## CS70: Jean Walrand: Lecture 23.

What do we learn from observations?

1. Probability Basics Review
2. Examples
3. Conditional Probability

## Probability Basics Review

**Setup:**

- Random Experiment.
  Flip a coin twice.
- Probability Space.
  - **Sample Space:** Set of outcomes, $\Omega$.
    $\Omega = \{HH, HT, TH, TT\}$
  - **Probability:** $Pr[\omega]$ for all $\omega \in \Omega$.
    $Pr[HH] = \cdots = Pr[TT] = 1/4$
    1. $0 \leq Pr[\omega] \leq 1$.
    2. $\sum_{\omega \in \Omega} Pr[\omega] = 1$.
  - **Event:** $A \subseteq \Omega, Pr[A] = \sum_{\omega \in \Omega} Pr[\omega]$.
    $Pr[\text{at least one H out of two tosses}] = Pr[HT, TH, HH] = 3/4$

## Exactly 50 heads in 100 coin tosses.

Sample space: $\Omega =$ set of 100 coin tosses $= \{H, T\}^{100}$.
$|\Omega| = 2 \times 2 \times \cdots \times 2 = 2^{100}$.

Uniform probability space: $Pr[\omega] = \frac{1}{2^{100}}$.

Event $E =$ "100 coin tosses with exactly 50 heads"

$|E|$?
Choose 50 positions out of 100 to be heads.
$|E| = \binom{100}{50}$.

$$Pr[E] = \frac{\binom{100}{50}}{2^{100}}.$$

**Calculation.**
Stirling formula (for large $n$):

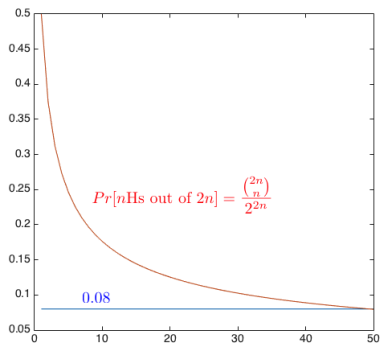
$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

$$\binom{2n}{n} \approx \frac{\sqrt{4\pi n}(2n/e)^{2n}}{[\sqrt{2\pi n}(n/e)^n]^2} \approx \frac{4^n}{\sqrt{\pi n}}.$$

$$Pr[E] \approx \frac{\frac{4^{50}}{\sqrt{50\pi}}}{2^{100}} = \frac{1}{\sqrt{50\pi}} \approx .08.$$

```
irb(main):006:0> (50*3.1415926)**(0.5)
=> 12.5331412662588
irb(main):007:0> 1/12.5331412662588
=> 0.0797884567608089
```

## Exactly 50 heads in 100 coin tosses.

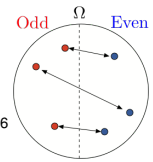



## Probability of even number of heads in 57 coin tosses

$\Omega$ = 57 coin tosses. $|\Omega| = 2^{57}$.

Let $E = \{\omega \in \Omega \mid$ number of Hs in $\omega$ is even$\}$.

**Fact:** $P(E) = 1/2$.




**Proof:**
Consider the correspondence:

$$HA_1A_2 \cdots A_{56} \leftrightarrow TA_1A_2 \cdots A_{56}$$

Here, $A_1, \ldots, A_{56}$ are 56 coin flips.

It matches every even sequence to an odd sequence, and conversely.

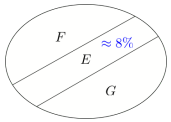
Hence, there are exactly as many odd as even sequences.

$$\Rightarrow |E| = |\Omega|/2 \Rightarrow P(E) = \frac{|E|}{|\Omega|} = \frac{1}{2}$$

## Probability more heads than tails in 100 coin tosses.

$\Omega$ = 100 coin tosses.
$|\Omega| = 2^{100}$.

Recall event $E$ = 'equal heads and tails'
Event $F$ = 'more heads than tails'
Event $G$ = 'more tails than heads'
A 1-to-1 correspondence between outcomes in F and G!
$|F| = |G|$.
$E$, $F$ and $G$ are disjoint.
$Pr[E] \approx 8\%$.




$|\Omega| = |E| + |F| + |G|$.
$\Rightarrow 1 = Pr[\Omega] = Pr[E] + 2Pr[F] \approx 8\% + 2Pr[F]$.

