THEORY PROBAB. APPL.
Vol. 60, No. 1, pp. 62–87

A POSTERIORI INTEGRATION OF PROBABILITIES. ELEMENTARY THEORY*

A. V. KRYAZHIMSKIY†

(*Translated by V. I. Khokhlov*)

Abstract. An approach to a posteriori integration of probability distributions serving as independent a priori models of observed elementary events from a given finite set of elementary events is proposed. A posteriori integration is understood as an improvement of data given by a priori probabilities. The approach is based on the concept of an a posteriori event in the product of probability spaces associated with a priori probabilities. The conditional probability on the product space that is specified by an a posteriori event determines in a natural way the probability on the set of initial elementary events; the latter is recognized as the result of a posteriori integration of a priori models. Conditions under which the integration improves the informativeness of a priori probabilities are established, algebraic properties of integration as a binary operation on the set of probabilities are studied, and the problem of integral convergence of infinite probability sequences is considered.

Key words. consistent observational methods, max-measure of concentration, max-compatibility, marginal compatibility, max-concentrator, integration convergence, integration concentration

DOI. 10.1137/S0040585X97T987466

Introduction. In the study of complex poorly observable systems, especially socio-economic and environmental systems, we often need to compare data obtained from alternative sources. As a rule, they are in a poor agreement. In typical cases, data are represented in the form of probability distributions reflecting observational or simulation noise (see, for example, [10]). In concrete applied research papers, certain specific (known or assumed) properties of objects under study are used to synthesize (integrate) unmatched distributions (see, for example, [6]). In practice, universal methods for integration of unmatched a priori distributions employ the idea of taking their convex combinations with coefficients chosen on the basis of information on the reliability of sources of these distributions (see, for example, [7], [8], [10]). These methods are generally empirical in nature, and specialists recognize the need to work out appropriate formalized approaches (see, for example, [5]). The lack of methods for integration of a priori distributions is even more obvious in the cases of absence of information on the reliability of their sources.

In this paper, which is motivated by the above-mentioned problems of processing the results of observations of complex socio-economic and environmental systems, we propose a unified approach to the integration of a priori distributions given by independent sources—"observational methods"—in the absence of any prioritization of these sources by reliability. It is assumed that some determined element—the value of an index of a socio-economic or environmental system—is observed by several alternative methods and that the result of each method is a probability distribution on the set of all admissible elements (elementary events). These probability distributions are hereinafter referred to as a priori probability distributions or, briefly, a priori probabilities. The problem is to construct an a posteriori probability that gives more

*Received by the editors March 25, 2014. This work was supported by RFBR grants 13-01-00685 and 13-01-12446-ofi-m2 and DG Research COMPLEX, Project 308601. Originally published in the Russian journal *Teoriya Veroyatnostei i ee Primeneniya*, 60 (2015), No. 1, pp. 45–79.
http://www.siam.org/journals/tvp/60-1/T98746.html

†Steklov Mathematical Institute RAS, Moscow, Russia, and International Institute for Applied Systems Analysis, Laxenburg, Austria.

precise information on the observed element by synthesizing data derived from a priori probabilities.

This problem is close in a sense to the problem of estimating an unknown parameter of a probability distribution by observing experimental outcomes (see [1, Chap. I, section 7, p. 97]). However, in our case, admissible elements do not serve as parameters of a priori distributions, and a family of alternative a priori distributions, rather than a collection of empirical results of observations, is used to identify the observed admissible element. To emphasize the peculiarity of the case, we can assume that a priori distributions are empirical frequencies obtained as the result of multiple experiments, and thus they are indistinguishable from probability distributions characterizing errors of the respective observational methods.

Sets of probability distributions are studied in the literature from different points of view. To see this, Wald's theory of statistical decisions [11] focuses on the optimization of decisions with undetermined distribution of "states of nature"; in the theory of comparison of experiments (see [2], [3], [9]), sets of probability distributions serve as models of experiments to be compared by the criterion of informativeness; some studies are devoted to analysis of statistical data generated by different sources (see [4]). In this paper, sets of probability distributions play the role of material for synthesis of integral information on the observed element.

The proposed approach is based on the notion of an a posteriori event in the product of probability spaces corresponding to a priori probabilities. The definition of an a posteriori event is based on the fact that all a priori probabilities are descriptions of the same determined element—the observed elementary event; consequently, an elementary event in the product space can be classified as an a posteriori admissible event only if all its components are identical. The collection of all a posteriori admissible elementary events forms the a posteriori event—the "diagonal" of the product space. The conditional probability given the a posteriori event, which is specified on the product space, is concentrated on the latter and naturally determines probability on the set of initial elementary events; the latter probability is taken as the result of a posteriori integration of a priori probabilities.

For the sake of simplicity, we consider here the case of a finite set of admissible elementary events. Section 1 contains basic definitions and an informal discussion of the approach. Section 2 is devoted to the comparison of informativeness of a priori and a posteriori probabilities. In section 3, we study algebraic properties of a posteriori integration as a binary operation on the set of probabilities. In section 4, studies of asymptotic behavior of the results of a posteriori integration of infinite probability sequences are outlined.

1. Basic definitions and informal discussion.

1.1. Basic elements. Here Z is a nonempty finite set with more than one element; its elements are interpreted as admissible elementary events. The set of all (elementary) probabilities on Z that are understood as nonnegative functions on Z with all the values summing up to the unity is denoted by Π. Any probability $\pi \in \Pi$ specifies the probability space understood as the pair (Z, π). The set of probabilities $\pi \in \Pi$ assuming positive values is denoted by Π^+. The uniform probability on Z that assumes the constant value $1/|Z|$ (hereinafter $|E|$ is the number of elements of a finite set E) is denoted by $\overline{\pi}$. We say that probability $\pi \in \Pi$ is concentrated at the point $z \in Z$ if $\pi(z) = 1$. We say that probability $\pi \in \Pi$ is concentrated if it is concentrated at some point. We put $Z^+(\pi) = \{z \in Z\colon\ \pi(z) > 0\}$ for any probability $\pi \in \Pi$. The set Π is considered to be a metric space with the natural mean square metric $(\pi_1, \pi_2) \mapsto [\sum_{z \in Z} |\pi_1(z) - \pi_2(z)|^2]^{1/2}$; it is clear that Π is a compactum. For any natural number k, we consider Π^k as a product of k copies of the metric space Π; any subset of Π^k is interpreted as its metric subspace. The continuity of functions on subsets of Π^k that assume either real values or values from Π is understood in the sense of the above metric spaces.

1.2. Informal discussion. We assume that a determined, a priori fixed unknown element $z^0 \in Z$ is observed by a researcher using n observational methods numbered $1, \ldots, n$. Each method i is inexact in the sense that it presents z^0 in the form of a probability distribution (probability) π_i on Z. For any $z \in Z$, it is natural to interpret the value $\pi_i(z)$ of the probability π_i as the empirical frequency of detection of element z (as z^0) in a large series of observations with method i. The probabilities $\pi_i, \ldots, \pi_n$ are called *a priori probabilities* or *a priori probability estimates* of the observed element z^0.

Now we consider the a posteriori situation after the observation. The problem is to synthesize more precise, integral information on element z^0 relying on a priori probability estimates.

The proposed approach of solving this nonstrictly posed problem is based on the assumption of mutual independence of observational methods, to be precise, on the assumption that the distribution of observed results $(z_1, \ldots, z_n) \in Z^n$ obtained with methods $1, \ldots, n$ is described by the product space $(Z^n, P) = (Z, \pi_1) \times \cdots \times (Z, \pi_n)$, where $P = \pi_1 \times \cdots \times \pi_n$.

This assumption requires some explanation. It may seem to be unjustified in the case where the observed element has a statistical nature, to be exact, if it is a variable elementary event in some nontrivial probability space. We assume that this case is beyond the scope of this work: as mentioned above, the observed element $z^0 \in Z$ is regarded as being a priori fixed but unknown to the researcher. In this case, the assumption of independence of observational methods in the above meaning reflects a rather typical variety of situations (and is meant in many applied works, some of which are cited in the introduction—see [5], [6], [7], [8], [10]). The corresponding model of observation can be described in the following way. A pair "player–noise" is placed inside a device used for multiple observations of the same fixed element. Every time the device is used, the pair "player–noise" performs a trial to "disturb" the observation, that is, to change from the actually observed element to generally another element—the observational result; the pairs "player–noise" placed in different devices act independently. Let us assume, for example, that the actually observed element is 2; the pair "player–noise" from the first device forms deviations from the true element of sizes 0 and 1 with probabilities $p_1(0)$ and $p_1(1) = 1 - p_1(0)$, respectively, and the pair "player–noise" from the second device, which is independent of the first pair, forms the same deviations from the true element with probabilities $p_2(0)$ and $p_2(1) = 1 - p_2(0)$, respectively. Then the probability spaces $(Z, \overline{p}_1)$ and $(Z, \overline{p}_2)$, where $Z = \{2, 3\}$, $\overline{p}_1(z) = p_1(z-2)$, $\overline{p}_2(z) = p_2(z-2)$ $(z \in Z)$, describe the distributions of the results of observations of the true element with the help of the first and second devices, respectively, and their product $(Z^2, \overline{p}) = (Z, \overline{p}_1) \times (Z, \overline{p}_2)$, where $\overline{p} = \overline{p}_1 \times \overline{p}_2$, describes the distribution of pairs (z_1, z_2) of observational results obtained with the two devices (here the result of observations with the first device goes first, and the result of observations with the second device goes second). The researcher does not know which of the two elements of the set Z is true; nor does he know values $\overline{p}_1(2)$ and $\overline{p}_1(3)$, so he approximates these values by the frequencies $\pi_1(2)$ and $\pi_1(3)$ of occurrence of values 2 and 3 in a large series of observations with the first device;

similarly, he approximates unknown values $\overline{p}_2(2)$ and $\overline{p}_2(2)$ by the frequencies $\pi_2(2)$ and $\pi_2(3)$ of occurrence of values 2 and 3 in a large series of observations with the second device. The functions π_1 and π_2 regarded as being probabilities on Z are taken as a priori probability estimates of the unknown observed true element, and the product space $(Z^2, P) = (Z, \pi_1) \times (Z, \pi_2)$ is taken as a probability (approximating) model describing the distribution of pairs (z_1, z_2) of observational results given by the two devices. The probability P serves as a rather accurate approximation of the above-mentioned probability $\overline{p}$, which characterizes the true distribution of pairs of observational results. On these grounds, the researcher takes the product space (Z^2, P) as a fairly accurate description of the true distribution of pairs (z_1, z_2) of observational results given by the first and second devices. In this scheme (which is roughly sketched or, on the contrary, idealized, for it corresponds to infinite series of observations), the researcher abstracts himself from the approximation nature of the probabilities π_1, π_2, and $P = \pi_1 \times \pi_2$ and uses them to solve the problem of a posteriori integration of a priori probability estimates given by the first and second devices.

The proposed approach to solving the problem of integration of a priori probability estimates is based on the trivial circumstance that, in the a posteriori situation, the results $z_1, \dots, z_n \in Z$ of single observations made with methods $1, \dots, n$, respectively, are true if and only if $z^0 = z_1 = \cdots = z_n$. Since the element z^0 is unknown, the equality $z_1 = \cdots = z_n$ is a necessary condition of a posteriori consistency of the results $z_1, \dots, z_n$. The event

$$A = \{(z_1, \dots, z_n) \in Z^n \colon\ z_1 = \cdots = z_n\} = \{(z, \dots, z) \colon\ z \in Z\}$$

in the product space $(Z^n, P) = (Z, \pi_1) \times \cdots \times (Z, \pi_n)$, where $P = \pi_1 \times \cdots \times \pi_n$, selects all of the a posteriori consistent combinations of observational results; all other combinations of observational results $(z_1, \dots, z_n)$ are mutually inconsistent and thus give false information on the observed element. From here we conclude that, in the a posteriori situation, the event A in the product space (Z^n, P) is realized with certainty. We call it the *a posteriori event*. We have

$$P(A) = \sum_{z \in Z} \pi_1(z) \cdots \pi_n(z).$$

If $P(A) = 0$, then methods $1, \dots, n$ are *inconsistent* in the sense that for any $z \in Z$ at least one method i allows the zero probability that $z = z^0 \colon\ \pi_i(z) = 0$.

Let methods $1, \dots, n$ be *consistent*, that is, $P(A) > 0$. We consider the conditional probability $P(\cdot \mid A)$ on the product space (Z^n, P) given the a posteriori event A:

$$P((z, \dots, z) \,|\, A) = \frac{\pi_1(z) \cdots \pi_n(z)}{P(A)} \qquad (z \in Z).$$

Since the conditional probability $P(\cdot \,|\, A)$ is concentrated on A—the "diagonal" in Z^n, all elements of which have identical components—we identify $P(\cdot \,|\, A)$ with a probability on Z; the latter is denoted by $\pi_1 \cdot \ldots \cdot \pi_n$, so we have

$$(\pi_1 \cdot \ldots \cdot \pi_n)(z) = P((z, \dots, z) \,|\, A) \qquad (z \in Z).$$

We call $\pi_1 \cdot \ldots \cdot \pi_n$ the result of *a posteriori integration* of a priori probabilities $\pi_1, \dots, \pi_n$, and we call the change from $\pi_1, \dots, \pi_n$ to $\pi_1 \cdot \ldots \cdot \pi_n$ the *a posteriori integration* of $\pi_1, \dots, \pi_n$.

