What is measure?

The concept of measure does not appear in elementary calculus, but it is a fundamental and important concept. It is not very difficult to understand, since it is important. Besides, the introduction of the Lebesgue measure by Lebesgue is a good example of conceptual analysis, so let us look at its elementary part. A good introductory book for this topic is Kolmogorov and Fomin: *Elements of the Theory of Functions and Functional Analysis* (MartinoFine Books, 2012).[1] Those who wish to study fundamental aspects of statistical mechanics and dynamical systems have proper understanding of the subject.

⟪What is the volume?⟫
For simplicity, let us confine ourselves to 2-space. Thus, the question is: what is the area? Extension to higher dimensions should not be hard. If the shape of a figure is complicated, whether it has an area could be a problem,[2] so let us begin with an apparently trivial case.

"The area of the rectangle $[0, a] \times [0, b]$ is ab."

Is this really so? If so, why is this true? Isn't it strange that we can ask such a question before defining 'area'?

If we wish to be logically conscientious, we must accept the following definition:
Definition. The area of a figure congruent to the rectangle $\langle 0, a\rangle$ (along the x-axis) $\times$ $\langle 0, b\rangle$ (along the y-axis) is *defined* as ab. Here, '$\langle$' implies '[' or '(', '$\rangle$' is ']' or ')', that is, we do not care whether the boundary is included or not.

Notice that the area of an rectangle does not depend on whether its boundary is included or not. This is already incorporated in the definition.

⟪The area of a fundamental set⟫
A figure made as the direct sum (that is, join without overlap except at edges and vertices) of a finite number of rectangles (whose edges are parallel to the coordinate axes and whose boundaries may or may not be included) is called a fundamental set (Fig. 0.0.1). It should be obvious that the join and the intersection (common set) of two fundamental sets are both fundamental sets. The area of a fundamental set is defined as the total sum of the areas of the constituent rectangles.

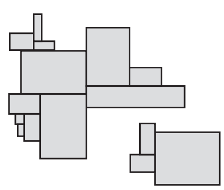

Figure 0.0.1: Fundamental set: it is a figure made of a finite number of rectangles whose edges are parallel to a certain Cartesian coordinate axes and the rectangles do not overlap except at edges and vertices. Its area is the total sum of the areas of the constituent rectangles

⟪How to define the area of more complicated figures; a strategy⟫
For a more complicated figure, a good strategy must be to approximate it by a sequence of fundamental sets allowing increasingly smaller rectangles. Therefore, following Archimedes, we approximate the figure from inside and from outside (that is, the figure is approximated by a sequence of fundamental sets enclosed by the figure and by a sequence of fundamental sets enclosing the figure).

[1] Do not buy the Dover edition of the book.
[2] (Under the usual axioms of mathematics = the ZFC axiomatic system) we encounter figures without areas.

If the areas of the inside and the outside approximate sequences agree in the limit of refinements, it is rational to define the area of the figure by the limit.

Let us start from outside.

⟪Outer measure⟫
Let A be a given bounded set (that is, a set that may be enclosed in a sufficiently large disk). Using a finite number of (or countably many) rectangles P_k $(k = 1, 2, \cdots)$, we cover A, where the boundaries of the rectangles may or may not be included, appropriately. If $P_i \cap P_j = \emptyset$ $(i \neq j)$ and $\cup P_k \supset A$, $P = \{P_k\}$ is called a finite (or countable) cover of A by rectangles (Fig. 0.0.2O). Let the area of the rectangle P_k be $m(P_k)$. We define the *outer measure* $m^*(A)$ of A as

$$m^*(A) \equiv \inf \sum_k m(P_k). \tag{0.0.1}$$

Here, inf[3] is taken over all the possible finite or countable covers by rectangles.