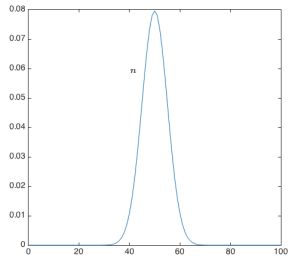
Solve for $|F|$:

$$|F| \approx 46\%$$

## Probability of $n$ heads in 100 coin tosses.

$\Omega$ = 100 coin tosses.
$|\Omega| = 2^{100}$.

Event $E_n$ = '$n$ heads'; $|E_n| = \binom{100}{n}$




$$p_n = Pr[E_n] = \frac{|E_n|}{|\Omega|} = \frac{\binom{100}{n}}{2^{100}}$$

## Roll a red and a blue die.

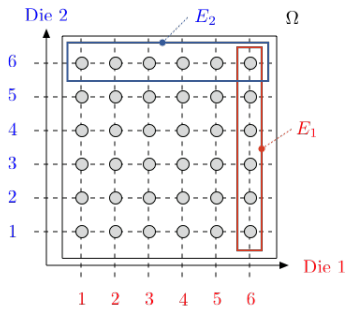
$$\Omega = \{(a,b) : 1 \le a,b \le 6\} = \{1,2,\ldots,6\}^2.$$

Uniform: $Pr[\omega] = \frac{1}{|\Omega|} = \frac{1}{36}$ for all $\omega$.

What is the probability of

1. the red die showing 6?
   $E_1 = \{(6,b) : 1 \le b \le 6\}$, $|E_1| = 6$, $Pr[E_1] = |E_1|/|\Omega| = 1/6$
2. at least one die showing 6?
   $E_1 = \{(6,b) : 1 \le b \le 6\}$ = red die shows 6
   $E_2 = \{(a,6) : 1 \le a \le 6\}$ = blue die shows 6
   $E = E_1 \cup E_2$ = red or blue die (or both) show 6
   $|E| = |E_1| + |E_2| - |E_1 \cap E_2|$ [Inclusion/Exclusion]
   $|E_1 \cap E_2| = \{(6,6)\}$
   $|E| = 6 + 6 - 1 = 11$
   $Pr[E] = 11/36$

## Roll a red and a blue die.




$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

$E_1$ = 'Red die shows 6'; $E_2$ = 'Blue die shows 6'

$E_1 \cup E_2$ = 'At least one die shows 6'

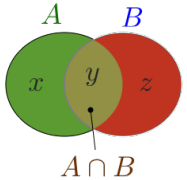
$Pr[E_1] = \frac{6}{36}, Pr[E_2] = \frac{6}{36}, Pr[E_1 \cup E_2] = \frac{11}{36}$.

## Inclusion/Exclusion

Note that,

$$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$$

whether or not the sample space has uniform distribution.




$$\begin{aligned} Pr[A] &= x + y \\ Pr[B] &= y + z \\ Pr[A \cap B] &= y \\ Pr[A \cup B] &= x + y + z \end{aligned}$$

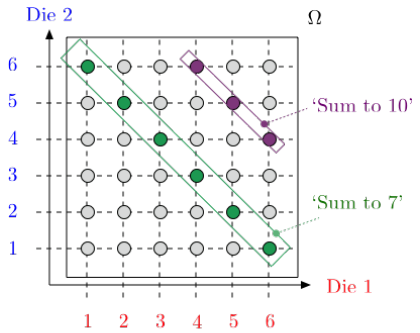
## Roll a red and a blue die.

$$\Omega = \{(a,b) : 1 \le a,b \le 6\} = \{1,2,\ldots,6\}^2.$$

Uniform: $Pr[\omega] = \frac{1}{|\Omega|} = \frac{1}{36}$ for all $\omega$.
What is the probability of

1. the dice sum to 7?
   $E = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$; $|E| = 6$.
   Counting argument:
   for each choice $a$ of the value of the red die,
   there is exactly one choice $b = 7 - a$ for the blue die,
   so there are 6 total choices.
   $Pr[E] = |E|/|\Omega| = 6/36 = 1/6$.
2. the dice sum to 10?
   $E = \{(4,6),(5,5),(6,4)\}$
   $Pr[E] = |E|/|\Omega| = 3/36 = 1/12$.

## Roll a red and a blue die.




$Pr[\text{Sum to } 7] = \frac{6}{36}$ $\quad$ $Pr[\text{Sum to } 10] = \frac{3}{36}$

## Roll two blue dice.

The key idea is that we do not distinguish the dice.

