(B|A) = \frac{P(A\cap B)}{P(A)}$$

---

② Let's prove this by induction.

The base case is $n=2$, which is

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 | A_1)$$

which is true since $P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}$.

Suppose the statement is true for $n=k$ sets.

Then,

$$P(A_1 \cap A_2 \cap \dots \cap A_{k+1}) = P((A_1 \cap A_2 \cap \dots \cap A_k) \cap A_{k+1})$$

n=2 base case

$$= P(A_1 \cap A_2 \cap \dots \cap A_k) \cdot P(A_{k+1} | A_1 \cap A_2 \cap \dots \cap A_k)$$

n=k inductive case

$$= P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1 \cap A_2) \cdots$$
$$\cdots P(A_k | A_1 \cap A_2 \cap \dots \cap A_{k-1}) \cdot$$
$$\cdot P(A_{k+1} | A_1 \cap A_2 \cap \dots \cap A_k)$$

So, the statement is true for $n=k+1$ given that is true for $n=k$. Thus, by induction, the statement is true for all $n \geq 2$.

<u>Note:</u> Since

$$A_1 \cap A_2 \cap \cdots \cap A_n \subseteq A_1,\ A_1 \cap A_2,\ A_1 \cap A_2 \cap A_3, \ldots$$

and $P(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$ this ensures that $P(A_1) > 0$, $P(A_1 \cap A_2) > 0$, $P(A_1 \cap A_2 \cap A_3) > 0$ etc and thus all the conditional probabilites above are well-defined.

---

(3) We have

$$P(E) = P\big((E \cap E_1) \cup (E \cap E_2) \cup \cdots \cup (E \cap E_n)\big)$$

$$\overset{\star}{=} \sum_{i=1}^{n} P(E \cap E_i)$$

$$\overset{\star\star}{=} \sum_{i=1}^{n} P(E \mid E_i) \cdot P(E_i) \qquad \blacksquare$$

axiom 6 of probability spaces

from (1) of this theorem

Ex: Suppose there are three boxes. In box 1, there are two 4-sided dice. In box 2, there are two 6-sided dice. In box 3, there are two 8-sided dice.

Suppose you randomly pick a box (each box is equally likely to be chosen), then you take the dice out of that box and roll them.

What is the probability that the sum of the dice is 8?

Solution:

Picked box 1 → sum of dice is 8: $\{(4,4)\}$ → probability of sum 8 when box 1 is picked: $1/16$

picked box 2 → $\dfrac{\text{sum of dice is 8}}{\{(2,6),(3,5),(4,4),(5,3),(6,2)\}}$ → $\dfrac{\text{probability of sum 8 when box 2 is picked}}{5/36}$

