Philosophy of Science

March, 1994

A RESOLUTION OF BERTRAND'S PARADOX*

LOUIS MARINOFF†‡

Department of Philosophy
University of British Columbia

Bertrand's random-chord paradox purports to illustrate the inconsistency of the principle of indifference when applied to problems in which the number of possible cases is infinite. This paper shows that Bertrand's original problem is vaguely posed, but demonstrates that clearly stated variations lead to different, but theoretically and empirically self-consistent solutions. The resolution of the paradox lies in appreciating how different geometric entities, represented by uniformly distributed random variables, give rise to respectively different nonuniform distributions of random chords, and hence to different probabilities. The principle of indifference appears consistently applicable to infinite sets provided that problems can be formulated unambiguously.

1. Introduction.

> In the theory of geometrical probabilities the random elements are not quantities but geometrical objects such as points, lines and rotations. Since the ascription of a measure to such elements is not quite an obvious procedure, a number of "paradoxes" can be produced by failure to distinguish the reference set. (Kendall and Moran 1963, 9)

In 1889, J. Bertrand published a collection of paradoxes, the most celebrated of which is his "random-chord" problem. This particular problem, often labeled simply "Bertrand's Paradox" (e.g., Uspensky 1937,

*Received June 1992; revised November 1992.

†I would like to express my gratitude to Bill Boos, Roy Douglas, Dorothy Edgington, J. Howard Sobel, and the referee for *Philosophy of Science* for their insightful commentaries. This paper was originally read in March 1992, at a Colloquium in the Department of Philosophy, University of British Columbia.

‡Send reprint requests to the author, Department of Philosophy, University of British Columbia, 1866 Main Mall E-370, Vancouver, B.C. V6T 1Z1, Canada.

Philosophy of Science, 61 (1994) pp. 1–24

251; Lucas 1970, 117; Weatherford 1982, 56) has beguiled mathematicians and philosophers since its inception. It has spawned a literature that spans a range of partly articulated insights, and fully articulated fallacies. This paper shows that the putative paradox can be unequivocally resolved.

The problem arises from Bertrand's (1889) three answers to the following question, which we label Q, "A chord is drawn *randomly* in a circle. What is the probability that it is shorter than the side of the inscribed equilateral triangle?" (pp. 4–5; emphasis in original; translation mine). Bertrand formulated his solutions in terms of the probability that the random chord is longer—as opposed to shorter—than the side of the inscribed equilateral triangle.

Let us call Bertrand's first solution B_1. Consider all chords generated from a vertex of the inscribed equilateral triangle. Those chords lying within the arc subtended by the angle at that vertex satisfy Bertrand's condition. That angle is $\pi/3$ radians. Since chords from the vertex range through an angular interval of π radians, the probability in question is $1/3$.

Let us call Bertrand's second solution B_2. Consider a diameter that is the right bisector of one side of the triangle. The intersection of the chords is a point of quadrisection of the diameter. Now consider all chords perpendicular to that diameter. Those lying within half a radius of either side of the center satisfy Bertrand's condition. Since this linear interval is just one-half the diameter, the probability in question is $1/2$.

Let us call Bertrand's third solution B_3. Consider a circle inscribed in the equilateral triangle. Any chord (except a diameter) is uniquely identified by its midpoint. Any chord whose midpoint falls either on or within the inscribed circle satisfies Bertrand's condition. The radius of the inscribed circle is half that of the large circle. The probability in question is just the ratio of their areas, which is $1/4$.

Bertrand reasoned, "Among these three answers, which one is proper? None of the three is incorrect, none is correct, the question is ill-posed" (ibid., 5; translation mine).

According to Jaynes (1973), this paradox "has been cited to generations of students to demonstrate that Laplace's 'principle of indifference' contains logical inconsistencies" (p. 478). Jaynes's claim begs two preliminary clarifications. The first pertains to the origin and appellation of said principle; the second, to the original intent and relevance of Bertrand's paradoxes themselves.

