CHAPTER 11
PRINCIPLE OF INDIFFERENCE

11.1. In order to epitomize my reasoning, I specify
- that a cognition is a piece of information belonging to the statute of reference *k*,
- that, once the piece of information derived from a subjective intervention is assumed as a part of a statute, it achieves the full status of a cognition (actually, in the current practice, distinguishing between subjective interventions and objective acquirements is often a difficult task)
- that an assignation is *k*-equiprobable or *k*-uniform (*k*-heteroprobable or *k*-non-uniform) iff it does (not) assign the same *k*-probability (the same *k*-measure) to all the *k*-compatible alternatives
- that, under a *k* entailing a uniform assignation, a further cognition *k'* privileges a hypothesis *h* concerning the possibility space under scrutiny iff $P(h|k\&k') > P(h|k)$.
        Therefore
(11.i)   in absence of privileging cognitions the assignation must be equiprobable
is a formulation of the Principle of Indifference as currently intended.
        In any theory of probability I know, far from being a theorem; such a Principle represents an integrative criterion which ought to overcome the circularity affecting the classical definition of probability. Yet its validity is highly controversial. Indeed there are many contexts where the uniformity is a very unreasonable integrative criterion. A trivial example: under a defective statute telling us only that the heights of a crew vary from 1,55m to 1,95m, are we induced to a uniform assignation (according to which 1,55 and 1.75 are equiprobable heights) or rather to a bell-shaped one, according to which the average height is more probable than the extreme ones?

        11.2. I refuse (11.i). And since
(11.ii)   every assignation is legitimated only by the presence of confirming cognitions
is an obvious rule,
(11.iii)   a heteroprobable assignation is legitimated only by the presence of privileging cognitions
is the unquestionable particularization of (11.ii).
        At first sight refusing (11.i) and accepting (11.iii) seems a pathetic contradiction. In fact, since *heteroprobable* is the opposite of *equiprobable* and *absence* is the opposite of *presence*,
(11.iv)   heteroprobability ⊃ presence of privileging cognitions
seems to imply
(11.v)   absence of privileging cognitions ⊃ equiprobability
by *Modus Tollens*. Yet this schematic deduction is misleading. In fact the apodosis of (11.iv) concerns two oppositions: *presence* vs. *absence*, and *privileging* vs. *non-privileging*. So, since the opposition *equiprobable* vs. *heteroprobable* concerns two assignations, and since both of them must be based upon cognitions, the opposition to consider for the correct application of *Modus Tollens* cannot be *known* vs. *unknown* (cannot be *presence* vs. *absence*), but rather *privileging* vs. *non-privileging*. Therefore
(11.vi)   presence of non-privileging cognitions ⊃ equiprobability
is the right inference by which (11.v) must be replaced.
        Aphoristically: since ignorance is not evidence, an ignorance-supported equiprobability is a logical abuse.

        11.2.1. In other words. Every assignation is grounded on a statute. If it is adequate, the absence of privileging cognitions implies the presence of non-privileging cognitions (otherwise the statute could not be adequate); therefore in condition of adequateness (11.i) is unobjectionable simply because it is equivalent to (11.vi). But of course if the statute is adequate, the Principle of Indifference is superfluous, as no integrative intervention is needed in order to perform the assignation complying with the same statute.

        11.3. This conclusion is immediately evident in ®: while a really well-grounded assignation follows from an adequate statute and is represented by a specific partition of the circle, ignorance is ignorance of the partition, and as such is not represented by a uniform partition but by a defective partition.
        In Chapter 12 this matter will be more carefully approached. Here I examine some (misleading) arguments which might be proposed in order to sustain the Principle of Indifference in its current interpretation.

        11.4. A first argument *pro* Principle of Indifference might run as follows. Usually the text of a problem adduces all of the information necessary to make it a well determinate problem, that is all of the information necessary to solve it. So, when some datum is lacking, we are implicitly induced to think that such a lack can be filled by an intuitive additional cognition clearly suggested by the context. And actually in thousands of ordinary applications such an additional cognition follows from some kind of uniformity. For instance

(11.vii)   Given a circle C and an inscribed equilateral triangle T,
                                      what is the probability p that a point of C belong to T?
is spontaneously interpreted as a well determinate problem whose solution is the ratio between the areas of T and C, that is $p=3\sqrt{3}/4\pi$. But this solution is right only under the presupposition of equiprobability (uniform density) for the points of C, therefore only under a tacit appeal to the Principle of Indifference. If the circle were the target of Robin Hood, the probability would be nearly 1, because all his arrows would run into the immediate nearness of the centre. In other words. Since
- we presuppose that (11.vii) formulates a determinate problem,
- without a density function the problem is indeterminate,
- no density function is indicated in (11.vii),
we are induced to think that the omitted density function is the uniform one. Why the uniform one? Not only because the uniform density function is the simplest one, but also and mainly because it is the only one which can be tacitly understood (while there is only one uniform density function, there are infinite non-uniform density functions, so if we refuse the simplest choice, it becomes impossible to determine the problem, as it is impossible to divine what of the non-uniform density functions must be assumed in order to fill the omission).