Roll die 1, then die 2. Then forget the order.

For instance, we consider that $(2,5)$ and $(5,2)$ are the same outcome.

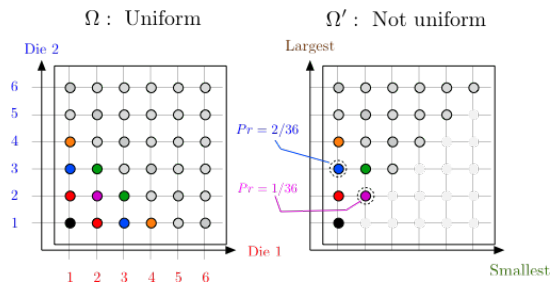
We designate this outcome by $(2,5)$.

Thus,

$$\Omega' = \{(a,b) \mid 1 \le a \le b \le 6\}.$$

We see that $Pr[(1,3)] = \frac{2}{36}$ and $Pr[(2,2)] = \frac{1}{36}$.

## Roll two blue dice

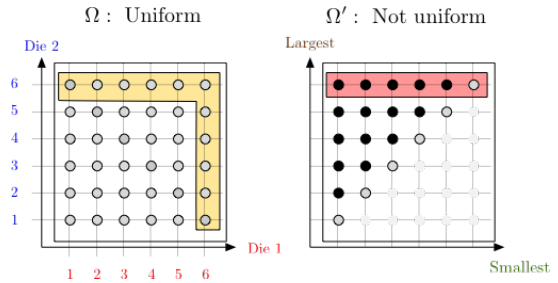



Two different models of the same random experiment.

In $\Omega'$, $Pr[(1,3)] = \dfrac{2}{36}$ and $Pr[(2,2)] = \dfrac{1}{36}$

## Roll two blue dice.

Now what is the probability of at least one die showing 6?



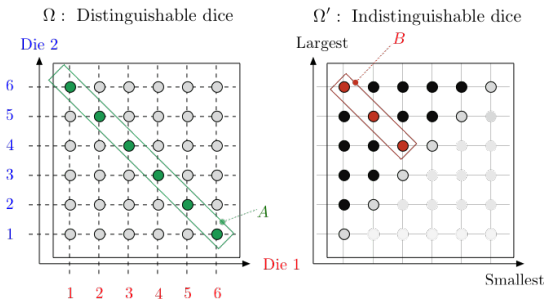

In $\Omega$, $Pr[A] = \frac{11}{36}$; in $\Omega'$, $Pr[B] = 5 \times \frac{2}{36} + 1 \times \frac{1}{36}$.

Of course, this is the same as for distinguishable dice!

The event does not depend on the dice being distinguishable.

## Roll two blue dice.

What is the probability of the dice sum to 7?




In $\Omega$, $Pr[A] = \frac{6}{36}$; in $\Omega'$, $Pr[B] = 3 \times \frac{2}{36}$.

Of course, this is the same as for distinguishable dice!

The event does not depend on the dice being distinguishable.

## Really not uniform! and not finite!

- Experiment: Toss three times a coin with $Pr[H] = 2/3$.
  - $\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$.
  - $Pr[HHH] = (\frac{2}{3})^3$; $Pr[HHT] = (\frac{2}{3})^2(\frac{1}{3})$; ...
- Toss a fair coin until you get a heads.
  - $\Omega = \{H, TH, TTH, TTTH, \ldots\}$
  - $Pr[H] = \frac{1}{2}, Pr[TH] = \frac{1}{4}, Pr[TTH] = \frac{1}{8}$
  - Still sums to 1. Indeed $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots = 1$.

## Set notation review

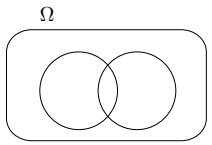



Figure : Two events

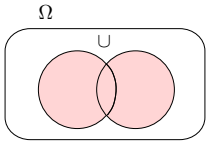



Figure : Union (or)

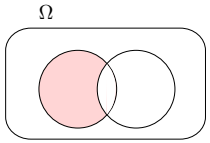



Figure : Difference ($A$, not $B$)

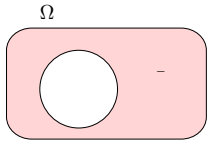



Figure : Complement (not)

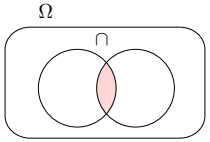



Figure : Intersection (and)

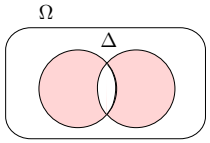
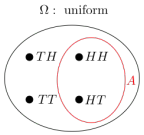


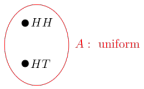

Figure : Symmetric difference (only one)

## Conditional probability: example.

Two coin flips. First flip is heads. Probability of two heads?
$\Omega = \{HH, HT, TH, TT\}$; Uniform probability space.
Event $A =$ first flip is heads: $A = \{HH, HT\}$.




