

3 PROBABILITY

3.01 Randomness and us

Recognizing and understanding randomness as well as the ability to reason about it, are important skills – not only to apply statistical analyses, but also to make sense of things happening around us every day. Here I will explain that humans are pathologically bad at dealing with randomness. I will use an example along the way.

Imagine, you are on a beach, watching the waves roll in. And then your attention is caught by a beautiful shell which is distinctive in shape and larger than its neighbouring shells. So you start to search for another one....

This will be an unpredictable enterprise. The shells may be distributed at random at this huge beach, hence the time it will take you to find another will be uncertain - you may not be able to find a similar shell at all.

The more you think about it, the more you will realize that randomness is pervasive in everyday life. It is therefore not surprising that we have a rich vocabulary to communicate it, with terms like uncertainty, chance, risk and likelihood.

Also, degrees of variability (and uncertainty) can be expressed quite subtly. Consider for example this sequence: rarely, seldom, sometimes, common, frequent, often.

Importantly, whether something is random is not just a property of that phenomenon, but very much also a consequence of our knowledge about it. If you'd have been at this beach before, you might have spotted the special shell previously so this may change your search strategy and increase your chance of finding more of them. Also the scale of your search matters. While you may not be very certain about finding another shell within a few minutes at a short stretch of beach, the chance to find one when you take a bit more time and cover a larger stretch of beach will increase.

But in spite of many words, our ability to memorize, and our daily experience with randomness, we are not very good at assessing it quantitatively at all. On one hand we see all sorts of patterns in what is really random data, there is a word for it: apophenia. And on the other hand we are unable to make up random data ourselves. An example of a failed attempt to create random data is this fabricated map of random shell locations, which turn out to be spaced too regular: this is how a realistic random point pattern looks like, with much more clusters.

Another example of over-interpreting randomness is the so-called gambler's fallacy, the false idea that a random phenomenon can be predicted from a series of preceding random phenomena. And who doesn't recognize this? If you have thrown a six four times in a row with a die, it feels as if it is going to be very unlikely to throw six again the fifth time. Yet the probability of this outcome was and continues to be one sixth.

The reason that we have a bad head for randomness is that our brain tends measure randomness as the effort it takes to memorize a pattern. And it turns out that memorizing frequently changing, short sequences is harder than longer sequences.

Given this state of affairs, it is really important to learn about formal ways for quantifying randomness, reasoning about it and generating realistic random patterns. It will help to avoid mistakes, predict more accurate and be more efficient when we try to make sense of the world around you.

Let me summarize what I hope you understood from this video:

- Randomness it is not an intrinsic property of a phenomenon. It also depends, amongst others, on prior knowledge, observation method and the scale at which the phenomenon is considered.
- While there are many words to express aspects of randomness, humans are not very good in assessing it quantitatively.
- We suffer from apophenia: the over-interpretation of what are purely random patterns; and are also bad in constructing randomness.