CHAPTER XXIII.

OF APPROXIMATE GENERALIZATIONS, AND PROBABLE EVIDENCE.

Sec. 1. In our inquiries into the nature of the inductive process, we must not confine our notice to such generalizations from experience as profess to be universally true. There is a class of inductive truths avowedly not universal; in which it is not pretended that the predicate is always true of the subject; but the value of which, as generalizations, is nevertheless extremely great. An important portion of the field of inductive knowledge does not consist of universal truths, but of approximations to such truths; and when a conclusion is said to rest on probable evidence, the premises it is drawn from are usually generalizations of this sort.

As every certain inference respecting a particular case, implies that there is ground for a general proposition, of the form, Every A is B; so does every probable inference suppose that there is ground for a proposition of the form, Most A are B: and the degree of probability of the inference in an average case, will depend on the proportion between the number of instances existing in nature which accord with the generalization, and the number of those which conflict with it.

Sec. 2. Propositions in the form, Most A are B, are of a very different degree of importance in science, and in the practice of life. To the scientific inquirer they are valuable chiefly as materials for, and steps towards, universal truths. The discovery of these is the proper end of science: its work is not done if it stops at the proposition that a majority of A are B, without circumscribing that majority by some common character, fitted to distinguish them from the minority. Independently of the inferior precision of such imperfect generalizations, and the inferior assurance with which they can be applied to individual cases, it is plain that, compared with exact generalizations, they are almost useless as means of discovering ulterior truths by way of deduction. We may, it is true, by combining the proposition Most A are B, with an universal proposition, Every B is C, arrive at the conclusion that Most A are C. But when a second proposition of the approximate kind is introduced,--or even when there is but one, if that one be the major premise,--nothing can in general be positively concluded. When the major is Most B are D, then, even if the minor be Every A is B, we cannot infer that most A are D, or with any certainty that even some A are D. Though the majority of the class B have the attribute signified by D, the whole of the sub-class A may belong to the minority.[34]

Though so little use can be made, in science, of approximate generalizations, except as a stage on the road to something better, for practical guidance they are often all we have to rely on. Even when science has really determined the universal laws of any phenomenon, not only are those laws generally too much encumbered with conditions to be adapted for every-day use, but the cases which present themselves in life are too complicated, and our decisions require to be taken too rapidly, to admit of waiting till the existence of a phenomenon can be proved by what have been scientifically ascertained to be universal marks of it. To be indecisive and reluctant to act, because we have not evidence of a perfectly conclusive character to act on, is a defect sometimes incident to scientific minds, but which, wherever it exists, renders them unfit for practical emergencies. If we would succeed in action, we must judge by indications which, though they do not generally mislead us, sometimes do; and must make up, as far as possible, for the incomplete conclusiveness of any one indication, by obtaining others to corroborate it. The principles of induction applicable to approximate generalization are therefore a not less important subject of inquiry, than the rules for the investigation of universal truths; and might reasonably be expected to detain us almost as long, were it not that these principles are mere corollaries from those which have been already treated of.

Sec. 3. There are two sorts of cases in which we are forced to guide ourselves by generalizations of the imperfect form, Most A are B. The first is, when we have no others; when we have not been able to carry our investigation of the laws of the phenomena any farther; as in the following propositions: Most dark-eyed persons have dark hair; Most springs contain mineral substances; Most stratified formations contain fossils. The importance of this class of generalizations is not very great; for, though it frequently happens that we see no reason why that which is true of most individuals of a class is not true of the remainder, nor are able to bring the former under any general description which can distinguish them from the latter, yet if we are willing to be satisfied with propositions of a less degree of generality, and to break down the class A into subclasses, we may generally obtain a collection of propositions exactly true. We do not know why most wood is lighter than water, nor can we point out any general property which discriminates wood that is lighter than water from that which is heavier. But we know exactly what species are the one and what the other. And if we meet with a specimen not conformable to any known species (the only case in which our previous knowledge affords no other guidance than the approximate generalization), we can generally make a specific experiment, which is a surer resource.

