Probability theory

Prof.dr.hab. Viorel Bostan

Technical University of Moldova

viorel.bostan@adm.utm.md

Lecture 2

Previous Lecture Outline

- General concept of probability;
- Sample space;
- Probability distribution;
- Axioms of probability;
- Properties of probability;
- Tables and trees to represent outcomes;
- Simple examples.

Lecture 2 Outline

- Monty Hall Problem;
- Simulation of probability;
- Uniform distribution;
- Infinite sample space;
- Geometric distribution;
- Conditional Probability;
- Product rule (why tree diagrams work).

Basic steps in solving discrete probability problems

Distinguish 4 basic steps in the solution process of discrete probability problems:

1. Find the Sample Space Ω.
2. Define Events of Interest.
3. Determine Outcome Probabilities (find distribution function $m(\omega)$).
4. Compute Event Probabilities.

Monty Hall Problem

Let's make some assumptions in order to model the game formally:

1. The car is equally likely to be hidden behind each of the three doors.
2. The player is equally likely to pick each of the three doors, regardless of the car's location.
3. After the player picks a door, the host must open a different door with a goat behind it and offer the player the choice of staying with the original door or switching.
4. If the host has a choice of which door to open, then he is equally likely to select each of them.

What is the probability that a player who switches wins the car?

Monty Hall Problem

First objective is to identify all possible outcomes of the experiment.

A typical experiment involves several randomly-determined quantities.

For example, the Monty Hall game involves three such quantities:

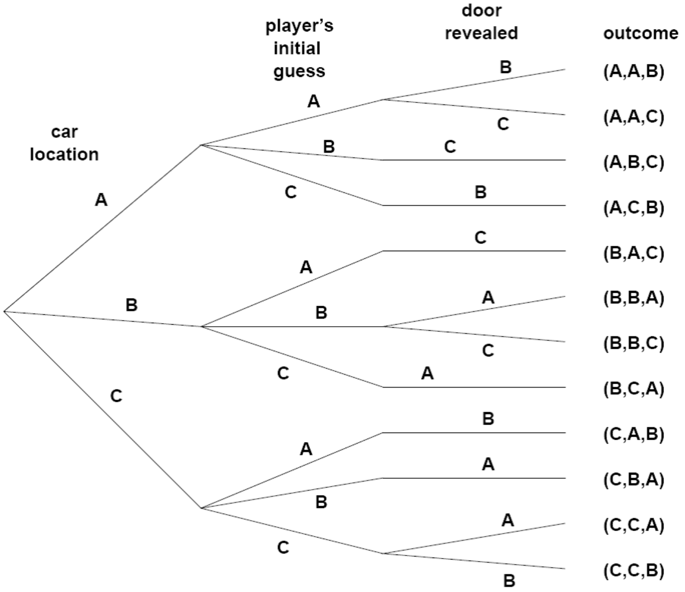
1. The door concealing the car.
2. The door initially chosen by the player.
3. The door that the host opens to reveal a goat.

Use tree diagrams. First stage indicates the car location.

For each possible location of the prize, the player could initially choose any of the three doors. Represent this in a second layer added to the tree.

Then a third layer represents the possibilities of the final step when the host opens a door to reveal a goat.

MHP. Step1: Find sample space



MHP. Step 2: Define events of interest

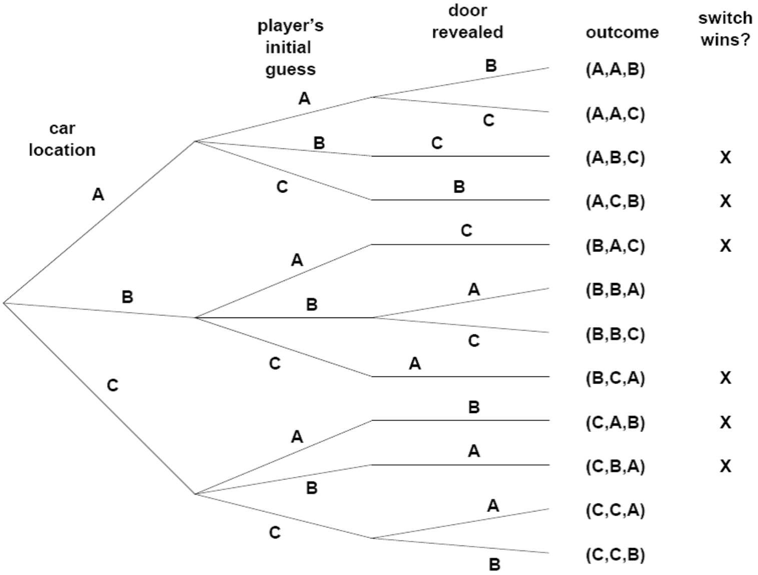
Objective is to answer questions of the form **"What is the probability that . . . ?"**, where the missing phrase might be:

- **"the player wins by switching"**,
- **"the player initially picked the door concealing the prize"**,
- **"the prize is behind door C"**.