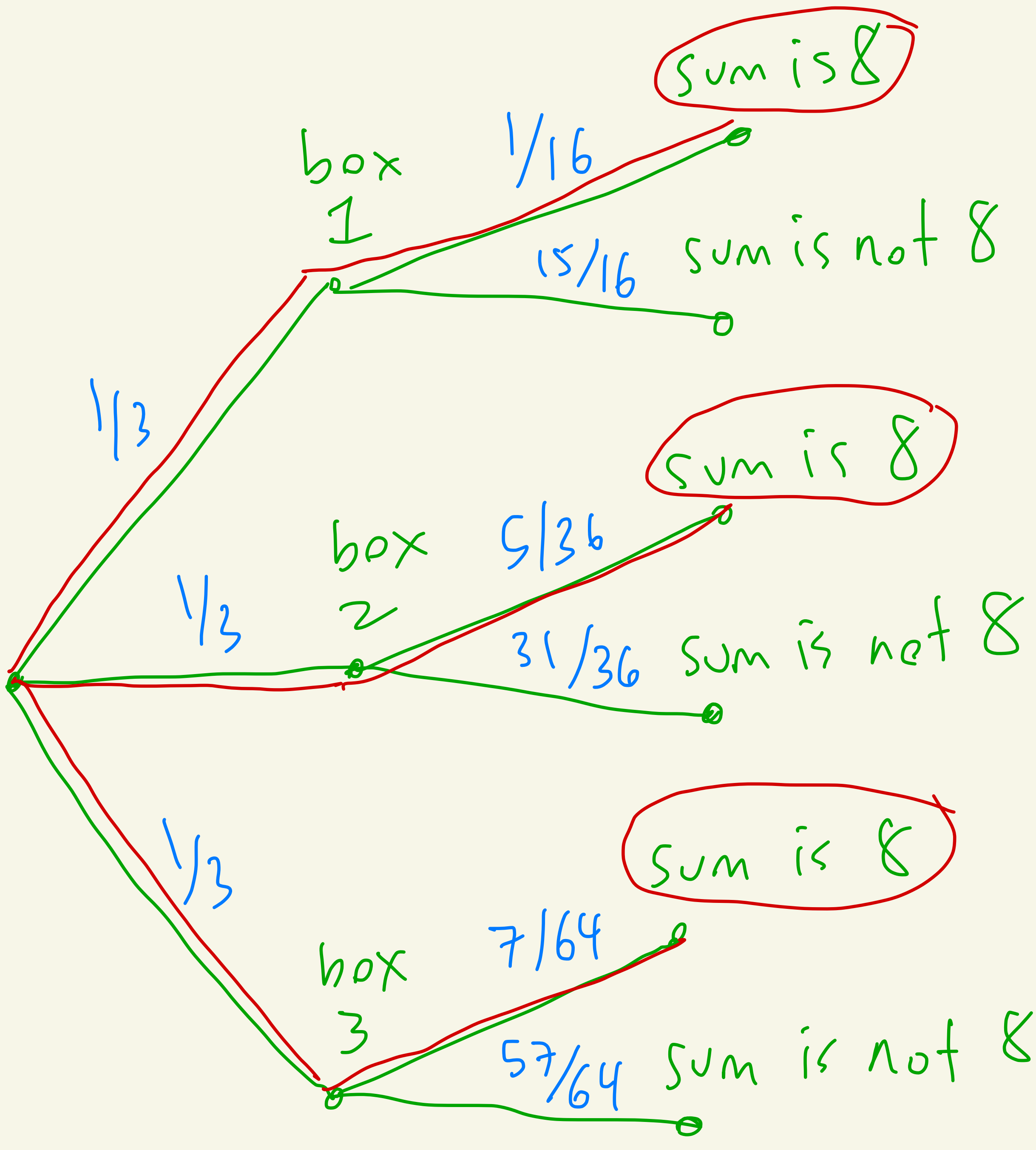
---

picked box 3 → $\dfrac{\text{sum of dice is 8}}{\{(1,7),(2,6),(3,5),(4,4),(5,3),(6,2),(7,1)\}}$ → $\dfrac{\text{probability of sum 8 when box 3 is picked}}{7/64}$

$$
\begin{aligned}
&P(\text{sum of dice is 8}) \\
&= P\left(\begin{matrix}\text{sum of}\\ \text{dice is 8}\end{matrix} \,\middle|\, \text{box 1 picked}\right) \cdot P\left(\begin{matrix}\text{box 1}\\ \text{picked}\end{matrix}\right) \\
&+ P\left(\begin{matrix}\text{sum of}\\ \text{dice is 8}\end{matrix} \,\middle|\, \text{box 2 picked}\right) \cdot P\left(\begin{matrix}\text{box 2}\\ \text{picked}\end{matrix}\right) \\
&+ P\left(\begin{matrix}\text{sum of}\\ \text{dice is 8}\end{matrix} \,\middle|\, \text{box 3 picked}\right) \cdot P\left(\begin{matrix}\text{box 3}\\ \text{picked}\end{matrix}\right)
\end{aligned}
$$

$$= \left(\frac{1}{16}\right)\left(\frac{1}{3}\right) + \left(\frac{5}{36}\right)\left(\frac{1}{3}\right) + \left(\frac{7}{64}\right)\left(\frac{1}{3}\right)$$

$$= \frac{11,456}{110,592} \approx 0.1036\ldots \approx \boxed{10.36\%}$$



Ex: (Monty Hall)

