# Statistics 3657 : Independent Events

Recall the definition of independence from class.

Definition of (statistical) independence of events

1. Events $A_1, A_2, \ldots, A_n$ are said to be be independent if and only if for every subset $J \subset \{1, 2, \ldots, n\}$
$$P(\cap_{i \in J} A_i) = \prod_{i \in J} P(A_i)$$

2. A collection of events $\mathcal{A}$ is said to be a collection of independent events if and only if for every finite sub collection of events $A_1, \ldots, A_n \in \mathcal{A}$, the events $A_1, \ldots, A_n$ are independent.

*end of definition*

Note that when the subset $J$ of the indices is a set of size 1, that is a single element, the product formula is automatically and trivially true. Similarly the product formula is trivial if $J = \emptyset$. Thus it is really only of interest to check these products when $|J|$, the cardinality of the set $J$ or the number of elements in the set $J$, is at least 2.

Consider two events $A, B$. There is only one condition, so $A$ and $B$ are independent if and only if
$$P(A \cap B) = P(A)P(B) \ .$$

If we consider 3 events, $A_1, A_2, A_3$, then there are $\binom{3}{2} = 3$ pairs to check and $\binom{3}{3} = 1$ triple to check, that is 4 conditions to check. Is it okay to just check the single triple condition? An example below shows that in general this is not true.

What happens if there are $n$ events. Then to check independence there are $2^n$ conditions, although the $\binom{n}{1} = n$ singleton sets and $J = \emptyset$ are trivial. Thus there are $2^n - n - 1$ actual conditions to check.

Fortunately if we start with independence, these $2^n$ product conditions are automatically true. We will have some results about this later.

In our course we refer to statistical independence more simply as independence. There are other notions of independence such as linear independence. We will only use the the qualifying term *statistical* with dependence when the context requires this extra clarification.

If events $A_1, A_2$ and $A_3$ are pairwise independent are they (mutually) independent?

This means, supposing that $P(A_i \cap A_j) = P(A_i)P(A_j)$ for every distinct pair $i, j \in \{1, 2, 3\}$ is it true that $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$?

Somewhat surprisingly the answer is no. A simple game of tossing coins will furnish a counterexample.

**A Coin Tossing Game Example**

Consider tossing two fair coins. Consider the following events.

- $A_1$ = event that the first coin is a head (H)
- $A_2$ = event that the second coin is a H
- $A_3$ = event that the number of heads is odd, that is the number of heads is 1.

Verify that for each of these events $P(A_i) = \frac{1}{2}$.
Verify that for each pair of these events we have

$$P(A_i \cap A_j) = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = P(A_i)P(A_j) \ .$$

Verify that $A_1 \cap A_2 \cap A_3$ is the empty set, so that

$$P(A_1 \cap A_2 \cap A_3) = 0 \neq P(A_1)P(A_2)P(A_3) = \frac{1}{8} \ .$$

For this game it may be easier to understand if we had proceeded using a sample space.
Take $\Omega = \{00, 01, 10, 11\}$ so that

- $A_1 = \{10, 11\}$
- $A_2 = \{01, 11\}$
- $A_3 = \{10, 01\}$

Also we have the probability measure $P(\omega) = \frac{1}{4}$ for each $\omega \in \Omega$. The student should now verify this sample space example really has the properties above, so that it is the formal example of non independence but pairwise independence.

Similar examples can be constructed with $n$ events. In particular it is possible to construct examples of $n$ events, $A_1, A_2, \ldots, A_n$ such that for any $n-1$ these $n-1$ events are independent, but all $n$ together are not (mutually) independent.

More about independence

If we perform an experiment with $n$ independent trials, then events $A_i$ from the $i$-th trial are independent. Therefore the $A_i$'s obey the product formula conditions for all subsets $J \subset \{1, 2, \ldots, n\}$, that is all subsets of the indices $1, \ldots, n$.

How can this be done? If we perform physically independent experimental trials, then each trial is also statistically independent. For this reason the notion of statistical independence results from physically independent trials in an experiment. This is what is often done in scientific experiments. It is also guaranteed by many of the commonly used sampling mechanisms, such as used for random samples in an opinion poll.

Some other consequences

If $A$ and $B$ are independent events then $A^c$ and $B$ are independent events.

Proof :

$$\begin{aligned} P(B) &= P\left((A \cap B) \cup (A^c \cap B)\right) \\ &= P\left(A \cap B\right) + P\left(A^c \cap B\right) \quad \text{using axiom 3} \\ &= P(A)P(B) + P\left(A^c \cap B\right) \quad \text{by independence of } A \text{ and } B \end{aligned}$$

Therefore

$$P\left(A^c \cap B\right) = P(B) - P(A)P(B) = \left(1 - P(A)\right)P(B) = P(A^c)P(B)$$

and therefore by the definition of independence $A^c$ and $B$ are independent events.