Reply. The argument shows only that in some contexts we appeal to the Principle of Indifference. Nevertheless in other contexts we appeal to other criteria (for instance in the example of the crew we could appeal to a standard bell-shaped distribution) or even (where the context does not suggest any reasonable way out) we accept the indeterminateness of the problem, thus renouncing any assignation. In this sense claiming that (11.i) constitutes a general rule is a logical abuse.

11.5.   A second argument *pro* Principle of Indifference might be based on the example proposed in §7.13.2. Bob has just purchased a die about which he possesses no information; nevertheless there is nothing abusive in the fact that he assigns the same probability to the six possible outcomes.

Reply. His equiprobable assignation is legitimate, but the legitimation follows from statistical cognitions supporting it. In fact, almost in their totality, the dice on sale are well balanced. If Bob were to have purchased the die at *www.Falsaria.com*, he would have immediately rejected the equiprobable assignation.

11.6.   A third and more subtle argument *pro* Principle of Indifference runs as follows. Thousands of frequencies relative to honestly performed games of chance prove that their outcomes are equiprobable. Yet (§10.8) it cannot be a compensative equiprobability: in fact, since the parametric differentials are uniform (an honest die is cubic and well balanced, the rotating disk of a honest roulette is partitioned in 37 equal sectors et cetera) the density function too, although unknown, must be uniform.

The concise reply is that a radical discrepancy forbids assimilating the context concerning the slider (§6.11) with the contexts concerning a die or a roulette, since only in the former case a uniformity of the parametric differential and the equiprobability of the upshots imply a uniform density function. The detailed reply below directs us towards a momentous theorization of the matter.

11.6.1.   First of all, let me summarise the general frame of the analysis. A probabilistic problem concerns factual contexts where at least one quantity can assume different values (the parametric range is the domain of a function whose co-domain is the set of possible upshots). Usually the different values of the parametric quantity are infinite, exactly as the final configurations they cause. For instance, the infinite different impulses to the slider and consequently the infinite different positions where it stops; analogously, the infinite different tosses to a die and consequently its infinite dynamic destiny (that is the complex sequence of rototranslations ending in its final position on the green baize). Yet a classification (a partition) is agreed upon whose ground physically different final configurations are made equal (in this sense I speak of conventional upshots). For instance a conventional upshot of the slider is constituted by all the punctual positions belonging to the same segment of the tract; analogously a conventional upshot of the die is constituted by all the outcomes where the die, quite independently on its specific rototranslative process and on its final position, stops with a certain side up. The parametric differential (relative to a certain conventional partition) is constituted by all the values of the parametric quantity determining a punctual final configuration which belongs to the conventional upshot under scrutiny (§10.6.1). But at this point we have to recognize that the same notion of parametric differential must be refined in order to distinguish a theoretically momentous discrepancy between two situations, both of them respecting the proposed condition. I clarify, and for the sake of concision henceforth I will speak of upshots to mean final configurations belonging to a conventional partition.

Let $\mathfrak{S}_1$ and $\mathfrak{S}_2$ be two values leading to a same upshot u_j. I say that $\mathfrak{S}_2 - \mathfrak{S}_1$ is a compact parametric differential iff every $\mathfrak{S}_3$ such that $\mathfrak{S}_1 < \mathfrak{S}_3 < \mathfrak{S}_2$ leads to u_j. But if the number of upshots is less than the number of parametric differentials, this means that there are distinct compact parametric differentials leading to the same upshot. In this sense I will speak of composite parametric differentials to mean the union of all the compact ones leading to the same upshot. That is, symbolically, $\Delta_j\mathfrak{S} = \Sigma_m \delta_{j,m}\mathfrak{S}$, where $\Delta_j\mathfrak{S}$ is the composite differential leading to u_j and $\delta_{j,m}$ is the generic compact differential leading to u_j.

Accordingly, I will speak of macropartitions to mean partitions where all parametric differentials are compact (that is where the number of upshots is also the number of parametric differentials) and of micropartitions in the contrary case. The radical discrepancy mentioned in the concise reply of §6 above is that while a rail partitioned in eight segments determines a macropartition, a rolled die or a roulette determines micropartitions. In order to emphasize the theoretical importance of this discrepancy I insert the slider in a new context.

11.6.2. The possible upshots are always 8, each of them identified by a specific colour (so for instance u_1 is red, u_2 is yellow and so forth until u_8, blue). Yet the tract of the rail is partitioned in eight million equal sections (the parametrical range is uniformly partitioned in eight million compact microdifferentials $\delta\mathcal{G}$) so that the sections 1, 9, 17, ...7999993 are red, the sections 2, 10, 18, ...7999994 are yellow, and so forth until the sections 8, 16, 24, ... 8000000 which are blue. The crucial point is that, under this partition, it is not necessary to suppose a uniform density function in order to get a uniform assignation. In fact, for instance, even a density function like the one represented in Figure 10.5 compensates statistically (I would dare to write "by entropy") the one million occurrences of any colour. Let me insist: the same (and non-uniform) density function that under a macropartition would entail a non-uniform assignation, under a micropartition entails a uniform assignation.