Jeffreys ([1948] 1961, 33–34), among others, refers to the Laplacean "Principle of Insufficient Reason" (equivalently, "the equal distribution of ignorance"), with regard to the following proposition, set forth by Laplace both in the introduction and at the beginning of Book Two of his *Théorie*

Analytique des Probabilités ([1820] 1886), "We have seen in the Introduction that the probability of an event is the ratio of the number of favourable cases to the number of all possible cases, when nothing engenders a belief that any one of these cases should occur rather than any other, which renders them, for us, equally possible" (p. 181; translation mine), to which Laplace appends, "The accurate assessment of these various cases is one of the most delicate points in the analysis of chance" (ibid.; translation mine).

Hacking (1975, 122–132) relates that both Laplace and Jacques Bernoulli owe a debt to Leibniz, who advanced a similar definition of probability in 1678. Laplace, in the above-mentioned introduction (itself reprinted and celebrated as a classic essáy), refers to Leibniz and his well-known principle of sufficient reason at the very outset. Hacking further points out that while von Kries aptly introduced the term "Principle of Insufficient Reason" in 1871, Keynes (1921, 42) coined the phrase "Principle of Indifference".

Both principles are frequently but inexactly attributed to Laplace (e.g., Jeffreys [1948] 1961, 34; Carnap [1950] 1962, 341; and Jaynes 1973, 478, respectively). For our present purposes, these principles are synonymous. Although their conceptual origin predates Laplace and their appellations postdate him, their meaning bears directly upon him, for he is rightly regarded as the outstanding formulator of and contributor to the classical theory of probability.

This leads to the second preliminary issue. Bertrand belonged to the informal school of French finitism (whose later exponents included Borel and Poincaré). Bertrand (1889, 2) rehearsed Laplace's definition of probability verbatim, but sought to restrict its application to problems whose number of possible cases is finite, "Another remark is necessary: infinity is not a number; one should not, without explanation, introduce it into arguments. The illusory precision of words can give rise to contradictions. To choose *randomly*, among an infinite number of possible cases, is not a sufficient specification" (p. 4; translation mine). Bertrand's "paradoxes" were set down to illustrate the contradictions that ensue from inadequate specification of various infinities of choice.

On the whole, Bertrand's criticism appears to be leveled against the subject of "local probability" (that is, geometrical probability). Cajori (1913) called this subject "the only noteworthy recent addition to probability" (p. 340). The first recorded question on local probability was Buffon's needle problem, described in his 1777 treatise "Essai d'Arithmetique morale", solved by both Buffon and Laplace (e.g., see Bulsenko et al. 1966, 4–5); the second was Sylvester's four-point problem (see Crofton 1875). Subsequently, a recognizable body of inquiry emerged, shaped largely by the work of Crofton. The geometric entities typically randomized, such as points on a line, lines in a plane, or angles in a rotational interval, are not discrete but continuous, and thus the reference sets (i.e., the number of possible outcomes) associated with them are of infinite, not finite, cardinality.

With respect to the quotation from Kendall and Moran at the beginning of this paper, it appears that Bertrand deliberately failed to distinguish the reference sets in his formulation of the random-chord problem, the better to advance his argument against applying Laplacean probabilism to infinities of possible outcomes. This in turn begs another question: If the reference sets could be properly distinguished, would the paradox then be resolved? In other words, can Bertrand's problem become "well posed"? This paper replies affirmatively, but plurally.

2. Bertrand's Several Questions. Our first task consists in clearly distinguishing three cases, which Bertrand's vague question Q (perhaps deliberately) conflates. A unit circle centered at the origin of a Cartesian plane defines two regions separated by a boundary. The curve $x^2 + y^2 = 1$ constitutes the boundary, which separates the region not enclosed by the curve ($x^2 + y^2 > 1$) from the region enclosed by the curve ($x^2 + y^2 < 1$). When generating random chords, one clearly faces methodological alternatives since, to begin with, the randomizing procedure can take place in either of these two regions, or on their boundary. Thus Bertrand's three answers can be construed initially as replies to three different questions: What is the probability that a chord drawn randomly in a circle is longer than the side of the inscribed equilateral triangle, given that the random chord is generated

(Q_1) by a procedure on the circumference of the circle?