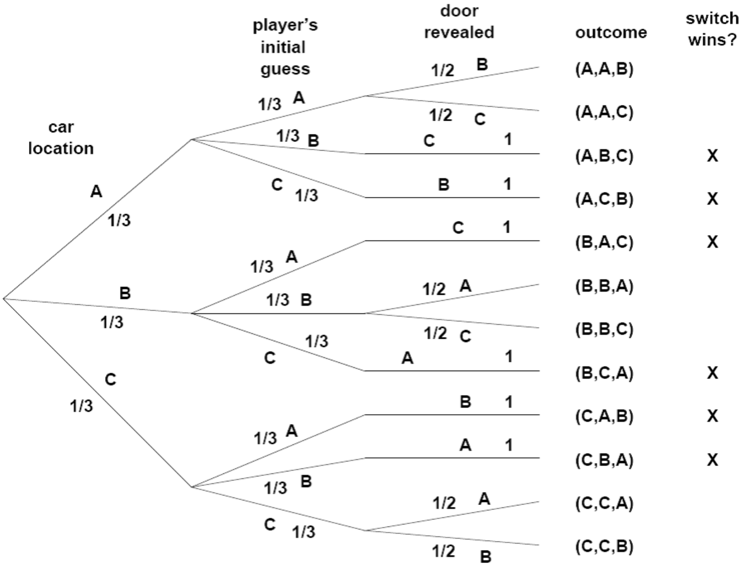
Each of these phrases characterizes a set of outcomes (events):

$$
\begin{aligned}
\text{"prize is behind door C"} &= \big\{(C, A, B), (C, B, A), (C, C, A), (C, C, B)\big\}; \\
\text{"player initially picked the door with prize"} &= \big\{(A, A, B), (A, A, C), (B, B, A), \\
&\qquad (B, B, C), (C, C, A), (C, C, B)\big\}; \\
\text{"player wins by switching"} &= \big\{(A, B, C), (A, C, B), (B, A, C), \\
&\qquad (B, C, A), (C, A, B), (C, B, A)\big\}.
\end{aligned}
$$

MHP. Step 2: Define events of interest



MHP. Step 3: Determine the outcomes probabilities



MHP. Step 3: Determine the outcome probabilities

Next convert edge probabilities into outcome probabilities.

This is a purely mechanical process: the probability of an outcome is equal to the product of the edge-probabilities on the path from the root to that outcome.

For example,

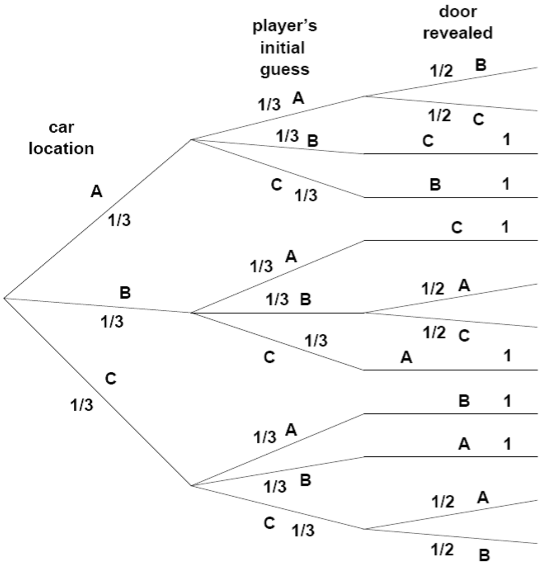
$$P\big((A, A, B)\big) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{18}.$$

There's an easy, intuitive justification for this rule. As the steps in an experiment progress randomly along a path from the root of the tree to a leaf, the probabilities on the edges indicate how likely the walk is to proceed along each branch.

For example, a path starting at the root in our example is equally likely to go down each of the three top-level branches.

Repeat this process for every outcome from the sample space.

MHP. Step 3: Determine the outcome probabilities



outcome	switch wins?	probability
(A,A,B)		1/18
(A,A,C)		1/18
(A,B,C)	X	1/9
(A,C,B)	X	1/9
(B,A,C)	X	1/9
(B,B,A)		1/18
(B,B,C)		1/18
(B,C,A)	X	1/9
(C,A,B)	X	1/9
(C,B,A)	X	1/9
(C,C,A)		1/18
(C,C,B)		1/18

MHP. Step 4: Compute event probability

$$
\begin{aligned}
P\{\text{switching wins}\} &= P\{(A, B, C), (A, C, B), (B, A, C), \\
&\qquad (B, C, A), (C, A, B), (C, B, A)\} \\
&= P\big((A, B, C)\big) + P\big((A, C, B)\big) + P\big((B, A, C)\big) \\
&\quad + P\big((B, C, A)\big) + P\big((C, A, B)\big) + P\big((C, B, A)\big) \\
&= \tfrac{1}{9} + \tfrac{1}{9} + \tfrac{1}{9} + \tfrac{1}{9} + \tfrac{1}{9} + \tfrac{1}{9} = \tfrac{2}{3}.
\end{aligned}
$$

Correct answer is: player who switches doors wins the car with probability $\frac{2}{3}$.

In contrast, a player who stays with his or her original door wins with probability $\frac{1}{3}$, since staying wins if and only if switching loses.

Simulation of probability

Goal: to simulate on a computer a probabilistic experiment.

Example

Simulate an experiment with three possible outcomes $\Omega = \{\omega_1, \omega_2, \omega_3\}$ such that

$$m(\omega_1) = \tfrac{1}{2},\ m(\omega_2) = \tfrac{1}{3},\ m(\omega_3) = \tfrac{1}{6}.$$

If we have a die, such an event can be simulated by marking three faces of a six-sided die with ω_1, two faces with ω_2 and one face with ω_3.

How to simulate this experiment on a computer?

Find a computer analog of rolling a die. This is done on the computer by using a **random number generator**. Actually, it is a pseudo-random number generator.

If you type in MATLAB/Octave

```
>> x=rand;
```

then x will be a random number between 0 and 1.

