Independence

Previously we considered the following experiment: A card is drawn at random from a standard deck of cards.

- Let H be the event that a heart is drawn,
- let R be the event that a red card is drawn and
- let F be the event that a face card is drawn (where the face cards are the kings, queens and jacks).

We found that

$$\mathrm{P}\left(H\middle|R\right) = \frac{1}{2} \neq \mathbf{P}(H) = \frac{1}{4}\ .$$

On the other hand

$$\mathrm{P}\left(F\middle|R\right) = \frac{6}{26} = \mathbf{P}(F) = \frac{12}{52}\ .$$

Independence

Since $\mathrm{P}\left(F \middle| R\right) = \mathbf{P}(F)$, we see that $\mathbf{P}(F)$ is **not** influenced by the prior knowledge that the card is red. In this case, we say that the events F and R are **independent**.

Definition: Two events A and B are said to be **independent** if

$$\boxed{\mathrm{P}\left(A \middle| B\right) = \mathbf{P}(A).}$$

This is the same as

$$\boxed{\mathrm{P}\left(B \middle| A\right) = \mathbf{P}(B)}$$

and

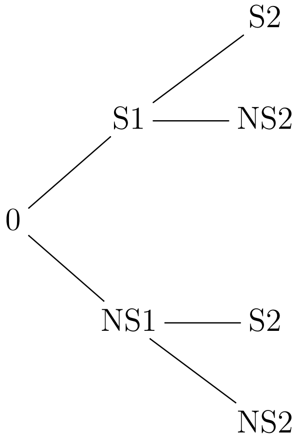
$$\boxed{\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B).}$$

If A and B are independent, the chance that one will occur is not influenced in any way by the knowledge that the other has occurred.

Independence

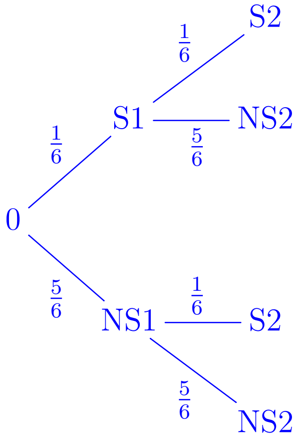
Example If we roll a fair six sided die twice and observe the numbers appearing on the uppermost face of each, it is reasonable to expect that the number appearing on the second dice is not influenced in any way by the number appearing on the first. In this case the probability of a six on the second roll should equal the probability of a six on the second given a six on the first. Using a tree diagram we can determine the probability of a six on both rolls, where Si denotes the event that we get a six on roll i and NSi denotes the event that we do not get a six on roll i.

The tree



Two sixes

$\mathbf{P}(Si) = \frac{1}{6}$ and $\mathbf{P}(NSi) = 1 - \frac{1}{6} = \frac{5}{6}$.



To get a two sixes there is only one path so the probability is $\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$.

Drawing from a bag

Example: A bag has 6 red marbles and 4 blue marbles. I draw a marble at random from the bag and replace it, then I draw a second marble. What is the probability that at least one of the marbles is blue?

As with the two dice, the two draws are independent. The probability of drawing a blue on the first draw is .4, the probability of drawing a blue on the second is .4, and the probability of drawing a blue on both is $.4 \times .4 = .16$ (this is where we use independence). By inclusion-exclusion, the probability of at least one blue is

$$.4 + .4 - .16 = .64.$$

Unconnected events

Example The Toddlers of the Lough soccer team in Cork, Ireland has no known connection to the Notre Dame Lacrosse team. The chances that the toddlers will win their game this weekend is 0.7 and the chances that the Notre Dame Lacrosse team will win their game this weekend is 0.999. It is reasonable to assume that the events that each team will win are independent, based on this assumption calculate the probability that both teams will win their games this weekend.

Let $\mathbf{P}(T) = 0.7$ be the chance that the Toddlers will win and let $\mathbf{P}(L) = 0.999$ be the chance that the ND Lacrosse team will win. If the events are independent, $\mathbf{P}(T \cap L) = 0.7 \cdot 0.999 = 0.6993$. $\mathbf{P}(T \cap L)$ is the probability that both teams will win.

Unconnected events

Warning sometimes our assumptions that seemingly unrelated events are independent can be wrong. For an example where independence was assumed leading to serious consequences, see the reference to the trial of Sally Clark in the following video:

`Ted Talks: How Statistics Fool Juries`

Union of Independent Events

If two events, A and B, are independent we can substitute the identity $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$ into the formula for $\mathbf{P}(A \cup B)$.