(Q_2) by a procedure outside the circle?

(Q_3) by a procedure inside the circle?

As even a cursory glance through the literature on Bertrand's problem reveals, these three questions have suffered a singular fate; namely, there has been little recognition of the distinction between them. Appeals to reason (i.e., to clear and unambiguous treatment of the problem) have been made by philosophers and mathematicians alike, but to little avail. The three questions have remained conflated. For example, van Fraassen (1989) makes one such sensible appeal:

> Most writers commenting on Bertrand have described the problems set by his paradoxical examples as not well posed. In such a case, the problem as initially stated is really not one problem but many. To solve it we must be told *what* is random; which means, *which* events are equiprobable; which means, *which* parameter should be assumed to be uniformly distributed. (P. 305)

By implementing van Fraassen's recommended method (and not the traditional departure from it), we will be able to tell precisely what is random, what is equiprobable, and what is uniformly distributed in each of Q_1, Q_2 and Q_3.

One further clarification is necessary. We would like to distinguish between the "answer" to a question, and the "solution" to a problem. By "answer", we mean simply "a reply", consisting (in this probabilistic context) of a number. By "solution", we mean "the argument or derivation that gives rise to the answer". Thus a "question" and a "problem" are viewed as a single interrogative that may be either answered or solved. This paper associates solutions with problems, not merely answers with questions. Since a given number n may be the correct answer to a plurality of questions, n bears little relevance to our understanding of a problem when dissociated from the particular solution that gives rise to it. For example, we will show that although Bertrand's respective answers are correct when applied to Q_1, Q_2 and Q_3, two of his three solutions are inappropriate in this context; that is, they partially solve a problem that he never posed. That problem, labeled Q_4 and solved herein, is as follows: With what probability does a random chord intersecting *a fixed diameter of a circle* have length equal to or greater than the side of the inscribed equilateral triangle?

3. A Solution to Q_1. Bertrand's solution B_1 solves the following problem: Given an equilateral triangle inscribed in a circle, with what probability does a point "tossed" randomly onto the circumference lie on the arc between any two given vertices of the triangle? If the vertices of the triangle are a, b and c, then any point tossed randomly onto the circumference lies between a and b, or between b and c, or between c and a, with a probability equal to the ratio of the respective arc length (ab, bc or ca) to the arc length of the circumference itself. This ratio is just $1/3$. Note that this problem is solved by the theory of a one-dimensional, uniformly-distributed random variable (e.g., see Rozanov 1969, 40–41).

Solution B_1 attempts to apply this one-dimensional solution to a two-dimensional problem by fixing one endpoint of the chord at any vertex of the triangle (say, vertex a), and allowing the other endpoint of the chord to represent the random variable (the point tossed onto the circumference). Now B_1 implicitly requires that we rotate our fixed point (vertex a) around the entire circumference (through an angle of 2π) in order to accommodate all possible chords of the circle. Then again, by circular symmetry, argument B_1 remains invariant no matter where we fix vertex

a. So on the one hand, we require a rotational coordinate in order to specify all possible chords; while on the other, this rotational coordinate never figures in the calculation of the probability itself. Thus B_1's appeal to circular symmetry creates the illusion that Q_1 is being solved in two dimensions, whereas B_1's actual probability calculus unfolds strictly in one dimension (that of a single random variable). In other words, B_1 solves the corresponding one-dimensional problem (above).

Question Q_1, however, is a two-dimensional problem, and must be solved in a commensurate space. The following solution to Q_1 is by topologist R. Douglas (private communication). Every chord of a circle is uniquely defined by one pair of points on its circumference; conversely, every pair of points on its circumference uniquely defines one chord. Thus an isomorphism obtains between chords and circumferential pairs of points.

