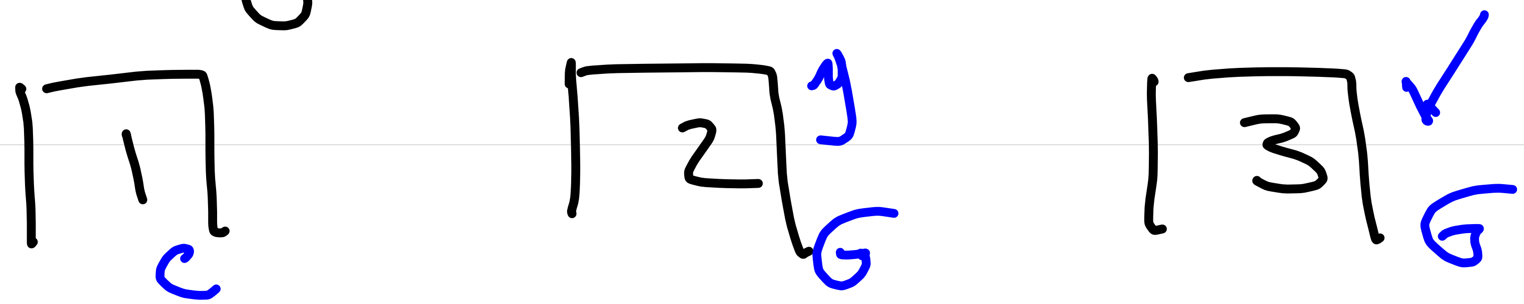
Week 12-Fri

- Conditional Probability
- Independence

Monty Hall Problem

The game show has three doors.

