Example 8: Probabilistic statements about the 1997 HERA high-Q^2 events.

A very instructive example of the misinterpretation of probability can be found in the statements which commented on the 'excess' of events observed by the e–p experiments at the HERA collider (DESY Laboratory, Hamburg, Germany) in the high-Q^2 region. For example, the official DESY statement [18] was:[18]

> *"The two HERA experiments, H1 and ZEUS, observe an excess of events above expectations at high x (or $M = \sqrt{x\,s}$), y, and Q^2. For* $Q^2 > 15\,000$ GeV^2 the joint distribution has a probability of less than one per cent to come from Standard Model NC DIS processes."

(Standard Model refers to the most believed and successful model of particle physics; NC and DIS stand for Neutral Current and Deep inelastic Scattering, respectively; Q^2 is inversely proportional to the region of space inside the proton probed by the electron beam.) Similar statements were spread out in the scientific community, and finally to the press. For example, a message circulated by INFN stated (it can be understood even in Italian)

> *"La probabilità che gli eventi osservati siano una fluttuazione statistica è inferiore all' 1%."*

Obviously these two statements led the press (e.g. Corriere della Sera, 23 Feb. 1998) to announce that scientists were highly confident that a great discovery was just around the corner.[19]

[18] One might think that the misleading meaning of that sentence was due to unfortunate wording, but this possibility is ruled out by other statements which show clearly a quite odd point of view of probabilistic matter. In fact the 1998 activity report [19] insists that *"the likelihood that the data produced are the result of a statistical fluctuation . . . is equivalent to that of tossing a coin and throwing seven heads or tails in a row"* (replacing 'probability' by 'likelihood' does not change the sense of the message). Then, trying to explain the meaning of a statistical fluctuation, the following example is given: *"This process can be simulated with a die. If the number of times a die is thrown is sufficiently large, the die falls equally often on all faces, i.e. all six numbers occur equally often. The probability for each face is exactly a sixth or 16.66%, assuming the die is not loaded. If the die is thrown less often, then the probability curve for the distribution of the six die values is no longer a straight line but has peaks and troughs. The probability distribution obtained by throwing the die varies about the theoretical value of 16.66% depending on how many times it is thrown."*

[19] One of the odd claims related to these events was on a poster of an INFN exhibition at Palazzo delle Esposizioni in Rome: *"These events are absolutely impossible within the current theory . . . If they will be confirmed, it will imply that. . . ."* Some friends of mine who visited the exhibition asked me what it meant that "something impossible needs to be confirmed".

The experiments, on the other hand, did not mention this probability. Their published results [20] can be summarized, more or less, as *"there is a* $\lessapprox 1\%$ *probability of observing such events or rarer ones within the Standard Model".*

To sketch the flow of consecutive statements, let us indicate by SM *"the Standard Model is the only cause which can produce these events"* and by tail the *"possible observations which are rarer than the configuration of data actually observed".*

(1) Experimental result: $P(data + tail \mid SM) \lesssim 1\%$.
(2) Official statements: $P(SM \mid data) \lesssim 1\%$.
(3) Press: $P(\overline{SM} \mid data) \gtrsim 99\%$, simply applying standard logic to the outcome of step 2. They deduce, correctly, that the hypothesis $\overline{SM}$ (= hint of new physics) is almost certain.

One can recognize an arbitrary inversion of probability. But now there is also something else, which is more subtle, and suspicious: *"why should we also take into account data which have not been observed?"*[20] Stated in a schematic way, it seems natural to draw conclusions on the basis of the observed data:

$$\mathbf{data} \longrightarrow P(H \mid data)\,,$$

although $P(H \mid data)$ differs from $P(data \mid H)$. But it appears strange that unobserved data should also play a role. Nevertheless, because of our educational background, we are so used to the tacit inferential scheme of the kind

$$\mathbf{data} \longrightarrow P(H \mid data + tail)\,,$$

that we even have difficulty in understanding the meaning of this objection (see Ref. [13] for an extensive discussion).

I have considered this case in detail because I was personally involved in one of the HERA experiments. There are countless examples of this kind of claim in the scientific community, and I am very worried when I think that this kind of logical mistake might be applied in other fields of research on which our health and the future of the Planet depends. Recent frontier

[20] This is as if the conclusion from the AIDS test depended not only on $P(Positive \mid \overline{HIV})$ and on the prior probability of being infected, but also on the probability that this poor guy experienced events rarer than a mistaken analysis, like sitting next to Claudia Schiffer on an international flight, or winning the lottery, or being hit by a meteorite.