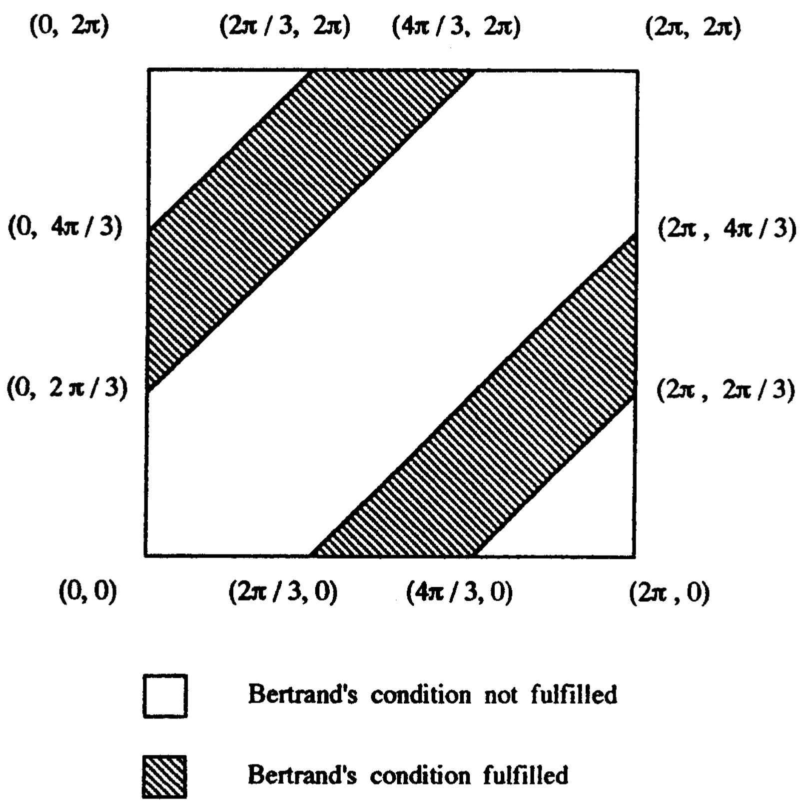

Figure 3.1. Solution to Q_1 (Douglas's Torus).

Since each random point on the circumference lies in the angular interval $[0, 2\pi]$, the probability density is $1/(4\pi^2)$. Euclidean methods show that Bertrand's condition is satisfied when the two points subtend an angle at the center that is greater than or equal to $2\pi/3$, and less than or equal to $4\pi/3$. We can choose ordered pairs of points (x,y) randomly and simultaneously by throwing darts at a Cartesian square of length 2π and height 2π (see figure 3.1). The shaded areas of the square are those within which Bertrand's condition is satisfied; moreover, the total shaded area is just one-third of the whole area. Given a highly unskilled dart thrower, who ensures a uniform distribution of darts on the square, then the probability in question is $1/3$. As Douglas indicates, the selection of a random ordered pair (x,y) on this square is topologically equivalent to the selection of a random point on the surface of a torus (formed by joining opposite sides of the square).

Note that although Bertrand's and Douglas's answers to Q_1 are equivalent, their solutions are not. Bertrand's B_1 fixes one endpoint of a chord, then selects the other endpoint randomly, in one dimension. Douglas's solution selects ordered pairs of endpoints simultaneously, randomly, and commensurately, in two dimensions.

4. A Solution to Q_2. Similarly, Bertrand's proposed solution B_2 solves the following problem: Given a linear interval $[-r,r]$, what is the probability that a point tossed randomly onto the interval falls within the subinterval $[-r/2,r/2]$? The probability is once again the ratio of the lengths of the intervals in question, in this case just $1/2$. Again, B_2 implicitly requires that we rotate our arbitrarily chosen diameter through an angle, this time of π radians, in order to accommodate all possible chords of the circle. But neither does this rotational coordinate figure in the calculation of the actual probability; it is once again dismissed by reason of circular symmetry. So B_2 also creates the illusion of a two-dimensional solution, whereas in fact it solves the one-dimensional problem above.

Jaynes (1973) and van Fraassen (1989, 305–317) set about solving Q_2. However, both Jaynes and van Fraassen erroneously claim that they are answering Bertrand's original question Q, which they both mistranslate as Q_2.