Thus, in the probability space $(Z, \pi_1 \cdot \ldots \cdot \pi_n)$, for any $z \in Z$ the probability that z is the true observed element (the probability that $z^0 = z$) is proportional to the probability $\pi_1(z) \cdots \pi_n(z)$ that all observational methods simultaneously allow that z is the true observed element (that $z^0 = z$). The value $\pi_1(z) \cdots \pi_n(z)$ is a "measure of consensus" among methods $1, \ldots, n$ regarding the fact that $z^0 = z$. All methods have equal rights in formation of the "measure of consensus" $\pi_1(z) \cdots \pi_n(z)$, and every method i has the "veto power" in the sense that with $\pi_i(z) = 0$ the "measure of consensus" takes the zero value. We believe that the a posteriori probability $\pi_1 \cdot \ldots \cdot \pi_n$ gives desired integral information on the observed element z^0 that is obtained by a posteriori analysis of the results of its observations with methods $1, \ldots, n$.

The proposed method of a posteriori integration of a priori probability estimates is based on the evident logical fact that the above-mentioned a posteriori event A is certainly realized. This differs from integration methods traditionally used in studies of socio-economic and environmental systems, which often got reduced to believable justification of choice of coefficients of convex combinations of a priori probability estimates. In this sense, the proposed method of a posteriori integration can be more effective that the method of convex combinations. To illustrate this, we give an example.

Example 1.1. This example is inspired by the research on classification of land areas by type (forest, grass, ploughland, desert, etc.) using satellite images that do not provide necessary information (see http://www.geo-wiki.org/). Let Z be a finite set of land types and $z^0 \in Z$ be the type of a particular land area. To estimate the unknown type z^0 of this area, n independent groups of experts are involved; these experts are assumed to have additional knowledge allowing them to make an informed opinion on the land type. The distribution of conclusions of experts from a group numbered i ($i = 1, \ldots, n$) is a probability on Z; we take it as an a priori probability estimate π_i; we assume that $\pi_i(z) > 0$ for any $z \in Z$. Since the expert groups $1, \ldots, n$ are independent, we assume that the distribution of all collections $(z_1, \ldots, z_n)$ of conclusions made by these groups is described by the product space $(Z^n, P) = (Z, \pi_1) \times \cdots \times (Z, \pi_n)$. We consider the integration result $\pi_1 \cdot \ldots \cdot \pi_n$ of a priori probability estimates $\pi_1, \ldots, \pi_n$ as the result of their a posteriori processing. We assume that among the expert groups there are "correctly recognizing" groups i in which the percent of experts concluding that the type of land area is z^0 is maximum: $\pi_i(z^0) > \pi_i(z)$ for any $z \in Z$ not equal to z^0. We denote the set of all such groups by G^+ and the set of all other groups by G^-. Let us assume that for any group $i \in G^+$ for all $z \in Z$ not equal to z^0 the inequality $\pi_i(z^0) > q\pi_i(z)$, where $q > 1$, holds true, and that for any group $j \in G^-$ for all $z \in Z$ not equal to z^0 the inequality $\pi_j(z^0) > r\pi_j(z)$, where $r \in (0, 1)$, holds true. Let m be the number of groups in G^+. We put $\pi^+(z) = \prod_{i \in G^+} \pi_i(z)$, $\pi^-(z) = \prod_{i \in G^-} \pi_i(z)$ $(z \in Z)$. For any $z \in Z$ we have

$$
\begin{aligned}
(\pi_1 \cdot \ldots \cdot \pi_n)(z) &= \frac{\pi_1(z) \cdots \pi_n(z)}{\sum_{y \in Z} \pi_1(y) \cdots \pi_n(y)} \\
&= \left(1 + \sum_{y \in Z \setminus \{z\}} \frac{\pi^+(y)\pi^-(y)}{\pi^+(z)\pi^-(z)}\right)^{-1}
\end{aligned}
$$

and

$$
(\pi_1 \cdot \ldots \cdot \pi_n)(z^0) > \frac{1}{1 + (n-1)/q^m r^{n-m}} > \frac{1}{1+\varepsilon} > 1 - \varepsilon
$$

for arbitrary small $\varepsilon > 0$ if $q^m r^{n-m} > (n-1)/\varepsilon$; for example, the latter inequality is true if $qr^\alpha > 1$, where $\alpha > 0$, the number n of expert groups is sufficiently large, and $n - m < \alpha m$; we have in this case

$$\frac{q^m r^{n-m}}{n-1} = \frac{q^m r^{\alpha(n-m)/\alpha}}{n-1} > \frac{q^m r^{\alpha m}}{n-1} = \frac{(qr^\alpha)^m}{n-1} > \frac{(qr^\alpha)^m}{(\alpha+1)m-1},$$

which tends to infinity as n (along with m) tends to infinity.

Now we consider an arbitrary convex combination $\pi = a_1\pi_1 + \cdots + a_n\pi_n$ a priori probability estimates $\pi_1, \ldots, \pi_n$ as the result of their a posteriori processing; here $a_1, \ldots, a_n \geqq 0$, $a_1 + \cdots + a_n = 1$. For arbitrary $i = 1, \ldots, n$ we have

$$\pi_i(z_0) \leqq 1 - (N-1)\gamma_i,$$

where N is a number of elements of Z and $\gamma_i = \min_{z \in Z \setminus \{z^0\}} \pi_i(z)$; it is evident that $\gamma_i(N - 1 + r) \leqq \sum_{z \in Z} \pi_i(z) = 1$. Consistent with the latter restriction, we assume that for any $i = 1, \ldots, n$ the inequality $\gamma_i > \beta/(N - 1 + r)$, where $\beta \in (0, 1)$, holds true. Then we have

$$\pi(z_0) \leqq \max_{i=1,\ldots,n} \pi_i(z_0) < 1 - \frac{N-1}{N-1+r}\,\beta.$$

Under the above assumptions made with respect to q, r, and γ_i $(i = 1, \ldots, n)$, the right-hand side of this upper estimate is smaller than the right-hand side of the above lower estimate $1 - \varepsilon$ for $(\pi_1 \cdot \ldots \cdot \pi_n)(z^0)$ if ε is sufficiently small (this means that n is sufficiently large).

So, in this example, the result of processing a priori probability estimates by the proposed method is preferable when compared to the result of their processing by the method of convex combinations.

1.3. Definitions. Let us give rigorous definitions. For an arbitrary positive integer $n > 1$, probabilities $\pi_1, \ldots, \pi_n \in \Pi$ are called *inconsistent* if $\pi_1(z) \cdots \pi_n(z) = 0$ for all $z \in Z$; otherwise probabilities $\pi_1, \ldots, \pi_n$ are called *consistent*; the set of all $(\pi_1, \ldots, \pi_n) \in \Pi^n$ such that $\pi_1, \ldots, \pi_n$ are consistent is denoted by $\Pi^{(n)}$.

Remark 1.1. The following statements are obviously true:

(i) $(\pi_1, \ldots, \pi_n) \in \Pi^{(n)}$ for any probability $\pi \in \Pi$ and any positive integer $n > 1$;

(ii) $(\pi_1, \pi_2) \in \Pi^{(2)}$ for all probabilities $\pi_1 \in \Pi$ and $\pi_2 \in \Pi^+$;

(iii) if probabilities $(\pi_1, \ldots, \pi_n) \in \Pi^{(n)}$, then $(\pi_{i_1}, \ldots, \pi_{i_n}) \in \Pi^{(n)}$ for any permutation $(i_1, \ldots, i_n)$ in $(1, \ldots, n)$;

(iv) $(\Pi^+)^n \subset \Pi^{(n)}$ for any positive integer $n > 1$.

Consistent with the preliminary definition given above, for any positive integer $n \geqq 2$ we consider a mapping $(\pi_1, \ldots, \pi_n) \mapsto \pi_1 \cdot \ldots \cdot \pi_n \colon \Pi^{(n)} \mapsto \Pi$ such that

$$(\pi_1 \cdot \ldots \cdot \pi_n)(z) = \frac{\pi_1(z) \cdots \pi_n(z)}{\sum_{z' \in Z} \pi_1(z') \cdots \pi_n(z')}, \qquad (z \in Z)$$

for any collection $(\pi_1, \ldots, \pi_n) \in \Pi^{(n)}$. This mapping is called the n-fold *a posteriori integration* (briefly, *integration*); for any collection $(\pi_1, \ldots, \pi_n) \in \Pi^{(n)}$ the probability $\pi_1 \cdot \ldots \cdot \pi_n$ is called *the a posteriori integration result* (briefly, *integration result*) of probabilities $\pi_1, \ldots, \pi_n$.

Remark 1.2. It is easy to see that the n-fold a posteriori integration is continuous for any positive integer $n \geqq 2$.

2. Integration and informativeness estimates.

2.1. Integration and extreme elements. Let us consider the change from an a priori probability estimate $\pi_1 \in \Pi$ to an a posteriori probability estimate $\pi_1 \cdot \pi_2$ obtained by integration of π_1 with some probability $\pi_2 \in \Pi$. We can state the following.

Remark 2.1. (i) Integration of an arbitrary probability $\pi \in \Pi$ with the uniform probability $\overline{\pi}$ does not change π: $\pi \cdot \overline{\pi} = \overline{\pi} \cdot \pi = \pi$. (ii) If $(\pi, \pi_*) \in \Pi^{(2)}$ and the probability π_* is concentrated, then the integration of π with π_* converts π into π_*: $\pi \cdot \pi_* = \pi_* \cdot \pi = \pi_*$.

We also note that, in typical cases, upon integration of an a priori probability π_1 with π_2, the probability of an element that is most likely with respect to the probability space (Z, π_2) increases, and the probability of a least likely element decreases.

LEMMA 2.1. *Let $(\pi_1, \pi_2) \in \Pi^{(2)}$, let z^* be a maximum point of π_2, $\pi_1(z^*) > 0$, and let there exist an element $z \in Z$ such that $\pi_2(z) < \pi_2(z^*)$ and $\pi_1(z)\pi_2(z) > 0$. Then $(\pi_1 \cdot \pi_2)(z^*) > \pi_1(z^*)$.*

Proof. We put $Z^* = \{z' \in Z\colon \pi_2(z') = \pi_2(z^*)\}$. We note that $z \in Z \setminus Z^*$. By definition, we have

$$
\begin{aligned}
(\pi_1 \cdot \pi_2)(z^*) &= \frac{\pi_1(z^*)\pi_2(z^*)}{\sum_{z' \in Z} \pi_1(z')\pi_2(z')} \\
&= \frac{\pi_1(z^*)}{\sum_{z' \in Z^*} \pi_1(z') + \pi_1(z)q(z) + \sum_{z' \in Z \setminus (Z^* \cup \{z\})} \pi_1(z')q(z')},
\end{aligned}
\tag{2.1}
$$

where

$$
q(z') = \frac{\pi_2(z')}{\pi_2(z^*)} \qquad (z' \in Z).
$$

Since $\pi_2(z^*) = \max_{z \in Z}$, we have $q(z') < 1$ for all $z' \in Z \setminus Z^*$, which, along with the inequality $\pi_1(z)\pi_2(z) > 0$, implies the estimate $\pi_1(z)q(z) < \pi_1(z)$. Consequently, the denominator on the right-hand side of (2.1) is smaller than $\sum_{z' \in Z} \pi_1(z') = 1$. Now (2.1) and the assumption $\pi_1(z^*) > 0$ yield $(\pi_1 \cdot \pi_2)(z^*) < \pi_1(z^*)$. The lemma is proved.

Similar reasoning leads to the following symmetric statement.

LEMMA 2.2. *Let $(\pi_1, \pi_2) \in \Pi^{(2)}$, let z_* be a minimum point of π_2, $\pi_1(z_*) > 0$, and let there exist an element $z \in Z$ such that $\pi_2(z) > \pi_2(z_*)$ and $\pi_1(z)\pi_2(z) > 0$. Then $(\pi_1 \cdot \pi_2)(z_*) < \pi_1(z_*)$.*

2.2. Measures of concentration. For probabilities from Π, we consider numeric indicators that assume the largest values on concentrated probabilities; we call them measures of concentration. Measures of concentration can be interpreted as indices of informativeness of probabilities. If the result of integration of two a priori probabilities (a priori estimates given by independent methods) has a larger measure of concentration than each of them, then we have reason to believe that the a priori probability estimates are consistent: when interacting, they carry more information than each of them separately. In the opposite situation, when the result of integration of two a priori probability estimates has a smaller measure of concentration than each of them, the a priori estimates are in conflict with each other, and one of them is likely to be rejected. Finally, in the intermediate situation, when the measure of concentration of the probability resulting from integration of two a priori probabilities is larger than the minimum and smaller than the maximum of their measures of concentration, the a priori probability models have a dissymmetric interrelation; that is, one of them makes the other more precise, but not vice versa. From the practical point of view, of greatest interest is the first of the above situations, where the result of integration of a priori probabilities π_1 and π_2 has a larger measure of concentration than each of them; in this case, we say that a pair (π_1, π_2) of probabilities is compatible (with respect to the given measure of concentration).

So, a *measure of concentration* is an arbitrary continuous function $\mu\colon \Pi \mapsto (-\infty, 1]$ such that $\mu(\pi) = 1$ if and only if the probability π is concentrated. A pair $(\pi_1, \pi_2) \in \Pi^2$ is called *compatible* with respect to a measure of concentration μ if $(\pi_1, \pi_2) \in \Pi^{(2)}$ and $\mu(\pi_1 \cdot \pi_2) > \max\{\mu(\pi_1), \mu(\pi_2)\}$, and it is *incompatible* with respect to μ if $(\pi_1, \pi_2) \in \Pi^{(2)}$ and $\mu(\pi_1 \cdot \pi_2) < \min\{\mu(\pi_1), \mu(\pi_2)\}$.

