Equally Likely outcomes

For many experiements it is reasonable to assume that all possible outcomes are equally likely. For example:

- Draw a random sample of size n from a population. The assumption that the sample is drawn at random means that all samples of size n have an equal chance of being chosen (much of statistical analysis depends on the assumption that samples are chosen randomly).
- Flip a fair coin n times and observe the sequence of heads and tails that results.
- Roll n dice, die 1, die 2, die 3, $\ldots$, die n, and observe the ordered sequence of numbers on the uppermost faces.

Equally Likely outcomes

Equally Likely Outcomes

For any sample space with N equally likely outcomes, we assign the probability $\frac{1}{N}$ to each outcome.

Example Experiment: Flip a fair coin. The sample space for this experiment has two equally likely outcomes: $S = \{H, T\}$. Assign probabilities to these outcomes.

Example Experiment: Flip a fair coin twice and record the sequence of Heads and tails. Each of the four outcomes: {HH, HT, TH, TT } have the same probability. What is the probability that TT is the outcome of this experiment?

Equally Likely outcomes

If E is an **event** in a sample space, S, with N **equally likely (simple) outcomes,** the probability that E will occur is the sum of the probabilities of the outcomes in E, which gives

$$\boxed{\mathbf{P}(E) = \frac{\text{the number of outcomes in } E}{\text{the number of outcomes in S}}} = \frac{n(E)}{n(S)} = \frac{n(E)}{N}$$

Notice that this formula displays the probability as the quotient of the answers to two counting problems.

Equally Likely outcomes

Example A pair of fair six sided dice, one red and one green, are rolled and the pair of numbers on the uppermost face is observed. We record red first and then green. The sample space for the experiment is shown below:

Equally Likely outcomes

$$\text{Sample Space:} = \left\{\begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix}\right\}$$

(a) Let E be the event that the numbers observed add to 7. What is the probability of E?

$$\mathrm{E} = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$$

$n(S) = 36$, $n(E) = 6$ so $\mathbf{P}(E) = \dfrac{6}{36}$.

Equally Likely outcomes

(b) Let F be the event that the numbers observed add to 11. List the elements of the set F and calculate $\mathbf{P}(F)$.

$F = \{(5,6),(6,5)\}$; $\mathbf{P}(F) = \dfrac{2}{36}$.

(c) Let G be the event that the numbers on both dice are the same. What is $\mathbf{P}(G)$?

$G = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$; $\mathbf{P}(G) = \dfrac{6}{36}$.

Equally Likely outcomes

Example A pair of fair dice, one six sided and one four sided are rolled and the pair of numbers on the uppermost face is observed. The sample space shown below has equally likely outcomes. Calculate the probability of the event "The numbers observed add to 4".

$$\text{Sample Space:} = \begin{matrix} \{(1,1) & (1,2) & (1,3) & (1,4) \\ (2,1) & (2,2) & (2,3) & (2,4) \\ (3,1) & (3,2) & (3,3) & (3,4) \\ (4,1) & (4,2) & (4,3) & (4,4) \\ (5,1) & (5,2) & (5,3) & (5,4) \\ (6,1) & (6,2) & (6,3) & (6,4)\} \end{matrix}$$

$E = \{(1,3),(2,2),(3,1)\}$ so $n(E) = 3$ and $n(Sample\ Space) = 4 \cdot 6 = 24$. Hence $\mathbf{P}(E) = \dfrac{3}{24}$.

Random Selection or Random Outcomes

When we say that outcomes are selected randomly, it implies (by definition) that individual outcomes in the sample space are equally likely. For example, if we say we drew a random sample of size n from a population, we are assuming that our selection process gave all samples of size n an equal chance of being drawn.

If we know the composition of a population, it's fairly easy to calculate probabilities when we draw a random sample of size one from the population. For bigger sample sizes things are more complicated, and the time we invested in learning to count pays off.

Test your intuition!

Each person in the room today selects a whole number between 1 and 100, at random (with no collaboration!).

How likely is it that at least two people pick the same number?

A: less than 20%

B: close to 40%

C: close to 60%

D: more than 80%

Probability

If I take a sample of size K from an bag with N objects, the probability that the sample is a sample of Type X is

$$\frac{\text{The number of samples of Type X}}{\text{The total number of samples}} =$$

$$\frac{\text{The number of samples of Type X}}{\mathbf{C}(N, K)}$$

Probability

Example Suppose I have a bag containing twelve numbered marbles, 8 of which are red and 4 of which are white. If I take a sample of two marbles (observing number and color) from the bag,

(a) What is the probability of getting two red marbles?