11.6.2.1. Objection. There are peculiar density functions which present different average values for the various upshots. For instance if we miniaturize Figure 10.5 so that its new range is $\delta\mathcal{G} = \Delta\mathcal{G}/10^6$, the density function reproducing one million times that same function assigns to the various upshots the same values of Figure 10.5.

Reply. These *ad hoc* density functions constitute only an infinitesimal part of the possible ones, therefore, again for statistic reasons, the assumption of a common average density is highly justified. Furthermore it is totally implausible to claim that a croupier is able to realize manually one of these ad hoc functions. Both considerations represent a strong cognition supporting the assumption of equiprobability, which therefore is perfectly compatible with a non-uniform density function.

Comment to the reply. Even supposing that the lengths of the eight million sections, far from being equal, are casually chosen, does not influence the probabilistic conclusions, since here too the average length is anyway the same. In order to exhibit a context where a micropartition leads to a non-uniform assignation we must assure a peculiar length to all the microsections of the same color: but this is so anti-entropic a case (obviously the more micro is the micropartition, the more anti-entropic is the anti-entropic case) that our old world continues coherently running under the opposite hypothesis.

11.7. I epitomise the topic through a schematic paradigm enriched by some remarks. Once agreed
- that n is the number of converging compact differentials (for the sake of simplicity I assume that every composite differential results from the same number of compact differentials),
- that $\Delta_j\mathcal{G} = \Sigma_{m(1,n)}\delta_{j,m}\mathcal{G}$
(the composite differential relative to a generic upshot is the sum of the respective compact differentials),
- that $\psi_j = \Sigma_{m(1,n)}\psi_{j,m}/n$
(the density relative to the composite differential is the average of the respective densities relative to the correspondent n compact differentials), the general formula is
(11.viii) $\mu(u_j) = c(\psi_j\Delta_j\mathcal{G})$
(that is: the measure of an upshot is directly proportional to the product of its composite differential and the correspondent average density).

Some particular formulas are obtainable by simplifying (11.viii) in accordance with some particular contexts.

11.7.1. Under a non-uniform macropartition and an uniform density function (that is, concisely, under $n=1$, $\sim(\Delta_j\mathcal{G} =\Delta_l\mathcal{G})$, $\psi_j= \psi_l= \psi$), (11.viii) becomes
(11.ix) $\mu(u_j) = c^*(\Delta_j\mathcal{G})$
where of cours c^* is $c\psi$ and any $\Delta_j\mathcal{G}$ is a compact differential. The upshots are heteroprobable in accordance with $\Delta_j\mathcal{G}$. It is the situation represented in Figure 10.6.

11.7.1.1. Just as (11.ix) is a particular case of (11.viii),
(11.x) $\mu(u_j) = c^{**}$
is the particular case of (11.ix) where, besides the density, the macropartition too is uniform ($\Delta_j\mathcal{G}=\Delta_l\mathcal{G}=\Delta\mathcal{G}$ and $c^{**}=c^*\Delta\mathcal{G}$. The upshots are equiprobable. It is the situation represented in Figure 10.8.

11.7.2. Under a non-uniform macropartition and a non-uniform density function (that is, concisely, under $n=1$, $\sim(\Delta_j\mathcal{G} =\Delta_l\mathcal{G})$, $\sim(\psi_j= \psi_l)$), the only simplification of (11.viii) is that no sum and no average are needed. The probability of the upshots is directly proportional to $\psi_j\Delta_j\mathcal{G}$. Then, in general, the upshots will be heteroprobable (it is the situation represented in 10.5); yet a compensative equiprobability follows from an extremely *ad hoc* and non-uniform density

function counterbalancing the non-uniformity of the parametric differentials, so assuring the invariance of (11.viii). It is the situation represented in Figure 10.9, where all the mixtilinear rectangles have the same area.

11.7.2.1. Under the particular case of an uniform macropartition, (11.viii) becomes
(11.xi) $\mu(u_j) = c^{***}\,\psi_j$
where of course $c^{***} = c\Delta\mathcal{S}$. The upshots are heteroprobable in accordance with the density function. It is the situation represented in Figure 10.7.

11.7.3. Uniform micropartition. The upshots are analytically equiprobable under a uniform ψ and statistically equiprobable under a non-uniform ψ because, statistically, also in the latter case $\psi_j = \psi_l$, that is because. statistically, also in the latter case $\Sigma_{m(1,n)}\psi_{j,m} = \Sigma_{m(1,n)}\psi_{l,m}$.

11.7.4. Contrary to uniform micropartitions, which are always non-privileging, both privileging and non-privileging non-uniform micropartitions do exist. I make my point clearer with reference to our canonical example.