Jaynes and a colleague conducted an experiment designed to answer Q_2. One of the experimenters stood erect and tossed broom straws onto a five-inch-diameter circle drawn on the floor. Jaynes (1973) reports that 128 "successful" tosses of a broom straw (i.e., tosses for which the straw intersected the circle, nonintersecting straws being null trials) met Bertrand's condition with a statistical frequency, or empirical probability, of $1/2$. So one-half of the chords formed by straws which intersected the circle had length equal to or greater than $\sqrt{3}r$. Moreover, a chi-squared test performed on the range of chord lengths grouped into ten categories yielded, in Jaynes's words, "an embarrassingly low value of chi-squared" (ibid., 487).

It is a novel but ironic disclosure that the "embarrassingly low value" (ibid.) failed to arouse Jaynes's suspicions—not as to the reliability of his data, but as to which question his data was actually answering.[1] Bertrand asked Q; Jaynes specifically answered Q_2, heedless of the distinction. The irony is that while Jaynes's experiment furnished precise empirical corroboration of the theoretical solution to Q_2 (which we will shortly derive), Jaynes himself misapplied the data to support B_2. We have seen that B_2 correctly answers, but does not generally solve, Q_2. We will later see that while B_2 actually solves one very special case of Q_2, it most generally provides a partial solution to Q_4.

Jaynes reports the results of his experiment near the end of his paper, following a lengthy appeal to the virtues of invariant probability density. Poincaré proved that the sole probability density which remains invariant under the group of rotations and translations in the Euclidean plane is that density which corresponds to Bertrand's solution B_2 (see Kendall and Moran 1963, 16). It also happens that the correct answer to Q_2, which Jaynes found out by experiment, has the same numerical value as the answer generated by solution B_2. So Jaynes understandably but mistakenly concluded that he had empirically corroborated the theoretical construct B_2, whereas in fact he had corroborated a theoretical construct hitherto unconstructed. We now derive the classical answer to Q_2, which Jaynes's empirical data so richly corroborates.

Consider a circle of radius r lying in the plane. Let any diameter be produced in one direction. Now consider any randomly drawn line in the plane, which intersects the produced diameter at point p, some distance h from the center. Now draw the two tangents, t and t', from p. Obviously, any straight line through p that lies within the acute angle tt' intersects the circle (see figure 4.1). Call this acute angle β. Now consider the two straight lines through p, labeled c and c', which just fulfill Bertrand's condition. That is, the segments of c and c' that lie within the circle have length $\sqrt{3}r$. Then, as figure 4.1 illustrates, c and c' define a critical acute angle α. Any line through p that lies within α clearly fulfills Bertrand's condition in that its segment within the circle has length equal to or greater than $\sqrt{3}r$.

Now we invoke the principle of indifference which, via the theory of uniformly distributed random variables, asserts the following: Given any random line through p lying within angle β, the probability that it lies within angle α is just α/β. In other words, the probability $P_\alpha = \alpha/\beta$.

[1] The lower the chi-squared value (for a given number of degrees of freedom), the less suspect is the hypothesis under consideration.

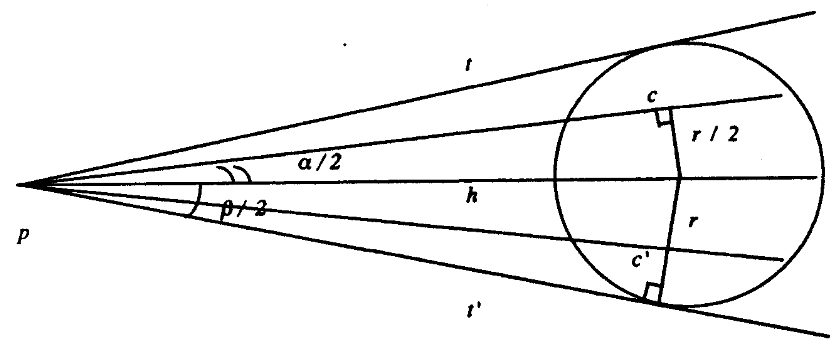

Figure 4.1. Solution to Q_2 (Classical Broom Straw Theory).