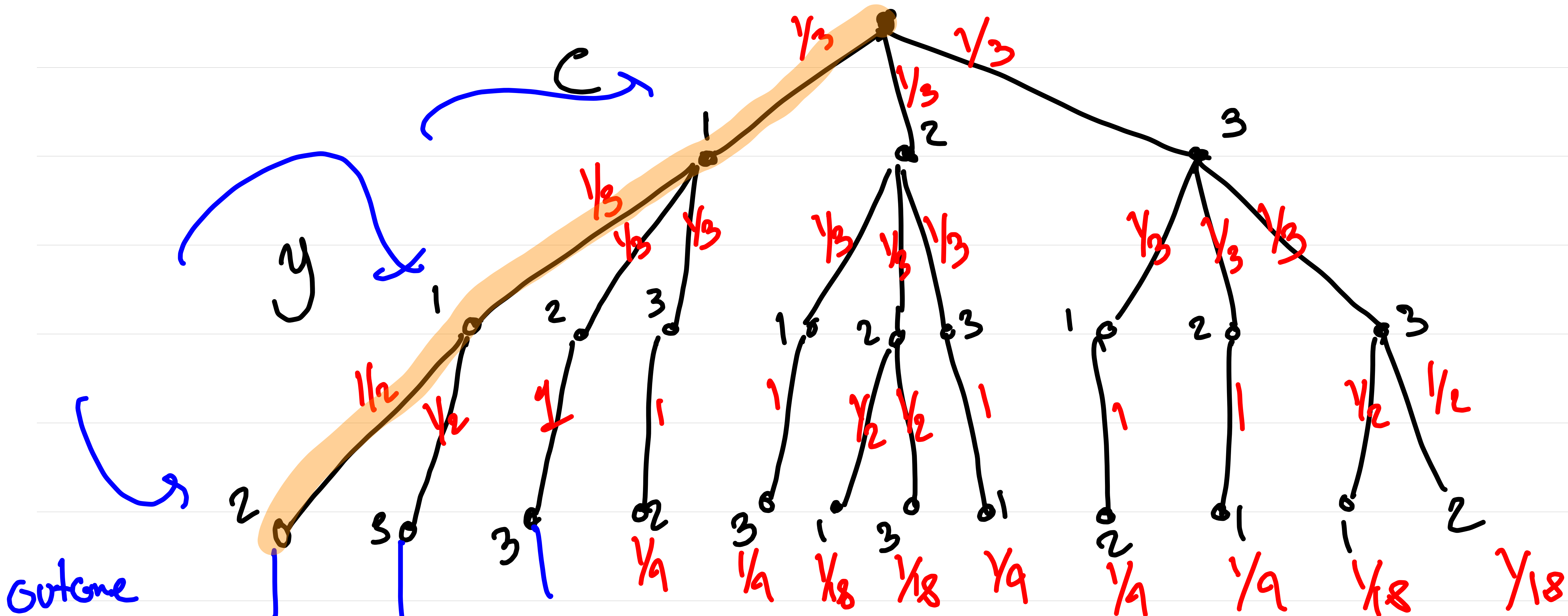


1- a Car is randomly placed behind one door uniformly at random. Goats placed behind the other doors

2- you choose a door randomly.

3- Monty opens one of the other doors with a goat behind it uniformly at random and reveals a goat.

- You are allowed to now switch the door or not. What should you do?



Outcome

(1, 1, 2) (1, 1, 3) (1, 2, 3)

P(S) $\frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{18}$

Why?

Conditional Probability:

Motivation: Suppose that we have two events E and F and you know that F happened. How does that change your belief about the likelihood of E happening?

Defn: If E and F are events and $\mathbb{P}(F) > 0$ then the probability of E given F is defined as

$$\mathbb{P}(E \mid F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

"the probability of E given F"

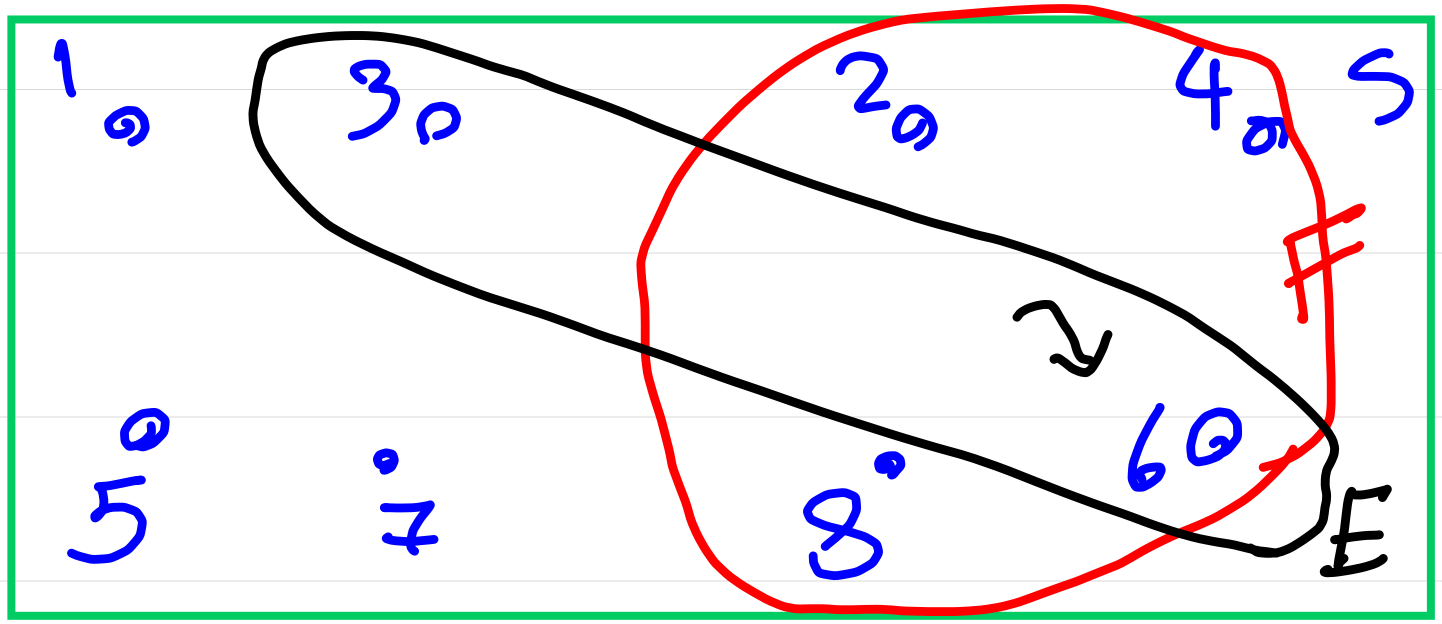
or "the probability of E conditioned on F"