Simulation of probability

Example

```
>> x=rand(2);
```

$$\begin{bmatrix} 0.3167 & 0.1175 \\ 0.6612 & 0.8711 \end{bmatrix}_{2\times 2},$$

```
>> rand(3,4)
```

$$\begin{bmatrix} 0.2785 & 0.9649 & 0.9572 & 0.1419 \\ 0.5469 & 0.1576 & 0.4854 & 0.4218 \\ 0.9575 & 0.9706 & 0.8003 & 0.9157 \end{bmatrix}_{3\times 4},$$

```
>> a=rand(1,6)
```

$$a = \begin{bmatrix} 0.7922 & 0.9595 & 0.6557 & 0.0357 & 0.8491 & 0.9340 \end{bmatrix}_{1\times 6},$$

```
>> c=rand(3,1)
```

$$c = \begin{bmatrix} 0.6787 \\ 0.7548 \\ 0.2431 \end{bmatrix}_{3\times 1}$$

Simulation of probability

What if we need a random real number in an arbitrary interval $[a, b]$?

In this case the the following MATLAB/Octave commands do the job:

```
>> x=rand;
>> y=(b-a)*x+a;
```

Clearly, if $x \in [0, 1]$, then $y = (b - a)x + a \in [a, b]$.

The following command is even better:

```
>> y=(b-a)*rand+a;
```

How about not a real number, but an integer?

Want to simulate tossing a die $\Rightarrow$ need a random variable with values from $\{1, 2, 3, 4, 5, 6\}$.

```
>> x=fix(6*rand)+1;
```

If $0 \leqslant r < 1$, then $0 \leqslant 6r < 6$, and consequently $[6r] + 1 \in \{1, 2, 3, 4, 5, 6\}$.

Note that the probability of any of six outcomes is the same: $m(i) = \frac{1}{6}$, $i = 1, 2, \ldots, 6$.

Simulation of probability

Example

Need to simulate an experiment with 3 possible outcomes$\Omega = \{\omega_1, \omega_2, \omega_3\}$ such that

$$m(\omega_1) = \tfrac{1}{2}, m(\omega_2) = \tfrac{1}{3}, m(\omega_3) = \tfrac{1}{6}.$$

Let $x = [6r] + 1$, with $r \in [0, 1]$ a random number.

If $x \in \{1, 2, 3\}$, then ω_1 has happened; If $x \in \{4, 5\}$, then ω_2; If $x \in \{6\}$, then ω_3.

Example (Coin Tossing)

Want to simulate coin tossing. Intuition suggests that the probability of obtaining a head on a single toss of a coin is $1/2$.

To have the computer toss a coin, can pick a random real number $r \in [0, 1]$.

If $r < 1/2$ call the outcome **heads (H)**; if not, call it **tails (T)**.

Another way to proceed would be to ask the computer to pick a random integer from set $\{0, 1\}$, and let **H** be 0, and **T** be 1.

Simulation of probability

Example (Coin Tossing (contd))

Running a simulation program 20 times results for ex. in:

$$HHHTTTHTTTTHTTTTTHTT.$$

Note that in 20 tosses, 6 heads and 14 tails have occurred.

Run it again, and results most probably will be different.

Toss a coin n times, where n is much larger than 20, and see if the proportion of heads is closer to intuitive guess of $1/2$.

The simulation program will keep track of the number of heads.

Run program with $n = 1000$, to obtain 489 heads. Proportion of heads is $\frac{489}{1000} = 0.489$.
Run it with $n = 10000$, to get 5067 heads. Proportion of heads is $\frac{5067}{10000} = 0.5067$.

This illustrates the frequency interpretation of probability.
The **Law of Large Numbers** will be studied later in this course.

Example (Heads or Tails Game)

Two friends, Petru and Paul play a game called "Heads or Tails".
In this game, a fair coin is tossed 40 times. Each time a head comes up Petru wins 1 cent from Paul, and each time a tail comes up Petru loses 1 cent to Paul.

For example, the results of one particular game with 40 coin tosses is

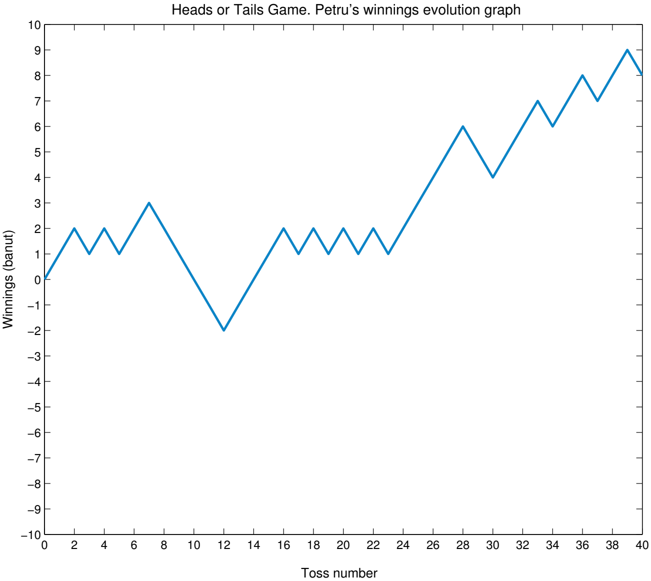
$$HHTHTHHTTTTTHHHHTHTHTHTHHHHHTTHHHTHHTHHT$$

with Petru winning 8 cents at the end of the game.