The "total number of samples" is $\mathbf{C}(12,2) = \dfrac{12 \cdot 11}{2} = 66.$
The "number of samples with 2 red marbles" is
$\mathbf{C}(8,2) = \dfrac{8 \cdot 7}{2} = 28.$
The probability that a sample of two marbles are both red is $\dfrac{28}{66} = 0.4242\cdots.$

Probability

(b) What the probability of getting one red and one white marble in the sample?

The "total number of samples" is still $\mathbf{C}(12, 2) = 66$.
The "number of samples with 1 red marble and 1 white marble" is $\mathbf{C}(8, 1) \cdot \mathbf{C}(4, 1) = 32$.
The probability that a sample of two marbles is 1 red and 1 white is $\frac{32}{66} = 0.4848\cdots$.

Probability

Example (The Hoosier Lottery) When you buy a Powerball ticket, you select 5 different white numbers from among the numbers 1 through 59 (order of selection does not matter), and one red number from among the numbers 1 through 35. What is the probability that your selection will be the winning one?

The "total number of samples" is
$\mathbf{C}(59, 5) \cdot \mathbf{C}(35, 1) = 175, 223, 510.$
The "number of samples of Type X" in this case is 1.
Your probability of winning the lottery is
$\dfrac{1}{175, 223, 510} \approx 5.7 \cdot 10^{-9}.$

Probability

Example A `poker hand` is dealt fairly (randomly). What is the probability of a hand with three cards from one denomination and two from another (a full house)?

The "total number of samples" is $\mathbf{C}(52, 5) = 2,598,960$. One way to count the "number of samples of Type X" is to first pick the two denominations. Since one of them will have 3 cards and the other 2 the two denominations can be distinguished so you can do this $\mathbf{P}(13, 2)$ ways. Then you can pick 3 cards from the first denomination in $\mathbf{C}(4, 3)$ ways and 2 cards from the second denomination in $\mathbf{C}(4, 2)$ ways. Hence "number of samples of Type X" is
$\mathbf{P}(13, 2) \cdot \mathbf{C}(4, 3) \cdot \mathbf{C}(4, 2) = 3,744$.
Hence the probability that a hand with three cards from one denomination and two from another is dealt is
$$\frac{3,744}{2,598,960} = 0.00144\ldots.$$

Probability

You can do a little research on the probabilities of all types of poker hands here

Example (more respectable than lotteries & poker) A box ready for shipment contains 100 light bulbs, 10 of which are defective. The quality control test is to take a random sample of 5 light bulbs, without replacement, from the box. If one is defective, the box will not be shipped. What is the probability that the box will be shipped?

The "total number of samples" is $\mathbf{C}(100, 5) = 75,287,520$.
The "number of samples of Type X" is
$\mathbf{C}(90, 5) = 43,949,268$.
The probability that the box will be shipped is
$\dfrac{43,949,268}{75,287,520} \approx 0.584$.

Probability

Example A coin is flipped 4 times and the sequence of heads and tails is recorded. All of these sequences are equally likely.

(a) How many elements are there in this sample space?
$2^4 = 16.$

(b) How many outcomes with exactly 3 heads?
$\mathbf{C}(4,3) = 4.$

(c) Let E be the event "we get exactly 3 heads", what is $\mathbf{P}(E)$?

$$\mathbf{P}(E) = \frac{4}{16} = 0.25.$$

The Complement rule

If E is an event in a sample space S, then
$n(E) + n(E') = n(S)$, therefore

$$\frac{n(E)}{n(S)} + \frac{n(E')}{n(S)} = \frac{n(S)}{n(S)} = 1.$$

Thus we have **the complement rule:**

$$\mathbf{P}(E) + \mathbf{P}(E') = 1 \quad \text{or} \quad \boxed{\mathbf{P}(E') = 1 - \mathbf{P}(E)}$$

Note: If we define "success" to be the event that we get an outcome in E and "failure to be the event that we do not get an outcome in E, then $\mathbf{P}(\text{success}) = 1 - \mathbf{P}(\text{failure})$. (We will use this terminology later when studying the Binomial distribution.)

The Complement rule

Example Flip a coin 10 times and observe the sequence of heads and tails.

(a) How many outcomes are in this sample space?

$2^{10} = 1,024.$

(b) What is the probability that you observe 5 heads?

$$\frac{\mathbf{C}(10,5)}{2^{10}} = \frac{252}{1,024} \approx 0.24609375$$

(c) What is the probability that you will observe at least one tail?

