

# Events and Probability

Consider an experiment with a finite (or countably infinite) number of outcomes.

Each outcome is a simple event (or a sample point).

The **sample space** is the set of all possible simple (elementary) events.

An **event**  $\mathcal{E}$  is a union of simple events - a subset of the sample space.

Two events are **mutually exclusive** if  $A \cap B = \emptyset$ .

With each simple event  $s$  we associate a number  $Pr(s)$  which is the **probability** of  $s$ .

# Probability Space

A probability distribution  $Pr$  on a sample space  $S$  is a mapping from events of  $S$  to real numbers such that

- $Pr(A) \geq 0$  for any event  $A$ .
- $Pr\{S\} = 1$ .
- For any (finite or countably infinite) sequence of pairwise mutually exclusive events  $A_1, A_2, \dots$ :

$$Pr\{\cup_i A_i\} = \sum_i Pr\{A_i\}$$

The pair  $(S, Pr)$  is called a discrete **probability space**.

$$Pr(\mathcal{E}) = \sum_{s \in \mathcal{E}} Pr(s).$$

## Examples

Consider the random process defined by the outcome of rolling a dice.

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

We assume that all “facets” have equal probability, thus

$$Pr(1) = Pr(2) = \dots Pr(6) = 1/6.$$

The probability of the event “odd outcome”

$$= Pr(\{1, 3, 5\}) = 1/2$$

# An Infinite Sample Space

Flip an unbiased coin until HEADS appears for the first time. Here the sample space is

$$\{H, TH, TTH, TTTH, \dots\}.$$

The event that “the number of TAILS seen is odd” is give by the infinite set

$$\{TH, TTTH, TTTTTH, \dots\}.$$

Assume that we roll two dice:

$\mathcal{S} =$  all ordered pairs  $\{(i, j), 1 \leq i, j \leq 6\}$ .

We assume that each (ordered) combination has probability  $1/36$ .

Probability of the event “sum = 2” =

$$Pr((1, 1)) = 1/36.$$

Probability of the event “sum = 3”

$$Pr(\{(1, 2), (2, 1)\}) = 2/36.$$

Let  $E_1 =$  “sum bounded by 6” ,

$$E_1 = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 2),$$

$$(2, 3), (2, 4), (3, 1), (3, 2), (3, 3), (4, 1), (4, 2), (5, 1)\}$$

$$Pr(E_1) = 15/36$$

Let  $E_2 =$  “both dice have odd numbers” ,  
 $Pr(E_2) = 1/4$ .

$$Pr(E_1 \cap E_2) =$$

$$Pr(\{(1, 1), (1, 3), (1, 5), (3, 1), (3, 3), (5, 1)\}) =$$

$$6/36 = 1/6.$$

# Principle of Inclusion-Exclusion

Let  $E_1, E_2, \dots, E_n$  be arbitrary events. Then

$$\Pr(\cup_{i=1}^n E_i) = \sum_i \Pr(E_i) - \sum_{i < j} \Pr(E_i \cap E_j) + \sum_{i < j < k} \Pr(E_i \cap E_j \cap E_k) - \dots + (-1)^{l+1} \sum_{i_1 < i_2 < \dots < i_l} \Pr(\cap_{r=1}^l E_{i_r}) + \dots$$

**Boole's inequality (union bound):** For any arbitrary sequence of events  $E_1, E_2, \dots, E_n$ :

$$\Pr(\cup_{i=1}^n E_i) \leq \sum_i \Pr(E_i)$$

# Back to the Polynomial Identity Checking

A simple event = a choice of  $r$ .

Sample Space = all integers in  $[1, \dots, 100d]$ .

We assume that all integers in the range have equal probability, thus the probability of a simple event  $r$  is  $Pr(r) = \frac{1}{100d}$ .

The “**bad**” events: choosing a root of the polynomial. There are no more than  $d$  simple events in the bad event.

$$Pr(\text{“bad” event}) \leq \frac{d}{100d}.$$

Assume that we repeat the algorithm  $k$  times.

If any iteration returns FALSE output FALSE, else output CORRECT.

A simple event = A sequence of  $k$  choices  $r_1, \dots, r_k$ .

The sample space = All sequences of  $k$  numbers in the range  $[1, \dots, 100d]$ .

The probability of a simple event =  $(\frac{1}{100d})^k$ .

The **bad** event = all  $k$  choices are roots of the polynomial, there are no more than  $d^k$  such simple events.

Probability of the bad event  $\leq d^k (\frac{1}{100d})^k$ .

# Conditional Probability

What is the probability that a random person born in Indiana is a student at Purdue.

$E_1$  = the event that a random person in the world is born in Indiana.

$E_2$  = the event that a random person in the world is a student at Purdue.

The conditional probability that a random person born in Indiana is a student at Purdue is denoted

$$Pr(E_2 | E_1).$$

# Computing Conditional Probabilities

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

By conditioning on  $B$  we restrict the sample space to the set  $B$ .

Thus we are interested in  $Pr(A \cap B)$  “normalized” by  $Pr(B)$ .

## Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

$$E_1 = \text{"sum is 8"},$$

$$Pr(E_1) = Pr((2, 6), (3, 5), (4, 4), (5, 3), (6, 2)) = 5/36$$

$$E_2 = \text{"sum even"},$$

$$Pr(E_2) = 1/2 = 18/36.$$

$$Pr(E_1 | E_2) = \frac{Pr(E_1 \cap E_2)}{Pr(E_2)} = \frac{5/36}{1/2} = 5/18.$$

## Example - a posteriori probability

We are given 2 coins. One is a fair coin  $A$ , the other coin,  $B$ , has head on both sides  $B$ .

We choose a coin at random (i.e. each coin is chosen with probability  $1/2$ ) and toss.

