Module 7.1: Probability Space and Events

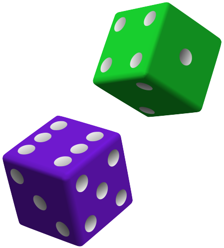
MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Rolling dice

A die has six faces, each marked with dots representing the numbers 1, 2, 3, 4, 5, or 6. Thus, when we roll a die we get (on top) one of six possible **outcomes**: the numbers 1 through 6.

With a **fair** die each of the six faces is **equally likely** to show up on top. Thus, when we roll a fair die the **chance** or **probability** of getting, for example, a 4, is $1/6$.



Distinguishable dice

Problem. We roll a pair of distinguishable fair dice (one die is green, the other is purple) together. What are the chances that we get a double (1-1, 2-2, ..., 6-6)?

Answer. A roll of the dice shows 6 possible green sides and 6 purple ones.

By the multiplication rule, that's $6 \cdot 6 = 36$ possible outcomes.

6 of these outcomes are doubles.

The chances of getting a double are therefore $6/36 = 1/6$.

Indistinguishable dice

Problem. We roll a pair of indistinguishable fair dice (both are beige) together. What are the chances that we get a double (1-1, 2-2, ..., 6-6)?

Answer. There are two kinds of outcomes: (1) dice show same number, and (2) dice show different numbers.

There are 6 outcomes of the first kind (the doubles) and $\binom{6}{2} = 15$ outcomes of the second kind for a total of $6 + 15 = 21$.

The chances of getting a double seem to be $6/21 = 2/7$.

But $2/7 \simeq 0.28$ is significantly bigger than $1/6 \simeq 0.17$. Backgammon (tavla) players love doubles. Must they always play with indistinguishable dice?

Distinguishable vs. indistinguishable

No, it does not matter whether the pair consists of distinguishable dice or not.

Explanation. The key is to consider **how likely** each outcome is.

For the green-purple dice, since the dice are fair, and since we roll them in the same way, each of the 36 outcomes is **equally likely**.

However, for the beige dice some of the outcomes are more likely than others. For example rolling a 5 and a 6 cannot be distinguished from rolling a 6 and a 5. Therefore this is twice as likely than rolling a 6 and a 6.

This discussion motivates the definition of **probability space**.

Probability space

A **probability space** (Ω, Pr) consists of

- a finite non-empty set Ω of **outcomes** and
- a **probability distribution** function $\Pr : \Omega \to [0, 1]$ that associates with each outcome $w \in \Omega$ its **probability** $\Pr[w]$ which is a real number between 0 and 1 (inclusive), such that

$$\sum_{w \in \Omega} \Pr[w] = 1$$

What we cover in this course is **finite** probability theory. It suffices for much of computer science and it frees us from the need to use calculus.

Quiz

A **fair** coin is equally likely to show heads (abbreviated `H`) as it is to show tails (abbreviated `T`) when it is flipped (or tossed).

We flip a fair coin three times in a row. How many outcomes are there in the resulting probability space?

A. 3
B. 6
C. 8

Answer

A. 3

Incorrect. There are 2 possibilities in each flip.

B. 6

Incorrect. You need to multiply the 2 possibilities in each flip.

C. 8

Correct. There are 2 possibilities in each flip and by the multiplication rule $2 \cdot 2 \cdot 2 = 8$.

More Information

Here are the 8 outcomes (`THH` means "tails in the first round followed by heads in rounds two and three):

`HHH` `HHT` `HTH` `HTT` `THH` `THT` `TTH` `TTT`

Examples of probability space I

When we roll the green-purple dice the outcomes are
$\Omega = \{(1,1),(1,2),\ldots,(6,5),(6,6)\}$ with $|\Omega| = 36$.

Following the "equally likely" intuition, the probability distribution is
$\Pr[(1,1)] = \Pr[(1,2)] = \cdots = \Pr[(6,5)] = \Pr[(6,6)] = 1/36$.

When we flip a fair coin three times the outcomes are
$\{\mathtt{HHH}, \mathtt{HHT}, \ldots, \mathtt{TTH}, \mathtt{TTT}\}$ with $|\Omega| = 8$.

Following again the "equally likely" intuition
$\Pr[\mathtt{HHH}] = \Pr[\mathtt{HHT}] = \cdots = \Pr[\mathtt{TTH}] = \Pr[\mathtt{TTT}] = 1/8$.