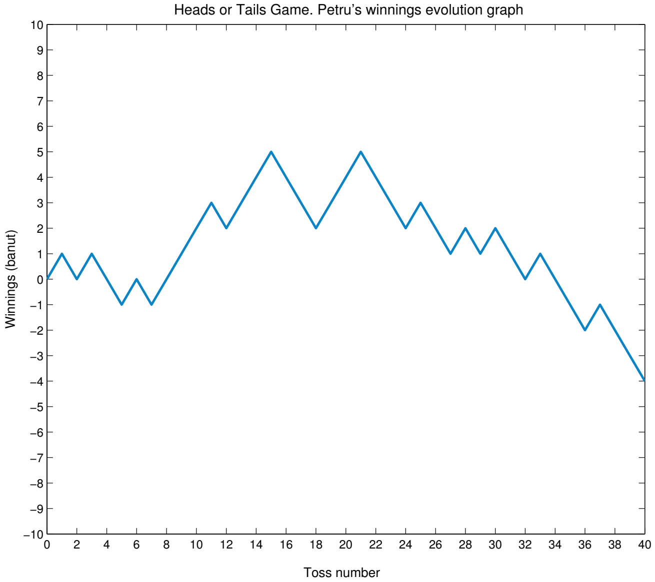
Clearly, for another run, the results will be different.

We can plot the dynamics (winning's evolution) of such a game. (Do it yourself!)

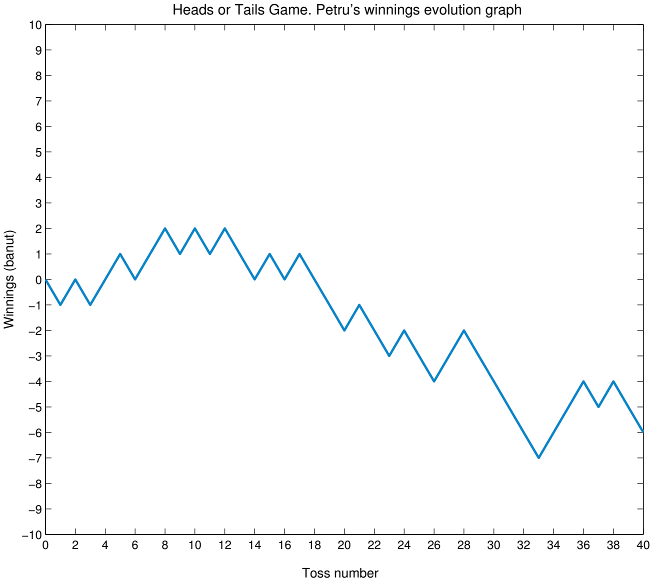
Heads or Tails Game



Heads or Tails Game



Heads or Tails Game



It is natural to ask for the probability that Petru will win j cents.
Let random variable $W =$ amount of cents Petru won in one game. What is $P(W = j)$?

Here, j could be any even number (Why even?) from -40 to 40. It is reasonable to guess that the value of j with the highest probability is $j = 0$, since this occurs when the number of heads equals the number of tails.

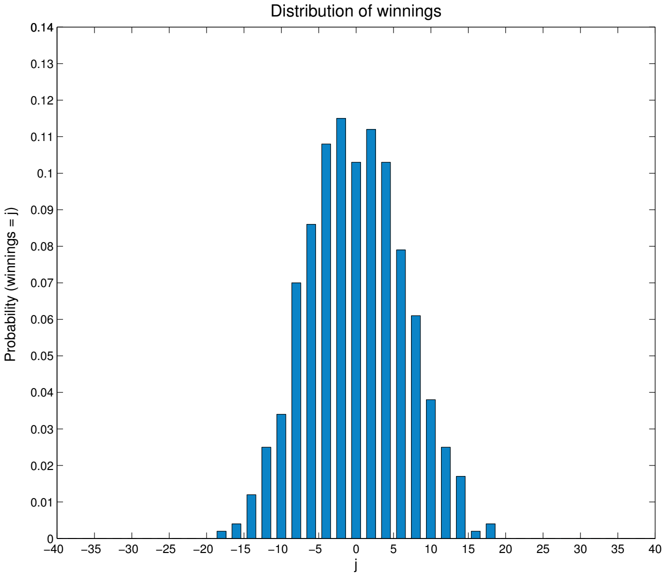
Similarly, we would guess that the values of j with the lowest probabilities are $j = \pm 40$. We would like to see if our intuition is correct (this time).

Simulate the game a large number of times and keep track of the number of times that Petru's final winnings are j. The proportions over the number of games played give estimates for the corresponding probabilities.

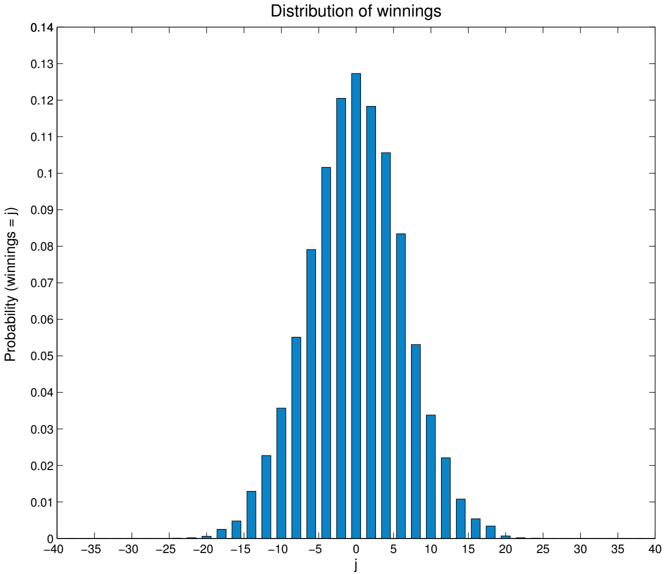
For example, if in 10000 games Petru has won six cents 754 times, then
$P(W = 6) \approx \frac{754}{10000} = 0.0754$.