Given that we got head, what is the probability that we chose the fair coin  $A$ ???

Define a sample space of ordered pairs  $(\textit{coin}, \textit{outcome})$ .

The sample space has three points

$$\{(A, h), (A, t), (B, h)\}$$

$$\begin{aligned} Pr((A, h)) &= Pr((A, t)) = 1/4 \\ Pr((B, h)) &= 1/2 \end{aligned}$$

Define two events:

$$E_1 = \text{“Chose coin } A\text{”}.$$

$$E_2 = \text{“Outcome is head”}.$$

$$Pr(E_1 \mid E_2) = \frac{Pr(E_1 \cap E_2)}{Pr(E_2)} = \frac{1/4}{1/4 + 1/2} = 1/3.$$

Bayes' Law.

## Useful identities:

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

$$Pr(A \cap B) = Pr(A | B)Pr(B)$$

$$Pr(A \cap B \cap C) = Pr(A | B \cap C)Pr(B \cap C)$$

$$= Pr(A | B \cap C)Pr(B | C)Pr(C)$$

Let  $A_1, \dots, A_n$  be a sequence of events.

Let  $E_i = \bigcap_{j=1}^i A_j$

$$Pr(E_n) = Pr(A_n | E_{n-1})Pr(E_{n-1}) =$$

$$Pr(A_n | E_{n-1})Pr(A_{n-1} | E_{n-2}) \dots Pr(A_2 | E_1)Pr(A_1)$$

# Independence

Two events  $A$  and  $B$  are independent if

$$Pr(A \cap B) = Pr(A) \times Pr(B),$$

or (when  $Pr(B) > 0$ )

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

Independent events do not have to be related to independent physical processes.

Example: the probability that the outcome of a dice roll is *even* ( $= \frac{3}{6}$ ) is independent of the event "the outcome is  $\leq 4$ " ( $= \frac{4}{6}$ ).

The probability of "an even outcome  $\leq 4$ " is

$$\frac{2}{6} = \frac{12}{36} = \frac{3}{6} \cdot \frac{4}{6}$$

The "intuition" here is that there are the same number of odd and even outcomes that are  $\leq 4$ . Thus, the "information" that the outcome is  $\leq 4$  does not "help" in deciding if it is odd or even.

## Back to the Polynomial Identity Test

The probability of a wrong answer in one run of the algorithm is  $1/100$ . Runs of the algorithms are **independent** (i.e., the probability of error in each run is independent of other runs.) It's enough to get a correct answer in one of the runs.

The probability of a wrong answer in  $k$  runs of the algorithm is  $(\frac{1}{100})^k$ .

# The Birthday Paradox

What is the probability that among  $m$  people no two have the same birthday?

Assumptions:

1. All birthdays are equally likely.
2. Birthdays are independent events.

The sample space is the set of all vectors  $S = \{(b_1, \dots, b_m) | b_i \in [1, \dots, N]\}$ .

We need to compute  $Pr(E)$  where  $E = \{(b_1, \dots, b_m) | b_i \neq b_j \text{ for all } i \neq j\}$ .

How many different atomic events are counted in  $E$ ?

The number of possible  $m$  different birthdays is  $N.(N - 1).(N - 2) \dots (N - m + 1)$

$$\Pr(E) = \frac{N.(N-1).(N-2)\dots(N-m+1)}{N^m}$$

$$= \prod_{i=0}^{m-1} (1 - i/N)$$

$$\leq \prod_{i=0}^{m-1} e^{-i/N} = e^{-\sum_{i=0}^{m-1} i/N} = e^{-m(m-1)/2N}$$

For  $m = \sqrt{2N} + 1 \leq 28$ ,

$$\Pr(E) < 1/e < 1/2.$$

## Alternate Analysis

Assume that we choose one birthday after the other independently and uniformly at random from  $[1 \dots N]$ .

Let  $E_i$ : "the  $i$ th choice is different from the first  $i - 1$  choices".

$$\Pr(\cap_{i=1}^m E_i) = \Pr(E_1) \Pr(E_2|E_1) \Pr(E_3|E_2 \cap E_1) \dots \Pr(E_m | \cap_{i=1}^{m-1} E_i) = \prod_{i=1}^m (1 - \frac{i-1}{N})$$

### Principle of Deferred Decisions.

# Verifying Matrix Multiplication

Given three  $n \times n$  integer matrices  $A$ ,  $B$ , and  $C$ , verify whether

$$AB = C$$

Simple matrix multiplication takes  $\Theta(n^3)$  steps. There exist a  $\Theta(n^{2.37})$  algorithm.

# Randomized Algorithm

Choose a random vector  $\bar{r} = (r_1, r_2, \dots, r_n) \in \{0, 1\}^n$ .

Compute  $A(B\bar{r})$  and  $C\bar{r}$ .

If  $A(B\bar{r}) \neq C\bar{r}$ , then output  $AB \neq C$ .

Else output  $AB = C$ .

Takes  $\Theta(n^2)$  time.

## Bounding the Error Probability

**Theorem 1.** *If  $AB \neq C$ , and  $\bar{r}$  is chosen uniformly at random from  $\{0, 1\}$ , then*

$$\Pr(AB\bar{r} = C\bar{r}) \leq 1/2$$

**Proof.** Let  $D = AB - C \neq 0$ .  $D$  has some non-zero entry. W.l.o.g let it be  $d_{11}$ .

For  $D\bar{r} = 0$ , we should have

$$\sum_{j=1}^n d_{1j}r_j = 0, \text{ i.e.,}$$

$$r_1 = -\frac{\sum_{j=2}^n d_{1j}r_j}{d_{11}}$$

Set  $r_n, \dots, r_2, r_1$  one by one.

The equality holds in at most one choice of  $r_1$ .  $\square$