It often happens, however, that the proposition, Most A are B, is not the ultimatum of our scientific progress, though the knowledge we possess beyond it cannot conveniently be brought to bear upon the particular instance. In such a case we know well enough what circumstances distinguish the portion of A which has the attribute B from the portion which has it not, but have no means, or have not time, to examine whether those characteristic circumstances exist or not in the individual case. This is the situation we are generally in when the inquiry is of the kind called moral, that is, of the kind which has in view to predict human actions. To enable us to affirm anything universally concerning the actions of classes of human beings, the classification must be grounded on the circumstances of their mental culture and habits, which in an individual case are seldom exactly known; and classes grounded on these distinctions would never precisely accord with those into which mankind are divided for social purposes. All propositions which can be framed respecting the actions of human beings as ordinarily classified, or as classified according to any kind of outward indications, are merely approximate. We can only say, Most persons of a particular age, profession, country, or rank in society, have such and such qualities; or, Most persons when placed in certain circumstances act in such and such a way. Not that we do not often know well enough on what causes the qualities depend, or what sort of persons they are who act in that particular way; but we have seldom the means of knowing whether any individual person has been under the influence of those causes, or is a person of that particular sort. We could replace the approximate generalizations by propositions universally true; but these would hardly ever be capable of being applied to practice. We should be sure of our majors, but we should not be able to get minors to fit: we are forced, therefore, to draw our conclusions from coarser and more fallible indications.

Sec. 4. Proceeding now to consider, what is to be regarded as sufficient evidence of an approximate generalization; we can have no difficulty in at once recognising that when admissible at all, it is admissible only as an empirical law. Propositions of the form, Every A is B, are not necessarily laws of causation, or ultimate uniformities of coexistence; propositions like Most A are B *cannot* be so. Propositions hitherto found true in every observed instance, may yet be no necessary consequence of laws of causation, or of ultimate uniformities, and unless they are so, may, for aught we know, be false beyond the limits of actual observation: still more evidently must this be the case with propositions which are only true in a mere majority of the observed instances.

There is some difference, however, in the degree of certainty of the proposition, Most A are B, according as that approximate generalization composes the whole of our knowledge of the subject, or not. Suppose, first, that the former is the case. We know only that most A are B, not why they are so, nor in what respect those which are, differ from those which are not. How then did we learn that most A are B? Precisely in the manner in which we should have learnt, had such happened to be the fact, that all A are B. We collected a number of instances sufficient to eliminate chance, and having done so, compared the number of instances in the affirmative with the number in the negative. The result, like other unresolved derivative laws, can be relied on solely within the limits not only of place and time, but also of circumstance, under which its truth has been actually observed; for as we are supposed to be ignorant of the causes which make the proposition true, we cannot tell in what manner any new circumstance might perhaps affect it. The proposition, Most judges are inaccessible to bribes, would be found true of Englishmen, Frenchmen, Germans, North Americans, and so forth; but if on this evidence alone we extended the assertion to Orientals, we should step beyond the limits, not only of place but of circumstance, within which the fact had been observed, and should let in possibilities of the absence of the determining causes, or the presence of counteracting ones, which might be fatal to the approximate generalization.