If A and B **are independent**, then

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A) \cdot \mathbf{P}(B).$$

Example If E and F are independent events, with $\mathbf{P}(E) = 0.2$ and $\mathbf{P}(F) = 0.4$, what is $\mathbf{P}(E \cup F)$?

Since the events are *independent*,
$\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F) = 0.2 \cdot 0.4 = 0.08$ and
$\mathbf{P}(E \cup F) = \mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F) = 0.2 + 0.4 - 0.08 = 0.52.$

Union of Independent Events

Example In an experiment I draw a card at random from a standard deck of cards and then I draw a second card at random from a different deck of cards. What is the probability that both cards will be aces?

The probability of drawing an ace is $\frac{4}{52} = \frac{1}{13}$. Let A be the event that I draw an ace from the first deck and B the event that draw an ace from the second deck. What is $\mathrm{P}\left(B \middle| A\right)$? It is the probability that I draw an ace from the second deck given that I drew an ace from the first deck. But this is $\frac{4}{52} = \frac{1}{13}$ again, so $\mathrm{P}\left(B \middle| A\right) = \mathbf{P}(B)$ and the events are independent. Hence
$\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B) = \frac{1}{13} \cdot \frac{1}{13} = \frac{1}{169}$.

Union of Independent Events

Note If two events, E and F, are independent, then their complements E' and F' are also independent.

$$\begin{aligned}\mathbf{P}(E' \cap F') =& \mathbf{P}((E \cup F)') = 1 - \mathbf{P}(E \cup F) \\ &= 1 - [\mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E \cap F)] \\ & 1 - [\mathbf{P}(E) + \mathbf{P}(F) - \mathbf{P}(E) \cdot \mathbf{P}(F)] = \\ & 1 - \mathbf{P}(E) - \mathbf{P}(F) + \mathbf{P}(E) \cdot \mathbf{P}(F) = \\ & \big(1 - \mathbf{P}(E)\big) \cdot \big(1 - \mathbf{P}(F)\big) = \mathbf{P}(E') \cdot \mathbf{P}(F').\end{aligned}$$

Union of Independent Events

Example Mary is taking a multiple choice quiz with two questions. Each question has 5 possible solutions (a) - (e). Mary has no idea what the right answers might be, so she takes a random guess for each answer.

(a) What are the chances that she gets both questions wrong?

Since there are 5 choices for an answer and only 1 correct answer, the probability of getting 1 question wrong is $\frac{4}{5} = 80\%$. The events are independent so the probability of getting 2 questions wrong is $\frac{4}{5} \cdot \frac{4}{5} = \frac{16}{25} = 64\%$.

Union of Independent Events

(b) What are the chances that she gets at least one question right?

The easiest way to answer this question is to observe that event in (b) is the complement of the event in (a) so the answer is $100\% - 64\% = 36\%$.

Or notice that her chances of getting the first question right and the second wrong is $\frac{1}{5} \cdot \frac{4}{5}$; her chances of getting the first question wrong and the second right is $\frac{4}{5} \cdot \frac{1}{5}$; and her chances of getting the both questions right is $\frac{1}{5} \cdot \frac{1}{5}$ for a total of $\frac{4+4+1}{25} = \frac{9}{25}$.

Many Independent Events

A collection $\{E_1, \ldots, E_n\}$ of events is independent if nothing you can learn about whether some of the events occurred or not tells you anything about whether any of the other have occurred. For collections of independent events, we have

$$\mathbf{P}(E_1 \cap E_2 \cap \cdots \cap E_n) = \mathbf{P}(E_1) \cdot \mathbf{P}(E_2) \cdots \mathbf{P}(E_n).$$

Example If there were 3 questions on Mary's quiz, and Mary makes a random guess for each question.

(a) what are the chances that she gets all three correct?

The events are independent so the answer is $\frac{1}{5} \cdot \frac{1}{5} \cdot \frac{1}{5}$.

(b) What are the chances that she gets none correct?

$\frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5}$.

Many Independent Events

(c) What are the chances that she gets at least one correct?