New sample space: $A$; uniform still.



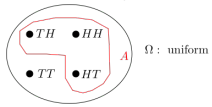

Event $B =$ two heads.

The probability of two heads if the first flip is heads.
**The probability of $B$ given $A$** is $1/2$.

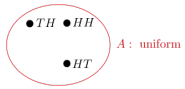
## A similar example.

Two coin flips. One of the flips is heads.
Probability of two heads?

$\Omega = \{HH, HT, TH, TT\}$; uniform.
Event $A =$ one flip is heads. $A = \{HH, HT, TH\}$.




New sample space: $A$; uniform still.

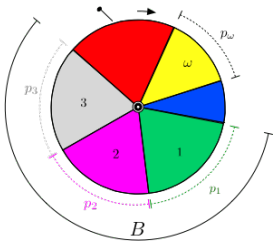



Event $B =$ two heads.

The probability of two heads if at least one flip is heads.
**The probability of $B$ given $A$** is $1/3$.

## Conditional Probability: A non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.

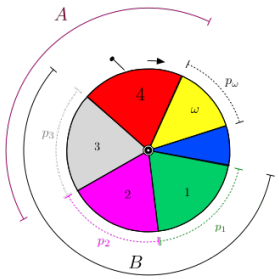



$$Pr[3|B] = \frac{p_3}{p_1 + p_2 + p_3} = \frac{p_3}{Pr[B]}.$$

$$\omega \notin B \Rightarrow Pr[\omega|B] = 0.$$

## Another non-uniform example

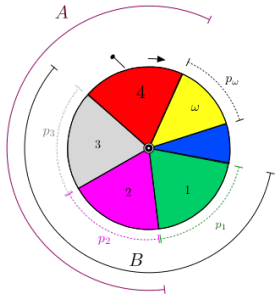
Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.




$$Pr[A|B] = \frac{p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

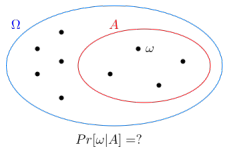
## Yet another non-uniform example

Consider $\Omega = \{1, 2, \ldots, N\}$ with $Pr[n] = p_n$.




$$Pr[A|B] = \frac{p_2 + p_3}{p_1 + p_2 + p_3} = \frac{Pr[A \cap B]}{Pr[B]}.$$

## Conditional Probability.




If in $A$, what is the probability of outcome $\omega$?

If $\omega \notin A$, probability is 0
Otherwise: Ratio of probability of $\omega$ to total probability of $A$

$$Pr[\omega|A] = \frac{Pr[\omega]}{Pr[A]}$$

Uniform Probability Space:
Ratio of $1/|\Omega|$ to $|A|/|\Omega| \implies 1/|A|$. (Makes sense!)

What do we learn from observations?

You observe that the event $B$ occurs.

That changes your information about the probability of every event $A$.

The conditional probability of $A$ given $B$ is

$$Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}.$$

Note: $Pr[A \cap B] = Pr[B] \times Pr[A|B] = Pr[A] \times Pr[B|A]$.

## Conditional Probability.

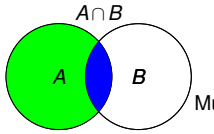
**Definition:** The **conditional probability** of $B$ given $A$ is

$$Pr[B|A] = \sum_{\omega \in B} Pr[\omega|A] = \frac{\sum_{\omega \in A \cap B} Pr[\omega]}{Pr[A]} = \frac{Pr[A \cap B]}{Pr[A]}$$

## Conditional Probability.

**Definition:** The **conditional probability** of $B$ given $A$ is

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}$$




In $A$!
In $B$?
Must be in $A \cap B$.

$$Pr[B|A] = \frac{Pr[A \cap B]}{Pr[A]}.$$