Count the number of times Petru has won j cents and plot these results to get the **distribution** of probabilities.

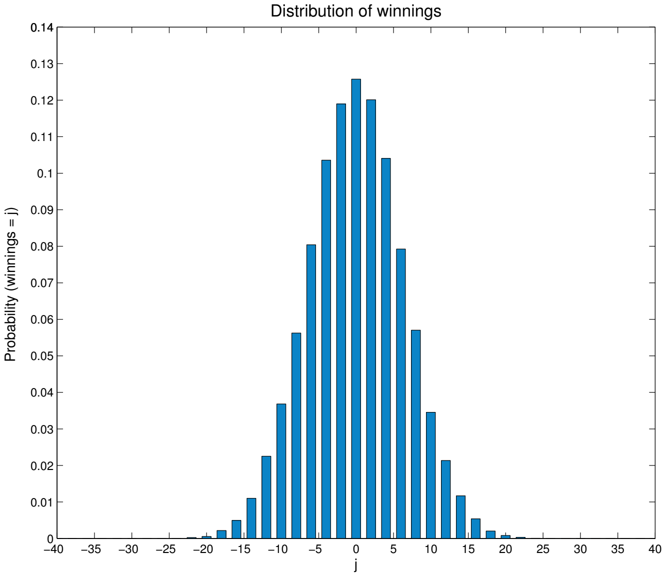
Heads or Tails Game: 1000 games



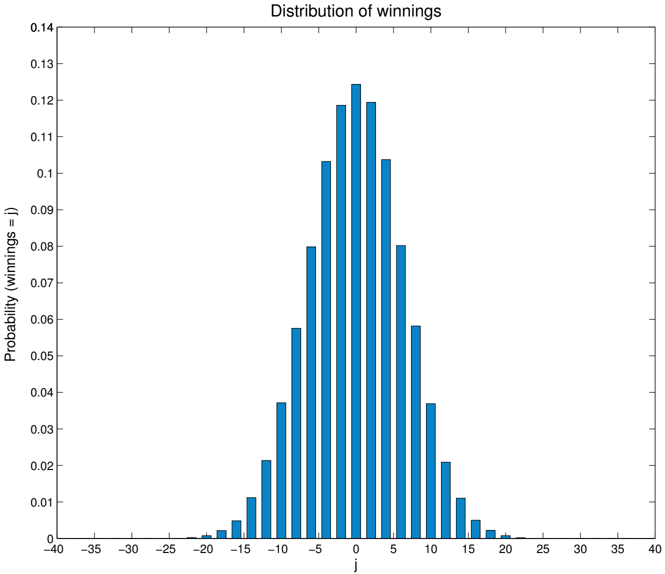
Heads or Tails Game: 10000 games



Heads or Tails Game: 50000 games



Heads or Tails Game: 1000000 games



Such graphs we call bar graphs.

The vertical line, or bar, at position j on the horizontal axis, has a height equal to the proportion of outcomes which equal j.

Such a graph can be plotted using the MATLAB command: `bar`

Note that the more games are played the "smoother" will be the resulting curve above the bars.

Intuition about Petru's final winnings was quite correct.

The highest probability will have the event consisting in Petru winning 0 cents, while the lowest probability (close to 0) will have the event *"Petru's winnings are* $\pm 40, \pm 38, \pm 36$", with such events corresponding to almost all of the 40 tosses being of the same type, which intuitively is highly not probable.

Uniform distribution

Definition

The **uniform distribution** on a sample space Ω containing n elements is the function m defined by

$$m(\omega) = \frac{1}{n}$$

for every $\omega \in \Omega$.

Uniform distribution assigns to each outcome the same probability.

Important!

When an experiment is analyzed to describe its possible outcomes, there is no single correct choice of sample space. In other words, for the same experiment we can have several choices for sample space.
But keep in mind that distribution functions can be different!

Uniform distribution

Example

Consider the experiment of tossing a coin twice. Sample space is $\Omega_1 = \{HH, HT, TH, TT\}$ and assign the uniform distribution function m:

$$m(HH) = m(HT) = m(TH) = m(TT) = \frac{1}{4}.$$

On the other hand, for some purposes, it may be more useful to consider the 3−element sample space for the same experiment $\Omega_2 = \{0, 1, 2\}$ in which i is the outcome i heads turn up. The distribution function m_2 on Ω_2 defined by :

$$m_2(0) = \frac{1}{4}, \quad m_2(1) = \frac{1}{2}, \quad m_2(2) = \frac{1}{4}$$

by definition is a distribution on Ω_2, but it is not uniform.
It is possible to choose another distribution function $\widetilde{m}$ on Ω_2 that will be uniform: $\widetilde{m}(i) = \frac{1}{3}$. Although $\widetilde{m}$ is a perfectly good **uniform** distribution function, it is not consistent with observed data on coin tossing. So, it is not appropriate for the considered experiment.

