

Future Statistics Education

Takaharu Ishikawa - Okazaki, Japan
Takeyuki Hida - Nagoya, Japan

1. Introduction

Statistics has rapidly developed with the help of modern probability theory and with the effective use of computers. Such an accomplishment marks a new era for statistics education. It is now time to think of *how to teach statistics*, taking into account both the subjects to be chosen and the actual method of education.

At present we can see many places where mathematical statistics is efficiently used and where people are requested to learn statistics, not only in academic institutions, but also in daily life. The remarkable fact is that, compared to the past, the need for statistics has changed, and its appearances in actual subject fields have become highly modernised. We are therefore led to discuss how to teach statistics and to think of what topics should be taught.

At this juncture, we are going to look over the present stage of practical use of statistics and to propose some ideas of statistics education by focussing our vision on high school level mathematics.

2. Three key topics

The first part of our report is concerned with three important subjects in statistics that we believe should now be considered at school.

2.1 *A better understanding of the notion "randomness"*

One might say "Randomness! Can it be a subject of mathematics?" or "Random means uncertain, so that it cannot be a topic in school mathematics, which never allows any kind of ambiguity."

In teaching statistics, specifically when the problems are motivated by some actual phenomena, it is most important to explain that even random events can be good subjects of probability theory under a suitable mathematical setting. Moreover, teachers should show that the mathematical theory established in this manner can well be applied to the original problems to see a beautiful interplay between theory and practice.

For this purpose, as soon as we meet actual random events, it is most important to let students see why the *elementary event*, often denoted by ω , should first be introduced. Perhaps some simple paradoxes, although not quite paradoxes in reality, would illustrate the reason. We do not need the full Kolmogorov axioms, but we must appeal to measure theory at an intuitive level, as in counting models, where we can see the idea of a probability space. Thus, in order to develop a mathematical theory we should first form a probability space in a suitable way, perhaps in terms of elementary events.

2.2 *A correct understanding of "mean value"*

A probability distribution arises whenever some random phenomenon is involved. Hence, the first step in the statistical approach is to understand the meaning of certain characteristic values of the distribution. There are many such characteristic values; among others, the *mean value* or *expectation* is the most important to be well understood. Also, we should know how much the mean value tells us about the distribution in discussion. It does not tell us quite so much as one expects; however, what is certain is that with some additional information the mean value is the most useful characteristic. We may say that statistics should begin with learning the role of the mean value.

2.3 *The Gaussian distribution (normal distribution)*

The central position among all the probability distributions is occupied by the Gaussian distribution. Its importance cannot be overestimated. For students to understand this fact well, it would be better to tell them about some characterisations and heuristic interpretations of this distribution. Here are some typical examples.

- (i) The Gaussian distribution was discovered by C F Gauss as the distribution of errors, which is one of his important discoveries. He found the distribution when he observed the positions of planets or surveyed land numerically. He recognised that if the maximum likelihood estimator of the mean value is always realised by the arithmetic mean, then the distribution has to be Gaussian. This is not only a fine theorem, but also an important suggestion for statistical thinking, so that the story of his discovery is most pedagogically valuable.
- (ii) Consider the class of probability distributions with density and a common, fixed variance, say V . Information theory tells us that it is the Gaussian distribution that attains the maximum entropy in this class. We should think of ways to illustrate this fact.
- (iii) The Gaussian distribution exists in the natural world! It was Albert Einstein who saw that the Gaussian distribution arises from the biological phenomenon of "Brownian motion". The distribution he thus obtained theoretically is in

agreement with the results of observations of this motion, which makes the distribution familiar for us.

3. Teaching approaches

We now come to the practice of teaching, namely how to teach in class to realise the ideas illustrated in Section 2. Our catchword is "Good examples are essential".

Here are some examples that we propose; indeed, they have come from our actual experiences at school and in other activities.

3.1 *Brownian motion*

One can observe this phenomenon through a microscope, as is well-known, and (depending on the level of the pupils) we can give a reasonable interpretation as to why such phenomenon can be observed. To give this importance, we should show that the randomness involved is Gaussian in distribution. Our present question is to ask how such a randomness can be made an object of mathematical statistics in class.

Through simulation, students are able to understand and appreciate

- (i) limit theorems;
- (ii) two-dimensional standard Gaussian distributions;
- (iii) accuracy of approximation;

among other ideas.

We add some more interpretation that comes from our simulation. Since the two-dimensional random walk can be approximated by a two-dimensional Brownian motion, the figure obtained may be thought of as producing a distribution close to a two-dimensional Gaussian distribution with no correlation. A student might not be satisfied by the actual figure, because the picture is, in a sense, far from a two-dimensional Gaussian distribution. Such an unsatisfactory approximation to the actual statistics could prompt the teacher to remind the class of the following diagram:

Gaussian distribution => binomial distribution => actual statistic

The game "Heads or Tails" using a coin, provides statistics for the uniform distribution on the set $\{0,1\}$ involving only two points; yet "the central limit theorem" is hidden behind. We have used this game in the simulation, where we see unsatisfactory accuracy of the approximation. Now follows a very important remark. When we apply mathematical formulae to actual problems, we can assume that the Gaussian approximation is acceptable if large numbers are involved. However, our simulation shows that the approximation is not as good as expected, particularly for the tail of the distribution. We should keep this fact in mind, particularly when testing statistical hypotheses being discussed.

Another remark is that the figure obtained presents two independent quantities: x - and y -marginal distributions that are still Gaussian. Such a visual illustration is

helpful for the students.

3.2 *The Poisson distribution*

This is another good example. For one thing, it often appears in our daily life, so students can feel it is close to them. Another reason is that a Poisson distribution has intimate connections with binomial distributions.

Again, we can report on our actual experiences. We were once involved in a survey of car accidents under the organisation of the traffic control department of our local government. Those accidents are subject to a Poisson distribution if they are suitably classified; in fact, they were classified into three levels according to how serious they were. Such an experience suggests to us how to classify the data in order for the Poisson distribution to emerge. We can, of course, give plausible interpretations as to why classification is necessary, also a point where students can find interest.

It is noted that a Poisson distribution has a single parameter. In the above example, we can associate a value of this parameter with each place where accidents are observed, so that the parameter characterises the geometric structure of the places where accidents happen. With such an observation we can think of ways to control the traffic flow.

3.3 *Succession of trials (time series)*

Since modern statistics has developed with the mathematical help, not only of probability theory, but also of analysis, it is important to apply analysis to random events that change as time goes by.

When we discuss a succession or repetition of random trials, the basic tool from probability theory is *conditional probability*, where *causality* should be taken into account. There are cases, like drawing lots, where we can ignore the influence of the order of the trials in drawing a favourite lot. On the other hand, we often meet examples where we must carefully consider the order of the events in question, in addition to their statistical characteristics. Such contrasts mean that the students will understand well what role is really played by conditional distributions and why the order of the trials is effective.

Such a consideration of causality is important and can be understood even at an elementary level if suitable examples are taken.

4. *Concluding remark*

Recalling how statistics education is organised at present, sticking to the traditional approaches, we must emphasise the importance of introducing new methods and new, good examples. Teachers are recommended to discover suitable examples, to let the students see how such examples stimulate a mathematical way of thinking, and so to let students feel a real interest in statistics.