In the case where the approximate proposition is not the ultimatum of our scientific knowledge, but only the most available form of it for practical guidance; where we know, not only that most A have the attribute B, but also the causes of B, or some properties by which the portion of A which has that attribute is distinguished from the portion which has it not; we are rather more favourably situated than in the preceding case. For we have now a double mode of ascertaining whether it be true that most A are B; the direct mode, as before, and an indirect one, that of examining whether the proposition admits of being deduced from the known cause, or from any known criterion, of B. Let the question, for example, be whether most Scotchmen can read? We may not have observed, or received the testimony of others respecting, a sufficient number and variety of Scotchmen to ascertain this fact; but when we consider that the cause of being able to read is the having been taught it, another mode of determining the question presents itself, namely, by inquiring whether most Scotchmen have been sent to schools where reading is effectually taught. Of these two modes, sometimes one and sometimes the other is the more available. In some cases, the frequency of the effect is the more accessible to that extensive and varied observation which is indispensable to the establishment of an empirical law; at other times, the frequency of the causes, or of some collateral indications. It commonly happens that neither is susceptible of so satisfactory an induction as could be desired, and that the grounds on which the conclusion is received are compounded of both. Thus a person may believe that most Scotchmen can read, because, so far as his information extends, most Scotchmen have been sent to school, and most Scotch schools teach reading effectually; and also because most of the Scotchmen whom he has known or heard of, could read; though neither of these two sets of observations may by itself fulfil the necessary conditions of extent and variety.

Although the approximate generalization may in most cases be indispensable for our guidance, even when we know the cause, or some certain mark, of the attribute predicated; it needs hardly be observed that we may always replace the uncertain indication by a certain one, in any case in which we can actually recognise the existence of the cause or mark. For example, an assertion is made by a witness, and the question is, whether to believe it. If we do not look to any of the individual circumstances of the case, we have nothing to direct us but the approximate generalization, that truth is more common than falsehood, or, in other words, that most persons, on most occasions, speak truth. But if we consider in what circumstances the cases where truth is spoken differ from those in which it is not, we find, for instance, the following: the witness's being an honest person or not; his being an accurate observer or not; his having an interest to serve in the matter or not. Now, not only may we be able to obtain other approximate generalizations respecting the degree of frequency of these various possibilities, but we may know which of them is positively realized in the individual case. That the witness has or has not an interest to serve, we perhaps know directly; and the other two points indirectly, by means of marks; as, for example, from his conduct on some former occasion; or from his reputation, which, though a very uncertain mark, affords an approximate generalization (as, for instance, Most persons who are believed to be honest by those with whom they have had frequent dealings, are really so) which approaches nearer to an universal truth than the approximate general proposition with which we set out, viz. Most persons on most occasions speak truth.

As it seems unnecessary to dwell further on the question of the evidence of approximate generalizations, we shall proceed to a not less important topic, that of the cautions to be observed in arguing from these incompletely universal propositions to particular cases.

Sec. 5. So far as regards the direct application of an approximate generalization to an individual instance, this question presents no difficulty. If the proposition, Most A are B, has been established, by a sufficient induction, as an empirical law, we may conclude that any particular A is B with a probability proportioned to the preponderance of the number of affirmative instances over the number of exceptions. If it has been found practicable to attain numerical precision in the data, a corresponding degree of precision may be given to the evaluation of the chances of error in the conclusion. If it can be established as an empirical law that nine out of every ten A are B, there will be one chance in ten of error in assuming that any A, not individually known to us, is a B: but this of course holds only within the limits of time, place, and circumstance, embraced in the observations, and therefore cannot be counted on for any sub-class or variety of A (or for A in any set of external circumstances) which were not included in the average. It must be added, that we can guide ourselves by the proposition, Nine out of every ten A are B, only in cases of which we know nothing except that they fall within the class A. For if we know, of any particular instance *i*, not only that it falls under A, but to what species or variety of A it belongs, we shall generally err in applying to *i* the average struck for the whole genus, from which the average corresponding to that species alone would, in all probability, materially differ. And so if *i*, instead of being a particular sort of instance, is an instance known to be under the influence of a particular set of circumstances. The presumption drawn from the numerical proportions in the whole genus would probably, in such a case, only mislead. A general average should only be applied to cases which are neither known, nor can be presumed, to be other than average cases. Such averages, therefore, are commonly of little use for the practical guidance of any affairs but those which concern large numbers. Tables of the chances of life are useful to insurance offices, but they go a very little way towards informing any one of the chances of his own life, or any other life in which he is interested, since almost every life is either better or worse than the average. Such averages can only be considered as supplying the first term in a series of approximations; the subsequent terms proceeding on an appreciation of the circumstances belonging to the particular case.