Either 1 minus the answer in (b) or

$$\frac{1}{5}\cdot\frac{4}{5}\cdot\frac{4}{5}+\frac{4}{5}\cdot\frac{1}{5}\cdot\frac{4}{5}+\frac{4}{5}\cdot\frac{4}{5}\cdot\frac{1}{5}+\frac{1}{5}\cdot\frac{1}{5}\cdot\frac{4}{5}+\frac{1}{5}\cdot\frac{4}{5}\cdot\frac{1}{5}+\frac{4}{5}\cdot\frac{1}{5}\cdot\frac{1}{5}+\frac{1}{5}\cdot\frac{1}{5}\cdot\frac{1}{5}=$$

$$\frac{16+16+16+4+4+4+1}{125}=\frac{61}{125}$$

OR (think trees)

She gets the first correct; or she gets the first wrong and the second correct; or she gets the first two wrong and the third one correct:

$$\frac{1}{5}+\frac{4}{5}\cdot\frac{1}{5}+\frac{4}{5}\cdot\frac{4}{5}\cdot\frac{1}{5}=\frac{25+20+16}{125}$$

Reliability Theory

(a) Suppose a new phone has 4 independent electronic components of type B. Suppose each component of type B has a probability of .01 of failure within 10 years. What are the chances that at least one of these components will last more than 10 years?

To say that the components are independent is to say that the failure of one does not influence the chance of the failure of another. The probability of one component failing within ten years is $\mathbf{P}(\text{fail}) = 0.01$ and the probability of its not failing in ten years is $\mathbf{P}(\text{not-fail}) = 0.99$. The complement to "at least one of these components will last more than 10 years" is "all four components fail within ten years" and since these failures are independent, the chance of this happening is $0.01^4 = 0.00000001$ so the chances that at least one of these components will last more than 10 years is 0.99999999.

Reliability Theory

(b) The phone company want to make sure that at least one of the components of type B in a new phone will still be working after 10 years. They know that each component of type B has a probability of .01 of failure within 10 years and they know that the failure of components of type B are independent events. What is the minimum number of these components in a new phone that will ensure that at least one will still be operating after 10 years with a 99.99% probability?

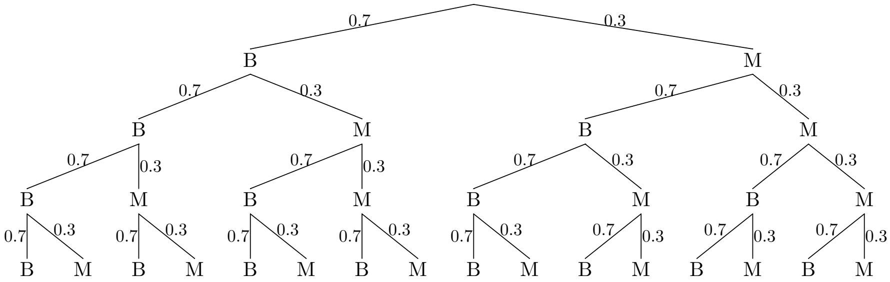
The probabilities are the same as in the previous part but the problem is now asking for the smallest integer m so that if the company puts in m independent components, the chance of failure within 10 years is 99.99%.

Reliability Theory

We saw that if $m = 4$, the probability is 99.999999% which is certainly OK but maybe a smaller m will do just as well and save the company a bunch of money. The probability that with m components there will be a failure within 10 years is $1 - (0.01)^m$. For $m = 1$ this 99% which is not good enough. For $m = 2$ this 99.99% which is good enough so the company should use two components.

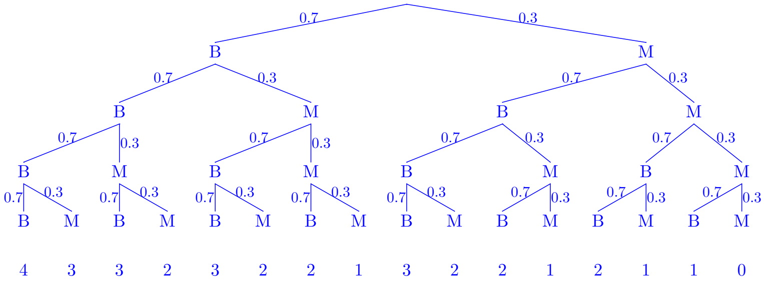
Repeating a trial many times

Example A basketball player takes 4 independent free throws with a probability of .7 of getting a basket on each shot. Use the tree diagram below to find the probability that he gets exactly 2 baskets. B = gets a basket, M = misses.

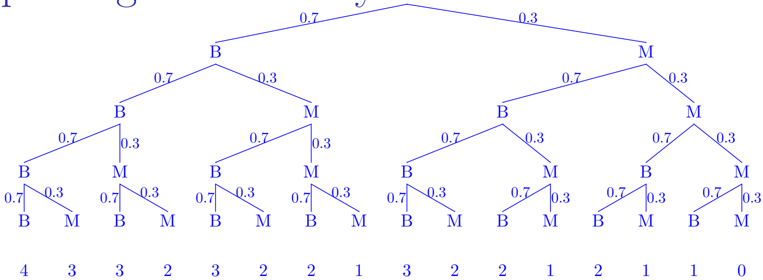


Repeating a trial many times

We need to find the paths from the top to the bottom that give precisely 2 made baskets. There help avoid error error (in missing a path or forgetting we already counted a path) we add a row to the diagram which counts the number of made baskets along each branch. The probability assigned to a path with b made baskets is $(0.7)^b(0.3)^{4-b}$.