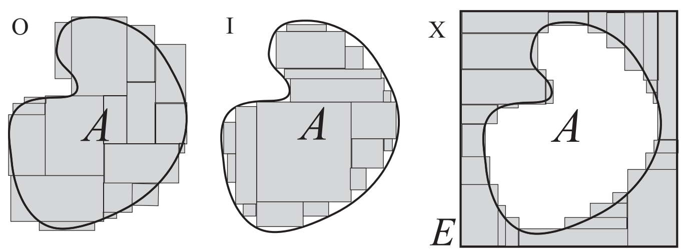

Figure 0.0.2: Let A be the set enclosed by a closed curve. O denotes a finite cover by rectangles. If there is an area of A, it is smaller than the sum of the areas of these rectangles. The outer measure is defined by approximating the area from outside. In contrast, the inner measure is computed by the approximation shown in I by the rectangles included in the figure A. In the text, by using a large rectangle E containing A, $E \setminus A$ is made and its outer measure is computed with the aid of finite covers; the situation is illustrated in X. The relation between I and X is just the relation between negative and positive films. If the approximation O from outside and the approximation I from inside agree in the limit of refinement, we may say that A has an area. In this case, we say A is measurable, and the agreed area is called the area of A

⟪Inner measure⟫
Take a sufficiently large rectangle E that can enclose A. Of course, we know the area of E is $m(E)$. The *inner measure* of A is defined as[4]

$$m_*(A) = m(E) - m^*(E \setminus A). \tag{0.0.2}$$

It is easy to see that this is equivalent to the approximation from inside (Fig. 0.0.2I). Clearly, for any bounded set A $m^*(A) \geq m_*(A)$ holds.

⟪Area of figure, Lebesgue measure⟫
Let A be a bounded set. If $m^*(A) = m_*(A)$, A is said to be a *measurable set* (in the present case, a

[3] The infimum of a set of numbers is the largest number among all the numbers that are not larger than any number in the set. For example, the infimum of positive numbers is 0. As is illustrated by this example, the infimum of a set need not be an element of the set. When the infimum is included in the set, it is called the minimum of the set. The above example tells us that the minimum need not exist for a given set.

[4] $A \setminus B$ in the following formula denotes the set of points in A but not in B, that is, $A \cap B^c$.

set for which its area is definable) and $\mu(A) = m^*(A)$ is called its area (two-dimensional *Lebesgue measure*.

At last the area is defined.

⟪Abstraction possible?⟫

The properties of a fundamental set we have used are the following two:
(i) It is written as a (countable) direct sum of the sets whose areas are defined.
(ii) The family of fundamental sets is closed under $\cap$, $\cup$ and $\setminus$ (we say that the family of the fundamental sets makes a set ring.[5])

An important property of the area is its additivity: If P_i are mutually non-overlapping rectangles, $\mu(\cup P_i) = \sum \mu(P_i)$. Furthermore, the *σ-additivity* for countably many summands also holds.[6]

Notice that such a summary as that the area is a translationally symmetric σ-additive set-theoretical function which is normalized to give unity for a unit square does not work, because this does not tell us on what family of sets this set-theoretical function is defined.[7] The above abstract summary does not state the operational detail about how to measure the areas of various shapes, so no means to judge is explicitly given what figures are measurable. Lebesgue's definition of the area outlined above explicitly designates how to obtain the area of a given figure.

⟪General measure (abstract Lebesgue measure)⟫

The essence of characterization of the area is that there is a family of sets closed under certain 'combination rules' and that there is a σ-additive set-theoretical function on it. Therefore, we start with a σ-additive family $\mathcal{M}$ consisting of subsets of a set X: A family of sets satisfying the following conditions is called a *σ-additive family*:
(s1) $X, \emptyset \in \mathcal{M}$,
(s2) If $A \in \mathcal{M}$, then $X \setminus A \in \mathcal{M}$,
(s3) If $A_n \in \mathcal{M}$ $(n = 1, 2, \cdots)$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{M}$.
$(X, \mathcal{M})$ is called a *measurable space*. A non-negative and σ-additive set-theoretical function m defined on a measurable set that assigns zero to an empty set is called a *measure*, and $(X, \mathcal{M}, m)$ is called a *measure space*. Starting with this measure m, we can define the outer measure on a general set $A \subset X$, mimicking the procedure already discussed above. The inner measure can also be constructed. When these two agree, we can define a set-theoretical function μ as $\mu(A) = m^*(A)$, and we say A is μ-measurable. Thus, we can define μ that corresponds to the Lebesgue measure explained above in the context of the area. μ is called the *Lebesgue extension* of m (this is called an abstract Lebesgue measure, but often this is also called a Lebesgue measure). This construction of μ is called the completion of m. In summary, if $(X, \mathcal{M}, m)$ is a measure space, we define a new family of subsets of X based on $\mathcal{M}$ as

$$\overline{\mathcal{M}} = \{A \subset X : \exists B_1, B_2 \in \mathcal{M} \text{ where } B_1 \subset A \subset B_2, m(B_2 \setminus B_1) = 0\}. \tag{0.0.3}$$

If μ is defined as $\mu(A) \equiv m(B_2)$ for $A \in \overline{\mathcal{M}}$, $(X, \overline{\mathcal{M}}, \mu)$ is a measure space, and is called the com-

[5] More precisely, that a family $\mathcal{S}$ of sets makes a ring implies the following two:
(i) $\mathcal{S}$ includes $\emptyset$,
(ii) if $A, B \in \mathcal{S}$, then both $A \cap B$ and $A \cup B$ are included in $\mathcal{S}$.

