In order to prove this claim, it suffices to repeat the previous arguments replacing the set Z with the set $Z_\theta = \{ z \mid z = k + \theta,\ k \in Z \}$, where θ is an arbitrary vector in E^n.

Remark 4 can be applied to derive from the identities for P(x, k) and $M_b(x, k)$ analogous identities for the special functions and orthogonal polynomials listed above with any real parameters, and not only integer-valued parameters.

LITERATURE CITED

1. L. A. Lyusternik, "On one problem of queueing theory and its relation to generalized cylindrical functions $\mathcal{J}_n$," Dokl. Akad. Nauk SSSR, 177, No. 5 (1967).
2. L. A. Lyusternik, "On one problem of queueing theory and cylindrical functions," Usp. Mat. Nauk, 25, No. 4 (1970).
3. V. I. Zalyapin and L. A. Lyusternik, "On a system of functions for the Poisson random walk," Dokl. Akad. Nauk SSSR, 207, No. 1 (1972).
4. L. A. Lyusternik, "On some new studies in probabilistic methods in the theory of special functions," Usp. Mat. Nauk, 32, No. 5 (197) (1977).
5. Yu. K. Belyaev and Ya. P. Lumel'skii, "Many-dimensional Poisson random walks," in: Statisticheskie Metody, Perm (1978).
6. H. Bateman and A. Erdelyi, Higher Transcendental Functions [Russian translation], Vol. 2, Nauka, Moscow (1966).
7. H. Bateman and A. Erdelyi, Higher Transcendental Functions [Russian translation], Vol. 2, Nauka, Moscow (1967).
8. G. A. Korn and T. M. Korn, Manual of Mathematics for Scientists and Engineers, McGraw-Hill (1970).
9. E. Janke, F. Emde, and F. Losch, Special Functions (Formulas, Graphs, Tables) [Russian translation], Nauka, Moscow (1977).
10. I. S. Gradshtein and I. M. Ryzhik, Tables of Integrals, Series, and Products, Academic Press (1970).
11. H. Bateman and A. Erdelyi, Higher Transcendental Functions [Russian translation], Vol. 1, Nauka, Moscow (1965).

SOME ISSUES IN THE EXPOSITION OF THE THEORY OF PROBABILITY: A METHODOLOGICAL NOTE

I. G. Shaposhnikov

UDC 519.2

The paper examines some methodological aspects of the exposition of a number of basic concepts and propositions using the frequency definition of probability.

In this paper we briefly discuss the content of the basic concepts of probability theory in the context of the traditional exposition of this theory which, in our opinion, leaves much to be desired.

1. Probability theory originated as a mathematical science which nevertheless operated with nonmathematical concepts lacking rigorous definition — such as the occurrence of a head when a coin is tossed (an event) and the numerical characteristic of the likelihood or the expectation that a head will occur (the probability of an event). The "purely mathematical part" of probability theory was identified, analyzed, and rigorously formulated at a much

Perm University.

Translated from Statisticheskie Metody, pp. 216-220, 1980.

0090-4104/87/3902-2678$12.50

later stage: this is actually a branch of set theory, although sometimes it also goes under the name of probability theory. Here events are sets which are subsets of some abstract set, and the probability of an event is the measure of a set endowed with certain properties. The axiomatic approach based on this analysis is combined in virtually all the modern expositions of probability theory (with the exception of those relying on subjective interpretation of probability) with the notion of events as nonmathematical objects and the notion of probability of an event as a number approximating in most cases the frequency of occurrence of the event in a sufficiently long series of trials. The term "probability theory" is thus used in two different meanings. In what follows, we focus on the "ordinary" (and not the "purely mathematical") probability theory.

2. In our opinion, the traditional exposition of this probability theory suffers from a number of methodological weaknesses. We can identify two groups of such weaknesses: the use of unnecessary concepts (such as conditional and unconditional probability, prior and posterior probability), which are a consequence of insufficiently rigorous definitions of the notion of event; and false assertions concerning the relation between the frequency and the probability of an event (for instance, the claim that probability theory links with practice through the law of large numbers).

3. Let us start with the notion of event. In probability theory, the event (a; S) is the occurrence of the outcome a subject to the condition S — for instance, the occurrence of a head (a) on condition that the coin is tossed according to definite rules (S); the realization of S is called a trial of the event (a; S). All this is pretty standard. However, it is hardly ever emphasized that only a subject to the condition s constitutes an event. This immediately implies that not only the events (a; S) and (a'; S) are different, but also the events (a; S) and (a; S') — for instance, the occurrence of a head while drawing lots or the occurrence of a head given some other gaming rule. It is also clear that a and S themselves may be events. For instance, consider the events $(a_1; S_1) \equiv A$ and $(a_2; S_2) \equiv B$; then (A; B) is also an event, i.e., (a; S) with $a \equiv A$ and $S \equiv B$. It is important to emphasize that, first, speaking of the events A and (A; B) we use the same notion of event and, second, A and (A; B) are two different events, and not "the same even under different conditions."