Repeating a trial many times



There are 6 paths with 2 made baskets so the answer is $6 \cdot (0.7)^2 \cdot (0.3)^2 = 26.46\%$.

Since we've done all the work, here are all the probabilities.

Probability	Number made	Number paths
0.81%	0	1
7.56%	1	4
26.46%	2	6
41.16%	3	4
24.01%	4	1

Checking for Independence

For events E and F whose probabilities are not 0, we can use our formulas:

$$\mathbf{P}(E \cap F) = \mathbf{P}(E) \cdot \mathbf{P}(F)$$

or equivalently

$$\mathrm{P}\left(E \middle| F\right) = \mathbf{P}(E)$$

or equivalently

$$\mathrm{P}\left(F \middle| E\right) = \mathbf{P}(F)$$

to check if given events are independent. If any one of the above 3 formulas hold true, then the other two are automatically true and E and F are independent.

To **verify that two events are independent** we need only check one of the above 3 formulas. We choose the most suitable one, depending on the information we are given.

Checking for Independence

Example Of the students at a certain college, it is known that 50% of all students regularly attend football games and 60% of the first year students regularly attend football games. We choose a student at random. Are the events A = "The student attends football games regularly" and FY = "The student is a first year student" independent?

We are given $\mathbf{P}(A) = 0.5$ and $\mathrm{P}\left(A \middle| FY\right) = 0.6$. The probability $\mathbf{P}(FY)$ is not given but at any reasonable university it is not 0 (there are first year students). Since $\mathrm{P}\left(A \middle| FY\right) \neq \mathbf{P}(A)$ and $\mathbf{P}(FY) \neq 0$ the events are not independent.

Example If $\mathbf{P}(E) = .3$ and $\mathbf{P}(F) = .5$ and $\mathbf{P}(E \cap F) = .2$, are E and F independent events?

This time it seems easier to check $\mathbf{P}(E \cap F) = .2$ and $\mathbf{P}(E) \cdot \mathbf{P}(F) = 0.3 \cdot 0.4 = 0.12$ so these events are not independent.

Checking for Independence

Example 300 students were asked if they thought that their online homework for Elvish 101 was too easy. The results are shown in the table below.

	Yes	No	Neutral
Male	75	39	36
Female	91	16	43

Let M denote the event that an individual selected at random is male and let Ne denote the event that the answer of an individual selected at random is "Neutral". Let Y denote the event that the answer of an individual selected at random says "Yes".

(a) What is $\mathbf{P}(Ne)$?

There are 300 students and 79 neutral responses so $\mathbf{P}(Ne) = \frac{79}{300}$.

Checking for Independence

(b) What is $\mathrm{P}\left(Ne \middle| M\right)$?

There are 150 male students, $75 + 39 + 36$, and 36 neutral male responses so $\mathrm{P}\left(Ne \middle| M\right) = 36/150 = .24$.

(c) Are the events Ne and M independent?

The question asks if $\mathrm{P}\left(Ne \middle| M\right)$ and $\mathbf{P}(M)$ are equal. Given the last two calculations, this equivalent to the question of whether $79/300 = .263\ldots$ and .24 equal, which they are not. So Ne and M are not independent — knowing that a student is male gives you some genuine new information about how likely they are to be neutral. **BUT**: $\mathrm{P}\left(Ne \middle| M\right)$ and $\mathbf{P}(M)$ are quite close to each other; a statistician might infer that there is no statistically significant difference between the two answers.

Mutually exclusive events and independence

Events E and F are **mutually exclusive** if $\mathbf{P}(E \cap F) = 0$; I.e., E and F can't happen at the same time.
(d) Are the events Y and Ne from the precious example mutually exclusive?

The events Y and Ne are mutually exclusive since you cannot answer both yes and neutral at the same time.

Mutually exclusive events are typically **not** independent: Knowing that E occurs tells you **for certain** that F can't occur. In terms of formulas, $\mathbf{P}(E \cap F) = 0$, and the **only** way this could equal 0 is if one (or both) of $\mathbf{P}(E) = 0$, $\mathbf{P}(F) = 0$ holds. But since we are seldom interested in events which can never happen, we seldom encounter events which are both mutually exclusive and independent.