11.7.4.1. A non-uniform and privileging micropartition results from a miniaturization similar to the one proposed in §11.6.2.1, provided that here we are under condition of non-uniformity. Since the ratio δ_j/δ_l is always the same, it reproduces itself in the ratio Δ_j/Δ_l; and since the micropartition entails $\Sigma_{m(1,n)}\psi_{j,m} = \Sigma_{m(1,n)}\psi_{l,m}$, we are in a situation analogous to §11.7.1 (non uniform macropartitions with uniform density). A non-uniform and privileging micropartition could also be easily realized by a deformed die, or by a cyclic context (for instance by a roulette where the rotating disk is fixed, its various spaces are not the same for every number, and the parametric range of the rolled ball is wide).
Only for the sake of completeness I evoke the theoretical possibility of a compensative equiprobability.

11.7.4.2. A non-uniform and non-privileging micropartition is realized where the microdifferentials are different, but casually determined (for instance: the rail is partitioned in eight million microsegments following the chromatic order, the possible lengths are eight, but the eight million assignations of a length to any microsegment are casual). In such a context, though $\sim(\delta_j = \delta_l)$, $\Delta_j = \Delta_l$.

11.8. At this point the solution of the puzzle proposed in §10.9 ought to be clear: owing to the micropartition of impulses transmitted by the croupier to the rotating disk and to the rolling ball, the density function is statistically uniform, then the upshots are equiprobable. Yet I hope that some more detailed reflections will be welcomed, also because the same notion of probability, historically, was born by questions about games of chance.
First of all I focus on the notions of compact and composite parametric differentials making reference to the simplest game of chance, that is to a coin rolled by a mechanized apparatus (the extrapolation to others games of chance is immediate). The range of the possible impulses transmitted to the coin is wide, but it can be partitioned in many compact microdifferentials, each of them delimitated (in accordance with the definition) by those impulsive microvariations which should determine the opposite upshot (that is, roughly, those impulsive microvariations which would modify the number of rototranslations of the coin) and the union of all the microdifferentials converging on a single upshot is its composite impulsive differential.
A serious game of chance must assure a basic equiprobability. Yet a compensative equiprobability is practically unrealizable (for instance, in order to realize it we should compensate a deceptive coin through an impulsive program privileging at the same ratio the unfavoured side). Therefore the equiprobability is uniform. The paradigm shows that a uniform equiprobability can be realized
a) either by the context of §11.7.1.1 (uniform macropartition and uniform density).
b) or by the context of §11.7.3, that is
b1) uniform micropartition and uniform density
b2) uniform micropartition and non-uniform density.
Then, since no usual game of chance can assure a uniform density, the equiprobability of the upshots follows from b2). Why? Because b2) is the simplest and best context to avoid the previous programmability of certain upshots (trivially: to avoid tricks). In fact the various trials are not realized through mechanized apparatus, but through manual interventions, and within a macropartition a short training would be sufficient to instruct the croupier in obtaining certain upshots (for instance, to instruct him in opportunely impulsing the slider on the basis of the actual stakes). In other words. The uniformity of the parametric differentials is usually assured by the fact that a normal coin is well balanced, or that a normal die is a cube whose geometrical centre is also its barycentre, or that the various numbers in the disk of a roulette have the same room, or that the sections of our rail have the same length and resistance, et cetera). But under a macropartition of the parametric range, the equiprobability can be assured only by a uniform density, and to alter cunningly this condition is a rather easy task; therefore a) is rejected.
As for b), so to say, the more micro the partition, the more difficult it is to alter cunningly the density function: in fact
- the microdifferentials render more difficult to perform a pre-programmed value belonging to one of them,

- the uniformity of the density function is not a condition for the equiprobability of the upshots.

In other words; the same empirical structure of the games of chance has been specifically studied in order to render practically impossible any trick (there is no human croupier able to transmit to the rotating disk of a well balanced roulette and to the rolled ball two impulses leading to a specific upshot). Of course there are other possibilities to cheat (for instance a hidden magnet et cetera), but while these eventualities concern a documentable modification of the context, a diabolically able croupier would not modify it. In this sense I say that in a micropartition the upshots are not pre-programmable.

Indeed the distinction *micro* vs. *macro* is not focused on the pre-programmability of the parametric values, but on the number of distinct compact differentials converging on the same upshot (that is: on the ratio $1/n$). The criterion of the pre-programmability is important with regards to the possibility to cheat, but nothing more.

11.8.1. Of course the mentioned statistical reasons weaken with the decreasing of n; for instance if the segments of the tract were only 16 ($n=2$) the uniformity of the average density values would be a quite debatable conclusion. Yet this is only a theoretical specification; in the actual game of chances n is extremely high and consequently the statistical reasons do not give any pretext to dissent.

11.9. The best evidence showing that the Principle of Indifference in its canonical formulation is untenable consists in the lot of paradoxes following from its acritical application.

11.9.1. A first family of paradoxes involving the Principle of Indifference concerns polyadic partitions (that is: the possibility spaces whose alternatives are more than two), and can be exemplified as follows.