*End of Proof.*

Can you prove that if $A_1, A_2, A_3$ are independent then $A_1, A_2, A_3^c$ are independent?

Circuit Example

This example is similar to circuit examples in the problems at the end of Chapters 1 and 3.

A circuit in Figure 1 consists of two subcircuits, the upper one consisting of components A and B, and the lower one consisting of components C and D.

In our example we suppose each of the basic components are independent. What is the probability of the circuit system working? This means that a message (electricity, water, email, etc) is able to pass from the left end to the right end. We think of the basic components as essentially lift bridges, that is a message passes over or through it if the component is *working* and not if the component *fails* or is not working.

Call the top subcomponent $SC1$ and bottom subcomponent $SC2$. Suppose the basic components $A, B, C, D$ are independent and each has probability $p$ of working and probability $q = 1 - p$ of failing.

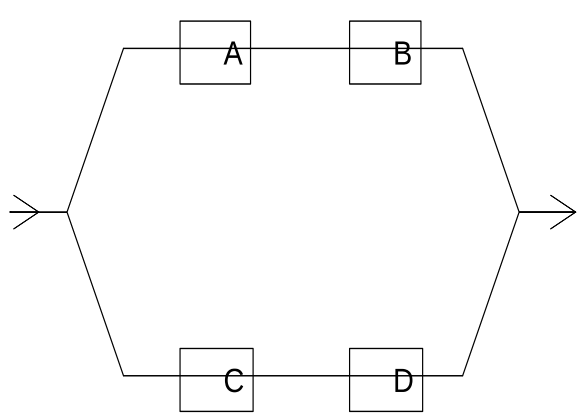


Figure 1: Circuit Diagram

Below we use set notation when convenient and needed to explicitly remind one that we are dealing with events. Notice here we have not explicitly written out the sample space, but are still using subset notation to help us in our careful logical calculation.

Thus

$$
\begin{aligned}
P(\{SC1\ fails\}) &= 1 - P(S1\ works) \\
&= 1 - P(\{A\ works\} \cap \{B\ works\}) \\
&= 1 - p^2
\end{aligned}
$$

Similarly

$$P(SC2\ fails) = 1 - p^2\ .$$

$$
\begin{aligned}
P(\{circuit\ works\}) &= 1 - P(\{circuit\ fails\}) \\
&= 1 - P(\{SC1\ fails\} \cap \{SC2\ fails\}) \\
&= 1 - P(\{SC1\ fails\})P(\{SC2\ fails\}) \\
&= 1 - (1 - p^2)(1 - p^2) \\
&= 2p^2 - p^4\ .
\end{aligned}
$$

One could also solve this using an appropriate sample space. This method is summarized in the following table. The first column gives all possible elementary outcomes, and hence the sample space. The second column gives an indicator to tell us if the circuit works, indicated by 1. The third column gives the probability for each of these elementary outcomes. Since the events which are the singleton sets of each of the elementary outcomes are disjoint, the Axioms of Probability tell us that $P(Circuit\ works)$ is equal to the sum of the probabilities of the elementary outcomes, summed over all elementary outcomes for which the circuit works.

The student should verify this gives the same answer as the method above.

| elementary outcome $ABCD$ | circuit works | probability |
|---|---|---|
| 0000 | 0 | |
| 0001 | 0 | |
| 0010 | 0 | |
| 0011 | 1 | $q^2p^2$ |
| 0100 | 0 | |
| 0101 | 0 | |
| 0110 | 0 | |
| 0111 | 1 | $qp^3$ |
| 1000 | 0 | |
| 1001 | 0 | |
| 1010 | 0 | |
| 1011 | 1 | $qp^3$ |
| 1100 | 1 | $p^2q^2$ |
| 1101 | 1 | $p^3q$ |
| 1110 | 1 | $p^3q$ |
| 1111 | 1 | $p^4$ |

The difference in the two methods is that the second uses an explicit sample space. The first method relies on properties of combining more complicated independent events. The later method is often easier once these properties are recognized. However method 1 is often useful when it is not clear what these corresponding more complicated events are and how they are related.

Finally we remind students of some helpful notation. For any event $A$, $P(A)$ is a number. Thus for events $A, B$ it makes sense to consider $A \cap B$ or $A \cup B$. However $\cap$ and $\cup$ do not apply to real numbers. For example if $A$ and $B$ are independent events we can write

$$P(A \cap B) = P(A)P(B)$$

makes sense.

It does not make sense to write

$$P(A \cap B) = P(A) \cap P(B) \ .$$

Why not? If it did, and for example $P(A) = \frac{1}{2}$ and $P(B) = \frac{1}{3}$ then we would need to make sense of

$$\frac{1}{2} \cap \frac{1}{3} \ .$$

The operations $\cap$ and $\cup$ apply to sets. Also arithmetic operations apply to numbers but not sets, unless some specific meaning can be defined for such operations.