What does this mean ?

$$\mathbb{P}(E) =$$

$$\mathbb{P}(E|S) = \mathbb{P}(E) = \frac{\mathbb{P}(E)}{1} = \frac{\mathbb{P}(E)}{\mathbb{P}(S)} = \frac{\mathbb{P}(E\cap S)}{\mathbb{P}(S)}$$

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

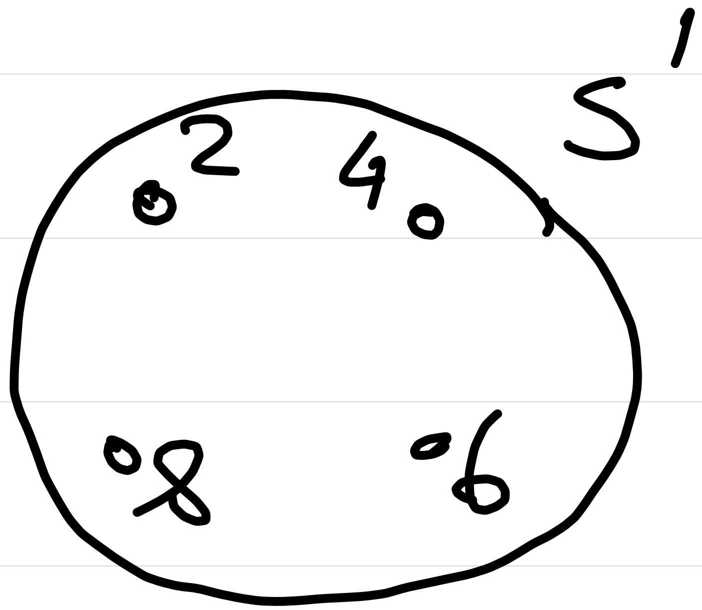


$\mathbb{P}(E|F)$

E: multiple of 3

F: even number

$P(2) = P(4) = P(6) = P(8) = 1/8$



$P'(2) = P'(4) = P'(6) = P'(8)$

$$P'(x) = \frac{P(x)}{\mathbb{P}(F)}$$

$$\mathbb{P}(E|F) = \sum_{x \in E \cap F} P'(x) = \sum_{x \in E \cap F} \frac{P(x)}{\mathbb{P}(F)} = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}$$

EX: Pick a uniformly random person in Canada $U = \{1, \dots, 30 \times 10^6\}$

Event C is "person is UTSC CS student"

and Event S is "person lives in Scarborough"

Let say $|C| = 10^3$ and $|S| = 20 \times 10^3$.

$C \cap S$ = UTSC cs student and lives in Scarborough.