A beautiful blonde enters into the railway compartment where Tom is growing weary. Ten minutes pass and he already knows she is a thirty-five-years-old and unmarried Norwegian pediatrist named "Greta". -Is Greta single?- Tom asks himself, and for want of privileging cognitions, the Principle of Indifference teaches him that the respective probability is 50%. Analogously, if his question were -Is Greta a divorcee?-, for want of privileging cognitions the Principle would teach him that the respective probability is 50%. Exactly as if his question were -Is Greta a widow?- Then, absurdly, the probability that she is a single or a divorcee or a widow ought to be 150%.

The solution is trivial. Everyone has the right to reason on the hypotheses he prefers, but if at least one of the two following conditions
- the coherence of his hypotheses
- the coherence of his reasoning
is not respected, the conclusions he draws are manifestly and completely valueless. In the case to assign $P=50\%$ to one of the three alternatives is to assign $P=50\%$ to the sum of the remaining two.

Since whatever assignation is based on an (objective or subjective) partition of the possibility space, and since all the theories of probability converge on the same theorems, an anti-theoremic partition is incoherent. Then, assuming that Tom has no information about the percentages of singles, divorcees and widows relative to the thirty-five-years-old and unmarried Norwegian pediatrists, either he recognizes the impossibility of an assignation, or he must appeal to some informational integration (either ruled by the Principle of Indifference or by some subjective opinion). Then, if Tom decides to follow the Principle, since the possibility space relative to *unmarried* is partitioned in *single*, *divorcee* and *widow*, he must assign the same probability (1/3) to the three alternatives. And if (much more plausibly) Tom refuses the Principle and accepts the subjective integration suggested by his common sense (according to which singles are much more numerous than widows among the thirty-five-years-old and unmarried Norwegian pediatrists) he will assign to the first hypothesis a greater probability et cetera, respecting anyhow the theoremic condition $P_1 + P_2 + P_3 = 1$).

In other words. While
$$P(h|k) + P(\sim h|k) > 1$$
is absurd (antitheoremic)
$$P(h|k_1) + P(\sim h|k_2) > 1$$
is perfectly possible, provided that k_1 and k_2 are distinct statutes. And the three previous assignations can only be justified by a reference to three incompatible statutes.

11.9.2. A second family of paradoxes involving the Principle of Indifference concerns the inverse quantities and can be exemplified as follows.

In this store there are one hundred well piled heaps of different timbers: we know only that each of them weighs exactly 2 tons and that the specific weight sw of the hundred timbers varies from 0.50 to 1,00. The Principle of Indifference (in its naive formulation) assigns the same probability to any given sw of the range, therefore, taken a generic heap, its probability of having a $sw>0.75$ is equal to its probability of having a $sw<0.75$. Yet the Principle can also be applied to the volumes v, which obviously vary from $4,00m^3$ to $2,00m^3$, therefore if we choose to reason on volumes, since the same probability is assigned to each of them, the probability for a generic heap of having a $v>3,00m^2$ is equal to the probability of having a $v<3,00m^2$. The contradiction is that two inferences applied to the same situation lead to incompatible conclusions, as the specific weight of a 2 ton heap whose volume is $3,00m^3$ is not 0.75. but 0.66.

The solution is easy. Since *sw* and *v* are inverse quantities, their link is not linear, but hyperbolic, so that as soon as we assume a uniform distribution of the specific weights we are assuming a non-uniform distribution of the volumes (and vice versa).

Let me insist through a concise appeal to analytic geometry:

$$v=2/sw$$

is the equation of a hyperbole, that is of a curve whose average abscissa, owing to strictly mathematical reasons, cannot correspond to its average ordinate.

Here too, then, the correct procedure is sufficient in order to avoid the paradox. Since every assignation must be based on (objective or subjective) cognitions supporting it, an equiprobable assignation relative to *sw* presupposes cognitions supporting it; and these same cognitions, necessarily, sustain a non-uniform distribution of the respective volumes (or vice versa) since specific weights and volumes are non-linearly related parameters.

11.9.2.1. Indeed the hypothesis that the distribution of the specific weights is uniform between 0.50 and 1.00 has been dictated by the opportunity to deal with the paradox in its usual formulation, nevertheless it is rather implausible; here too (as in the example of the crew) the most plausible hypothesis suggests a normal distribution, that is a bell-shaped curve). Anyhow the solution holds even in this case, since as soon as we suppose the symmetry of the *sw*-curve with respect to the value $0.75t/m^3$ (so legitimating the conclusion that the probabilities of being less and of being more are equal), we are supposing an asymmetry (a deformation) of the *v*-curve with respect to the value $3.00m^3$ (so entailing the conclusion that the probabilities of being $<3.00m^3$ and of being $>3.00m^3$ are different) et cetera.

11.10. A deeper analysis is needed to evaluate satisfactorily a family of geometrical paradoxes involving the Principle of Indifference. I summarize the most renowned one, that is Bertrand's paradox, as follows.