The simplest measure of concentration is the function $\pi \mapsto \max_{z\in Z} \pi(z)$; we call it the max-*measure of concentration.* In this study, we consider only the max-measure of concentration.

Remark 2.2. Now we give several examples of other measures of concentration: (i) $\pi \mapsto \max_{z\in Z} \pi(z) - \min_{z\in Z} \pi(z)$; (ii) $\pi \mapsto \sum_{z\in Z} \pi^k(z)$, where $k > 1$; (iii) $\pi \mapsto 1-\sum_{z\in Z}[\xi(z)-\sum_{z'\in Z} \xi(z')\pi(z')]^2\pi(z)$, where ξ is an arbitrary real-valued one-to-one function on Z; the value of this measure of concentration at $\pi \in \Pi$ is variance of the random variable ξ on (Z, π); (iv) $\pi \mapsto 1 + \sum_{z\in Z} \pi(z) \log \pi(z)$ (for $\pi(z) = 0$ we put $\pi(z) \log \pi(z) = 0$); the latter sum with the opposite sign is known as the entropy of π.

Remark 2.3. It is clear that the minimum value of the max-measure of concentration is $1/|Z|$; it is assumed on the uniform probability $\overline{\pi}$ only.

For brevity, a pair $(\pi_1, \pi_2) \in \Pi^2$ that is compatible (incompatible) with respect to the max-measure of concentration is called max-*compatible* (max-*incompatible*).

Using Lemma 2.1, we derive a typical case when a pair $(\pi_1, \pi_2) \in \Pi^{(2)}$ is max-compatible. In this case, it is assumed that sets of elementary events that are most likely in probability spaces (Z, π_1) and (Z, π_2) have a nonempty intersection.

THEOREM 2.1. *Let $(\pi_1, \pi_2) \in \Pi^{(2)}$, and let there exist an element $z^* \in Z$ that maximizes each probability π_1, and let π_2 on Z. The following statements hold true.*

1. *The element z^* maximizes $\pi_1 \cdot \pi_2$ on Z.*

2. *If there exists an element $z \in Z$ such that $0 < \pi_1(z) < \max_{z'\in Z} \pi_1(z')$ and $0 < \pi_2(z) < \max_{z'\in Z} \pi_2(z')$, then the pair (π_1, π_2) is* max-*compatible.*

Proof. Statement 1 follows directly from definition of the result of integration $\pi_1 \cdot \pi_2$. So, let us prove statement 2. It is evident that $\pi_1(z^*) > 0$. By assumption, we have $\pi_2(z) < \pi_2(z^*)$ and $\pi_1(z)\pi_2(z) > 0$. Consequently, all the assumptions of Lemma 2.1 are satisfied. Using this lemma, we get that $(\pi_1 \cdot \pi_2)(z^*) > \pi_1(z^*) = \max_{z'\in Z} \pi_1(z')$. Interchanging π_1 and π_2, we arrive at the symmetric relations $(\pi_1 \cdot \pi_2)(z^*) > \pi_2(z^*) = \max_{z'\in Z} \pi_2(z')$. The proof is complete.

Theorem 2.1 immediately yields the following statement.

COROLLARY 2.1. *For any not concentrated not uniform probability $\pi \in \Pi$ the pair (π, π) is* max-*compatible.*

The notion of max-compatibility of pairs $(\pi_1, \pi_2) \in \Pi^{(2)}$ may be extended to include n-fold collections $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$. For any positive integer $n \geqq 2$, a collection $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$ is called max-*compatible* if

$$\max_{z\in Z}(\pi_1 \cdot \dots \cdot \pi_n)(z) > \max\{\max_{z\in Z} \pi_1(z), \dots, \max_{z\in Z} \pi_n(z)\}.$$

The following statement is based on Theorem 2.1.

COROLLARY 2.2. *Let $n \geqq 2$ be a positive integer, $(\pi_1, \dots, \pi_2) \in \Pi^{(n)}$, and there exists an element $z^* \in Z$ that maximizes each probability $\pi_1, \dots, \pi_n$ on Z. The following statements are true.*

1. *The element z^* maximizes $\pi_1 \cdot \dots \cdot \pi_n$ on Z.*

2. *If there exists an element $z \in Z$ such that $0 < \pi_i(z) < \max_{z' \in Z} \pi_i(z')$ for all $i \in \{1, \dots, n\}$, then the collection $(\pi_1, \dots, \pi_n)$ is* max-*compatible.*

The following statement derived from Corollary 2.1 is related to the topological structure of the set of integration-invariant probabilities. We say that a set $G \subset \Pi$ is *integration-invariant* if any $\pi_1, \pi_2 \in G$ are consistent and $\pi_1 \cdot \pi_2 \in G$.

THEOREM 2.2. *Let a nonempty set $G \subset \Pi$ be integration-invariant and not contain concentrated probabilities. Then one and only one of the following statements is true*: (i) *the set G has one element, and this element is the uniform probability $\overline{\pi}$*; (ii) *G is not closed in Π.*

Let the set G consist of one element π. Then we have $\pi \cdot \pi = \pi$. Let us assume that $\pi \neq \overline{\pi}$. From Corollary 2.1, the pair (π, π) is max-compatible, that is, $\max_{z \in Z}(\pi \cdot \pi)(z) > \max_{z \in Z} \pi(z)$. The latter cannot be true, since we have $\pi \cdot \pi = \pi$. Consequently, $\pi = \overline{\pi}$.

Let G consist of more than one element. Then G contains the uniform probability. Let us assume that G is closed. We put $p = \sup_{\pi \in G} \max_{z \in Z} \pi(z)$. Since G contains the uniform probability, the quantity p is larger than $1/|Z|$, the value of the uniform probability. Since G is closed, there exists a probability $\pi_* \in G$ such that $\max_{z \in Z} \pi_*(z) = p$. From the estimate $p > 1/|Z|$ it follows that π_* is not uniform. Since, by assumption, the set G does not contain concentrated probabilities, π_* is not concentrated. Then, from Corollary 2.1, the pair (π_*, π_*) is max-compatible, that is, $\max_{z \in Z}(\pi_* \cdot \pi_*)(z) > \max_{z \in Z} \pi_*(z) = p$. But we have $\pi_* \cdot \pi_* \in G$, and hence we infer that $\max_{z \in Z}(\pi_* \cdot \pi_*)(z) \leqq p$. This contradiction completes the proof.

Theorem 2.2 has the following direct corollary.

COROLLARY 2.3. *Let a nonempty set $G \subset \Pi$ be integration-invariant, not contain concentrated probabilities, and consist of more than one element. Then the set G is not finite.*

To conclude this subsection, we note that the situation when a pair $(\pi_1, \pi_2) \in \Pi^{(2)}$ is max-incompatible means that elements with high probabilities in the probability space (Z, π_1) have small probabilities in the probability space (Z, π_2) and vice versa, which suggests the qualitative inconsistency between the probability models π_1 and π_2. Below we give a simple example of a max-incompatible pair of probabilities.

Example 2.1. We take $Z = \{z_1, z_2\}$, $\pi_1(z_1) = 3/4$, $\pi_1(z_2) = 1/4$, $\pi_2(z_1) = 1/4$, $\pi_2(z_2) = 3/4$. Then $(\pi_1 \cdot \pi_2)(z_1) = (\pi_1 \cdot \pi_2)(z_2) = 1/2 < 3/4 = \max_{z \in Z} \pi_1(z) = \max_{z \in Z} \pi_2(z)$.

2.3. Marginal measure. For any not concentrated probability $\pi \in \Pi$, we define the *marginal measure* of $\pi \in \Pi$ as $\min_{z \in Z^+(\pi)} \pi(z)$.

It is clear that for any not concentrated probability $\pi \in \Pi$, its marginal measure does not exceed $1/|Z|$ and is equal to $1/|Z|$ if and only if the probability π is uniform ($\pi = \overline{\pi}$). In this context the smaller the marginal measure of a not concentrated probability π, the less it is uniform. As the uniform probability is least informative among all probabilities from Π, probabilities with small marginal measures may be interpreted as being more informative in a sense than probabilities with large (close to $1/|Z|$) marginal measures.

We say that a pair $(\pi_1, \pi_2) \in \Pi^{(2)}$ of not concentrated probabilities is *marginally compatible* if

$$\min_{z' \in Z^+(\pi_1,\pi_2)} (\pi_1 \cdot \pi_2)(z') < \min\left\{ \min_{z' \in Z^+(\pi_1)} \pi_1(z'), \min_{z' \in Z^+(\pi_2)} \pi_2(z') \right\}$$

and *marginally incompatible* if the opposite strict inequality holds.

Lemma 2.2 allows us to describe a typical situation where a pair $(\pi_1, \pi_2) \in \Pi^{(2)}$ is marginally compatible: this property occurs if the sets of the least likely elements in probability spaces (Z, π_1) and (Z, π_2) have a nonempty intersection.

THEOREM 2.3. *Let probabilities $\pi_1, \pi_2 \in \Pi$ not be concentrated, $(\pi_1, \pi_2) \in \Pi^{(2)}$, and there exists an element $z_* \in Z$ that minimizes the probability p_1 on $Z^+(\pi_1)$ and the probability p_2 on $Z^+(\pi_2)$. Then the following statements are true.*

1. *The element z_* minimizes $\pi_1 \cdot \pi_2$ on $Z^+(\pi_1 \cdot \pi_2)$.*
2. *It there exists an element $z \in Z$ such that*

$$\pi_1(z) > \min_{z' \in Z^+(\pi_1)} \pi_1(z') \quad \textit{and} \quad \pi_2(z) > \min_{z' \in Z^+(\pi_2)} \pi_2(z'),$$

then the pair (π_1, π_2) is marginally compatible.

The proof of this theorem is similar to that of Theorem 2.1. Theorem 2.3 implies the following.

COROLLARY 2.4. *For any not concentrated not uniform probability $\pi \in \Pi$, the pair (π, π) is marginally compatible.*

We say that an n-fold collection $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$ of not concentrated probabilities (here n is a positive integer exceeding unity) is *marginally compatible* if

$$\min_{z \in Z} (\pi_1 \cdot \dots \cdot \pi_n)(z) < \min\{\min_{z \in Z} \pi_1(z), \dots, \min_{z \in Z} \pi_n(z)\}.$$

We give without proof a natural extension of Theorem 2.3.

COROLLARY 2.5. *Let $n \geqq 2$ be a positive integer, let probabilities $\pi_1, \dots, \pi_n \in \Pi$ be not concentrated, $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$, and there exists an element $z_* \in Z$ such that for any $i = 1, \dots, n$ it minimizes π_i on the set $Z^+(\pi_i)$. Then the following statements are true.*

1. *The element z_* minimizes $\pi_1 \cdot \dots \cdot \pi_n$ on $Z^+(\pi_1 \cdot \dots \cdot \pi_n)$.*
2. *If there exists an element $z \in Z$ such that*

$$\pi_i(z) > \max_{z' \in Z} \pi_i(z') \quad \textit{for all } i \in \{1, \dots, n\},$$

then the collection $(\pi_1, \dots, \pi_n)$ is marginally compatible.

The situation where a pair $(\pi_1, \pi_2) \in \Pi^{(2)}$ of not concentrated probabilities is marginally incompatible is similar in a sense to the situation where this pair is not max-compatible: both situations mean that elementary events with large probabilities in the probability space (Z, π_1) have small probabilities in the probability space (Z, π_2) and vice versa. Example 2.1 illustrates this fact.

2.4. Max-concentrators. A probability $\pi \in \Pi$ is called a max-*concentrator* for a collection $(\pi_1, \dots, \pi_n) \in \Pi^n$ (here n is a positive integer exceeding unity) if the pair (π, π_i) is max-compatible for any $i \in \{1, \dots, n\}$. So, a probability estimate π that is a max-concentrator for a collection $(\pi_1, \dots, \pi_n)$ of estimates increases the max-measure of concentration of each of them through integration.

It is easy to detect max-concentrators for collections of pairwise max-compatible probabilities.

THEOREM 2.4. *Let $n \geqq 2$ be a positive integer, and let a collection $(\pi_1, \dots, \pi_n) \in \Pi^n$ of probabilities be such that for any different $i, j \in \{1, \dots, n\}$ the pair (π_i, π_j) is* max-*compatible. Then for any $i \in \{1, \dots, n\}$ the probability π_i is a* max-*concentrator for the collection $(\pi_1, \dots, \pi_n)$.*

Proof. We take $i, j \in \{1, \dots, n\}$, $j \neq i$. We have $(\pi_i, \pi_i) \in \Pi^{(2)}$ (see Remark 1.1(i)); besides, $(\pi_i, \pi_j) \in \Pi^{(2)}$ by assumption. It is also assumed that the pair (π_i, π_j) is max-compatible. This excludes the fact that the probability π_i is uniform (see Remark 2.1(i)). Consequently, from Corollary 2.1, the pair (π_i, π_i) is max-compatible. The proof is complete.

The following theorem demonstrates that, in typical cases, a probability giving sufficiently strong preference to an elementary event that has a nonzero probability in every probability space $(Z, \pi_1), \dots, (Z, \pi_n)$ is a max-concentrator for the collection $(\pi_1, \dots, \pi_n)$.

