MATH 180A: INTRO TO PROBABILITY
(FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 2.1 – 2.2

Next: § 2.2 – 2.3

HW. 1 due TONIGHT!

Lab.1 due MONDAY, 10/07

Screencast & video available after each lecture @ podcast.ucsd.edu

Before / After slides now available on course webpage.

Lots of active discussion on Piazza.

Conditional Probability

2.1

Eg. Your friend rolls two fair dice, and asks you what is the probability the sum is 10.

[$\Omega = \{(i,j) : 1 \le i,j \le 6\}$ $A = \{i+j=10\} = \{(4,6),(5,5),(6,4)\}$ $\therefore \mathbb{P}(A) = \frac{3}{36} = \boxed{\frac{1}{12}}$]

Before you answer, however, she reveals that the actual sum that came up was a two digit number. In light of this information, was your probability calculation correct?

"updated" $\tilde{\Omega} = \{\text{sum has 2 digits}\}$

$= \{i+j=10\} \cup \{i+j=11\} = \{i+j=12\}$

$= \{(4,6),(5,5),(6,4),(6,5),(5,6),(6,6)\}$

$$\tilde{\mathbb{P}}(A) = \frac{\#\tilde{A}}{\#\tilde{\Omega}} = \frac{3}{6} = \frac{1}{2}$$

$\tilde{A} = A \cap \tilde{\Omega}$

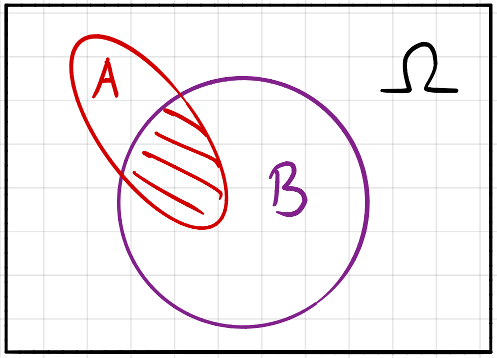
Conditional Probability

Moral: given information (i.e. that an event B is known to have happened), we **condition** on B; we make B the new sample space.

We must modify events afterward so they're "in" B:

$$\Omega \to B = \tilde{\Omega}$$

$$\mathcal{F} \to \mathcal{F}_B = \{AB : A \in \mathcal{F}\}$$



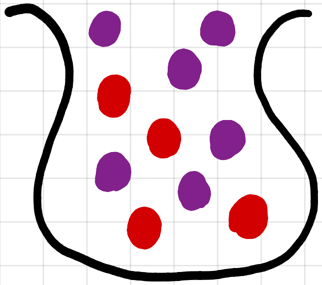
Problem: $\mathbb{P}(\tilde{\Omega}) = \mathbb{P}(B) < 1$

$\mathbb{P}(\cdot \mid B) = \tilde{\mathbb{P}} = \dfrac{\mathbb{P}}{\mathbb{P}(B)}$ (caveat: $\mathbb{P}(B) \neq 0$)

Def: Given an event B with $\mathbb{P}(B) > 0$, we define the conditional probability of an event A given B as

$$\mathbb{P}(A \mid B) = \mathbb{P}(AB) / \mathbb{P}(B).$$

E.g.



An urn contains 4 red balls and 6 blue balls. 3 are sampled, without replacement.

What is the probability that exactly two are red?

$\Omega = \{\{b_1, b_2, b_3\} : b_i \neq b_j \text{ if } i \neq j\}$ $\quad A = \{2 \text{ red}, 1 \text{ blue}\}$ $\quad \mathbb{P}(A) = \frac{6 \cdot 6}{120} = 30\%$

$\#\Omega = \binom{10}{3}$ $\quad \#A = \binom{4}{2}\binom{6}{1}$

Suppose we somehow know a priori that at least one red is sampled. What is the conditional probability that exactly two red balls are sampled?

$A = \{\text{exactly } 2 \text{ red}\}$ $\quad B = \{\text{at least one red}\}$

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(BA)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{3/10}{5/6} = \frac{9}{25}$$

$AB = A$

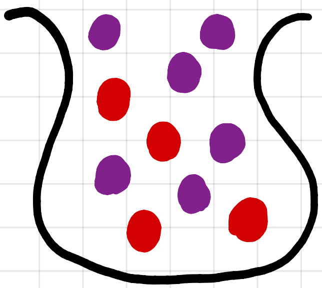
$$\mathbb{P}(B^c) = \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} = \frac{20}{120} = \frac{1}{6}, \quad \mathbb{P}(B) = 1 - \mathbb{P}(B^c) = 5/6$$

Recovering $\mathbb{P}$ from $\mathbb{P}(\cdot \mid B)$

By definition, $\mathbb{P}(B|A) = \dfrac{\mathbb{P}(AB)}{\mathbb{P}(A)}$; $\Rightarrow$ $\mathbb{P}(AB) = \mathbb{P}(A)\,\mathbb{P}(B|A)$

"multiplication rule"

Can generalize: $\mathbb{P}(ABC) = \mathbb{P}(AB)\,\mathbb{P}(C|AB) = \mathbb{P}(A)\,\mathbb{P}(B|A)\,\mathbb{P}(C|AB)$

Eg. 