Sec. 6. From the application of a single approximate generalization to individual cases, we proceed to the application of two or more of them together to the same case.

When a judgment applied to an individual instance is grounded on two approximate generalizations taken in conjunction, the propositions may co-operate towards the result in two different ways. In the one, each proposition is separately applicable to the case in hand, and our object in combining them is to give to the conclusion in that particular case the double probability arising from the two propositions separately. This may be called joining two probabilities by way of Addition; and the result is a probability greater than either. The other mode is, when only one of the propositions is directly applicable to the case, the second being only applicable to it by virtue of the application of the first. This is joining two probabilities by way of Ratiocination or Deduction; the result of which is a less probability than either. The type of the first argument is, Most A are B; most C are B; this thing is both an A and a C; therefore it is probably a B. The type of the second is, Most A are B; most C are A; this is a C; therefore it is probably an A, therefore it is probably a B. The first is exemplified when we prove a fact by the testimony of two unconnected witnesses; the second, when we adduce only the testimony of one witness that he has heard the thing asserted by another. Or again, in the first mode it may be argued that the accused committed the crime, because he concealed himself, and because his clothes were stained with blood; in the second, that he committed it because he washed or destroyed his clothes, which is supposed to render it probable that they were stained with blood. Instead of only two links, as in these instances, we may suppose chains of any length. A chain of the former kind was termed by Bentham[35] a self-corroborative chain of evidence; the second, a self-infirmative chain.

When approximate generalizations are joined by way of addition, we may deduce from the theory of probabilities laid down in a former chapter, in what manner each of them adds to the probability of a conclusion which has the warrant of them all.

In the early editions of this treatise, the joint probability arising from the sum of two independent probabilities was estimated in the following manner. If, on an average, two of every three As are Bs, and three of every four Cs are Bs, the probability that something which is both an A and a C is a B, will be more than two in three, or than three in four. Of every twelve things which are As, all except four are Bs by the supposition; and if the whole twelve, and consequently those four, have the characters of C likewise, three of these will be Bs on that ground. Therefore, out of twelve which are both As and Cs, eleven are Bs. To state the argument in another way; a thing which is both an A and a C, but which is not a B, is found in only one of three sections of the class A, and in only one of four sections of the class C; but this fourth of C being spread over the whole of A indiscriminately, only one-third part of it (or one-twelfth of the whole number) belongs to the third section of A; therefore a thing which is not a B occurs only once, among twelve things which are both As and Cs. The argument would in the language of the doctrine of chances, be thus expressed: the chance that an A is not a B is 1/3, the chance that a C is not a B is 1/4; hence if the thing be both an A and a C, the chance is 1/3 of 1/4 = 1/12.

It has, however, been pointed out to me by a mathematical friend, that in this statement the evaluation of the chances is erroneous. The correct mode of setting out the possibilities is as follows. If the thing (let us call it T) which is both an A and a C, is a B, something is true which is only true twice in every thrice, and something else which is only true thrice in every four times. The first fact being true eight times in twelve, and the second being true six times in every eight, and consequently six times in those eight; both facts will be true only six times in twelve. On the other hand if T, although it is both an A and a C, is not a B, something is true which is only true once in every thrice, and something else which is only true once in every four times. The former being true four times out of twelve, and the latter once in every four, and therefore once in those four; both are only true in one case out of twelve. So that T is a B six times in twelve, and T is not a B, only once: making the comparative probabilities, not eleven to one, as I had previously made them, but six to one.

It may be asked, what happens in the remaining cases? since in this calculation seven out of twelve cases seem to have exhausted the possibilities. If T is a B in only six cases of every twelve, and a not-B in only one, what is it in the other five? The only supposition remaining for those cases is that it is neither a B nor not a B, which is impossible. But this impossibility merely proves that the state of things supposed in the hypothesis does not exist in those cases. They are cases that do not furnish anything which is both an A and a C.