THEOREM 2.5. *Let $n \geqq 2$ be a positive integer, $(\pi_1, \dots, \pi_n) \in \Pi^n$, the probabilities $\pi_1, \dots, \pi_n$ be not concentrated, and an element $z_* \in Z$ be such that $\pi_i(z_*) > 0$ for any $i \in \{1, \dots, n\}$. Then any probability $\pi \in \Pi$ such that $\pi(z_*)$ is sufficiently close to unity is a* max-*concentrator for the collection $(\pi_1, \dots, \pi_n)$.*

Proof. We take the probability $\pi_* \in \Pi$ that is concentrated at z_*. It is clear that $(\pi_*, \pi_i) \in \Pi^{(2)}$ for any $i \in \{1, \dots, n\}$. According to Remark 2.1(ii), for any $i \in \{1, \dots, n\}$ we have $\pi_* \cdot \pi_i = \pi_*$ and, consequently, $\max_{z \in Z}(\pi_* \cdot \pi_i)(z) = 1 > \max_{z \in Z} \pi_i(z)$; the latter inequality is due to the fact that the probability π_i is not concentrated. Owing to continuity of integration (see Remark 1.2) and max-measure of concentration, the latter inequality holds true for all $i \in \{1, \dots, n\}$ when replacing the probability π_* by any $\pi \in \Pi$ such that $\pi(z_*)$ is sufficiently close to unity. The proof is thus complete.

The following theorem is related to collections of probabilities that are sufficiently close to the uniform probability.

THEOREM 2.6. *Let $n \geqq 2$ be a positive integer, and let a probability $\pi \in \Pi$ not be uniform. Then π is a* max-*concentrator for any collection $(\pi_1, \dots, \pi_n) \in \Pi^n$ such that the probabilities $\pi_1, \dots, \pi_n$ are sufficiently close to the uniform probability $\overline{\pi}$.*

Proof. According to Remark 2.1(i), the equality $\pi \cdot \overline{\pi} = \pi$ holds true. Then, given the fact that the probability π is not uniform, we have $\max_{z \in Z}(\pi \cdot \overline{\pi}) = \max_{z \in Z} \overline{\pi}(z) = 1/|Z|$. Because of continuity of integration (see Remark 1.1) and continuity of the max-measure of concentration, the latter inequality holds upon replacing the uniform probability $\overline{\pi}$ in it by any probabilities $\pi_1, \dots, \pi_n \in \Pi$ that are sufficiently close to $\overline{\pi}$. The proof is complete.

Now let us show that, if $n < |Z|$, in typical cases, for a given collection $(\pi_1, \dots, \pi_n)$ of probabilities there exists a max-concentrator that is sufficiently close to the uniform probability. Interpreting this property, we can say that an estimate given by any of n observational methods may be improved through integration with an estimate given by the same additional observational method with a rather low information quality.

THEOREM 2.7. *Let $N = |Z|$ and n be a positive integer such that $2 \leqq n \leqq N$. Let $\pi_1, \dots, \pi_n \in \Pi^n$, $Z = \{z_1, \dots, z_N\}$, z_{k_i} be a point of maximum of probability π_i for any $i \in \{1, \dots, n\}$,*

$$v_{ik} = \begin{cases} 0 & \textit{if } k \neq k_i, \\ 1 & \textit{if } k = k_i \end{cases} \qquad (i \in \{1, \dots, n\},\ k \in \{1, \dots, N\}),$$

and the rank of the matrix

$$A = \begin{pmatrix} \pi_1(z_1) - v_{11} & \pi_1(z_2) - v_{12} & \dots & \pi_1(z_N) - v_{1N} \\ \dots & \dots & \dots & \dots \\ \pi_n(z_1) - v_{n1} & \pi_n(z_2) - v_{n2} & \dots & \pi_n(z_N) - v_{nN} \\ 1 & 1 & \dots & 1 \end{pmatrix} \tag{2.2}$$

not be smaller than $n+1$*. Then, for any* $\varepsilon > 0$*, there exists a* max*-concentrator* π *for the collection* $(\pi_1, \dots, \pi_n)$ *such that the distance in* Π *from* π *to the uniform probability* $\overline{\pi}$ *is smaller than* ε*.*

Proof. For any probability $\pi \in \Pi$, we denote by π^* the vector of its values $(\pi(z_1), \dots, \pi(z_N)) \in \mathbf{R}^N$. For any vector $h = (h_1, \dots, h_N) \in \mathbf{R}^N$ such that

$$h_1 + \dots + h_N = 0 \tag{2.3}$$

and any sufficiently small $\lambda > 0$, we obviously have

$$\overline{\pi}^* + \lambda h \in \Pi^* = \{\pi^* \colon \ \pi \in \Pi\}. \tag{2.4}$$

For any vector $p = (p+1, \dots, p_N) \in \mathbf{R}^N$ with positive components, we put

$$g_{ik}(p) = \frac{\pi_i(z_k) p_k}{\sum_{j=1}^N \pi_i(z_j) p_j} \qquad (i \in \{1, \dots, n\}, \ k \in \{1, \dots, N\}). \tag{2.5}$$

It is clear that for any probability $\pi \in \Pi$ with positive values we have

$$g_{ik}(\pi^*) = (\pi_i \cdot \pi)(z_k) \qquad (i \in \{1, \dots, n\}, \ k \in \{1, \dots, N\}); \tag{2.6}$$

in particular,

$$g_{ik}(\overline{\pi}^*) = (\overline{\pi}_i \cdot \pi)(z_k) = \pi(z_k) \qquad (i \in \{1, \dots, n\}, \ k \in \{1, \dots, N\}). \tag{2.7}$$

Now, in view of the fact that the relations (2.6) and (2.4) hold for all $h \in \mathbf{R}^N$ satisfying (2.3) and all sufficiently small $\lambda > 0$, it remains to show that there exists a vector $h \in \mathbf{R}^N$ such that the inequalities

$$\max_{k=1,\dots,N} g_{ik}(\overline{\pi}^* + \lambda h) > \max_{k=1,\dots,N} \pi_i(z_k) \qquad (i \in \{1, \dots, n\})$$

are true for all sufficiently small $\lambda > 0$. Taking into account the fact that z_{k_i} for any $i \in \{1, \dots, n\}$ is a point of maximum of probability π_i, it is enough to establish that for some $h \in \mathbf{R}^N$ and all sufficiently small $\lambda > 0$ the inequalities

$$g_{ik_i}(\overline{\pi}^* + \lambda h) > \pi_i(z_{k_i}) \qquad (i \in \{1, \dots, n\}) \tag{2.8}$$

hold. Let us show this. We note that, with a given $h \in \mathbf{R}^N$ and all sufficiently small $\lambda > 0$, the inequalities (2.8) are equivalent to the inequalities

$$g_{ik_i}(\overline{\pi}^*) + \langle \operatorname{grad} g_{ik_i}(\overline{\pi}^*), h \rangle \lambda > \pi_i(z) i k_i \qquad (i \in \{1, \dots, n\})$$

or, in view of (2.7) (where $\pi = \pi_i$ for $i \in \{1, \dots, n\}$), to the inequalities

$$\langle \operatorname{grad} g_{ik_i}(\overline{\pi}^*), h \rangle > 0 \qquad (i \in \{1, \dots, n\}); \tag{2.9}$$

here $\operatorname{grad} g_{ik}(\overline{\pi}^*)$ is the gradient of the function $p \mapsto g_{ik}(p)$ at the point $\overline{\pi}^*$ and $\langle \cdot, \cdot \rangle$ is the scalar product in $\mathbf{R}^N$.

Now let us show that there exists a vector $h \in \mathbf{R}^N$ satisfying the equality (2.3) and the inequalities (2.9). In view of (2.5), we have

$$\operatorname{grad} g_{ik_i}(\overline{p}) = (\gamma_i^{(1)}, \dots, \gamma_i^{(N)}),$$

where

$$\gamma_i^{(k)} = -\frac{\pi_i(z_{k_i})\pi_i(z_k)(1/N)}{(\sum_{j=1}^N \pi_i(z_j)(1/N))^2} = -N\pi_i(z_{k_i})\pi_i(z_k) \quad \text{for } k \neq k_i,$$
$$\gamma_i^{(k_i)} = \frac{\pi_i(z_{k_i})\sum_{j=1}^N \pi_i(z_j)(1/N) - \pi_i(z)\pi_i(z_k)(1/N)}{(\sum_{j=1}^N \pi_i(z_j)(1/N))^2} = N[\pi_i(z_{k_i}) - \pi_i^2(z_{k_i})];$$

it is taken into account here that $\sum_{j=1}^N \pi_i(z_j) = 1$. So for an arbitrary vector $h = (h_1, \dots, h_N) \in \mathbf{R}^N$ we have

$$\begin{aligned}&\langle \operatorname{grad} g_{ik_i}(\overline{\pi}^*), h\rangle \\ &= -N\pi_i(z_{k_i})\Bigg(\sum_{k=1}^{k_i-1} \pi_i(z_k)h_k + (\pi_i(z_{k_i}) - 1)h_{k_i} + \sum_{k=k_i+1}^{N} \pi_i(z_k)h_k\Bigg),\end{aligned} \tag{2.10}$$

where $i \in \{1, \dots, n\}$. We take $a_1, \dots, a_n < 0$. Let us consider the following system of linear algebraic equations with $h_1, \dots, h_N$:

$$\sum_{k=1}^{k_i-1} \pi_i(z_k)h_k + (\pi_i(z_{k_i}) - 1)h_{k_i} + \sum_{k=k_i+1}^{N} \pi_i(z_k)h_k = a_k, \tag{2.11}$$
$$(i \in \{1, \dots, n\}),$$
$$h_1 + \dots + h_N = 0. \tag{2.12}$$

Its matrix form is as follows: $Ah^T = a^T$, where the matrix A is given by (2.2), $a = (a_1, \dots, a_n, 0)$, and T is the sign of transposition for row vectors. By assumption, in the first place, the rank of the matrix A is not smaller than the number of its rows $n + 1$ (the number of equations in the system (2.11), (2.12)); in the second place, the latter number does not exceed N, which is the number of columns of the matrix A (the number of unknowns in the system (2.11), (2.12)). Thus the system of equations (2.11), (2.12) has a solution. Let $h = (h_1, \dots, h_N)$ be its solution. As $a_1, \dots, a_n < 0$, the right-hand sides of the equalities (2.10) are positive, and so the inequalities (2.9) hold true. The proof is complete.

3. Algebraic properties of integration.

3.1. Integration as multiplication. The following theorem establishes that integration as a binary operation is commutative and associative, that is, it possesses characteristic algebraic properties of multiplication. First, we give the following obvious remark.

Remark 3.1. If $(\pi_1, \pi_2, \pi_3) \in \Pi^{(3)}$, then $((\pi_1 \cdot \pi_2), \pi_3), (\pi_1, (\pi_2, \pi_3)) \in \Pi^{(2)}$.

THEOREM 3.1. *The following statements are true.*

1. *Integration is commutative*, *i.e.*, $\pi_1 \cdot \pi_2 = \pi_2 \cdot \pi_1$ *for any* $(\pi_1, \pi_2) \in \Pi^{(2)}$.

2. *Integration is associative, i.e.,* $(\pi_1 \cdot \pi_2) \cdot \pi_3 = \pi_1 \cdot (\pi_2 \cdot \pi_3)$ *for any* $(\pi_1, \pi_2, \pi_3) \in \Pi^{(3)}$.

Proof. Statement 1 is obviously true. Let us prove statement 2. We take an arbitrary collection $(\pi_1, \pi_2, \pi_3) \in \Pi^{(3)}$. Let us consider an arbitrary element $z \in Z$. By definition, we have

$$(\pi_1 \cdot \pi_2)(z) = \pi_1(z)\pi_2(z)c_{12}, \qquad \text{where} \quad c_{12} = \frac{1}{\sum_{z' \in Z} \pi_1(z')\pi_2(z')},$$

and

$$((\pi_1 \cdot \pi_2) \cdot \pi_3)(z) = (\pi_1 \cdot \pi_2)(Z)\pi_3(z)c_{(12)3} = \pi_1(z)\pi_2(z)\pi_3(z)c_{12}c_{(12)3},$$

where

$$c_{(12)3} = \frac{1}{\sum_{z' \in Z}(\pi_1 \cdot \pi_2)(z')\pi_3(z')} = \frac{1}{\sum_{z' \in Z} \pi_1(z')\pi_2(z')\pi_3(z')c_{12}}.$$

This yields that

$$((\pi_1 \cdot \pi_2) \cdot \pi_3)(z) = \frac{\pi_1(z)\pi_2(z)\pi_3(z)}{\sum_{z' \in Z} \pi_1(z')\pi_2(z')\pi_3(z')}.$$

We can similarly establish that the right-hand side of the latter equality coincides with $(\pi_1 \cdot (\pi_2 \cdot \pi_3))(z)$. In view of the fact that $z \in Z$ is arbitrary, we have $(\pi_1 \cdot \pi_2) \cdot \pi_3 = \pi_1 \cdot (\pi_2 \cdot \pi_3)$. The proof is complete.

Remark 3.2. Remark 2.1 can be interpreted in the following way: the uniform probability $\overline{\pi}$ plays the role of unity with respect to integration as multiplication, and any concentrated probability plays the role of zero.

Theorem 3.1 yields the following statement.