As figure 4.1 illustrates, $\sin(\beta/2) = r/h$; $\beta = 2\arcsin(r/h)$. Similarly, $\sin(\alpha/2) = (r/2h)$; $\alpha = 2\arcsin(r/2h)$. Thus

$$P_\alpha = \alpha/\beta = [\arcsin(r/2h)]/[\arcsin(r/h)] \quad (\text{where } h \geq r).$$

Since the circle lies on an unbounded Euclidean plane, the value of h itself is unbounded. So it would be useful to know whether the expression P_α approaches some limit as h increases without bound. Applying L'Hospital's Rule, we find

$$\lim_{(h\to\infty)} P_\alpha = \lim_{(h\to\infty)} (d\alpha/dh)/(d\beta/dh) = 1/2.$$

Thus, according to classical probability theory, the answer to question Q_2 is 1/2. This result was corroborated, albeit before the fact, by Jaynes's experiment. I also performed the experiment, repeatedly tossing a whittled drinking straw onto a four-inch diameter circle drawn upon the floor. After 100 successful trials, I found that forty-nine of one hundred chords satisfied Bertrand's condition.

(A word might be said about "successful" versus "unsuccessful" trials. The set L of straight lines in the plane through a fixed point p at a distance h from the circle's center has cardinality c, the power of the continuum. The subset L' of L, containing lines that intersect the circle in two points [thus defining chords] also has cardinality c, as does the subset L'' of L', whose members are just those lines defining chords that satisfy Bertrand's condition. Members of L are randomly distributed in the angular interval $[0,2\pi]$; of L', in $[-\beta/2,\beta/2]$; of L'', in $[-\alpha/2,\alpha/2]$. A uniform distribution on the interval $[0,2\pi]$ entails a uniform distribution on any subintervals, i.e., on $[-\beta/2,\beta/2]$ and $[-\alpha/2,\alpha/2]$. In practice, this allows the experimenter to ignore unsuccessful trials, i.e., to discount straws that fail to intersect the circle in two points, without prejudicing the statistics on successful trials.)

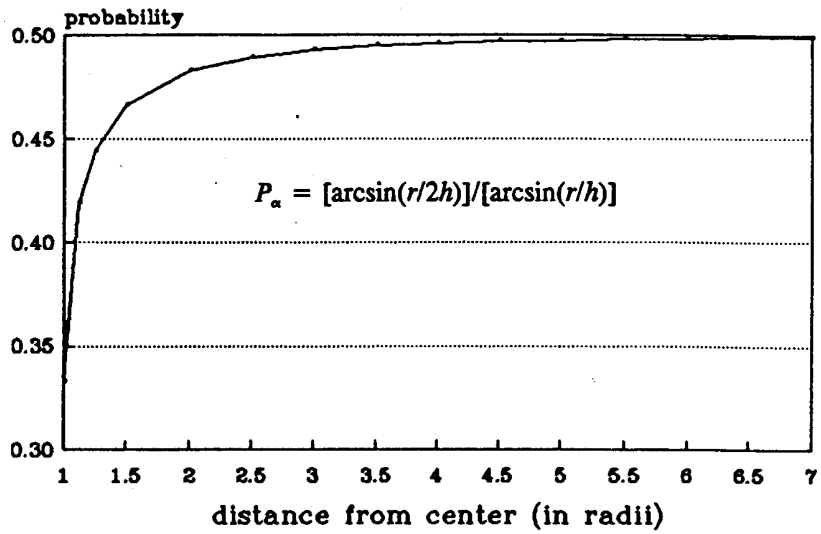

Figure 4.2. Limit of P_α as Length of Broom Straw Increases Without Bound.