[6] Indeed, if $A = \cup_{n=1}^{\infty} A_n$ and A_n are mutually exclusive (i.e., for $n \neq m$ $A_n \cap A_m = \emptyset$), for an arbitrary positive integer N $A \supset \cup_{n=1}^{N} A_n$, so $\mu(A) \geq \sum_{n=1}^{N} \mu(A_n)$. Taking the limit $N \to \infty$, we obtain $\mu(A) \geq \sum_{n=1}^{\infty} \mu(A_n)$. On the other hand, for the external measure $m^*(A) \leq \sum_{n=1}^{\infty} m^*(A_n)$, so $\mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$.

[7] If we assume that every set has an area, under the usual axiomatic system of mathematics, we are in trouble. See Discussion Discussion 2.4A.1.

pletion of $(X, \mathcal{M}, m)$.[8]

⟪What is the volume, after all?⟫
The final answer to the question, "What is the area?" is: the area is the completion of the Borel (1871-1956) measure, where the *Borel measure* is the σ-additive translation-symmetric measure that gives unity for a unit square and is defined on the Borel family of sets which is the smallest σ-additive family of sets including all the rectangles.

Generally speaking, a measure is something like a weighted volume. However, there is no guarantee that every set has a measure. It is instructive that a quite important part of the characterization of a concept is allocated to an 'operationally' explicit description (e.g., how to measure, how to compute). Recall that Riemann's definition of the integral was based on this operational spirit, so it can immediately be used to compute integrals numerically.

⟪Jordan measure⟫
Before Lebesgue, Jordan (1838-1922) defined his measure. His idea was to tessellate small elementary figures (e.g., squares of edge length ϵ) in a given set A as much as possible to estimate the area (from below). This estimate gives $a_\epsilon(A) =$ the upper bound of the number of the elementary figures in A $\times$ the area of the elementary figure. Then, take the limit $\epsilon \to 0$ to define the inner measure $\underline{a}(A)$ of A. The outer measure of A may be analogously defined. Then, if $\overline{a}(A) = \underline{a}(A)$, we say A is (Jordan) measurable, and the agreed value is the area of A. Since Lebesgue's method of covering is more flexible than that by Jordan, we have the following inequalities:

$$\underline{a}(A) \leq \underline{m}(A) \leq \overline{m}(A) \leq \overline{a}(A). \tag{0.0.4}$$

Therefore, Jordan measurability implies Lebesgue measurability. However, the converse is not generally true. Since Jordan does not allow the use of all the sizes at once, we cannot say, for example, that the outer measure of any countable set is zero within Jordan's framework. σ-additivity cannot be asserted, either.

It is very interesting to note that the argument based on a certain unit is, even if the unit size is taken infinitesimal eventually, definitely weaker than the argument that allows the use of all the sizes at once. Since, eventually, we allow indefinitely small units, we might expect that the conclusions must be the same, but it is not. Is there any implication of this observation for the physical world or for physics?

Incidentally, we must clearly pay attention to the fact that humankind recovered the refined mathematical and logical level of Archimedes (287-212 BCE) of more than 2000 years ago only around or slightly before the time of Jordan. We must not forget that culture can easily retrogress (medievalization occurs all too easily) with a sense of impending crisis.

What is probability?

⟪Kolmogorov's definition⟫
Kolmogorov (1903-1987) defined probability as a measure whose total mass is normalized to unity.[9]

[8] The completion is unique. In a complete measure space, if A is measure zero ($\mu(A) = 0$), then its subsets are all measure zero. Generally, a measure with this property is called a complete measure. Completion of $(X, \mathcal{M}, m)$ may be understood as the extension of the definition of measure m on the σ-additive family generated by all the sets in $\mathcal{M}$ + all the measure zero set with respect to m.