Given a circle of radius R, in condition of parametric uniformity the problem

(11.xii)   what is the probability *P* that a randomly drawn chord will have a length more than R√3?

admits three incompatible solutions. If we reason on a sheaf of parallel chords (parallel approach S_1) the answer is $P_1=1/2$. If we reason on the sheaf of chords rotating on a point of the circumference (polar approach S_2) the answer is $P_2=1/3$. If we reason on the mid-points of the chords (areal approach S_3) the answer is $P_3=1/4$.

Its current solution (Van Fraassen, Hajek) claims that the paradox is born by the application of the Principle of Indifference to non-linearly related parameters. In my opinion such a claim is not a solution because it does not realize
I) that S_1 and S_2 are only two different instances of the same application
II) that S_3 is vitiated by a geometrical mistake.

11.10.1. As for I) I start from a more punctual formulation of the probabilistic problem. For the sake of simplicity the discourse is limited to a semicircle. So, with reference to Figure 11.1,

Figure 11.1
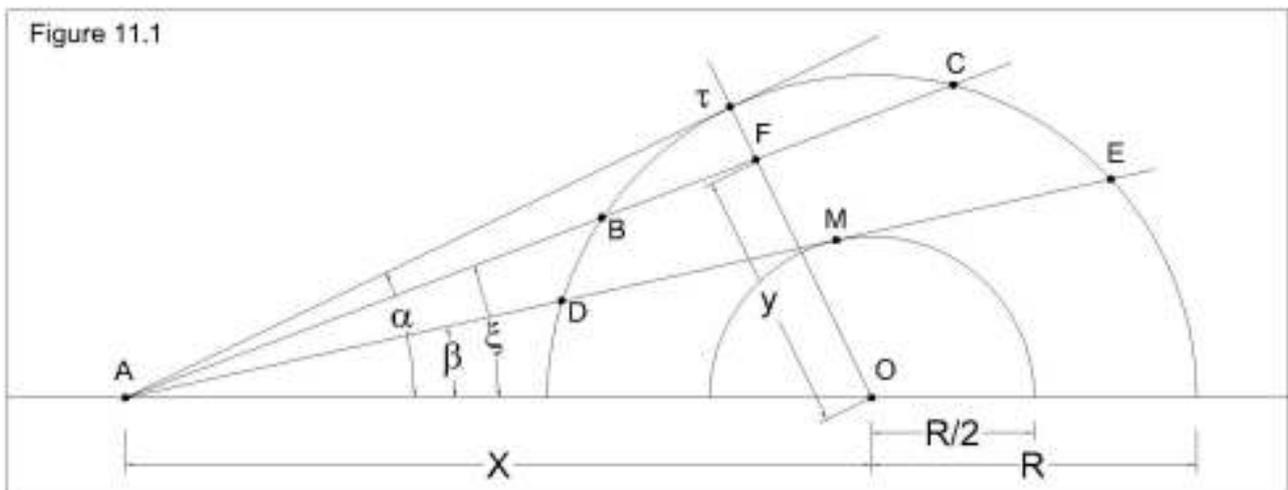
where
- O is the centre of a semicircle of radius R
- A is an external and generic point on the diametric straight of basic reference
- X is the distance between O and A (obviously X varies from R to ∞)
- AT is the tangent,
- ABC is a generic straight intersecting the semi-circumference in B and C
- ADE is the straight intersecting the chord of critical length (R√3)
- M is the mid-point of DE, therefore AM is the tangent to the concentric circle with radius R/2
- F is the intersection OT-BC

- Y is the distance OF
- α is the angle OAT
- ξ is the generic angle OAB
- β is the angle OAD

the probabilistic problem to solve finds in
(11.xiii)   what is the probability P that BC is longer than R√3?
its new formulation. Thus we have specified (11.xii) by considering the sheaf of chords passing for a generic point A of the plane. In such a context, ξ is the parameter to which the Principle of Indifference must be applied, so establishing a polar uniformity (equiprobability) for the direction of the various chords. Therefore
(11.xiv)   β/ α
is the probability that BC is longer than DE. Since elementary trigonometry teaches us that
          α = arc sin(R/X)
          β = arc sin(R/2X)
and since the lengths of OT and OM are respectively R and R/2,
          α=arc sin(2sinβ)
(AO sin α = R = 2AO sin β); the conclusion is that, on the ground of (11.xiv),
(11.xv)   β/arc sin(2sinβ)
or indifferently
(11.xvi)   arc sin(R/2X) / arc sin(R/X)
is the ratio expressing P. I wrote "indifferently" to mean that (11.xv) and (11.xvi) can be immediately and reciprocally transformed, for their discrepancy depends only on the variable we choose to express the generic position of A: exactly as we get (11.xv) if we choose "β", we get (11.xvi) if we choose "X".
          Anyhow the crucial conclusion is that P depends on the position of A. For instance
(11.xvii)   when X=R, α=90°, sin α=1, sin β=1/2, β=30°, β/ α=1/3 (that is P_2)
          when X=2R/√3, α=60°, sin α=√3/2, sin β =√3/4, β≈25°40', β/ α≈0,428
          when X=R√2 α=45°, sin α=√2/2, sin β =√2/4, β≈21°, β/ α≈0,466
          when X=2R, α=30°, sin α=1/2, sin β =1/4, β≈14°30', β/ α≈0,483
- when X→∞, α→0, β→0, sin2β→2sin β, β/ α→1/2 (that is P_1).
          All these instances are deduced by the same polar application of the Principle of Indifference (δξ constant). Thus it is mathematically proved that S_1 and S_2 are nothing but two particular cases of the same approach: while S_2 is nothing but the application of a ξ-uniformity when A is a point of the circumference, S_1 is nothing but the application of the same ξ-uniformity when A is a point at ∞, for in this case the ξ-uniformity implies the linear Y-uniformity). Therefore
- the difference between P_1 and P_2 (contrary to Van Fraassen's and Hajek's claim) cannot depend on the application of the Principle of Indifference to non-linearly related parameters, because both of them are simply two particular cases of the situation generically illustrated in Figure 11.1;
- the difference between P_1 and P_2 depends on the fact that while P_1 is the value of the function P_X when X=∞, P_2 is the value of the same function when X=R;
- there is nothing paradoxical in ascertaining that different values of a function correspond to different values of its argument.