Let say $|C \cap S| = 200$

(assumption: $P(x) = \frac{1}{|U|}$, for all $x \in U$)

$$\mathbb{P}(C) = \frac{|C|}{|U|} = \frac{10^3}{30 \times 10^6} = \frac{1}{30 \times 10^3}$$

$$\mathbb{P}(C \mid S) = \frac{\mathbb{P}(C \cap S)}{\mathbb{P}(S)} = \frac{|C \cap S| / |U|}{|S| / |U|} = \frac{|C \cap S|}{|S|} = \frac{200}{20 \times 10^3} = \frac{1}{100}$$

$$\mathbb{P}(S \mid C) = \frac{|S \cap C|}{|C|} = \frac{200}{10^3} = \frac{1}{5}$$

Two important Consequences of the definition

Consequnce #1:

Product Rule: $\mathbb{P}(E \cap F) = \mathbb{P}(E \mid F)\, \mathbb{P}(F)$

(also know as "chain rule")

More generally: $\mathbb{P}(E \cap F \cap G) = \mathbb{P}(E \mid F \cap G)\, \mathbb{P}(F \cap G)$

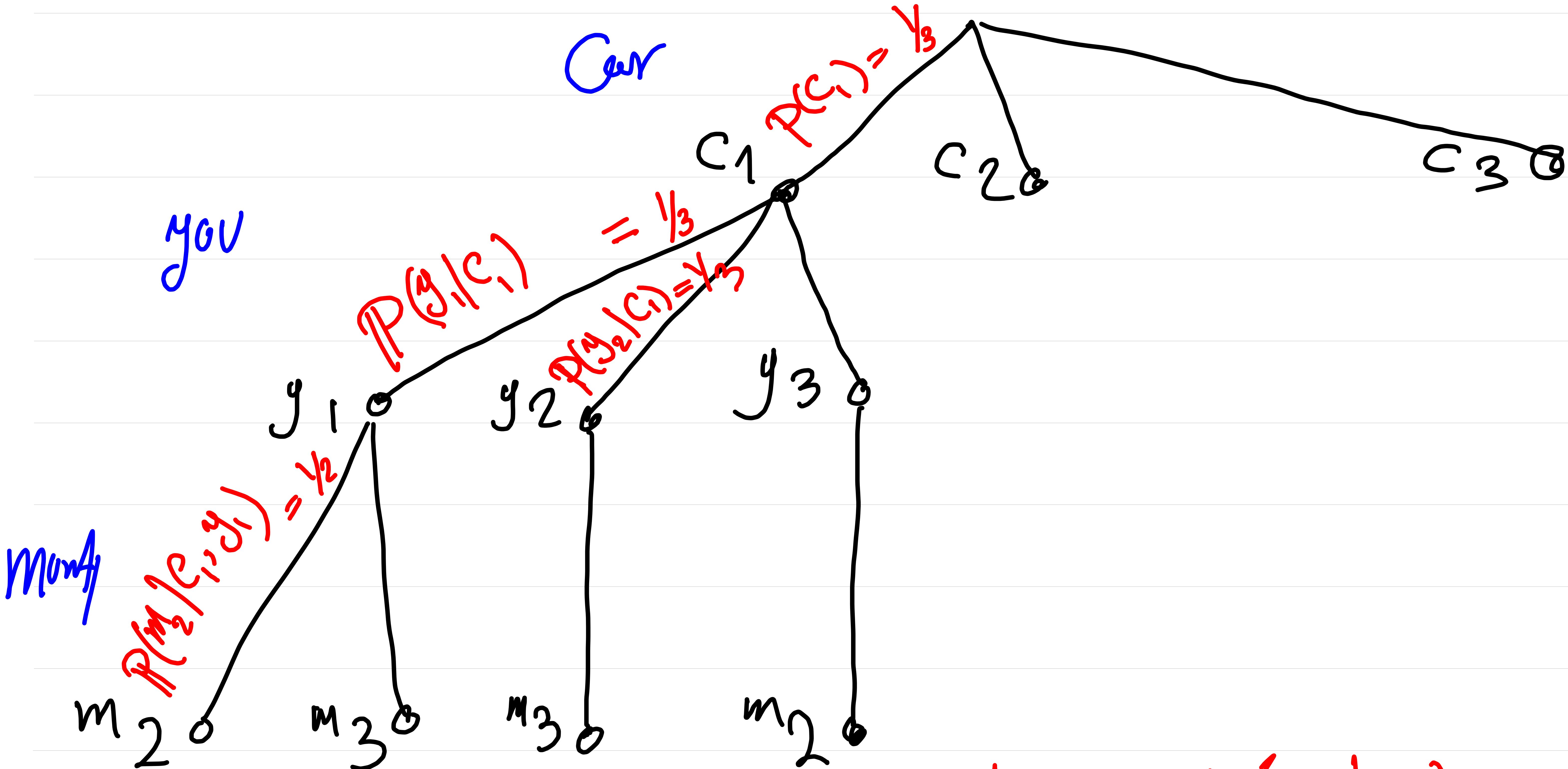
$$= \mathbb{P}(E \mid F \cap G)\, \mathbb{P}(F \mid G)\, \mathbb{P}(G).$$