[9] A. N. Kolmogorov, *Foundations of the Theory of Probability* (2nd edition) (Chelsea, 1957);

⟨⟨The axioms are 'handling rules of probability⟩⟩

Since a long time ago 'What is probability?' has been a difficult problem.[10] An interpretation of the probability of an event is that it is a measure of our confidence in the occurrence of the event (*subjective probability*). It is something like a weight of the event, if we express the event as a set of elementary events compatible with the event. That is, if an event is interpreted as a set, its probability should be handled just as a measure of the set. Therefore, without going further into the problem of interpretation of probability, to specify only how to handle it clearly is the approach adopted by Kolmogorov. This approach may not squarely answer the question: "What is the probability?". For example, there is no obvious relation to relative frequencies. An important lesson is that we can construct a theory of probability that is sufficiently rich and practical without answering any 'philosophically (apparently) deep' questions.

However, notice that even Kolmogorov's definition is not aloof from the interpretation of probability, although he apparently avoids it. This approach contains the crucial idea that for the subjective probability (the extent of confidence) to be rational it must be interpreted as a sort of volume, or as a measure.

⟨⟨LLN makes probability observable⟩⟩

For the cases of casting dice and tossing coins the numbers we call probability are based on our experience about frequencies and are consistent with the law of large numbers. It is not hard to accept intuitively that the empirical probabilities thus obtained obey the same logic as measures do. Such probabilities are understood as objective (and can be empirically confirmed with the aid of the law of large numbers). There is, however, a deep-rooted opinion that subjective probability is distinct from empirical probability (based on frequency). Such an opinion is a typical example of the humanistic fallacy that our logic and language are unrelated to our empirical world.

⟨⟨You cannot beat Darwinism⟩⟩

Suppose there are two mutually exclusive events 1 and 2 with objective probabilities (relative frequencies) p_1 and p_2 ($> p_1$), respectively. If the subjective probability p' of a gambler for these events becomes $p'_1 > p'_2$, then his gain on the average must be smaller than the gambler with $p = p'$ (i.e., whose subjective assessment is consistent with the objective reality).[11]

Thus, the agreement of subjective probability and empirical probability based on relative frequency is forced upon us (i.e., the subjects who choose), when we are subjected to natural selection. The probability based on relative frequency satisfies measure-theoretical axioms. Therefore, the subjective probability molded by natural selection follows, as long as it is useful for our survival, measure-theoretical axioms. Consequently, the assertion that subjective probability = extent of confidence behaves as volume or weight looks very natural. Or, we should say that our nervous system/emotion has been made to evolve so that this looks natural.[12] The essence of probability is

---

`http://www.mathematik.com/Kolmogorov/index.html`. Recommended. The third edition in Russian contains a reprint of this article with a nice outline of the history of probability theory by A. N. Siryaev.

[10]A summary can be found in D. Gillis, *Philosophical theories of probability* (Routledge, 2000), but the argument given here is not described in this book.

[11]The reader might say what matters is not the subjective probability, but the probability of the person to undertake appropriate behaviors, because actual behavior is important, not the belief. However, it must be disadvantageous that the correspondence between thoughts and actions is not simple.

[12]R. T. Cox, "Probability, frequency and reasonable expectation," Am. J. Phys. **14**, 1 (1946) is a highly interesting paper that deduces additivity of rational expectations axiomatically. N. Chapter, J. B. Tenenbaum and A. Yuille, "Probabilistic models of cognition: Conceptual foundations," Trends Cognitive Sci. **10**, 287 (2006) contains many related topics.

the amount of confidence backed by relative frequency, so even apparently subjective probabilities can be effective in empirical sciences.

⟪You cannot easily relate probability to randomness⟫

There have been numerous attempts to relate probability to randomness. The feeling of randomness comes from the experience when we choose items (events) in which there is no way to reduce damage by any suitable bias. Consequently, if the world is 'uniform,' it is equivalent to equal probabilities for the events. However, it is hard to define randomness precisely, so it is not easy to found probability on randomness.[13]

[13] G. Shafer and V. Vovk, *Probability and Finance: its only a game!* (Wiley-Interscience, 2001) is a notable attempt to develop probability theory not depending on the Kolmogorov axiomatic system. For example, it is possible to prove that if a gambler plays against Nature, and if there is no way to accumulate his wealth unboundedly, then the strategy of Nature (or the output of Nature) must satisfy the strong law of large numbers.