          11.10.1.1. Incidentally. The crucial conclusion that P depends on a free variable is also valid if we renounce the assumption of ξ-uniformity for an assumption of ξ-non-uniformity (that is if we privilege some directions). Such a new assumption entails simply the substitution of (11.xv) or (11.xvi) by more complicated trigonometric expressions where anyhow the free variable continues occurring because P continues depending on the ratio β/α, and this ratio continues depending on the position of A.

          11.10.2. Coming back to the assumption of ξ-uniformity I prove that S_3 is vitiated by a geometric mistake (claim II) of §9).
          First of all I reason in compliance with S_1 (X=∞). In this case OT is perpendicular to OA, the points M, F, T are aligned and M is also the mid-point of OT. If F belongs to OM, the length of the respective chord is >R√3, and if F belongs to MT, the length of the respective chord is <R√3. Therefore, since OM=MT and since the geometric situation holds for whatever basic direction (it does not depend on the specific direction OA), P_1=1/2.
          Now I reason in compliance with S_3. Since each chord whose length is >R√3 (<R√3) has a mid-point internal (external) to the concentric circle with radius R/2, the probability that a chord is longer than R√3 is the probability of its mid-point being in the internal circle. And since the ratio between the areas of two circles whose radii are respectively R/2 and R is 1/4, we must conclude that the respective probability P_3 is not 1/2, but 1/4.
          Here too Van Fraassen's and Hajek's claim is not a solution because here too both approaches follow from the same assumption of ξ-uniformity (the same sheaf leading to P_1 if we reason on intersections with OT leads to P_3 if we

reason on mid-points). The right solution follows immediately as soon as we realize the logical error affecting S_3. In fact S_3 grounds on an implicit postulate, that is a one-to-one correspondence between chords and mid-points: but such a postulate is wrong because actually every chord has one and only one mid-point, and every point of the circle (except O) is the mid-point of one and only one chord, but O is the mid-point of infinite chords (diameters), therefore P_3=1/4 is an erroneous and underestimated value (which from now on will be neglected).

11.11. All these considerations, strictly, could lead to the conclusion that the problem proposed by (11.xii) is indeterminate for lack of a datum. In fact

(11.xviii)   Let O be the centre and R the radius of a circle; in conditions of parametric uniformity, what is
               the probability P that a chord passing for a point A at a distance X from O be longer than R√3?

is a surely determinate problem and

(11.xix) P = arc sin(R/2X) / arc sin(R/X)

is its unobjectionable solution (since the parameter to which the Principle of Indifference can be applied is only one, that is ξ, no ambiguity affects the application of the same Principle). On the grounds of (11.xix) we can compute whatever specific value of P (for instance the values listed in (11.xvii)); so in particular we legitimate the perfect compatibility of S_1 and S_2, in spite of $P_1 \neq P_2$. (the indeterminateness of S_1 and S_2 is only apparent, since both of them, although tacitly, fix a distance, that is ∞ and R, so making computable the respective probabilistic values). Nevertheless, the following puzzle is pending.

11.11.1. Let us draw the most casual and rich interlacement of straight lines over a circle, and let us compute the per cent of chords whose length is >R√3. The author who performed systematically this experiment [Jaines 1973] reports he achieved results with an *embarassingly low* value of χ^2, that is he achieved highly converging arithmetical values of such a per cent, that is 50% (and, modestly, I can confirm both the convergence and the value). The puzzle is manifest: we empirically give a univocal and sound answer to an indeterminate problem.