To make this intelligible, we will substitute for our symbols a concrete case. Let there be two witnesses, M and N, whose probabilities of veracity correspond with the ratios of the preceding example: M speaks truth twice in every thrice, N thrice in every four times. The question is, what is the probability that a statement, in which they both concur, will be true. The cases may be classed as follows. Both the witnesses will speak truly six in every twelve times; both falsely once in twelve times. Therefore, if they both agree in an assertion, it will be true six times, for once that it will be false. What happens in the remaining cases is here evident; there will be five cases in every twelve in which the witnesses will not agree. M will speak truth and N falsehood in two cases of every twelve; N will speak truth and M falsehood in three cases, making in all five. In these cases, however, the witnesses will not agree in their testimony. But disagreement between them is excluded by the supposition. There are, therefore, only seven cases which are within the conditions of the hypothesis; of which seven, veracity exists in six, and falsehood in one. Resuming our former symbols, in five cases out of twelve T is not both an A and a C, but an A only, or a C only. The cases in which it is both are only seven, in six of which it is a B, in one not a B, making the chance six to one, or 6/7 and 1/7 respectively.

In this correct, as in the former incorrect computation, it is of course presupposed that the probabilities arising from A and C are independent of each other. There must not be any such connexion between A and C, that when a thing belongs to the one class it will therefore belong to the other, or even have a greater chance of doing so. Otherwise the not-Bs which are Cs may be, most or even all of them, identical with the not-Bs which are As; in which last case the probability arising from A and C together will be no greater than that arising from A alone.

When approximate generalizations are joined together in the other mode, that of deduction, the degree of probability of the inference, instead of increasing, diminishes at each step. From two such premises as Most A are B, Most B are C, we cannot with certainty conclude that even a single A is C; for the whole of the portion of A which in any way falls under B, may perhaps be comprised in the exceptional part of it. Still, the two propositions in question afford an appreciable probability that any given A is C, provided the average on which the second proposition is grounded, was taken fairly with reference to the first; provided the proposition, Most B are C, was arrived at in a manner leaving no suspicion that the probability arising from it is otherwise than fairly distributed over the section of B which belongs to A. For though the instances which are A *may* be all in the minority, they may, also, be all in the majority; and the one possibility is to be set against the other. On the whole, the probability arising from the two propositions taken together, will be correctly measured by the probability arising from the one, abated in the ratio of that arising from the other. If nine out of ten Swedes have light hair, and eight out of nine inhabitants of Stockholm are Swedes, the probability arising from these two propositions, that any given inhabitant of Stockholm is light-haired, will amount to eight in ten; though it is rigorously possible that the whole Swedish population of Stockholm might belong to that tenth section of the people of Sweden who are an exception to the rest.

If the premises are known to be true not of a bare majority, but of nearly the whole, of their respective subjects, we may go on joining one such proposition to another for several steps, before we reach a conclusion not presumably true even of a majority. The error of the conclusion will amount to the aggregate of the errors of all the premises. Let the proposition, Most A are B, be true of nine in ten; Most B are C, of eight in nine: then not only will one A in ten not be C, because not B, but even of the nine-tenths which are B, only eight-ninths will be C: that is, the cases of A which are C will be only 8/9 of 9/10, or four-fifths. Let us now add Most C are D, and suppose this to be true of seven cases out of eight; the proportion of A which is D will be only 7/8 of 8/9 of 9/10, or 7/10. Thus the probability progressively dwindles. The experience, however, on which our approximate generalizations are grounded, has so rarely been subjected to, or admits of, accurate numerical estimation, that we cannot in general apply any measurement to the diminution of probability which takes place at each illation; but must be content with remembering that it does diminish at every step, and that unless the premises approach very nearly indeed to being universally true, the conclusion after a very few steps is worth nothing. A hearsay of a hearsay, or an argument from presumptive evidence depending not on immediate marks but on marks of marks, is worthless at a very few removes from the first stage.