Depending on the objective relation between S and a, the realization of S may inevitably lead to the occurrence of a, or inevitably exclude its occurrence, or neither necessarily lead to nor necessarily exclude it; this produces the classification of events into certain, impossible, and random. Thus, the probability of the event (a; S) is an expression of the objective relation between S and a and is in no way connected with how detailed and complete is our knowledge of the relevant situation regarding the event. For instance, let S be the firing at a target under certain conditions and a the hit at a certain point of the target; the firing conditions do not determine uniquely all the parameters needed for the solution of the mechanical problem to find the trajectory of the bullet, and the event (a; S) is therefore a random event. If we could fix all the relevant parameters, then we would have the event (a; S'), where S' differs from S in that it includes also specific values of the "missing" parameters; the event (a; S') is not random — it is either certain or impossible, depending on what specific values of the parameters have been fixed. But the events (a; S) and (a; S') are different events; the event (a; S) of course remains random also after "augmentation of its values."

4. Let us now consider the notion of probability. Experience shows that among repeatable events there are such (we may call them statistically stable) that in almost all sufficiently long series of trials the frequencies of the event A are close to one another and thus to some common number P(A) uniquely determined by the event. Assuming that the numbers P(A) have the properties expressed by the standard axioms of the "purely mathematical probability theory" (it is argued in justification of this assumption that for frequencies the relevant properties are proved), we arrive at a definition of the probability P(A) of the event A as a numerical characteristic of the objective likelihood of its occurrence. Clearly, this definition of probability only applies to repeatable events. If a nonrepeatable event is one of the concrete realizations of a repeatable event (e.g., the occurrence of a head in a specific toss of a coin according to definite rules of the game, which is about to take place now), then we can naturally define the probability of such a nonrepeatable event as the probability of the corresponding repeatable event, but the resulting concept is useless and its application does not yield any new information about the events being studied. With regard to nonrepeatable events which are not realizations of repeatable events, every attempt to introduce the notion of probability for these events inevitably presumes a subjective approach to this concept, which cannot be used as a basis for the construction of a theory with practical relevance.

The definition of an event in Sec. 3 and this definition of probability naturally exclude the need for introducing the notions of conditional and unconditional, prior and posterior probabilities. When we say, for instance, that the unconditional probability of getting a 2 on a throw of a dice is 1/6, while the conditional probability of this event given that an even number has occurred is 1/3, we are actually referring not to two different probabilities of the same event but to the probabilities of two different events.

5. Let us now consider the meaning of the laws of large numbers and their relation to practice. We will only consider the simplest law of large numbers, known as the Bernoulli theorem. This will be quite sufficient for elucidating the fundamental aspects of the problem.

The Bernoulli theorem is stated as follows: as we increase indefinitely the number N of trials of the event A having the probability P(A) = p, the probability of the event ($|\frac{m}{N}-p|<\varepsilon$; N trials of A) goes to 1, where m is the number of occurrences of the event A in N trials of this event and $\varepsilon > 0$ is an arbitrary number. This theorem is usually cited as an argument justifying the closeness of the frequency of an event to its probability in almost all sufficient long series of trials, which in its turn is responsible for the linkage between probability theory and practice. This, however, is not so. If we consider probability theory in the "ordinary" and not "purely mathematical" sense (as we agreed to do in Sec. 1), then closeness of the frequency of an event to its probability is almost invariably incorporated, as we have done above, in the definition of probability, which naturally precedes the construction of the theory, including the proof of the Bernoulli theorem. If we visualize a construction of probability theory in which this is not so, then we must make sure that the proof of the Bernoulli theorem relies only on those properties of probability which are expressed by the axioms of the "purely mathematical probability theory" (and the usual definition of independence of events). Therefore, if we associate with each event A a number R(A) so that these numbers have the appropriate "purely mathematical" properties of probability and for independent events satisfy the multiplicative condition, then for the numbers R(A) we will have the "Bernoulli theorem" regardless of whether they are probabilities in the usual sense or not. With regard to the "purely mathematical theory of probability," it does not contain the notion of frequency and thus does not contain the Bernoulli theorem.

The Bernoulli theorem thus relates to frequencies of repeatable events A treated in nonmathematical sense (see Sec. 3) and also to numbers R(A) (see above) which are not necessarily the probabilities of these events. Therefore, the Bernoulli theorem in itself has absolutely no relation to practice. This relation is determined by the "nonmathematical meaning" assigned in one way or another to the numbers R(A) as they are transformed into probabilities P(A) and by the relation of probability to practice established by this meaning — factors outside the bounds of the Bernoulli theorem. If R(A) is taken as the probability P(A) and the probability itself is defined as suggested in Sec. 4, then the practical relevance of the Bernoulli theorem is apparently in its claim that any specified proximity of the frequency of the event A to its probability can be achieved in the great majority of series of trials of A after the same sufficiently large number of trials in each series.

The relation of the law of large numbers to practice has been discussed by many authors. We did not aim at presenting a comprehensive literature survey. We only note that the first critique of the traditional notions in this field was presented by von Mises.