Let's redo the probability of the switch strategy for Montey Hall (start with door 1 and switch after Monty reveals another door).

$$P(\text{win car})$$

$$= P\left(\begin{matrix}\text{win}\\ \text{car}\end{matrix} \,\middle|\, \begin{matrix}\text{car behind}\\ \text{door 1}\end{matrix}\right) \cdot P\left(\begin{matrix}\text{car behind}\\ \text{door 1}\end{matrix}\right)$$

$$+ P\left(\begin{matrix}\text{win}\\ \text{car}\end{matrix} \,\middle|\, \begin{matrix}\text{car behind}\\ \text{door 2}\end{matrix}\right) \cdot P\left(\begin{matrix}\text{car behind}\\ \text{door 2}\end{matrix}\right)$$

$$+ P\left(\begin{matrix}\text{win}\\ \text{car}\end{matrix} \,\middle|\, \begin{matrix}\text{car behind}\\ \text{door 3}\end{matrix}\right) \cdot P\left(\begin{matrix}\text{car behind}\\ \text{door 3}\end{matrix}\right)$$

$$= (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{3}\right)$$

$$= 2/3$$

Sometimes $P(E|F)$ is not equal to $P(E)$ and sometimes it is.

Suppose $P(E|F) = P(E)$.

Then, $\dfrac{P(E \cap F)}{P(F)} = P(E)$.

So, $P(E \cap F) = P(E) \cdot P(F)$

<u>Def:</u> We say that two events E and F are <u>independent</u> if

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

otherwise we say they are <u>dependent</u>.

$\boxed{\text{Note:}}$ $\boxed{\text{Suppose } P(E) > 0 \text{ and } P(F) > 0}$

$$\boxed{E \text{ and } F \text{ are independent}}$$

is equivalent to

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

is equivalent to

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

is equivalent to

$$\boxed{P(F \mid E) = P(F) \text{ and } P(E \mid F) = P(E)}$$

Ex: Suppose you roll two 6-sided die, one green and one red.

Let E be the event that the green die is 1.

Let F be the event that the red die is 3.

Are these events independent?

$$S = \left\{ (g,r) \;\middle|\; \begin{matrix} g = 1,2,3,4,5,6 \\ r = 1,2,3,4,5,6 \end{matrix} \right\} \leftarrow |S| = 36$$

$$E = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6)\}$$

$$F = \{(1,3),(2,3),(3,3),(4,3),(5,3),(6,3)\}$$

$$E \cap F = \{(1,3)\}$$

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

$$P(E) \cdot P(F) = \frac{6}{36} \cdot \frac{6}{36} = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

So, $P(E \cap F) = P(E) \cdot P(F)$

Thus, E and F are independent.

Ex: Suppose you roll two 6-sided die, one green and one red.

Let E be the event that the sum of the dice is 6.

Let F be the event that the red die equals 4.

Are E and F independent?

$S = \{(g,r) \mid g,r = 1,2,3,4,5,6\}$ ← $\boxed{|S| = 36}$

$E = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$

$F = \{(1,4),(2,4),(3,4),(4,4),(5,4),(6,4)\}$

$E \cap F = \{(2,4)\}$

$P(E \cap F) = 1/36 \approx 0.0278\ldots$

$P(E) \cdot P(F) = (5/36)(6/36) = \frac{5}{216} \approx 0.0231\ldots$

Thus, $P(E \cap F) \neq P(E) \cdot P(F)$.

So, E and F are not independent.