$1 -$ the probability that you will observe no tails.
The number of outcomes with no tails is $\mathbf{C}(10,0) = 1$.
The probability that you will observe at least one tail is
$1 - \dfrac{1}{1024} \approx 0.99902.$

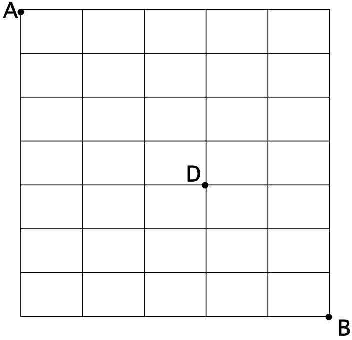
The Complement rule

(d) What is the probability that we will observe at least two heads?

$1-$ the probability that you will observe 0 or 1 heads.
The number of outcomes with no heads is 1.
The number of outcomes with one head is $\mathbf{C}(10, 1) = 10$.
The probability that we will observe at least two heads is
$1 - \dfrac{10+1}{1024} \approx 0.9892$.

Example (a) Kristina, on her morning run, wants to get from point A to point B. How many routes with no backtracking can she take (she always travels South or East)?

The Complement rule



During her run she needs to go 5 blocks east and 7 blocks south. Hence the number of routes is $\mathbf{C}(7+5,7) = \mathbf{C}(12,7) = \mathbf{C}(12,5) = 792.$

The Complement rule

(b) If Kristina chooses a route from among those with no backtracking at random, what is the probability that she will not run past the doberman at D?

It is easier to count the routes which do go by the doberman: $\mathbf{C}(3+4,3)\cdot\mathbf{C}(2+3,2) = 35\cdot 10 = 350$. Hence the probability that she will not run past the doberman is $1-\dfrac{350}{792}\approx 0.558$.

The Complement rule

Example A box ready, for shipment, contains 100 light bulbs, 10 of which are defective. The quality control test is to take a random sample of 5 light bulbs, without replacement, from the box. If one is defective, the box will not be shipped. What is the probability that the box will not be shipped.

We've already calculated the probability that the box will get shipped is $0.583\ldots$, so the probability that it will not get shipped in $1 - 0.583\ldots = 0.416\ldots$.

Extras

Example Harry Potter's closet contains 12 numbered brooms, of which 8 are Comet Two Sixty's(numbered 1-8) and 4 are Nimbus Two Thousand's(numbered 9 -12). Harry, Ron, George and Fred want to sneak out in the middle of the night for a game of Quidditch. They are afraid to turn on the light in case Filch catches them. Harry reaches into the closet and pull out a random sample 4 brooms.

(a) Calculate the probability that all of the brooms will be Comet Two Sixty's.

The number of samples is $\mathbf{C}(12, 4)$ and the number of samples in which all the brooms are Comet Two Sixty's is $\mathbf{C}(8, 4)$. (Since Harry is grabbing all four brooms there is no way to order the set he gets.) Hence the answer is $\dfrac{\mathbf{C}(8, 4)}{\mathbf{C}(12, 4)} = 0.14 \cdots$.

Extras

(b) What is the probability that Harry chooses a sample with exactly 4 Nimbus Two Thousand's?

$$\frac{\mathbf{C}(4,4)}{\mathbf{C}(12,4)} = 0.00202\ldots$$

(c) What is the probability that Harry will have at least one Nimbus Two Thousand in his sample?

1 − probability of 0 Nimbus Two Thousands =

$$1 - \frac{\mathbf{C}(8,4)}{\mathbf{C}(12,4)} \approx 0.858.$$

Old Exam Questions

1 An experiment consists of drawing 3 balls at random from a bag containing 2 red balls and 4 white balls. What is the probability of getting at least 2 white balls?

(a) $\frac{4}{5}$ (b) $\frac{2}{3}$ (c) $\frac{1}{5}$ (d) $\frac{2}{5}$ (e) $\frac{3}{5}$

$$\frac{\mathbf{C}(4,2)\cdot\mathbf{C}(2,1)+\mathbf{C}(4,3)\cdot\mathbf{C}(2,0)}{\mathbf{C}(6,3)}=\frac{16}{20}.$$

Old Exam Questions

2 Three out of 25 new cars are selected at random to check for steering defects. Suppose that 7 of the 25 cars have such defects. What is the probability that all 3 of the selected cars are defective?

(a) $\dfrac{\binom{7}{3}}{\binom{25}{3}}$ (b) $\dfrac{\binom{25}{3}}{\binom{25}{7}}$ (c) $\dfrac{\binom{18}{3}}{\binom{25}{3}}$ (d) $\dfrac{\binom{18}{7}}{\binom{25}{7}}$ (e) $\dfrac{3!}{\binom{25}{3}}$.

$$\frac{\mathbf{C}(7,3)}{\mathbf{C}(25,3)}.$$

Old Exam Questions

3 A fair coin is tossed 10 times. What is the probability of observing exactly 3 heads?

(a) $\frac{3}{10}$ (b) $\frac{\mathbf{P}(10,3)}{2^{10}}$ (c) $\frac{2^3}{2^{10}}$ (d) $\frac{1}{10^3}$ (e) $\frac{\mathbf{C}(10,3)}{2^{10}}$.

$$\frac{\mathbf{C}(10,3)}{2^{10}}$$

Coincidence: The Birthday Problem

What is the probability that at least two people in a group will share a birthday (month and day) under the assumption that birthdays are randomly distributed throughout the year, and that February 29 does not occur as a birthday?