COROLLARY 3.1. *For any positive integer* $n > 1$ *and any collection* $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$, *the integration result* $\pi_1 \cdot \ldots \cdot \pi_n$ *does not change when "multiplying" in any order and in any number of steps*; *to be precise,* $\pi_1 \cdot \ldots \cdot \pi_n = (\pi_{i_1} \cdot \ldots \cdot \pi_{i_{k_1}}) \cdot (\pi_{i_{k_1}+1} \cdot \ldots \cdot \pi_{i_{k_2}}) \cdot \ldots \cdot (\pi_{i_{k_m}} \cdot \ldots \cdot \pi_{i_n})$ *for any permutation* $(i_1, \dots, i_n)$ *of* $(1, \dots, n)$ *and any increasing sequence* $(k_j)_1^m$ *from* $\{2, \dots, n-1\}$.

3.2. Integration degrees. In view of Corollary 3.1, for any probability $\pi \in \Pi$ and any positive integer $n > 1$, we denote the result of integration of n copies of π by π^n; the probability π^n is called the nth *integration degree* of probability π; for consistency, probability π is called the first integration degree and is denoted by π^{-1}.

The next theorem, which follows directly from the definition of integration result, establishes that the nth integration degree of a probability inherits the order of its values and makes the difference between them larger as n increases.

THEOREM 3.2. *For any probability* $\pi \in \Pi$ *and any positive integer* n, *the following statements are true.*

1. $Z^+(\pi^n) = Z^+(\pi)$.
2. *For any elements* $z_1, z_2 \in Z^+(\pi)$ *the equality*

$$\frac{\pi^n(z_1)}{\pi^n(z_2)} = \left(\frac{\pi(z_1)}{\pi(z_2)}\right)^n$$

holds.

Theorem 3.2 obviously yields the following.

COROLLARY 3.2. *Let Z_* be the set of all points of maximum of a probability $\pi \in \Pi$, and let the probability $\pi_* \in \Pi$ be uniform on Z_*, i.e., $\pi_*(z) = 0$ for all $z \in Z \setminus Z^+(\pi)$ and $\pi_*(z) = 1/|Z_*|$ for all $z \in Z_*$. Then $\pi^n \to \pi_*$ in Π as $n \to \infty$.*

We call a probability $\pi_* \in \Pi$ an *integration root of the nth degree* of a probability $\pi \in \Pi$ if $\pi_*^n = \pi$ (here n is a positive integer). According to Theorem 3.2, an integration root π_* of the nth degree of a probability $\pi \in \Pi$ inherits the order of values of π and smoothes the difference between them; this fact allows for the interpretation that the estimate π_* is due to an imperfect prototype of the method giving the estimate π.

THEOREM 3.3. *For any probability $\pi \in \Pi$ and any positive integer n there exists a unique integration root of the nth degree of π.*

Proof. Let us consider some probability $\pi \in \Pi$. By definition, probability $\pi_* \in \Pi$ is an integration root of the nth degree of π if $\pi_*^n = \pi$ or, which is the same, the relation

$$\frac{\pi_*^n(z)}{\sum_{z' \in Z} \pi_*^n(z')} = \pi(z) \qquad (z \in Z)$$

is true. Let us put elements of the set Z in order; i.e., we put $Z = \{z_1, \dots, z_N\}$, where $N = |Z|$. Then the criterion that $\pi_* \in \Pi$ is an integration root of the nth degree of π is that the vector $(\pi_*^n(z_1), \dots, \pi_*^n(z_N), \sum_{i=1}^N \pi_*^n(z_i))$ solves the system of algebraic equations

$$\begin{aligned} x_1 - \pi(z_1)x_{N+1} &= 0, \\ \dots\dots\dots&\dots\dots\dots \\ x_N - \pi(z_N)x_{N+1} &= 0, \\ x_1 + \cdots + x_N - x_{N+1} &= 0 \end{aligned} \tag{3.1}$$

with additional restrictions

$$x_1, \dots, x_N \geqq 0, \quad x_1^{1/n} + \cdots + x_N^{1/n} = 1. \tag{3.2}$$

Let A denote the matrix of the system (3.1). We have

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & -\pi(z_1) \\ 0 & 1 & 0 & \dots & 0 & 0 & -\pi(z_2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 & -\pi(z_N) \\ 1 & 1 & 1 & \dots & 1 & 1 & -1 \end{pmatrix}.$$

The sum of the first N rows of the matrix A is equal to its $(N+1)$th row, and its left upper submatrix of size $N \times N$ is not degenerate. Consequently, the rank of the matrix A is equal to N. So the set of all solutions of the system (3.1) forms a one-dimensional subspace in $\mathbf{R}^{N+1}$. Let $(y_1, \dots, y_{N+1})$ be some nonzero solution of the system (3.1). Since $\pi(z_1), \dots, \pi(z_N)$ are nonnegative and there are nonzero values among them, then, as it can be seen from (3.1), $y_{N+1} \neq 0$, there are nonzero numbers among $y_1, \dots, y_N$, and the signs of all such nonzero numbers coincide with the sign of y_{N+1}. Without loss of generality, we assume that $y_{N+1} > 0$ (otherwise we multiply $y_1, \dots, y_{N+1}$ by -1). Then we have $y_1, \dots, y_{N+1} \geqq 0$.

We put

$$\lambda = \left(\frac{1}{y_1^{1/n} + \cdots + y_N^{1/N}} \right)^n, \tag{3.3}$$

$$x_i = \lambda y_i \qquad (i \in \{1, \dots, N+1\}). \tag{3.4}$$

It is evident that $x_1, \dots, x_N$ satisfy the inequality from (3.2). Further, we have

$$x_1^{1/n} + \cdots + x_N^{1/n} = \lambda^{1/n}(y_1^{1/n} + \cdots + y_N^{1/n}) = 1; \tag{3.5}$$

so the equality from (3.2) holds true for $x_1, \dots, x_N$. Hence, an integration root π_* of the nth degree of π exists and can be found from the relations

$$\pi_*(z_i) = x_i^{1/n} \qquad (i \in \{1, \dots, N\}). \tag{3.6}$$

If π_* is an integration root of the nth degree of π, numbers $x_1, \dots, x_N$ are specified by (3.6), and $x_{N+1} = x_1 + \cdots + x_N$; then, as established above, $x_1, \dots, x_{N+1}$ form a solution of the system (3.1) and satisfy the restrictions (3.2). Since the subspace of all solutions of (3.1) is one-dimensional, the relations (3.4) hold for some real number λ. Then the relations (3.2) imply (3.5); consequently, λ is determined from (3.3). So there exists a unique integration root of the nth degree of π. The proof is complete.

Since, according to Theorem 3.3, for any positive integer n there exists a unique integration root of the nth degree of an arbitrary probability $\pi \in \Pi$, hereinafter we denote it by $\pi^{1/n}$.

Remark 3.3. Theorem 3.2 yields that, for any probability $\pi \in \Pi$ and any element $z \in Z \setminus Z^+(\pi)$, the equality $\pi^{1/n}(z) = 0$ is true for any positive integer n, and for any $z_1, z_2 \in Z^+(\pi)$ we have the convergence

$$\frac{\pi^{1/n}(z_1)}{\pi^{1/n}(z_2)} = \left(\frac{\pi(z_1)}{\pi(z_2)}\right)^{1/n} \to 1 \quad \text{as } n \to \infty.$$

The latter means that $\pi^{1/n} \to \pi_*$ in Π, where the probability π_* is uniform on $Z^+(\pi)$, i.e., $\pi_*(z) = 0$ for all $z \in Z \setminus Z^+(\pi)$ and $\pi_*(z) = 1/|Z^+(\pi)|$ for all $z \in Z^+(\pi)$.

In view of Theorem 3.3, we now introduce rational integration degrees of probabilities. That is, for any probability $\pi \in \Pi$ and any positive integers n and m, we call the probability $\pi^{m/n} = (\pi^m)^{1/n}$ the *m/nth integration degree* of the probability π.

Remark 3.4. Standard arithmetic relations hold true for rational integration degrees of probabilities. To be precise, for any probability $\pi \in \Pi$ and any positive integers n and m, the probability $\pi^{m/n} = (\pi^m)^{1/n}$ can be also defined as $\pi^{m/n} = (\pi^{1/n})^m$. To see this, we have $((\pi^{1/n})^m)^n = (\pi^{1/n})^{mn} = (\pi^{1/n})^{nm} = ((\pi^{1/n})^n)^m = \pi^m$. Referring to the definition of the root of the nth degree of π^m, we obtain that $\pi^{m/n} = (\pi^{1/n})^m$.

3.3. Disintegration. From the definition of integration of probabilities, it follows that $Z^+(\pi \cdot \pi_1) = Z^+(\pi) \cap Z^+(\pi_1)$ for any probabilities $(\pi, \pi_1) \in \Pi^{(2)}$. With this fact in mind, we give the following definition. For probabilities $\pi_1, \pi_2 \in \Pi$ such that $Z^+(\pi_2) \subset Z^+(\pi_1)$, a probability $\pi \in \Pi$ is called a *result of disintegration of* π_2 *over* π_1 if $Z^+(\pi_2) = Z^+(\pi) \cap Z^+(\pi_1)$ (and thus $(\pi, \pi_1) \in \Pi^{(2)}$) and $\pi \cdot \pi_1 = \pi_2$.

THEOREM 3.4. *Let probabilities $\pi_1, \pi_2 \in \Pi$ be such that $Z^+(\pi_2) \subset Z^+(\pi_1)$. Then the following statements are true.*

1. *There exists a result of disintegration of the probability π_2 over the probability π_1.*
2. *If probability π is a result of disintegration of π_2 over π_1, then probability $\pi' \in \Pi$ is also a result of disintegration of π_2 over π_1 if and only if for some $\mu > 0$ the equality $\pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)}$, where $\pi|_{Z^+(\pi_1)}$ and $\pi'|_{Z^+(\pi_1)}$ are the restrictions of π and π' to $Z^+(\pi_1)$ respectively, holds true.*

Proof. Let us organize elements of the set Z and put $Z = \{z_1, \dots, z_N\}$, where $N = |Z|$; without loss of generality, we can assume that $Z^+(\pi_2) = \{z_1, \dots, z_k\}$ and $Z^+(\pi_1) = \{z_1, \dots, z_m\}$ for some $k, m \in \{1, \dots, N\}$, $m \geqq k$. Then we have

$$\pi_1(z_i)\pi_2(z_i) < 0 \qquad (i \in \{1, \dots, k\}), \tag{3.7}$$

$$\pi_1(z_i) < 0, \quad \pi_2(z_i) = 0 \qquad (i \in \{k+1, \dots, m\}), \tag{3.8}$$

$$\pi_1, \pi_2(z_i) = 0 \qquad (i \in \{m+1, \dots, N\}). \tag{3.9}$$

By definition, a probability $\pi \in \Pi$ is a result of disintegration of π_2 over π_1 if $Z^+(\pi) \cap Z^+(\pi_1)$ and

$$\frac{\pi(z)\pi_1(z)}{\sum_{z' \in Z} \pi(z')\pi_1(z')} = \pi_2(z) \qquad (z \in Z)$$

or, which is equivalent, the vector $(\pi(z_1), \dots, \pi(z_N), \ \sum_{i=1}^N \pi(z_i)\pi_1(z_i))$ is a solution of the system of algebraic equations

$$\begin{array}{l} \pi_1(z_1)x_1 - \pi_2(z_1)x_{N+1} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \pi_1(z_N)x_N - \pi_2(z_N)x_{N+1} = 0, \\ \pi_1(z_1)x_1 + \dots + \pi_1(z_N)x_N - x_{N+1} = 0 \end{array} \tag{3.10}$$

with additional restrictions

$$x_1, \dots, x_k > 0, \quad x_{k+1}, \dots, x_m = 0, \quad x_{m+1}, \dots, x_N \geqq 0, \quad x_1 + \dots + x_N = 1 \tag{3.11}$$

(in the cases $k = m$ and $m = N$, the second or third restriction should be omitted, respectively).

If $k < m$, we put

$$x_{k+1}, \dots, x_m = 0. \tag{3.12}$$

Then, in view of (3.8), $x_{k+1}, \dots, x_m$ satisfy the equations from (3.10) in the rows numbered $k+1, \dots, m$. Appealing to (3.9), we see that if $m < N$, then the equations from (3.10) in the rows numbered $m+1, \dots, N$ are satisfied by arbitrary numbers $x_{m+1}, \dots, x_N$.