<u>Def:</u> (General def of independence)

In a probability space (S, Ω, P) the events $E_1, E_2, \ldots, E_n$ are said to be <u>independent</u> if for every $2 \leq k \leq n$ we have that

$$P(E_{i_1} \cap E_{i_2} \cap \cdots \cap E_{i_k}) = P(E_{i_1}) \cdot P(E_{i_2}) \cdots P(E_{i_k})$$

whenever $1 \leq i_1 < i_2 < \cdots < i_k \leq n$

Ex: E_1, E_2, E_3 are independent if all of the following are true:

$$\boxed{\begin{aligned} P(E_1 \cap E_2) &= P(E_1) \cdot P(E_2) \\ P(E_1 \cap E_3) &= P(E_1) \cdot P(E_3) \\ P(E_2 \cap E_3) &= P(E_2) \cdot P(E_3) \\ P(E_1 \cap E_2 \cap E_3) &= P(E_1) \cdot P(E_2) \cdot P(E_3) \end{aligned}}$$

$k=2$, $i_1=1$, $i_2=2$

$k=2$, $i_1=1$, $i_2=3$

$k=2$, $i_1=2$, $i_2=3$

$k=3$, $i_1=1$, $i_2=2$, $i_3=3$

<u>Theorem</u>: Let S be a sample space of a repeatable experiment. Let A and B be events where $A \cap B = \phi$ [they don't overlap. This is called disjoint events] Suppose further that each time we repeat the experiment S, the experiment is independent of the previous times we did experiment S. Suppose we keep repeating S until either A or B occurs and then we stop. Then the probability that A occurs before B is given by

$$\frac{P(A)}{P(A)+P(B)}$$

Proof: Let E be the event that A occurs before B. Let A_1, B_1, N_1 be the events that A occurs on the first experiment, B occurs on the first experiment, or neither occurs on the first experiment. Then,

$$P(E) = P(E|A_1)\cdot P(A_1) + P(E|B_1)\cdot P(B_1) + P(E|N_1)\cdot P(N_1)$$

$$= 1\cdot P(A_1) + 0\cdot P(B_1) + P(E|N_1)\cdot \underbrace{[1 - P(A_1) - P(B_1)]}_{\text{because the sample space is the disjoint union of } A_1, B_1 \text{ and } N_1}$$

$$= P(A_1) + P(E)\cdot [1 - P(A_1) - P(B_1)]$$

$P(E|N_1) = P(E)$

Since the outcomes of succesive experiments are all independent of each other. When the second experiment begins the whole procedure probabilistically starts over again. Therefore if in the 1st experiment neither A nor B occurs, the probability of E before doing the 1st experiment and after doing the 1st experiment is the same

So,

$$P(E) - P(E)[1 - P(A_1) - P(B_1)] = P(A_1)$$

Thus,

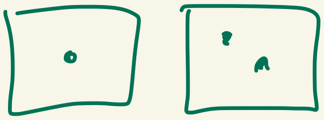
$$P(E) = \frac{P(A_1)}{P(A_1) + P(B_1)}$$

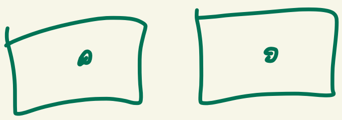
$$= \frac{P(A)}{P(A) + P(B)}$$

$\blacksquare$

Ex: Suppose we roll two 6-sided die over and over. Let A be the event that the sum of the dice is 5. Let B be the event that the sum of the dice is 7. We keep rolling the dice until either A or B happens and then we stop. Whats the probability that A occurs before B, ie that we roll sum of 5 before we roll sum of 7?

Ex:

roll 1 –  ← sum = 3

roll 2 –  ← sum = 2

roll 3 –  ← sum = 5

Sum is 5 occured before sum is 7

$$S = \{(a,b) \mid a,b = 1,2,3,4,5,6\} \leftarrow \boxed{|S| = 36}$$

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

$$B = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$$P(A) = 4/36$$

$$P(B) = 6/36$$

Probability sum is 5 occurs before

sum is 7 is [ie A before B]

$$\frac{P(A)}{P(A)+P(B)} = \frac{4/36}{4/36 + 6/36} = \frac{4}{10} = \boxed{\frac{2}{5}} = \boxed{40\%}$$

probability sum is 7 occurs before sum is 5 occurs is [ie B before A]

$$\frac{P(B)}{P(B)+P(A)} = \frac{6/36}{6/36 + 4/36} = \boxed{\frac{6}{10}} = \boxed{60\%}$$