The next table shows the approximate probability, according to group size.

Coincidence: The Birthday Problem

Group Size	Probability at least 2 Birthdays are the same
20	0.4114
30	0.7063
40	0.8912
50	0.9704
60	0.9941
70	0.9991
80	0.9999

Coincidence: The Birthday Problem

We will work through the calculation for a group of twenty people. Let E be the event that at least two people in the group share a birthday. It turns out to be much easier to calculate $\mathbf{P}(E^c)$, the probability that everybody has their birthday on a different day, and then use the complement rule

$$\mathbf{P}(E) = 1 - \mathbf{P}(E^c).$$

To calculate the probability of E^c, we consider the experiment of making a list of 20 possible birthdays. By our assumptions, all lists are equally likely. There are

$$365 \times 365 \times \ldots \times \ldots \times 365 = 365^{20}$$

possible lists. (This is a 52-digit number)

Coincidence: The Birthday Problem

The event E^c corresponds to the set of lists where all birthdays on the list are different.. By the multiplication principle there are

$$365 \cdot 364 \cdot 363 \cdot 362 \cdots 345 = \mathbf{P}(365, 20)$$

such lists (a 51-digit number).

Now $\mathbf{P}(E^c) = \mathbf{P}(365, 20)/365^{20} \approx 0.588$ and hence

$$\mathbf{P}(E) = 1 - \mathbf{P}(E^c) \approx 0.411.$$

The calculations for groups of different sizes are similar. In particular we can show that in a group of 23 people, there is more than a 50/50 chance of a coincidence.

Since each of the 32 teams in the 2014 World Cup had a squad of 23 players, we would expect about 16 teams to have two squad members with a shared birthday. This was in fact the case: `Birthday paradox at World Cup`

Coincidence: The Birthday Problem

Try the following problem, which is similar, using your calculator (First make a guess as to what you think the probability is)

Problem If 10 people each choose a number (secretly) between 1 and 50 (inclusive), what are the chances that at least two of the numbers will be the same if everyone chooses randomly?

$$1 - \left(\frac{\mathbf{P}(50, 10)}{50^{10}} \right) \approx 0.618.$$

Back to "Test your intuition"

Remember: each person in the room today has selected a whole number between 1 and 100, at random (with no collaboration!).

The question was, how likely is it that at least two people pick the same number?

Assuming 20 people are selecting,

$$1 - \left(\frac{\mathbf{P}(100, 20)}{100^{20}} \right) \approx 87\%.$$

So the correct answer was **D**.

More on coincidence

The following series of coincidences between the life events of Abraham Lincoln and John F. Kennedy often strike people as unusual or even spooky:

Lincoln was elected to Congress in 1846; Kennedy in 1946.

Lincoln was elected president in 1860; Kennedy in 1960.

Lincoln's secretary was named Kennedy; Kennedy's was named Lincoln.

Andrew Johnson, who succeeded Lincoln, was born in 1808; Lyndon Johnson, who succeeded Kennedy, was born in 1908.

John Wilkes Booth, who assassinated Lincoln, was born in 1839; Lee Harvey Oswald, who assassinated Kennedy, was born in 1939.

More on coincidence

Given the amount of information we have about these two men, is it really surprising that we might find 5 such coincidences in their parallel lives?

Experiment: Suppose I create 2 fictitious characters, Ms. A and Ms. B and I give each a profile by listing for each a randomly chosen year-of-occurrence, somewhere in the 20th century, of each of 1000 life events, such as:

- year born
- year died
- year they got their first dog
- year they got married
- year they visited Ireland
- ...

More on coincidence

(a) What is the probability that I assign the same year to both Ms. A and Ms. B for a life event?

Probability that I assign different years: $\dfrac{100 \cdot 99}{100 \cdot 100} = 0.99.$

Probability of coincidence: $1 - 0.99 = .01$.

(b) How many coincidences would you expect among the 1000 life events?

$\approx 1000 \cdot 0.01 = 10$

(c) How likely is it to have at least 5 coincidences?

$\approx .97$ (We'll see how I got this later)

The following video discusses some common misconceptions about coincidences:

`It could just be coincidence`