Uniform distribution

Example

Consider the experiment that consists of rolling a pair of dice.

$$\Omega = \{(i,j) \mid 1 \leqslant i,j \leqslant 6, i,j \in \mathbb{Z}\}$$

Sample space consists of 36 different outcomes.

Assume that the dice are not loaded. In mathematical terms, this means we adopt the uniform distribution function on Ω:

$$m((i,j)) = \frac{1}{36}, \quad 1 \leqslant i,j \leqslant 6.$$

Interested in answering the questions:
What is the probability of getting a sum of 8 on the roll of two dice?

What is the probability of getting neither snake eyes (double ones) nor boxcars (double sixes)?

Example (Contd.)

For the first event E (sum of 8),

$$\begin{aligned} P(E) &= P\big((2,6),(6,2),(3,5),(5,3),(4,4)\big) \\ &= m((2,6)) + m((6,2)) + m((3,5)) + m((5,3)) + m((4,4)) \\ &= 5 \cdot \tfrac{1}{36} = \tfrac{5}{36}. \end{aligned}$$

For the second event F, consider the complement $F^c =$ "get either $(1,1)$ or $(6,6)$":

$$\begin{aligned} P(F^c) &= P\big((1,1),(6,6)\big) \\ &= P\big((1,1)\big) + P\big((6,6)\big) = \tfrac{2}{36} \end{aligned}$$

and thus $P(F) = 1 - P(F^c) = 1 - \frac{1}{18} = \frac{17}{18}$.

Determination of probabilities

Until now, we assigned an equal probability to each outcome, i.e. the uniform distribution function was chosen. These are the natural choices provided the coin is a fair one and the dice are not loaded.

How to decide which probability distribution should be used in practice?

One way is by **symmetry**.

For example, in the case of the coin toss, there are no physical difference between the two sides of a coin that can affect the chance of one side or the other turning up.

Similarly, with an ordinary die there is no essential difference between any two sides of the die, and so by symmetry we assign the same probability for any possible outcome.

In general, considerations of symmetry often suggest the uniform distribution function.

Determination of probabilities

ATTENTION:

We should not always assume that, just because we do not know any reason to suggest that one outcome is more likely than another, it is appropriate to assign equal probabilities.

Consider the experiment of guessing the sex of a newborn child.

Your intuition induced by symmetry argument may say that the probability of a boy being born is the same as that of a girl, that is:

$$P(Boy) = 0.5 \quad and \quad P(Girl) = 0.5$$

In reality, it is more appropriate to assign a distribution function which assigns

$$P(Boy) = 0.513 \quad and \quad P(Girl) = 0.487$$

Infinite sample spaces

If a sample space Ω has an infinite number of points, then the way that a distribution function is defined depends upon whether or not the sample space is **countable**. If

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_{2020}, \omega_{2021}, \ldots\}$$

is a countably infinite sample space, then we are in **discrete** probability case.

Example

Consider the experiment of tossing a coin until the first head will turn up. The sample space in this case consists of the outcomes

$$\Omega = \{H, TH, TTH, TTTH, TTTTH, TTTTTH, \ldots\}.$$

Clearly, it is a infinite countable set.

So, if Ω is countable, then a distribution function m is defined exactly as previously, except that the sum $\sum_{\omega \in \Omega} m(\omega) = 1$ must now be a **convergent infinite sum**.

All results considered until now are true in this case.

Infinite sample spaces

One thing impossible on a **countably infinite sample space** (but possible on a finite sample space) is to define a uniform distribution function.

Indeed, if $\Omega = \{\omega_1, \omega_2, \omega_3, \dots\}$ and uniform distribution is to be used:

$$m(w_i) = \alpha, \quad \forall i \in \mathbb{N},$$

then, by definition of distribution function we should have:

$$1 = \sum_{\omega \in \Omega} m(\omega) = \sum_{\omega \in \Omega} \alpha.$$

On the other hand, no matter how small is α, the infinite sum of $\alpha' s$ will be always ∞.

For example, $\alpha = 0.000001 = \frac{1}{100000}$, a sum of one million of such $\alpha' s$ will be 10.

No uniform distribution on an infinite sample space!.

If the sample space is infinite, but NOT COUNTABLE (like $\mathbb{R}$), then it is a continuous probability case and other approach should be used.

Infinite Sample Space

Consider the experiment of tossing a coin until the first head appears:
$\Omega = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$.

$$\begin{aligned} m(H) &= \tfrac{1}{2}, \\ m(TH) &= \tfrac{1}{4}, \\ m(TTH) &= \tfrac{1}{8}, \\ m(TTTH) &= \tfrac{1}{16}, \\ &\cdots\cdots\cdots\cdots \\ m(\underbrace{TTT\ldots T}_{k-1}H) &= \tfrac{1}{2^k}. \end{aligned}$$

$$\sum_{\omega\in\Omega} m(\omega) = \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \tfrac{1}{16} + \cdots + \tfrac{1}{2^k} + \cdots = 1.$$

Since **geometric series** $q + q^2 + q^3 + \cdots + q^k + \cdots = \frac{q}{1-q}$, if $-1 < q < 1$.
This distribution is so-called **geometric distribution**.

Infinite Sample Space

Additivity Axiom

If $E \cap F = \emptyset$, then $P(E \cup F) = P(E) + P(F)$.

Theorem

If $A_1, \ldots, A_n$ are pairwise disjoint subsets of Ω, then

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Strong Additivity Axiom

If $A_1, \ldots, A_n, \ldots$ are (infinitely many) pairwise disjoint subsets of Ω, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Suppose we assign a distribution function to a sample space and **then** learn that an event E has occurred.