Sec. 7. There are, however, two cases in which reasonings depending on approximate generalizations may be carried to any length we please with as much assurance, and are as strictly scientific, as if they were composed of universal laws of nature. But these cases are exceptions of the sort which are currently said to prove the rule. The approximate generalizations are as suitable, in the cases in question, for purposes of ratiocination, as if they were complete generalizations, because they are capable of being transformed into complete generalizations exactly equivalent.

First: If the approximate generalization is of the class in which our reason for stopping at the approximation is not the impossibility, but only the inconvenience, of going further; if we are cognizant of the character which distinguishes the cases that accord with the generalization from those which are exceptions to it; we may then substitute for the approximate proposition, an universal proposition with a proviso. The proposition, Most persons who have uncontrolled power employ it ill, is a generalization of this class, and may be transformed into the following:--All persons who have uncontrolled power employ it ill, provided they are not persons of unusual strength of judgment and rectitude of purpose. The proposition, carrying the hypothesis or proviso with it, may then be dealt with no longer as an approximate, but as an universal proposition; and to whatever number of steps the reasoning may reach, the hypothesis, being carried forward to the conclusion, will exactly indicate how far that conclusion is from being applicable universally. If in the course of the argument other approximate generalizations are introduced, each of them being in like manner expressed as an universal proposition with a condition annexed, the sum of all the conditions will appear at the end as the sum of all the errors which affect the conclusion. Thus, to the proposition last cited, let us add the following:--All absolute monarchs have uncontrolled power, unless their position is such that they need the active support of their subjects (as was the case with Queen Elizabeth, Frederick of Prussia, and others). Combining these two propositions, we can deduce from them an universal conclusion, which will be subject to both the hypotheses in the premises; All absolute monarchs employ their power ill, unless their position makes them need the active support of their subjects, or unless they are persons of unusual strength of judgment and rectitude of purpose. It is of no consequence how rapidly the errors in our premises accumulate, if we are able in this manner to record each error, and keep an account of the aggregate as it swells up.

Secondly: there is a case in which approximate propositions, even without our taking note of the conditions under which they are not true of individual cases, are yet, for the purposes of science, universal ones; namely, in the inquiries which relate to the properties not of individuals, but of multitudes. The principal of these is the science of politics, or of human society. This science is principally concerned with the actions not of solitary individuals, but of masses; with the fortunes not of single persons, but of communities.

For the statesman, therefore, it is generally enough to know that *most* persons act or are acted upon in a particular way; since his speculations and his practical arrangements refer almost exclusively to cases in which the whole community, or some large portion of it, is acted upon at once, and in which, therefore, what is done or felt by *most* persons determines the result produced by or upon the body at large. He can get on well enough with approximate generalizations on human nature, since what is true approximately of all individuals is true absolutely of all masses. And even when the operations of individual men have a part to play in his deductions, as when he is reasoning of kings, or other single rulers, still, as he is providing for indefinite duration, involving an indefinite succession of such individuals, he must in general both reason and act as if what is true of most persons were true of all.

The two kinds of considerations above adduced are a sufficient refutation of the popular error, that speculations on society and government, as resting on merely probable evidence, must be inferior in certainty and scientific accuracy to the conclusions of what are called the exact sciences, and less to be relied on in practice. There are reasons enough why the moral sciences must remain inferior to at least the more perfect of the physical: why the laws of their more complicated phenomena cannot be so completely deciphered, nor the phenomena predicted with the same degree of assurance. But though we cannot attain to so many truths, there is no reason that those we can attain should deserve less reliance, or have less of a scientific character. Of this topic, however, I shall treat more systematically in the concluding Book, to which place any further consideration of it must be deferred.