$$\mathbb{P}(E_1 \cap E_2 \cap \dots \cap E_n)$$

$$= \mathbb{P}(E_1 \mid E_2 \cap \dots \cap E_n)\, \mathbb{P}(E_2 \mid E_3 \cap \dots \cap E_n) \cdots \mathbb{P}(E_{n-1} \mid E_n)$$

$$\times\, \mathbb{P}(E_n)$$

Let's Revisit The Monty Hall Problem



$$P(c_1 \cap y_1 \cap m_2) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = \mathbb{P}(m_2 | c_1 \cap y_1)\, \mathbb{P}(y_1 | c_1)\, \mathbb{P}(c_1)$$

Two important Consequences of the definition

$\mathbb{P}(E\cap F)$ $\qquad\qquad$ $\mathbb{P}(E\cap \bar{F})$

Consequnce #2:

The Law of total Probability:

$$\mathbb{P}(E) = \mathbb{P}(E|F)\,\mathbb{P}(F) + \mathbb{P}(E|\bar{F})\,\mathbb{P}(\bar{F})$$

Proof:

$$\mathbb{P}(E) = \mathbb{P}\big((E\cap F)\cup(E\cap\bar{F})\big) \overset{\text{disjoint}}{=} \mathbb{P}(E\cap F) + \mathbb{P}(E\cap \bar{F})$$

$$= \mathbb{P}(E|F)\,\mathbb{P}(F) + \mathbb{P}(E|\bar{F})\,\mathbb{P}(\bar{F})$$

Remark: $E = (E\cap F)\cup(E\cap \bar{F})$

and $(E\cap F)\cap(E\cap\bar{F}) = \emptyset$

EX: Medicine Testing.

Prostate Cancer (PC)

PSA test

NIH: 1) If patient has PC, test is positive 80% time,
negative 20% time.
false negative

2) If patient does not have PC,
test is positive 10% time, ← false positive
negative 90% time.

3) 0.16% of the male population has PC.

Q1) What is the likelihood of testing positive.

Answer:

C = chosen person has PC

T = chosen person tests positive

$$\mathbb{P}(T) = \mathbb{P}(T|C)\,\mathbb{P}(C) + \mathbb{P}(T|\overline{C})\,\mathbb{P}(\overline{C})$$

$$= 0.8 \times 0.0016 + 0.1 \times (1 - 0.0016) \approx 0.10$$

What If We Want to Answer the Following Question?

Q2) If I test positive, what is the likelihood that I have PC?

We know $\mathbb{P}(T|C)$ and we want $\mathbb{P}(C|T)$.

- To answer this, we need to introduce another rule.

Baye's Rule:

$$\mathbb{P}(E|F) = \mathbb{P}(F|E)\,\frac{\mathbb{P}(E)}{\mathbb{P}(F)}$$

Proof:

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

Answer to Q2)

$$\mathbb{P}(C \mid T) = \mathbb{P}(T \mid C)\,\frac{\mathbb{P}(C)}{\mathbb{P}(T)} = 0.8 \times \frac{0.0016}{0.1} \approx 0.013$$

about 1.3%