The theoretical result that we have just derived also accounts nicely for Jaynes's "embarrassingly low" chi-squared value for his data. We have shown that P_α approaches a limit of $1/2$ as h increases without bound. Significantly, we can also show that P_α approaches this asymptote very rapidly indeed; thus, h need not be very large compared with r in order that P_α get close to $1/2$. Figure 4.2 graphs values of P_α as a continuous function of h. As h becomes just one order of magnitude larger than r, P_α rapidly approaches $1/2$. Recall that this problem is posed on the unbounded plane where h may become as large as we please. Empirical data draw ever-closer to the limiting frequency of $1/2$ as experimental straws become longer (compared with the radius of the circle).

Note that at the circumference, when $h = r$, $P_\alpha = 2\arcsin(1/2)/2\arcsin(1) = 1/3$. Thus P_α is consistent with the solution to Q_1.

We must now quell a potential objection to this treatment of question Q_2. To raise and address this objection most clearly, let us reconfigure the problem in Cartesian coordinates. Let the circle of radius r be centered at the origin. All random lines on this plane are uniquely and completely defined either by the equation $y = \mu x + h$, where μ is the slope and h is the y-intercept, or where the slope is undefined (i.e., the line is parallel to the ordinate) by the equation $x = k$, where k is any real constant. The previous argument holds, and P_α is still given by α/β, which asymptotically approaches $1/2$.

The objection is that suppose a random line has a y-intercept such that $-r < y < r$. In other words, suppose the random line intersects the diameter itself, and not the production of the diameter? Then $h < r$, so the expression for P_α is undefined.

The first response to the objection is that a random straight line on this plane, if its slope is other than undefined, has some y-intercept that lies in the open interval $(-\infty,+\infty)$. The probability that the y-intercept falls within the open subinterval $(-r,r)$ is $\lim_{(h\to\infty)} 2r/h$, or zero. Hence the probability that the y-intercept falls on or outside the circle is unity. And the probability P_α for these y-intercepts rapidly approaches $1/2$, as we have seen.

This first response, however, readily fails. For now consider a random straight line on the unbounded plane, with any y-intercept $|h| > r$. If we now ask with what probability the line intersects a circle of radius r (this time along the x-axis) then, ceteris paribus, the same argument applies. The probability is $\lim_{(h\to\infty)} 2r/h$, or zero. But in that case, the probability P_α is indeterminate.

In order to circumvent indeterminacy, one allows the plane to be large, but demands that it be bounded. Then the limit of h is finite, and the probability P_α approaches $1/2$. But then the objection has not been answered, for P_α is undefined if h falls within the circle. On a bounded plane, there is a finite probability that h will do just that.

The second (and effective) response to the objection is to let the random line intersect the diameter of the circle with slope μ, where μ is other than undefined. Now we are obliged to answer the question, "With what probability does the chord formed by this line meet Bertrand's condition?", without recourse to P_α.

Consider the set of all random lines that intersect the diameter with slope μ. Now construct another diameter, with slope $-\mu^{-1}$. This new diameter is perpendicular to the given set of random lines. It follows that any chord formed from an element of this set meets Bertrand's condition with probability $1/2$. Furthermore, the same argument applies to the set of random lines of undefined slope with respect to which the perpendicular diameter is simply the abscissa. The objection is thereby answered, and the answer is consistent with P_α.

It follows that B_2 can be viewed as the solution to a very special (and extremely improbable) case of Q_2 in which all random lines intersecting the circle are perpendicular to a fixed diameter. We will see that B_2 is more generally appreciable as a partial solution to Q_4. Meanwhile, the most general solution to Q_2 is that derived above.

5. A Solution to Q_3. Bertrand's B_3 answers the question, "Given a circle of radius r and a concentric circle of radius $r/2$, with what probability does a point tossed randomly into the larger circle fall on or within the smaller one?"