How should we change the probabilities of the remaining events?

We call the new probability for an event F the **conditional probability** of F given E and denote it by $P(F \mid E)$.

Example 1. An experiment consists of rolling a die once. Let X be the outcome. Let $F = \{X = 6\}$, and $E = \{X > 4\}$.

Assign the distribution function $m(\omega) = 1/6$ for $\omega = 1, 2, ..., 6$. Thus, P(F)=1/6.

The die is rolled and we are told that the event E has occurred.
This leaves only two possible outcomes: 5 and 6.

In the absence of any other information, we would still regard these outcomes to be equally likely, so now F has probability $1/2$, in other words, $P(F \mid E) = 1/2$.

Example 2. One finds that in a population of $100,000$ females:

- 89.835% can expect to live to age 60,
- 57.062% can expect to live to age 80.

Given that a woman is 60, what is the probability that she lives to age 80?

The initial sample space can be thought of as a set of $100,000$ females.

Event $E = \{\text{women who live at least 60 years}\}$,
Event $F = \{\text{women who live at least 80 years}\}$.
The size of E is $89,835$, and the size of F is $57,062$.

Now, suppose event E **has happened**. We consider E to be the new sample space, and note that F is a subset of E.
So, the probability in question equals

$$\frac{57,062}{89,835} = 0.6352.$$

Thus, a woman who is 60 has a 63.52% chance of living to age 80.

Discrete conditional probability

Let $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ be a sample space with distribution function $m(\omega_j)$ assigned.

Suppose we learn that the event E has occurred.

We want to assign a new distribution function $m(\omega_j|E)$ to to reflect it.

Clearly, if a sample point ω_j is not in E, we want $m(\omega_j|E) = 0$.

Moreover, it is reasonable to assume that the probabilities for ω_j in E should have the same relative sizes that they had before we learned that E had occurred. We require that

$$m(\omega_j|E) = c \cdot m(\omega_j), \quad \forall \omega_j \in E,\, c > 0.$$

But, we must also have

$$1 = \sum_{\omega_j \in E} m(\omega_j|E) = \sum_{\omega_j \in E} c \cdot m(\omega_j) = c \sum_{\omega_j \in E} m(\omega_j) = c \cdot P(E).$$

$$c = \tfrac{1}{P(E)}.$$

Note that this requires us to assume that $P(E) > 0$.

Discrete conditional probability

Therefore, we define

$$m(\omega_j|E) = c \cdot m(\omega_j = \frac{m(\omega_j)}{P(E)}, \quad \text{for all } \omega_j \in E.$$

We will call this new distribution the **conditional distribution given event** E.
For a general event F, this gives

$$\begin{aligned} P(F|E) &= \sum_{\omega_j \in F \cap E} m(\omega_j|E) = \sum_{\omega_j \in F \cap E} \frac{m(\omega_j)}{P(E)} \\ &= \frac{1}{P(E)} \sum_{\omega_j \in F \cap E} m(\omega_j) = \frac{P(F \cap E)}{P(E)}. \end{aligned}$$

Conditional probability of F **given that** E **occurs** is

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

Rolling a die

Example

Return to the example of rolling a die.
Recall that F is the event $X = 6$, and E is the event $X > 4$.
Note that $E \cap F$ is the event F. So, the above formula gives

$$P(F \mid E) = \frac{P(F \cap E)}{P(E)} = \frac{1/6}{1/3} = \frac{1}{2}.$$

which is in agreement with previous arguments.

Conditional Probability

Probability of event F given event E is

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

Problem

We have two urns, I and II.

Urn I contains 2 black balls and 3 white balls.

Urn II contains 2 black balls and 1 white ball.

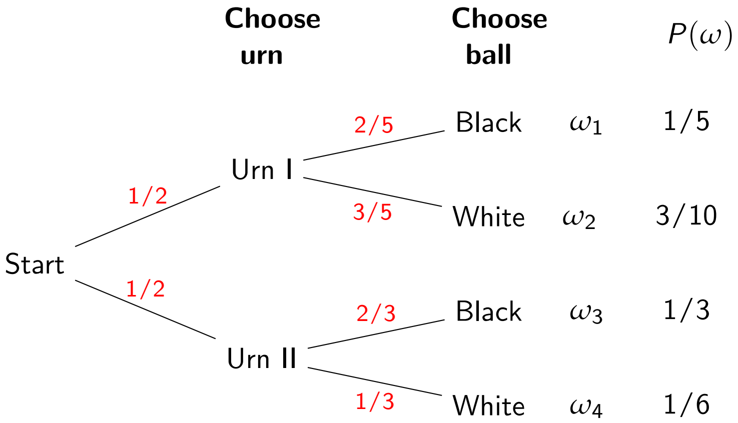
An urn is drawn at random and a ball is chosen at random from it.

Find sample space Ω. Represent the sample space of this experiment as the paths through a tree. Then assign probabilities to the paths (outcomes).

Discrete conditional probability

Urn I: 2 black balls $+3$ white balls;

Urn II: 2 black balls $+1$ white ball.