We flip a fair coin and then, without regard to what the coin showed, we roll a fair die. In this case the outcomes are:
$\Omega = \{(\mathtt{H},1),\ldots,(\mathtt{H},6),(\mathtt{T},1),\ldots,(\mathtt{T},6)\}$ with $|\Omega| = 2 \cdot 6 = 12$.

Again our intuition says the outcomes are "equally likely" so each of them will have probability $1/12$.

Another example of probability space

When we roll the indistinguishable (beige) dice the outcomes are
$\Omega \;= \{1-1,\ldots,6-6\} \cup \{\{1,2\},\{1,3\},\ldots,\{6,4\},\{6,5\}\}$
And $|\Omega| = 6 + 15 = 21$.

Outcomes $1-1,\ldots,6-6$ are equally likely, probability p.

Outcomes $\{1,2\},\{1,3\},\ldots,\{6,4\},\{6,5\}$ are also equally likely.
Probability q.

An outcome of the second kind is twice as likely as one of the first kind.

That is $$q \;=\; 2p.$$

But all probabilities in the space must add up to 1!

Hence $$6p \;+\; 15q \;=\; 1.$$

Solving, we get $p = 1/36$ and $q = 1/18$.

Events

Let (Ω, Pr) be a probability space. An **event** in this space is a subset $E \subseteq \Omega$. We extend the probability function from outcomes to events as follows

$$\mathrm{Pr}[E] = \sum_{w \in E} \mathrm{Pr}[w]$$

Note that

- $\mathrm{Pr}[E] \in [0, 1]$
- $\mathrm{Pr}[\emptyset] = 0$
- $\mathrm{Pr}[\{w\}] = \mathrm{Pr}[w]$

Now we can calculate the probability of getting a double with beige dice:

$$\mathrm{Pr}[\{1-1, \ldots, 6-6\}] = \mathrm{Pr}[\{1-1\}] + \cdots + \mathrm{Pr}[\{6-6\}] = 1/36 + \cdots + 1/36 =$$
$$= 6/36 = 1/6$$

Distinguishable or not (again)

Explanation (again). By setting up the correct probability space for the indistinguishable beige dice we have corrected the conclusion we drew earlier: that beige dice are better at getting doubles!

In fact one can formally relate the probability space of the beige dice to that of the green-purple dice. Then we can prove that for all events (not just the doubles) in the beige space the probability is the same as in the green-purple space.

Informally, we can treat the beige dice as if they are distinguishable!

Here is an intuitive argument for that: instead of rolling the beige dice simultaneously, roll them sequentially.

Event probability exercise

Problem. We roll a pair of fair dice. Compute the probability of the event "at least one of the dice shows a 6". Solve the problem in both the green-purple space and the beige space.

Answer. First in the green-purple space. The outcomes in the event of interest are of two kinds. Green 6 with any purple: six outcomes. And purple 6 with any green: six outcomes.

However, we double counted green 6 purple 6! So only 11 outcomes and the answer is $1/36 + \cdots + 1/36 = 11/36$

Now in the beige space the outcomes in the event of interest are $\{1, 6\}, \ldots, \{5, 6\}, 6-6$. The answer is

$$(1/18 + \cdots + 1/18) + 1/36 = 5/18 + 1/36 = 11/36$$

Module 7.2: Uniform Spaces

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Uniform probability spaces

A probability space (Ω, Pr) is called **uniform** if all the outcomes have the **same** probability.

Denote $n = |\Omega|$. Since the probabilities are equal and sum up to 1:

. $\mathrm{Pr}[w] = 1/n$ for each outcome $w \in \Omega$.

Proposition. In a uniform probability space $\mathrm{Pr}[E] = m/n$ where $m = |E|$ and $n = |\Omega|$.

Proof.

$$\mathrm{Pr}[E] \; = \; \sum_{w \in E} \mathrm{Pr}[w] \; = \; \sum_{w \in E} \frac{1}{n} \; = \; m \cdot \frac{1}{n} \; = \; \frac{m}{n}$$

Dice add up to an even number

Problem. Compute the probability of the event "when we roll a pair of fair dice the numbers add up to an even number".

Answer. W.l.o.g. we work in the green-purple dice probability space. This space is uniform with 36 outcomes.

Let g be the number shown by the green die and p the one shown by the purple one.
Each outcome corresponds to a pair (g, p) where $g, p \in \{1, \ldots, 6\}$

$g + p$ is even iff both g and p are even or both are odd.