Now we consider the rest of the system (3.10), i.e., its subsystem consisting of equations in the rows numbered $1, \dots, k$ and $N+1$:

$$\begin{array}{l} \pi_1(z_1)x_1 - \pi_2(z_1)x_{N+1} = 0, \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \pi_1(z_k)x_k - \pi_2(z_k)x_{N+1} = 0, \\ \pi_1(z_1)x_1 + \dots + \pi_1(z_k)x_k - x_{N+1} = 0. \end{array} \tag{3.13}$$

The latter equation from (3.13) is equivalent to the $(N+1)$th equation from (3.10). This follows from (3.12) and (3.9). Let A denote the matrix of the system (3.13). Then we have

$$A = \begin{pmatrix} \pi_1(z_1) & 0 & 0 & \dots & 0 & 0 & -\pi_2(z_1) \\ 0 & \pi_1(z_2) & 0 & \dots & 0 & 0 & -\pi_2(z_2) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0\dots & 0 & -\pi_2(z_1) & -\pi_2(z_k) & \\ \pi_1(z_1) & \pi_1(z_2) & \pi_1(z_3) & \dots & \pi_1(z_{k-2}) & \pi_1(z_{k-2}) & -1 \end{pmatrix}.$$

The first k rows of the matrix A add up to its $(k+1)$th row, and its upper left submatrix of size $(k \times k)$ does not degenerate in view of (3.7). So, the rank of matrix A of the size $((k+1) \times (k+1))$ is equal to k. Consequently, the set of all solutions of the system (3.13) is a one-dimensional subspace in $\mathbf{R}^{k+1}$. Let $(y_1, \dots, y_k, y_{N+1})$ be a nonzero solution of the system (3.13). It is evident that $y_{N+1} \neq 0$. Without loss of generality, we can assume that $y_{N+1} > 0$ (otherwise, we multiply $(y_1, \dots, y_k, y_{N+1})$ by -1). Then $y_1, \dots, y_k, y_{N+1} > 0$. We take $c \in (0, 1]$ and put

$$\lambda = \frac{c}{y_1 + \cdots + y_k}, \qquad x_i = \lambda y_i \qquad (i \in \{1, \dots, k, N+1\}). \tag{3.14}$$

It is obvious that $x_1, \dots, x_k > 0$, $x_1 + \cdots + x_k = c$. Combining this with (3.12), we obtain that $x_1, \dots, x_m$ satisfy the first and the last restrictions in (3.11).

If $m = N$, then we put $c = 1$ and (in the case $m > k$) refer to (3.12); as a result, we obtain that the collection $(x_1, \dots, x_N, x_{N+1})$ of nonnegative numbers solves the system (3.10) and satisfies the restrictions (3.11) (where the third restriction should be omitted, and the second restriction should be omitted in the case $m = k$). If $m < N$, then we take arbitrary numbers $x_{m+1}, \dots, x_N \geqq 0$ that add up to $1 - c$. It can be seen that the collection $(x_1, \dots, x_N, x_{N+1})$ solves the system (3.10) and satisfies the restrictions (3.11) (where the second restriction should be omitted in the case $m = k$). So, the probability $\pi \in \Pi$ determined by the relations

$$\pi(z_i) = x_i \qquad (i \in \{1, \dots, N\}) \tag{3.15}$$

is a result of disintegration of π_2 over π_1. Statement 1 is proved.

Let us prove statement 2. Let a probability $\pi' \in \Pi$ be such that $\pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)}$ for some $\mu > 0$ or, what is the same, its values

$$x'_i = \pi'(z_i) \qquad (i \in \{1, \dots, N\}) \tag{3.16}$$

satisfy the equalities

$$x'_i = \mu x_i \qquad (i \in \{1, \dots, m\}).$$

Then, in view of (3.12) (in the case $m > k$), we have

$$x'_{k+1}, \dots, x'_m = 0. \tag{3.17}$$

We put

$$x'_{N+1} = \pi_1(z_1)x'_1 + \cdots + \pi_1(z_N)x'_N. \tag{3.18}$$

It is obvious that $x'_{N+1} = \mu x_{N+1}$. Since a collection $(x_1, \dots, x_m, x''_{m+1}, \dots, x''_N, x_{N+1})$ with arbitrary numbers $x''_{m+1}, \dots, x''_N \geqq 0$ (in the case $m < N$) solves the system (3.10) with the restrictions (3.11), we obtain that $(\pi'(z_1), \dots, \pi'(z_N), x'_{N+1}) = (x'_1, \dots, x'_{N+1})$ possesses the same properties. Consequently, the probability π' is a result of disintegration of the probability π_2 over the probability π_1.

Conversely, let a probability π' be a result of disintegration of the probability π_2 over the probability π_1. Then the collection $(x'_1, \dots, x'_{N+1})$ determined by the formulas (3.16) and (3.18) is a solution of the system (3.10) with the restrictions (3.11); in particular, with $m > k$, we have (3.17). Then the relations

$$x'_i = \lambda y_i \qquad (i \in \{1, \dots, k, N+1\})$$

hold true for some $\lambda' > 0$. According to (3.14), we hence obtain

$$x'_i = \mu x_i \qquad (i \in \{1, \dots, k, N+1\}),$$

where $\mu = \lambda'/\lambda$. Consequently, in view of (3.15), (3.16), (3.12), and (3.17) (the two latter relations occur only in the case $m > k$), we conclude that

$$\pi'(z_i) = \mu\pi(z_i) \qquad (i \in \{1, \dots, m\}).$$

So we have $\pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)}$. Theorem 3.4 is proved.

For arbitrary probabilities $\pi_1, \pi_2 \in \Pi$ such that $Z^+(\pi_2) \subset Z^+(\pi_1)$, we denote the set of all results of disintegration of π_2 over π_1 by $[\pi_2/\pi_1]$. The multivalued mapping $(\pi_1, \pi_2) \mapsto [\pi_2/\pi_1]$ that is defined on the set of all $(\pi_1, \pi_2) \in \Pi^2$ such that $Z^+(\pi_2) \subset Z^+(\pi_1)$ is called *disintegration.*

COROLLARY 3.3. *Let probabilities* $\pi_1, \pi_2 \in \Pi$ *be such that* $Z^+(\pi_2) \subset Z^+(\pi_1)$. *The following statements are valid.*

1. *If* $Z^+(\pi_1) = Z$, *then the set* $[\pi_2/\pi_1]$ *consists of one element.*
2. *If* $Z^+(\pi_1) \neq Z$ *and* $\pi \in [\pi_2/\pi_1]$, *then the relation*

$$[\pi_2/\pi_1] = \left\{\pi' \in \Pi\colon\ \pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)},\ 0 < \mu \leqq \frac{1}{\sum_{z \in Z^+(\pi_1)} \pi(z)}\right\}$$

holds.

Proof. Let $Z^+(\pi_1) = Z$. If $\pi' \in [\pi_2/\pi_1]$, then, from statement 2 of Theorem 3.4, we have $\pi' = \pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)} = \mu\pi$ for some $\mu > 0$. As $\pi, \pi' \in \Pi$, it is necessary that $\mu = 1$. Statement 1 is thus proved.

Let $Z^+(\pi_1) \neq Z$. From statement 2 of Theorem 3.4, $\pi' \in [\pi_2/\pi_1]$ if and only if $\pi'|_{Z^+(\pi_1)} = \mu\pi|_{Z^+(\pi_1)}$ for some $\mu > 0$. The latter is, in turn, possible if and only if $\mu \sum_{z \in Z^+(\pi_1)} \pi(Z) \in (0, 1]$. Statement 2 is proved.

Remark 3.5. Under the assumptions of Corollary 3.3, there exists a unique probability $\pi' \in [\pi_2/\pi_1]$ such that $Z^+(\pi') = Z^+(\pi_2)$. Indeed, let a probability $\pi' \in [\pi_2/\pi_1]$ be such that

$$\pi'|_{Z^+(\pi_1)} = \frac{1}{\sum_{z \in Z^+(\pi_1)} \pi(z)}\, \pi|_{Z^+(\pi_1)}. \tag{3.19}$$

Then $\sum_{z \in Z^+(\pi_1)} \pi'(z) = 1$, whence it follows that $Z^+(\pi') \subset Z^+(\pi_1)$. Further, the fact that $\pi' \cdot \pi_1 = \pi_2$ implies the equality $\pi'(z) = 0$ for any $z \in Z^+(\pi_1) \setminus Z^+(\pi_2)$ and the inequality $\pi'(z) > 0$ for any $z \in Z^+(\pi_2)$. Consequently, $Z^+(\pi') = Z^+(\pi_2)$. Conversely, if a probability $\pi' \in [\pi_2/\pi_1]$ is such that $Z^+(\pi') = Z^+(\pi_2)$, then the relations (3.19) hold.

Based on Remark 3.5, for any probabilities $\pi_1, \pi_2 \in \Pi$ such that $Z^+(\pi_2) \subset Z^+(\pi_1)$, we denote the only probability $\pi \in [\pi_2/\pi_1]$ such that $Z^+(\pi) = Z^+(\pi_2)$ by π_2/π_1.

Remark 3.6. For any probability $\pi \in \Pi^+$ (with positive values; see notation in section 1), we have $Z^+(\pi) = Z$. Thus the reduction of disintegration on the product $\Pi^+ \times \Pi^+$ is correctly defined. From statement 1 of Corollary 3.3, the reduction of disintegration on the product $\Pi^+ \times \Pi^+$ is a one-to-one mapping; it evidently assumes values in the set Π^+. So, the set Π^+ is invariant with respect to both integration and disintegration. It can be easily seen that disintegration as a (one-to-one) function on $\Pi^+ \times \Pi^+$ with values in Π^+ is continuous. As noted above (see Remark 1.1), integration is also continuous. Consequently, with account for commutativity and associativity of integration (Theorem 3.1), we conclude that the set Π^+ with integration understood as an algebraic operation of multiplication is a topological Abelian group, in which the uniform probability $\overline{\pi}$ plays the role of unity (see Remark 2.1).

With respect to the operation of integration, the operation of disintegration plays the same role that the operation of division plays with respect to the operation of multiplication in arithmetics. The restriction $Z^+(\pi_2) \subset Z^+(\pi_1)$ in defining the "quotient" $[\pi_2/\pi_1]$ is an analogue of the standard arithmetic restriction that a divisor is not equal to zero. The following theorem states that the relationship between the operations of integration and disintegration is completely similar to that between the arithmetic operations of multiplication and division.

THEOREM 3.5. *The following statements are valid.*

1. *Let probabilities* $\pi_1, \pi_2, \pi_3 \in \Pi$ *be such that* $Z^+(\pi_3) \subset Z^+(\pi_2) \subset Z^+(\pi_1)$*. Then*

$$[[\pi_3/\pi_2]/\pi_1] = [\pi_3/(\pi_2 \cdot \pi_1)], \tag{3.20}$$

where

$$[[\pi_3/\pi_2]/\pi_1] = \bigcup_{\pi' \in [\pi_3/\pi_2]} [\pi'/\pi_1].$$

2. *Let probabilities* $\pi_1, \pi_2, \pi_3 \in \Pi$ *be such that* $Z^+(\pi_1) = Z^+(\pi_2) = Z^+(\pi_3)$ *and* $\pi_1 \cdot \pi_2 = \pi_1 \cdot \pi_3$*. Then* $\pi_2 = \pi_3$*.*

3. *Let probabilities* $\pi_1, \pi_2, \pi_3 \in \Pi$ *be such that* $Z^+(\pi_1) = Z^+(\pi_2) = Z^+(\pi_3)$*. Then*

$$(\pi_3 \cdot \pi_2)/\pi_1 = (\pi_3/\pi_1) \cdot \pi_2. \tag{3.21}$$

Proof. Let us prove statement 1. Let probabilities $\pi_1, \pi_2, \pi_3 \in \Pi$ be such that $Z^+(\pi_3) \subset Z^+(\pi_2) \subset Z^+(\pi_1)$. We take some probability π from $[[\pi_3/\pi_2]/\pi_1]$. Then $\pi \in [\pi'/\pi_1]$ for some $\pi' \in [\pi_3/\pi_2]$. Hence it follows that $\pi \cdot \pi_1 = \pi'$ and $\pi' \cdot \pi_2 = p_3$. Thus we have $\pi \cdot \pi_1 \cdot \pi_2 = \pi_3$ which is the same as $\pi \cdot (\pi_1 \cdot \pi_2) = \pi_3$. Consequently, $\pi \in [\pi_3/(\pi_2 \cdot \pi_1)]$. Now let us show the inverse inclusion. We take some probability $\pi \in [\pi_3/(\pi_2 \cdot \pi_1)]$. We have $\pi \cdot (\pi_2 \cdot \pi_1) = \pi_3$ which is the same as $(\pi \cdot \pi_1) \cdot \pi_2 = \pi_3$. Consequently, $\pi' = (\pi \cdot \pi_1) \in [\pi_3/\pi_2]$ and $\pi \in [\pi'/\pi_1]$. The necessary inverse inclusion is established. The equality (3.20) is valid.

Let us prove statement 2. Let probabilities $\pi_1, \pi_2, \pi_3 \in \Pi$ be such that $Z^+(\pi_3) = Z^+(\pi_2) = Z^+(\pi_1)$ and $\pi_1 \cdot \pi_2 = \pi_1 \cdot \pi_3$. It is evident that $\pi_2 = (\pi_1 \cdot \pi_2)/\pi_1$ and $\pi_3 = (\pi_1 \cdot \pi_3)/\pi_1$. This yields $\pi_2 = \pi_3$.

Let us prove statement 3. Let probabilities $\pi_1, \pi_2, \pi_3 \in \Pi$ be such that $Z^+(\pi_1) = Z^+(\pi_2) = Z^+(\pi_3)$. We put $\pi = (\pi_3 \cdot \pi_2)/\pi_1$. By definition and assumptions, we have

$$Z^+(\pi_3 \cdot \pi_2) = Z^+(\pi_1) = Z^+(\pi) \tag{3.22}$$

and $\pi \cdot \pi_1 = \pi_3 \cdot \pi_2$. Let $\pi' = \pi/\pi_2$. By definition and in view of (3.22), the equalities

$$Z^+(\pi') = Z^+(\pi) = Z^+(\pi_1) \tag{3.23}$$

and

$$\pi = \pi' \cdot \pi_2 \tag{3.24}$$

hold. From here, we have $\pi' \cdot \pi_2 \cdot \pi_1 = \pi_3 \cdot \pi_2$ or

$$(\pi' \cdot \pi_1) \cdot \pi_2 = \pi_3 \cdot \pi_2. \tag{3.25}$$

Using (3.23) and the assumption, we obtain the equalities $Z^+(\pi' \cdot \pi_1) = Z^+(\pi_1) = Z^+(\pi_2) = Z^+(\pi_3)$. Then from statement 2 of (3.25) it follows that $\pi' \cdot \pi_1 = \pi_3$. Hence $\pi' = \pi_3/\pi_1$. Consequently, in view of (3.24), the equalities $\pi = \pi' \cdot \pi_2 = (\pi_3/\pi_1) \cdot \pi_2$ are true. This proves the equality (3.21). The proof is complete.

