3 PROBABILITY

3.10 Independence between random events

In this video I will explain a probability-concept that you will encounter frequently: independence. If random events are independent, it means that knowing the outcome of one does not provide information about the others. It is an important bit of information because if random events are independent, it simplifies probability calculations enormously and is therefore an assumption that is often made.

Let me quickly introduce an example: you have made a count of different activities by the people on your beach, which you distinguished by gender. And you have turned it into table with probabilities: joint probabilities in the central block and marginal probabilities here and there.

Independence between random events means that knowing the outcome of one event does not influence your knowledge about the outcome for the others. In the example it means that if you'd know about the gender of a person, it would not influence the probabilities for that person's activities. Or, if you would consider a certain activity it would not influence the probabilities for gender.

Would that be reasonable? Let's have a look. Compare the probability of selecting a male person from the people who are resting with the probability of selecting a male when disregarding activity (the marginal probability for gender). It turns out that the two probabilities are not quite the same.

We can take a different approach as well to look at independence: if you could assume independence between the two variables, gender and activity, you could calculate the joint probabilities by multiplying the marginal probabilities for gender and activity.

Let's see whether this would be reasonable in our case. We multiply the marginal probabilities for gender with those for the activities. If gender and activity at the beach would be independent, you would see these probabilities. However, we did observe joint probabilities that are quite different, take a look. So the assumption of independence seems not to hold!

Here we have applied two ways to check for independence among random variables, and these two ways are in fact quite close to the formal conditions for independence. The first is: events A and B are independent if the probability of event A given any outcome of event B is equal to the marginal probability of A (and in this definition you could switch the labels A and B). The second definition is: events are independent if their joint probability equals the product of the separate marginal probabilities.

If any of the two conditions holds, then events A and B can be considered as independent. For many real life situations, independence among random events or variables of interest is not that common at all. Quite often there are unexpected and unknown factors that influence the events of interest so that a specific outcome for one event turns out to be related to another in a probabilistic sense – there does not have to be any direct or any understood relation between events to be dependent in terms of probabilities.

Another aspect to bear in mind when thinking about independence is that, while the definitions are clear, it is in practice not always that easy to find sufficient support for the assumption of independence. You might need a large random sample to gather enough evidence. And as a final note, it is important to be aware of the difference between disjoint versus independent events. If events are disjoint, it means that if one event occurs, the others cannot occur; so by definition their joint probability is zero. As a consequence, disjoint events are very dependent: once one event occurs you know that the probability of the others is zero. Contrastingly, if events are independent the occurrence of one event tells nothing about another and can thus also not be disjoint.

I hope you have learned the following from this video.

- random events are independent if:
  - the joint probability of the events is equal to the product of the marginal probabilities; which is equivalent to the situation where
  - the probability of an outcome for one event given any outcome for another event is equal to the marginal probability for that event
- For many real life situations, independence among random events is not that common, because unknown factors may influence the various events at the same time.
- And finally, you shouldn't confuse the terms disjoint event and independent event. If two events are disjoint it implies that they are dependent, whereas if two events are independent it implies that they cannot be disjoint.