In order to overcome such a puzzle, firstly we have to realize that under a more liberal interpretation, (11.xii) can be read as a shortened formulation appealing (implicitly, of course) to some average value of P. Under this more liberal interpretation, the problem proposed by (11.xii) is no longer indeterminate, since an average of various values is a single value. Yet in order to compute the average we need a density function for the various specific positions of A, that is for the various specific values resulting from the application of (11.xix). Here too the Principle of Indifference suggests us to adopt a criterion of uniformity; yet here the suggestion is insufficient. In §10.1, once fixed a certain position of A, we applied the Principle to the various possible directions of the chords passing through A, therefore to the parameter ξ. But any given position of A corresponds to a precise value both of the linear parameter X and of the polar parameter β (or indifferently α). As such we must decide if the parametric uniformity suggested by the Principle of Indifference is a linear uniformity (the various distances X are equiprobable) or a polar uniformity (the various angles β are equiprobable). The decision is momentous because, generally speaking, a linear uniformity does not imply a polar uniformity, and in this case it implies a polar non-uniformity. I make my point clearer through a geometrically similar example.

Let ABC be a right-angled triangle having the cathetus AB of unitary length and the angle γ in C of 60° (therefore the hypotenuse BC is 2/√3 and AC is 1/√3); let D be the point where the bisector of γ intersects AB. The probability that a generic point Y of AB belongs to AD is by hypothesis 1/2 if we reason on angles, that is if we apply the parametric uniformity to the generic angle θ between CA and CY, but it is 1/3 if we reason on segments, that is if we apply the parametric uniformity to the distance AY (AD=1/3, BD=2/3). No paradox, of course. Postulating the θ-equiprobability (*i.e.*: δθ=constant) means postulating the equiprobability of each non-constant linear interval tg(θ+δθ)-tgθ on AB. Thus an example of discrepancy between polar and linear uniformity is exhibited

Of course there are geometrical contexts where linear and polar uniformities correspond (an elementary example is a point running on a circumference and the direction of the respective radius). Yet, manifestly, Figure 11.1 shows that in the situation under scrutiny tg(β+δβ)-tgβ is a non-constant linear interval. Therefore the average we obtain by applying the linear uniformity must differ from the average we obtain by applying the polar uniformity..

Concisely: wherever a linear equiprobability implies a polar heteroprobability (and vice versa, obviously), a very paradox would arise only if those two different criteria should lead us to the same probabilistic assignation.

The three following arguments tell us that in this context the linear uniformity is the right choice.

11.11.1.1. A polar uniformity entails a sort of spacial rarefaction, so to say, because as X increases a progressively longer δX corresponds to the same δβ (until an infinite δX where β+δβ=0°) But when we draw the mentioned interlacement of straight lines, quite independently of its position on the plane every point has the same chance of becoming a point on a randomly drawn straight line (the space we move into does not rarefy).

11.11.1.2. While the empirical value is 50%, the mathematical average corresponding to a polar uniformity is nearly 46,6% (very roughly , in order to avoid any appeal to integral calculi, we can simply remark that the middle β is 15°). The discrepancy between 50% and 46,6% is just a consequence of the above denounced linear rarefaction entailed by the polar uniformity.

11.11.1.3. The mathematical average corresponding to a linear uniformity is just 50% since it identifies with the value for X→∞ (very roughly: the mid-point of a semi-straight is at ∞).

11.11.2. A subtle objection runs as follows. Jaines (and Gandolfi) were mistaken, as the same (11.xvii) shows that in an actually random set of chords $P=50\%$ is an unattainable average. In fact, once the interlacement of straight lines is drawn, let us classify the set of chords by couples of non-parallel straight lines; then if A_j is the intersection of the generic j-couple and X_j is its (finite) distance from O, a look at (11.xvii) shows that anyhow $P_j<50\%$; therefore it is mathematically impossible to obtain 50% as an average of values <50%.
Reply. The objection does not hold because
- either we do not account for the couples whose intersections are inside the circumference (x<R), therefore actually $P<50\%$, but it concerns only a subset of the chords we have drawn,
- or we do account for all the couples, and then we must recognize that the couples of chords whose A_j are inside the circumference have a $P_j>50\%$ ($P=66\%$ where $R>A_jO>R/2$, and $P=90\%$ where $R/2\geq A_jO$); therefore an average $P=50\%$ is quite possible.

11.11.3. In conclusion. Since every chord has a direction, since no chord has more than one direction, since these obvious evidences do not depend on some peculiar direction, to classify the set of chords by parallels means to exhaust it without omissions and duplications. Therefore S_1 is the unexceptionable approach and P_1 is the unexceptionable probabilistic value.
A last objection. If also from the experimental viewpoint P_1 is the unexceptionable probabilistic value, then P_2 is experimentally unjustifiable.
Reply. The geometric context leading to P_1 (§11.1: the most casual and rich interlacement of straight lines) is evidently unfit for S_2, where all the straight lines ought to intersect at a single point (belonging to the circumference, furthermore). And as soon as we draw a sheaf complying with S_2 we get the empiric confirmation of P_2. In other words: P_1 and P_2 are connected to two different 'metrics' (that is two incompatible geometric contexts), and it is not at all puzzling that two structurally different trials lead to different results.