4. Integration and probability sequences.

4.1. Integration limits. We call a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π *consistent* if for any positive integer n the probabilities $\pi_1, \dots, \pi_n$ are consistent or, what is the same, $(\pi_1, \dots, \pi_n) \in \Pi^{(n)}$. We note that, for a consistent sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π, the result $\pi_1 \cdot \ldots \cdot \pi_n$ of integration of the probabilities $\pi_1, \dots, \pi_n$ is determined for any positive integer n. For an arbitrary consistent sequence $(\pi_i)_{n=1}^{\infty}$ of probabilities from Π, we call each partial limit in Π of the sequence $(\pi_1 \cdot \ldots \cdot \pi_n)_{n=1}^{\infty}$ a *partial integration limit* of the sequence $(\pi_i)_{i=1}^{\infty}$; if there exists only a partial integration limit of a sequence $(\pi_i)_{i=1}^{\infty}$, we call it the *integration limit* of the sequence $(\pi_i)_{i=1}^{\infty}$.

Remark 4.1. As noted in section 1, the set Π as a topological space is a compactum. Thus, every consistent sequence of probabilities in Π has at least one partial limit.

We say that a consistent sequence of probabilities in Π *integrationally converges* if it has an integration limit, and it *integrationally diverges* in the opposite case. An integrationally converging sequence of probabilities in Π is called *integrationally concentrated* if its integration limit is a concentrated probability.

These definitions allow simple informal interpretations. If a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities is integrationally concentrated, then the corresponding observational methods $1, 2, \dots$ interact in the course of their sequential addition, improve information on the observed element, and give complete information on it in the limit. If a sequence $(\pi_i)_{i=1}^{\infty}$ integrationally converges but is not integrationally concentrated, then the methods $1, 2, \dots$ "find consensus" in the course of their sequential integration and finally give substantial, though not complete, information on the observed element. If a sequence $(\pi_i)_{i=1}^{\infty}$ integrationally diverges, then the methods $1, 2, \dots$ do not give consistent information on the observed element in the limit.

Now we give an example of a consistent integrationally diverging sequence of probabilities.

Example 4.1. Let $Z = \{z_1, z_2\}$, probabilities $\pi^{(1)}, \pi^{(2)} \in \Pi$ be such that

$$\pi^{(1)}(z_1) > \pi^{(1)}(z_2) > 0, \quad \pi^{(2)}(z_2) > \pi^{(2)}(z_1) > 0,$$

and a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be determined by the following relations:

$$\begin{aligned} \pi_i &= \pi^{(1)} \quad (i \in \{1, \dots, k_{2^j-1}\},\ j \in \{1, 2, \dots\}), \\ \pi_i &= \pi^{(2)} \quad (i \in \{k_{2^j-1}+1, \dots, k_{2^j}\},\ j \in \{1, 2, \dots\}), \end{aligned}$$

where $1 < k_1 < k_2 < k_3 < \cdots$. The sequence $(\pi_i)_{i=1}^{\infty}$ is obviously consistent. We put

$$\pi^* j = \pi_1 \cdot \ldots \cdot \pi_{k_j}, \quad \pi^*_{j,j+1} = \pi_{k_j+1} \cdot \ldots \cdot \pi_{k_{j+1}} \qquad (j = 1, 2, \dots), \tag{4.1}$$

$$q^{(1)} = \frac{\pi^{(1)}(z_2)}{\pi^{(1)}(z_1)}, \quad q^{(2)} = \frac{\pi^{(2)}(z_2)}{\pi^{(2)}(z_1)}.$$

It is obvious that $q^{(1)} < 1$, $a^{(2)} > 1$. We have

$$\pi_1^*(z_1) = \frac{\pi^{(1)k_1}(z_1)}{\pi^{(1)k_1}(z_1) + \pi^{(1)k_1}(z_2)} = \frac{1}{1 + q^{(1)k_1}}.$$

Let us take a sequence $(\varepsilon_j)_{j=1}^{\infty}$ of positive numbers that tends to zero. Let a positive integer k_1 be such that $q^{(1)k_1} < \varepsilon_1$. Then we have

$$\pi_1^*(z_1) > \frac{1}{1+\varepsilon_1}. \tag{4.2}$$

Furthermore, we obtain

$$\pi_{1,2}^*(z_1) = \frac{\pi^{(2)(k_2-k_1)}(z_1)}{\pi^{(2)(k_2-k_1)}(z_1) + \pi^{(2)(k_2-k_1)}(z_2)} = \frac{1}{1 + q^{(2)(k_2-k_1)}}, \tag{4.3}$$

$$\pi_2^*(z_1) = (\pi_1^* \cdot \pi_{1,2}^*)(z_1) = \frac{\pi_1^*(z_1)\pi_{1,2}^*(z_1)}{\pi_1^*(z_1)\pi_{1,2}^*(z_1) + \pi_1^*(z_2)\pi_{1,2}^*(z_2)}$$

$$< \frac{\pi_1^*(z_1)\pi_{1,2}^*(z_1)}{\pi_1^*(z_2)\pi_{1,2}^*(z_2)} = \frac{\pi_1^*(z_1)}{\pi_1^*(z_2)} \frac{\pi_{1,2}^*(z_1)}{1 - \pi_{1,2}^*(z_1)}. \tag{4.4}$$

Taking into account (4.3) and the inequality $q^{(2)} > 1$, we choose a positive integer $k_2 > k_1$ in such a way that the right-hand side of (4.4) is smaller than ε_2. This yields

$$\pi_2^*(z_1) < \varepsilon_2. \tag{4.5}$$

As with the estimates (4.2) and (4.5), we provide the validity of the inequalities

$$\pi_{2^j-1}^*(z_1) > \frac{1}{1+\varepsilon_{2^j-1}}, \quad \pi_{2^j}^*(z_1) < \varepsilon_{2^j} \qquad (j \in \{1, 2, \ldots\})$$

by choosing positive integers $k_3, k_4, \ldots$. It can be seen that probability $\lim_{j\to\infty} \pi_{2^j-1}^*$ is concentrated at z_1 and the probability $\lim_{j\to\infty} \pi_{2^j}^*$ is concentrated at z_2. According to (4.1), both probabilities are partial integration limits of the sequence $(\pi_i)_{i=1}^{\infty}$. Consequently, the sequence $(\pi_i)_{i=1}^{\infty}$ integrationally diverges.

Theorems 4.1, 4.2, and 4.3 specify simple conditions of integration concentration of probability sequences. Theorem 4.1 follows directly from the fact that concentrated probabilities play the role of zero elements with respect to the operation of integration understood as multiplication (see Remark 3.2).

THEOREM 4.1. *Let $(\pi_i)_{i=1}^{\infty}$ be a consistent sequence of probabilities from Π and a probability π_k be concentrated for some positive integer k. Then the sequence $(\pi_i)_{i=1}^{\infty}$ is integrationally concentrated and π_k is its integration limit.*

The next theorem following from the definition of a measure of concentration gives a criterion of concentration of a probability sequence.

THEOREM 4.2. *Let a measure μ be a measure of concentration. A consistent sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π is integrationally concentrated if and only if* $\lim_{n\to\infty} \mu(\pi_1 \cdot \ldots \cdot \pi_n) = 1$.

The following theorem indicates that if probabilities forming a consistent sequence give an unambiguous preference to the same element, then the sequence is integrationally concentrated at this element.

THEOREM 4.3. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent, and there exists an element $z_* \in Z$ and a positive number $q < 1$ such that $\pi_i(z)/\pi_i(z_*) < q$*

for all $z \in Z \setminus \{z_*\}$ and all positive integers i. Then the sequence $(\pi_i)_{i=1}^{\infty}$ is integrationally concentrated and its integration limit is concentrated at z_*.

Proof. For each element $z \in Z$ and each positive integer n, we put

$$v_n(z) = \frac{(\pi_1 \cdot \ldots \cdot \pi_n)(z)}{(\pi_1 \cdot \ldots \cdot \pi_n)(z_*)} = \frac{\pi_1(z) \cdots \pi_n(z)}{\pi_1(z_*) \cdots \pi_n(z_*)}. \tag{4.6}$$

It is obvious that, for each element $z \in Z \setminus \{z_*\}$, we have $v_n(z) \leqq q^n$ $(n \in \{1, 2, \ldots\})$ and, consequently, $(\pi_1 \cdot \ldots \cdot \pi_n)(z) \to 0$ as $n \to \infty$. So the sequence $(\pi_1 \cdot \ldots \cdot \pi_n)_{n=1}^{\infty}$ converges in Π to the probability concentrated at z_*. The proof is complete.

The following example demonstrates that, if the conditions of Theorem 4.3 are satisfied with $q = 1$, then the statement of the theorem is generally not true.

Example 4.2. Let $Z = \{z, z_*\}$, $\pi_1, \pi_2, \ldots \in \Pi$, and $q_i = \pi_i(z)/\pi_i(z_*) \in (0, 1)$ for all positive integers i. Furthermore, let $q_{i+1} > q_i$ for all positive integers i and the series $\sum_{i=1}^{\infty} |\log q_i|$ converge. Then the sequence $(v_n)_{n=1}^{\infty}$ specified by the formula (4.6) assumes values in $(0, 1)$ and decreases; consequently, $v_n \to v \in [0, 1)$ as $n \to \infty$. So, the sequence $(\pi_i)_{i=1}^{\infty}$ is consistent and integrationally converges. Since the series $\sum_{i=1}^{\infty} |\log q_i|$ converges, we have $v = \lim_{n\to\infty} q_1 \cdots q_n > 0$. Hence the integration limit of the sequence $(\pi_i)_{i=1}^{\infty}$ is not concentrated.

Now we give some sufficient conditions of integration convergence of a probability sequence. The statement below follows directly from the definition of integration and its continuity (see Remark 1.1).

THEOREM 4.4. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent, let there exist a positive integer $k \geqq 2$ such that the sequence $(\pi_i)_{i=k}^{\infty}$ integrationally converges, and let π be its integration limit. Then the sequence $(\pi_i)_{i=1}^{\infty}$ integrationally converges and $\pi_1 \cdot \ldots \cdot \pi_{k-1}\pi$ is its integration limit.*

The following theorem is a generalization of Theorem 4.3, and its proof is similar to that of the latter theorem.

THEOREM 4.5. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent, let the probabilities $\pi_1, \pi_2, \ldots$ have a common set Z_* of maximum points, and there exists a positive number $q < 1$ such that $\pi_i(z)/\pi_i(z_*) < q$ for all $z \in Z \setminus Z_*$, all $z_* \in Z_*$, and all positive integers i. Then the sequence $(\pi_i)_{i=1}^{\infty}$ integrationally converges and its integration limit π is the uniform probability on Z_*, i.e., $\pi(z) = 1/|Z_*|$ for all $z_* \in Z_*$ and $\pi(z) = 0$ for all $z \in Z \setminus Z_*$.*

The following theorem states that a probability sequence integrationally converges if the probabilities forming it put the elementary events in order in a similar way.