The event of interest contains exactly **half of the outcomes** because for each die there are as many even faces as there are odd ones.

The answer is $18/36 = 1/2$.

Three dice show the same number

Problem. We roll three fair dice. What is the probability that all three show the same number?

Answer. Each outcome corresponds to a triple $(d_1, d_2, d_3) \in \{1, \ldots, 6\}^3$. By the multiplication rule there are $6 \cdot 6 \cdot 6 = 216$ outcomes.

Since the dice are fair and rolled in the same way, each of the outcomes (d_1, d_2, d_3) is **equally likely**.

Therefore the space is uniform: each outcome has probability $1/216$.

The event of interest consists of outcomes
. $(1, 1, 1), \ldots, (6, 6, 6)$. That's six outcomes.

Hence the answer is $6/216 \ = \ 1/36$.

Quiz

We flip a fair coin 3 times. The probability that we get one heads and two tails (in some order) is

(A) $3/8$

(B) $1/3$

(C) $1/8$

Answer

The correct answer is (A) $3/8$. This is a uniform probability space with 8 outcomes `HHH` `HHT` `HTH` `HTT` `THH` `THT` `TTH` `TTT`

Of these, the outcomes in the event of interest are `HTT` `THT` `TTH` so 3 outcomes. Hence the probability is $3/8$.

Activity

We flip a fair coin n times. Compute the probability that we get at least one heads.

Describe an outcome as a sequence of length n of $\mathtt{H}$ and $\mathtt{T}$.

ASK HOW MANY OUTCOMES

By the multiplication rule there are $2 \cdot 2 \cdots 2 = 2^n$ outcomes.

This is a uniform probability space so we just need to count how many outcomes have at least one heads?

It's easier to count complementarily!

Only one outcome has no heads at all: $\mathtt{TT} \cdots \mathtt{T}$.

$$\text{Answer:} \qquad \frac{2^n - 1}{2^n}$$

Module 7.3: Biased Coins and Bernoulli Trials

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

Biased coins, Bernoulli trials

A **fair** coin corresponds to a uniform probability space with two outcomes, heads ($\mathtt{H}$) and tails ($\mathtt{T}$). Each of the outcomes has probability $1/2$.

A **biased** coin corresponds (in general) to a **non-uniform** space with the same outcomes, $\mathtt{H}$ and $\mathtt{T}$, that is parameterized by $\Pr[\mathtt{H}] = p \in [0, 1]$. It follows that $\Pr[\mathtt{T}] = 1 - p$.

The biased coin with parameter p is commonly invoked under a different name. A **Bernoulli trial** corresponds to a probability space with two outcomes, conventionally called "success" and "failure" that is parameterized by a probability of success which is p. The probability of failure is then $1 - p$.

Therefore flipping a biased coin is a Bernoulli trial in which heads is conventionally designated as success.

Biased coin flipped twice I

Problem. A biased coin has a probability $1/3$ of showing heads. We flip this coin twice. What is the probability that we obtain one tails and one heads (in either order)?

Answer. There are four outcomes $\mathtt{HH}$, $\mathtt{HT}$, $\mathtt{TH}$, $\mathtt{TT}$. However, they are **not** equally likely. So, what is the probability distribution?

Consider an urn holding three identical marbles. On one of them we write $\mathtt{H}$ and on the others $\mathtt{T}_1$ and $\mathtt{T}_2$. Assuming that each marble is equally likely to be extracted, sampling one marble corresponds to flipping our biased coin.

Now consider **two** such urns, U and U'. Sampling a marble from U then one from U' corresponds to flipping our biased coin **twice**.

Biased coin flipped twice II

Answer (continued). We now have $3 \cdot 3 = 9$ outcomes:

$$\mathtt{HH},\ \mathtt{HT}_1,\ \mathtt{HT}_2,\ \mathtt{T}_1\mathtt{H}, \ldots\ \mathtt{T}_1\mathtt{T}_2,\ \mathtt{T}_2\mathtt{H}, \ldots$$

Extractions from each urn happen in the same way so these 9 outcomes are **equally likely**. We have a **uniform space**.

The event of interest (a heads and a tails, in some order) is $\{\mathtt{HT}_1, \mathtt{HT}_2, \mathtt{T}_1\mathtt{H}, \mathtt{T}_2\mathtt{H}\}$.

That's 4 outcomes out of a total of 9, so its probability is $4/9$.