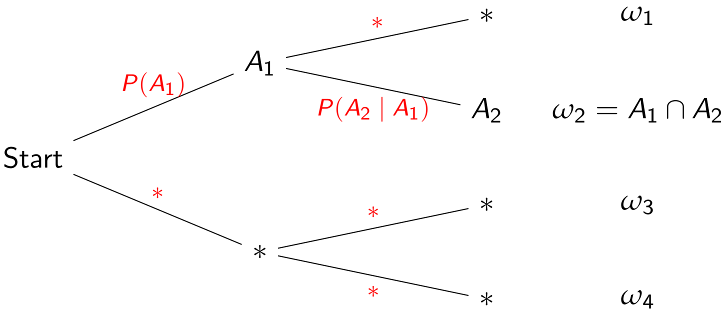


Why Tree Diagrams Work?

Theorem (Product rule for 2 events)

If $P(A_2) \neq 0$*, then:* $P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1)$.

Proof follows directly from definition of conditional probability: $P(A_2 \mid A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$.

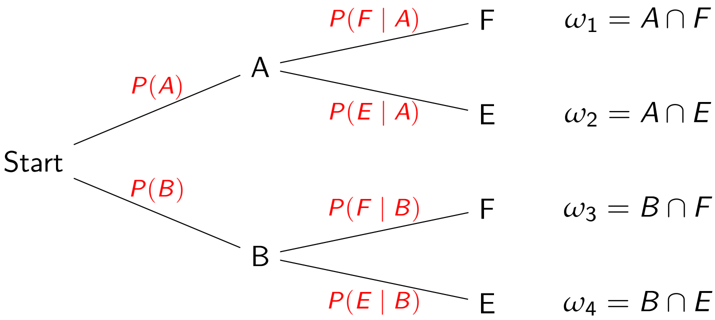


Why Tree Diagrams Work?

Theorem (Product rule for 2 events)

If $P(A_2) \neq 0$, then: $P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1)$.

Proof follows directly from definition of conditional probability: $P(A_2 \mid A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$.



Why Tree Diagrams Work?

Theorem (Product rule for n events)

If $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \neq 0$, then:

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \ldots$$
$$\ldots P(A_n \mid A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

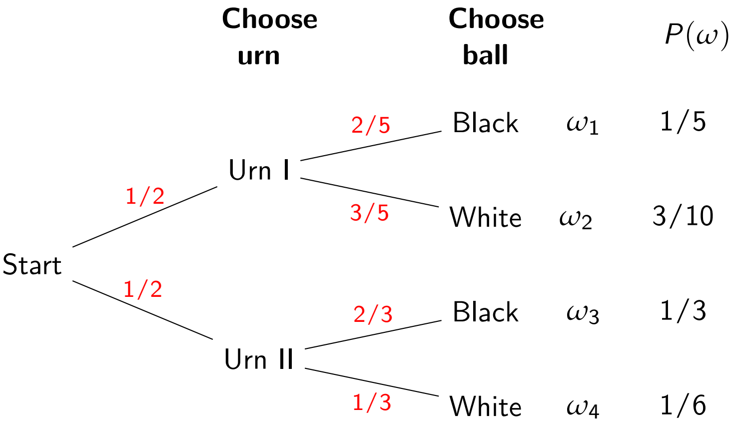
Compute the probability that an experiment traverses a particular "root to leaf" path of length n.

Let A_i be event that the experiment traverse the $i-$th edge of the path.

Then $A_1 \cap \ldots \cap A_n$ is event that the experiment traverse the whole path.

The Product Rule says that the probability of this is the probability that the experiment takes the first edge times the probability that it takes the second, given it takes the first edge, times the probability it takes the third, given it takes the first two edges, and so forth.

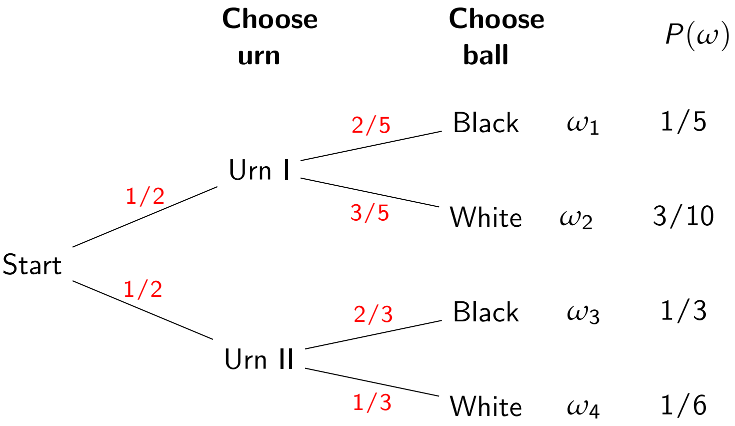
Discrete conditional probability



$$P(Black \mid \text{Urn I}) = \frac{2}{5}, \qquad P(White \mid \text{Urn I}) = \frac{3}{5},$$

$$P(Black \mid \text{Urn II}) = \frac{2}{3}, \qquad P(White \mid \text{Urn I}) = \frac{1}{3}.$$

Discrete conditional probability



$$P(Black) = P(\omega_1) + P(\omega_3) = \frac{1}{5} + \frac{1}{3} = \frac{8}{15},$$
$$P(White) = P(\omega_2) + P(\omega_4) = \frac{3}{10} + \frac{1}{6} = \frac{7}{15}.$$

Lecture 2 Summary

- Monty Hall Problem;
- Simulation of probability;
- Uniform distribution;
- Infinite sample space;
- Geometric distribution;
- Conditional Probability;
- Product rule (why tree diagrams work).

There will be a quiz next time. Don't be late!

Homework 1 posted on course page, Due ....

Probability Joke

A lottery is a tax on people ...

who don't know probability!