THEOREM 4.6. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent, and there exists a sequence $(z_k)_{k=1}^{N}$ from Z such that $\{z_1, \ldots, z_N\} = Z$ and $\pi_i(z_1) \leqq \cdots \leqq \pi_i(z_N)$ for all positive integers i. Then the following statements are true.*

1. *The sequence $(\pi_i)_{i=1}^{\infty}$ integrationally converges and its integration limit π satisfies the inequalities $\pi(z_1) \leqq \cdots \leqq \pi(z_N)$.*

2. *If there exist a positive number $q < 1$, a positive integer $l \in \{1, \ldots, N-1\}$, and a subsequence $(\pi_{i_m})_{m=1}^{\infty}$ of the sequence $(\pi_i)_{i=1}^{\infty}$ such that $\pi_{i_m}(z_l)/\pi_{i_m}(z_N) < q$ for all positive integers m, then $\pi(z_1) = \cdots = \pi(z_l) = 0$.*

Proof. We prove statement 1. Let $j \in \{1, \ldots, N\}$ be the minimum of all $k \in \{1, \ldots, N\}$ such that $\pi_i(z_k) > 0$ for all positive integers i. It is evident that, for any positive integer $k < j$, we have $(\pi_1 \cdot \ldots \cdot \pi_n)(z_k) = 0$ for all sufficiently large n. If $j = N$, then for all sufficiently large n we obtain $(\pi_1 \cdot \ldots \cdot \pi_n)(z_k) = 1$, which completes the proof. Let $j \leqq N - 1$. For any $k \in \{j, \ldots, N-1\}$ and any positive integer n, we put

$$v_{kn} = \frac{(\pi_1 \cdot \ldots \cdot \pi_n)(z_k)}{(\pi_1 \cdot \ldots \cdot \pi_n)(z_{k+1})} = \frac{\pi_1(z_k) \cdots \pi_n(z_k)}{\pi_1(z_{k+1}) \cdots \pi_n(z_{k+1})}.$$

By assumption, for any $k \in \{j, \ldots, N-1\}$ we have $v_{kn} \leqq 1$, and the sequence $(v_{kn})_{n=1}^{\infty}$ does not increase; hence $v_{kn} \leqq 1$ for all positive integers n. Consequently, $(\pi_1 \cdot \ldots \cdot \pi_n)(z_k) \leqq (\pi_1 \cdot \ldots \cdot \pi_n)(z_{k+1})$ for all $k \in \{j, \ldots, N-1\}$ and all positive integers n. To complete the proof, it is sufficient to show that the sequence $((\pi_1 \cdot \ldots \cdot \pi_n)(z_k))_{n=1}^{\infty}$ converges for any $k \in \{j, \ldots, N-1\}$. We note that, for any $k \in \{j, \ldots, N-1\}$, the not increasing sequence $(v_{kn})_{n=1}^{\infty}$ converges. Let us show that the sequence $((\pi_1 \cdot \ldots \cdot \pi_n)(z_N))_{n=1}^{\infty}$ does not decrease (here we slightly modify the reasoning used in the proof of Lemma 2.1). By definition, we have

$$(\pi_1 \cdot \pi_2)(z_N) = \frac{\pi_1(z_N)\pi_2(z_N)}{\sum_{k=1}^{N} \pi_1(z_k)\pi_2(z_k)} = \frac{\pi_1(z_N)}{\sum_{k=1}^{N-1} \pi_1(z_k)q(z_k) + \pi_1(z_N)},$$

where $q(z_k) = \pi_2(z_k)/\pi_2(z_N)$ for $k \in \{1, \ldots, N-1\}$. Since $q(z_k) \leqq 1$ for $k \in \{1, \ldots, N-1\}$, we have $(\pi_1 \cdot \pi_2)(z_N) \geqq \pi_1(z_N)$. Similarly, by induction, we establish that $(\pi_1 \cdot \ldots \cdot \pi_{n+1})(z_N) \geqq (\pi_1 \cdot \ldots \cdot \pi_n)(z_N)$ for any positive integer n. So, the sequence $((\pi_1 \cdot \ldots \cdot \pi_n)(z_N))_{n=1}^{\infty}$ does not decrease and, consequently, converges. For all positive integers n, we have $(\pi_1 \cdot \ldots \cdot \pi_n)(z_{N-1}) = v_{N-1,n}(\pi_1 \cdot \ldots \cdot \pi_n)(z_N)$. As each of the sequences $(v_{N-1,n})_{n=1}^{\infty}$ and $((\pi_1 \cdot \ldots \cdot \pi_n)(z_N))_{n=1}^{\infty}$ converges, the sequence $((\pi_1 \cdot \ldots \cdot \pi_n)(z_{N-1}))_{n=1}^{\infty}$ also converges. Similarly, by induction, we establish that the sequence $((\pi_1 \cdot \ldots \cdot \pi_n)(z_k))_{n=1}^{\infty}$ converges for any $k \in \{j, \ldots, N-1\}$. Statement 1 is proved.

Let us prove statement 2. Let its assumptions be satisfied. For all sufficiently large positive integers n, we have

$$w_n = \frac{(\pi_1 \cdot \ldots \cdot \pi_n)(z_l)}{(\pi_1 \cdot \ldots \cdot \pi_n)(z_N)} = \frac{\pi_1(z_l) \cdots \pi_n(z_l)}{\pi_1(z_N) \cdots \pi_n(z_N)} \leqq q^{s(n)},$$

where $s(n)$ is the maximum of all i_m ($m \in \{1, 2, \ldots\}$) such that $i_m \leqq n$. Since $s(n) \to \infty$ as $n \to \infty$, we have $w_n \to 0$ as $n \to \infty$. Consequently, $\pi(z_l) = \lim_{n\to\infty}(\pi_1 \cdot \ldots \cdot \pi_n)(z_l) = 0$. Since $\pi(z_1) \leqq \cdots \leqq \pi(z_l)$, we have $\pi(z_1) = \cdots = \pi(z_l) = 0$. Statement 2 is proved.

Taking into account Theorem 4.4, we immediately derive from Theorem 4.6 that, for integration convergence of a probability sequence, it is sufficient that all probabilities forming it, with the possible exception of a finite number, prioritize elementary events in the same order.

Corollary 4.1. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent and there exist a sequence $(z_k)_{k=1}^{N}$ in Z such that $\{z_1, \ldots, z_N\} = Z$ and $\pi_i(z_1) \leqq \cdots \leqq \pi_i(z_N)$ for all sufficiently large positive integers i. Then the sequence $(\pi_i)_{i=1}^{\infty}$ integrationally converges, and statement* 2 *of Theorem* 4.6 *is valid.*

4.2. Permutations in integration of sequences. As noted above (see Corollary 3.1), the result of integration of a finite number of probabilities does not depend on the order in which they are integrated. From here it follows immediately that finite permutations in infinite probability sequences do not change their partial limits. To be precise, the following is true.

Corollary 4.2. *Let a sequence $(\pi_i)_{i=1}^{\infty}$ of probabilities from Π be consistent, k be a positive integer, $(i_1, i_2, \ldots, i_k)$ be a permutation in the collection $(1, 2, \ldots, k)$,* *and*

$$\pi_j^* = \begin{cases} \pi_{i_j} & \text{if } j \in \{1, \ldots, k\}, \\ \pi_j & \text{if } j \in \{k+1, k+2, \ldots\}. \end{cases}$$

Then the sets of all partial integration limits of the sequences $(\pi_i)_{i=1}^{\infty}$ *and* $(\pi_i^*)_{i=1}^{\infty}$ *coincide.*

The following example demonstrates that this statement is not true for infinite permutations in probability sequences.

Example 4.3. Let $Z = \{z_1, z_2\}$ and probabilities $\pi_i^{(1)}, \pi_i^{(2)} \in \Pi$ ($i \in \{1, 2, \ldots\}$) be such that

$$\pi_i^{(1)}(z_1) = 1 - \varepsilon_i, \quad \pi_i^{(1)}(z_2) = \varepsilon_i, \quad \pi_i^{(2)}(z_1) = \varepsilon_i, \quad \pi_i^{(2)}(z_2) = 1 - \varepsilon_i,$$

where $\varepsilon_i \in (0, 1)$ and $\lim_{i\to\infty} \varepsilon_i = 0$. For arbitrary i and $j \geqq i$, we introduce notation

$$\pi_{ij}^{(1)} = \pi_i^{(1)} \cdot \ldots \cdot \pi_j^{(1)}, \quad \pi_{ij}^{(2)} = \pi_i^{(2)} \cdot \ldots \cdot \pi_j^{(2)}.$$

From Theorem 4.3 we have

$$\pi_{ij}^{(1)} \to \pi^{(1)} \quad \text{in } \Pi, \qquad \pi_{ij}^{(2)} \to \pi^{(2)} \quad \text{in } \Pi, \tag{4.7}$$

where $\pi^{(1)}$ and $\pi^{(2)}$ are concentrated at z_1 and z_2, respectively. We take positive integers $k_1, k_2, \ldots$ and put

$$\begin{aligned} \pi_j' &= \pi_j^{(1)}, \quad (j \in \{1, \ldots, k_1\}), \\ \pi_{k_1+1}' &= \pi_1^{(2)}, \\ \pi_j' &= \pi_j^{(1)}, \quad (j \in \{k_1+2, \ldots, k_1+2+k_2\}), \\ \pi_{k_1+k_2+3}' &= \pi_2^{(2)}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \pi_j' &= \pi_j^{(1)}, \quad (j \in \{m_s, \ldots, m_{s+1}\}), \\ \pi_{m_{s+1}+1}' &= \pi_{s+1}^{(2)}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \end{aligned}$$

where $m_s = \sum_{l=1}^{s} k_l + s + 1$. The sequence $(\pi_j')_{j=1}^{\infty}$ is obviously consistent. We take positive numbers $\delta_1, \delta_2, \ldots$ such that $\lim_{s\to\infty} \delta_s = 0$. Taking into account (4.7), the fact that $\pi^{(1)}$ is a zero with respect to integration as an operation of multiplication, and the continuity of integration, we sequentially choose positive integers $k_1, k_2, \ldots$ in such a way that the relations

$$\begin{aligned} (\pi_1' \cdot \ldots \cdot \pi_{m_{s+1}+1}')(z_1) &= (\pi_1' \cdot \ldots \cdot \pi_{m_s+1}' \cdots \pi_{m_s+1\, m_s+1}^{(1)} \cdot \pi_{m_{s+1}+1}^{(2)})(z_1) \\ &= \Big(\pi_{m_s+1\, m_{s+1}}^{(1)} \cdot \big(\pi_1' \cdot \ldots \cdot \pi_{m_s+1}' \cdot \pi_{m_{s+1}+1}^{(2)}\big)\Big)(z_1) > 1 - \delta_s \end{aligned}$$

are true. Then, in view of Theorem 2.1, for all $k \in \{m_{s+1}+2, \ldots, m_{s+3}\} = \{m_{s+2}, \ldots, m_{s+3}\}$, we obtain

$$\begin{aligned} (\pi_1' \cdot \ldots \cdot \pi_k')(z_1) &= ((\pi_1' \cdot \ldots \cdot \pi_{m_{s+1}+1}') \cdot \pi_{m_{s+2}} \cdot \ldots \cdot \pi_k')(z_1) \\ &= ((\pi_1' \cdot \ldots \cdot \pi_{m_{s+1}+1}') \cdot \pi_{m_{s+2}}^{(1)} \cdot \ldots \cdot \pi_k^{(1)})(z_1) \\ &> (\pi_1' \cdot \ldots \cdot \pi_{m_{s+1}+1}')(z_1) > 1 - \delta_s. \end{aligned}$$

So, $\pi^{(1)}$ is the integration limit of the sequence $(\pi'_j)_{j=1}^{\infty}$. Now we define probabilities $\pi''_1, \pi''_2, \ldots \in \Pi$ by interchanging the positions of $\pi_j^{(1)}$ and $\pi_j^{(2)}$ in the definition of $\pi'_1, \pi'_2, \ldots$. Reasoning in a similar way, we establish that $\pi^{(2)}$ is the integration limit of the sequence $(\pi''_j)_{j=1}^{\infty}$. It is clear that the sequence $(\pi''_j)_{j=1}^{\infty}$ is derived from the sequence $(\pi'_j)_{j=1}^{\infty}$ by infinite permutation. So, the infinite permutation of the sequence $(\pi'_j)_{j=1}^{\infty}$ changes its integration limit, though it remains integrationally converging.

Conclusion. This paper is inspired by problems arising in applied research in analysis of inexact data coming from alternative independent sources. A method for integration of data represented in the form of probability distributions is proposed. Initial research of the method in the simplest case of a finite set of elementary events is conducted. Among problems planned for further research, we mention the following: extension of this theory to the cases of infinite probability spaces, in particular, spaces with probabilities given by distribution densities; creation of a "multiplication table" of standard distributions, that is, of the results of their pairwise integration; comparison of informativeness of a priori probabilities and results of their a posteriori integration in terms of different measures of concentration; extension of the proposed method to stochastic processes; and problems of optimal choice of a priori probabilities for a posteriori integration.

Acknowledgments. Numerous discussions of the problem of integration of probability estimates of forest ecosystem primary production that were initiated by Shvidenko served as a direct impetus for this research; the author expresses his sincere gratitude to him. The author is grateful to Rovenskaya, who familiarized him with basic literature on empirical methods of distribution integration that are used in research on climatic changes. The author is grateful to Shiryaev, as well as to participants of the International school of young scientists "Modern problems of mathematics and its applications" (Ekaterinburg, February 2–8, 2014) for discussion of some parts of the work.

REFERENCES

[1] A. N. Shiryaev, *Probability* 1, MCCME, Moscow, 2011 (in Russian).
[2] D. Blackwell, *Comparisons of experiments*, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, CA, 1951, pp. 93–102.
[3] D. Blackwell, *Equivalent comparisons of experiments*, Ann. Math. Statist., 24 (1953), pp. 265–272.
[4] S. Frühwirth-Schnatte, *Finite Mixture and Markov Switching Models*, Springer, New York, 2006.
[5] R. Knutti, R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, *Challenges in combining projections from multiple climate models*, J. Climate, 23 (2010), pp. 2739–2758.
[6] S. Nilsson, A. Shvidenko, M. Jonas, I. McCallum, A. Thompson, and H. Balzter, *Uncertainties of a regional terrestrial biota full carbon account: A systems analysis*, Water Air Soil Pollution Focus, 7 (2007), pp. 425–441.
[7] B. Rajagopalan, U. Lall, and S. E. Zebiak, *Categorical climate forecasts through regularization and optimal combination of multiple GCM ensembles*, Monthly Weather Review, 130 (2002), pp. 1792–1811.
[8] A. W. Robertson, U. Lall, S. E. Zebiak, and L. Goddard, *Improved combination of multiple atmospheric GCM ensembles for seasonal prediction*, Monthly Weather Review, 132 (2004), pp. 2732–2744.
[9] E. Sulganik and I. Zilcha, *The value of information: The case of signal-dependent opportunity sets*, J. Economic. Dynam. Control, 21 (1997), pp. 1615–1625.
[10] C. Tebaldi and R. Knullti, *The use of the multi-model ensemble in probabilistic climate projections*, Philos. Trans. Roy. Soc. Lond. Ser. A, 365 (2007), pp. 2053–2075.
[11] A. Wald, *Statistical decision functions*, Ann. Math. Statist., 20 (1949), pp. 165–205.