Random permutations

Distinct objects $a_1, \ldots, a_n$. A **random permutation** of $a_1, \ldots, a_n$ is an element of the **uniform** probability space whose outcomes are all the permutations . Each outcome has probability $1/n!$.

Problem. Let $i, j \in [1..n]$ (not necessarily distinct). Calculate the probability that a_i occurs in position j in a random permutation.

Answer. Let E_{ij} be the event consisting of all outcomes in which a_i occurs in position j. The probability we want is $|E_{ij}|/n!$.

A permutation in E_{ij} can be constructed in two steps: (1) put a_i in position j (1 way) then (2) put any permutation of $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ in the rest of the positions ($(n-1)!$ ways).

By the multiplication rule $|E_{ij}| = 1 \cdot (n-1)! = (n-1)!$.

The answer is $(n-1)!/n! \ = \ 1/n$.

Module 7.4: Probability Properties

MCIT Online - CIT592 - Professor Val Tannen

LECTURE NOTES

Probability properties I

In this segment we give the statements of the properties, some intuition that justifies them, and some examples of putting the properties to use. Rigorous proofs of the properties appear in the segment entitled "Proofs of probability properties".

Consider an arbitrary probability space (Ω, Pr) and arbitrary events E, A, B in this space.

Property P0. $\mathrm{Pr}[E] \geq 0$

Since it's the sum of non-negative numbers.

Property P1. $\mathrm{Pr}[\Omega] = 1$

Since it adds up the probabilities of all the outcomes in the space.

Probability properties II

Property P2. If A, B are disjoint then $\Pr[A \cup B] = \Pr[A] + \Pr[B]$

This is called the **addition rule** and is analogous to the addition rule in counting applied to set cardinality: $A \cap B = \emptyset \;\Rightarrow\; |A \cup B| = |A| + |B|$.

Property P3. If $A \subseteq B$ then $\Pr[A] \leq \Pr[B]$

This is called **monotonicity** and it has an analogous property of set cardinality: $A \subseteq B \;\Rightarrow\; |A| \leq |B|$.

If $E \subseteq \Omega$ is an event then the **complement** of E is the event $\overline{E} = \Omega \setminus E$.

Property P4. $\Pr[\overline{E}] = 1 - \Pr[E]$

Some applications

Problem. We roll a pair of fair dice. Compute the probability of the event "none of the dice shows a 6".

Answer. In a previous segment we computed the probability of the event $G =$ "at least one of the dice shows a 6" to be $11/36$. The event here is $\overline{G}$, the **complement** of G. By **P4** the answer here is:

$\Pr[\overline{G}] = 1 - 11/36 = (36 - 11)/36 = 25/36.$

Problem. We roll a pair of fair dice. Compute the probability of the event "the numbers add up to 2, 3, 4, 6, 8, 10 or 12".

Answer. The event of interest is $C \cup D$ where $C =$ "the numbers add up to an even number" and $D =$" the numbers add up to 3" are **disjoint**. By **P2**:

$\Pr[C \cup D] = \Pr[C] + \Pr[D] = 1/2 + (1/36 + 1/36) = 5/9.$

More probability properties

Property P5. $\Pr[\emptyset] = 0$

By the definition of event probability, this is a sum with **no terms**! A common convention is that such a sum is 0. However, see the next activity.

Property P6. $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$

This is called (of course!) **inclusion-exclusion** for two events and is analogous to PIE for two sets.

Property P7. $\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$

This is called the **union bound**, and it plays a major role in the analysis of **probabilistic algorithms**. It's quite clear that it follows immediately from **P6**.

Activity : Two proofs of property P5

In this activity we will present two different ways of deriving:

Property P5. $\Pr[\emptyset] = 0$

from some of the other probability properties.

First proof. Recall

Property P1. $\Pr[\Omega] = 1$

and

Property P4. $\Pr[\overline{E}] = 1 - \Pr[E]$

Question. How should we choose E in P4 to help (together with P1) prove P5?

In the video, there is a box here for learners to put in an answer to the question above. As you read these notes, try it yourself using pen and paper!

ACTIVITY : Two proofs of property P5 (continued)

Answer. Choose $E = \Omega$.

Now replacing $E = \Omega$ in P4 and observing that $\overline{\Omega} = \emptyset$ we obtain

$$\Pr[\emptyset] \;=\; 1 \;-\; \Pr[\Omega]$$

Using P1 we obtain $\Pr[\emptyset] = 1 - 1 = 0$.