An urn contains 4 red balls and 6 blue balls.
2 are sampled, without replacement.
What is the probability that both are red?

$R_1 = \{1^{st} \text{ red}\}$
$R_2 = \{2^{nd} \text{ red}\}$

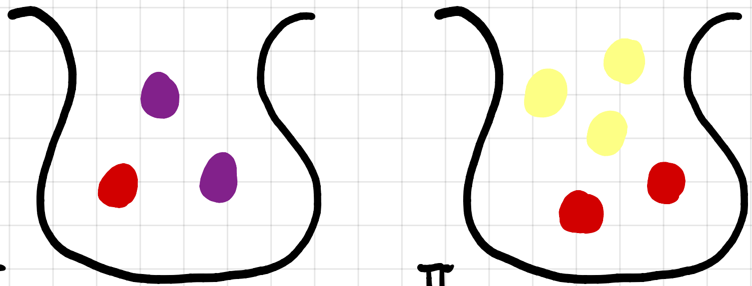
$$\mathbb{P}(R_1R_2) = \mathbb{P}(R_1)\,\mathbb{P}(R_2|R_1)$$
$$= (0.4)\left(\frac{3}{9}\right) = \frac{2}{15}$$

(old way: $\dfrac{\binom{4}{2}\binom{6}{0}}{\binom{10}{2}}$)

Two-Stage Experiments

* perform an experiment, measure a random outcome.
* perform a second experiment whose setup depends on the outcome of the first!

E.g.



I II

* First, choose an urn at random.
* Then, sample a ball at random from the chosen urn.

What is the probability it is red?

$$\mathbb{P}(R) = \mathbb{P}\big((R \cap \mathrm{I}) \cup (R \cap \mathrm{II})\big)$$
$$= \mathbb{P}(R\,\mathrm{I}) + \mathbb{P}(R\,\mathrm{II})$$
$$= \mathbb{P}(\mathrm{I})\,\mathbb{P}(R \mid \mathrm{I}) + \mathbb{P}(\mathrm{II})\,\mathbb{P}(R \mid \mathrm{II})$$
$$= \tfrac{1}{2}\left(\tfrac{1}{3}\right) + \tfrac{1}{2}\left(\tfrac{2}{5}\right) = \tfrac{1}{6} + \tfrac{1}{5} = \tfrac{11}{30}$$

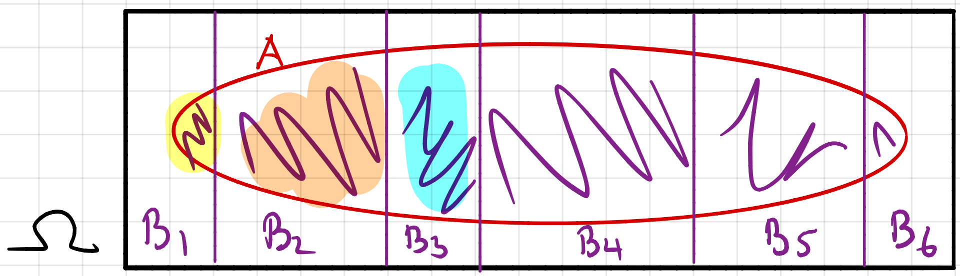
Law of Total Probability

If $B_1, B_2, \ldots, B_n$ partition Ω (disjoint, $B_1 \cup \cdots \cup B_n = \Omega$, $\mathbb{P}(B_j) > 0$)

then for any event A:

$$\mathbb{P}(A) = \mathbb{P}(AB_1 \cup AB_2 \cup \cdots \cup AB_n) = \sum_{j=1}^{n} \mathbb{P}(AB_j)$$

$$= \sum_{j=1}^{n} \mathbb{P}(B_j)\,\mathbb{P}(A \mid B_j)$$



Eg. 90% of coins are fair B_1

9% are biased to come up heads 60% B_2

1% are biased to come up heads 80% B_3

You find a coin on the street. How likely is it to come up heads?

$$\mathbb{P}(H) = \mathbb{P}(B_1)\,\mathbb{P}(H \mid B_1) + \mathbb{P}(B_2)\,\mathbb{P}(H \mid B_2) + \mathbb{P}(B_3)\,\mathbb{P}(H \mid B_3)$$

$$= (0.9)(50\%) + (0.09)(60\%) + (0.01)(80\%) = 0.512$$

Subtler question:

90% of coins are fair. 9% are biased to come up heads 60%.
1% are biased to come up heads 80%.

You find a coin on the street. You toss it, and it comes up heads.

How likely is it that this coin is heavily biased?

2.2

Eg. According to Forbes Magazine, as of April 10, 2019, there are
2208 billionaires in the world.
1964 of them are men.