Second proof. In this proof we use

Property P2. If A, B are disjoint then $\Pr[A \cup B] \;=\; \Pr[A] \;+\; \Pr[B]$

ACTIVITY : Two proofs of property P5 (continued)

We wish to take $A = B = \emptyset$ in P2. We can do this if A and B are disjoint. And they are: $\emptyset \cap \emptyset = \emptyset$! This can feel surprising. Any non-empty set is **not** disjoint from itself. But the empty set is disjoint from itself!

Taking $A = B = \emptyset$ in P2 we obtain $\Pr[\emptyset \cup \emptyset] = \Pr[\emptyset] + \Pr[\emptyset]$.

Since $\emptyset \cup \emptyset = \emptyset$ we get $\Pr[\emptyset] = \Pr[\emptyset] + \Pr[\emptyset]$.

It follows that $\Pr[\emptyset] = 0$.

An application and generalizations

Problem. We roll a pair of fair dice. Compute the probability of the event "at least one of the dice shows a 6" using **P6** (inclusion-exclusion).

Answer. In the green-purple dice space define $C =$ "green die rolls 6", $D =$ "purple die rolls 6". By **P6**:

$$\Pr[C \cup D] = \Pr[C] + \Pr[D] - \Pr[C \cap D] = 1/6 + 1/6 - 1/36 = 11/36$$

Finally, we note that **P2** and **P7** generalize:

Property P2gen. If $A_1, \ldots, A_n$ are pairwise disjoint then

$$\Pr[A_1 \cup \cdots \cup A_n] = \Pr[A_1] + \cdots + \Pr[A_n]$$

Property P7gen. $\Pr[A_1 \cup \cdots \cup A_n] \leq \Pr[A_1] + \cdots + \Pr[A_n]$

Module 7.5: Birthdays, Balls and Bins

MCIT Online - CIT592 - Professor Val Tannen

Lecture Notes

The birthday "paradox" I

Problem. Suppose there are k people in a room. What is the probability that at least two people in the room have the same birthday? What is the smallest value of k for which this probability is at least $1/2$?

Answer. We set up a probability space in which the outcomes are the birthdays of the k people: sequences of length k of elements from $[1..365]$. There are 365^k such sequences. Next, we make two assumptions:

- It is equally likely for any given person to be born on any of the 365 days of the year.
- We also assume that the birthdays of the different people in the room are unrelated (they are **independent**).

Based on the intuition supported by these two assumptions, we state that our probability space is **uniform**. Each outcome has probability $1/365^k$.

The birthday "paradox" II

Answer (continued). Let E be the event that **at least** two people in the room have the same birthday.

Its complement is $\overline{E} =$"all k people have distinct birthdays". The outcomes in $\overline{E}$ are the partial permutations of k out of 365.

Therefore, using **P4**:

$$\Pr[E] = 1 - \Pr[\overline{E}] = 1 - \frac{365!/(365-k)!}{365^k} = 1 - \frac{365!}{(365-k)! \cdot 365^k}$$

Using a "big integer" calculator we find that the smallest value of k for which $\Pr[E] \geq 0.5$ is ... **23**!

Taking $k = 60$ we obtain $\Pr[E] \simeq 0.99$. Therefore, with 60 people in the room it is **almost certain** that there are two sharing the same birthday!

Balls into bins I

We have $n \geq 1$ distinguishable bins into which we throw $k \geq 0$ distinguishable balls under two assumptions:

- Each ball is equally likely to land in each of the n bins.
- The k throws are independent of each other.

As with the birthday "paradox" we assume a **uniform** probability space whose outcomes are sequences of length k of elements from $[1..n]$. There are n^k outcomes so each outcome has probability $1/n^k$.

Problem. What is the probability that ball $i \in [1..k]$ lands in bin $j \in [1..n]$?

Answer. The number of outcomes with ball i in bin j is n^{k-1}.
Therefore the probability is $n^{k-1}/n^k = 1/n$.

Balls into bins II

Problem. We throw k balls into $n \geq 2$ bins. What is the probability that Bin 1 remains empty?

Answer. Recall that in this space the outcomes are sequences of length k of elements from $[1..n]$.

Event of interest: sequences whose elements are just from $[2..n]$.

The number of such sequences is $|[2..n]|^k = (n - 2 + 1)^k = (n - 1)^k$.

Therefore the probability is

$$\frac{(n-1)^k}{n^k} = \left(1 - \frac{1}{n}\right)^k$$