Every point in either circle is uniquely specifiable by an ordered pair of coordinates, that is, (x,y) in the Cartesian system, or (R,θ) in the polar system. Moreover, every chord of the large circle (diameters excepted) is uniquely defined by its midpoint. Conversely, every point in the large circle (center excepted) uniquely defines one chord. (There is a singularity at the center of the circle. Jaynes 1973, 485, tries unsuccessfully to "explain" it away. See f.n. 2.) That is, with every chord we uniquely associate one point; and with every point, we uniquely associate one chord. Thus (except for the singularity at the center) points and chords stand in an isomorphic relation.[2]

It is easy to show that any chord whose midpoint lies on or within the small circle has length equal to or greater than $\sqrt{3}r$. Owing to the isomorphism between points and chords, Q_3 is solved by B_3. The theory of uniformly distributed continuous random variables provides a consistent method for deriving B_3. In the polar coordinate system, the probability density is $\pi^{-1}r^{-2}$. Bertrand's condition is met when $R \leq r/2$. Therefore, the probability $P_{(R,\theta)}$ of meeting Bertrand's condition is just

$$P_{(R,\theta)} = 1/\pi r^2 \int_0^{r/2} \int_0^{2\pi} R dR d\theta = 1/4.$$

This result can be corroborated empirically. Suppose a dart board is constructed of the two concentric circles, and that our highly unskilled dart thrower tosses a large number of darts at it. Again, a uniform distribution of darts occurs. Of all darts which strike anywhere within the large circle (naturally, an unskilled dart thrower will also be expected to miss the target a fair number of times), that fraction which strikes within the small circle will approach one-fourth, statistically, as the number of throws increases. Recall, with every point struck by a dart we can uniquely associate the midpoint of one chord, and therefore the chord itself. Hence, Bertrand's condition in Q_3 is met with statistical frequency 1/4.

6. A Solution to Q_4. We would like to show that B_1 and B_2 are not only discrete solutions to separate problems of a one-dimensional random variable; they are also endpoints of a continuous curve of solutions to Q_4: With what probability does a random chord intersecting a fixed diameter of a circle meet Bertrand's condition?

[2] To avoid the singularity at the center, and thus to preserve the isomorphism, consider a concentric circle of infinitesimal radius ρ. Since the isomorphism between points and chords holds everywhere except at the center, we exclude the area in the neighborhood of the center from the probability calculation. Then the probability is $\lim_{(\rho\to 0)} [\pi(r/2)^2 - \pi\rho^2]/[\pi r^2 - \pi\rho^2] = 1/4$.

Let a circle of radius r be described in Cartesian coordinates, centered at the origin. Its equation is $x^2 + y^2 = r^2$. Consider any chord of the circle that intersects the vertical diameter (or y-axis) at $\pm h$ units from the center ($-r \leq h \leq r$) at any angle α ($0 \leq \alpha < \pi$) with the horizontal. The chord touches the circumference at two points: (x_1,y_1) and (x_2,y_2).

To express these points in terms of h and α, one solves the system of simultaneous equations of a line of slope $\tan\alpha$ and y-intercept h (i.e., $y = x\tan\alpha + h$), intersecting a circle of radius r. This yields the following equations:

$$x = (-h\tan\alpha \pm \sqrt{r^2\sec^2\alpha - h^2})/\sec^2\alpha.$$

$$y = (h \pm \tan\alpha\sqrt{r^2\sec^2\alpha - h^2})/\sec^2\alpha.$$

Now the length of the chord L can be expressed in terms of h and α:

$$L = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} = 2\sqrt{r^2 - h^2\cos^2\alpha}.$$

Bertrand's condition requires that $L \geq \sqrt{3}r$, which simplifies to

$$\cos\alpha \leq |r/2h|. \qquad (1)$$

In other words, if inequality (1) is satisfied by a particular (h,α), then the chord specified by that (h,α) will indeed have length greater than the side of the inscribed equilateral triangle.

From inequality (1) we can derive a probabilistic expression P_h that represents the angular probability with which, for a given h, a chord passing through h meets Bertrand's condition. That is, we first select some h, and then rotate the intersecting chord through an angular interval of π (rotation through 2π would duplicate each chord). We then ask, for a given h, what angular proportion of these chords has length greater than or equal to $\sqrt{3}r$. Given the principle of indifference (and assuming a uniform distribution of random angles), the desired proportion is the ratio of the angular subinterval that satisfies inequality (1) to the entire interval $[0,\pi]$.

That is,

$$P_h = (\alpha_{max} - \alpha_{min})/\pi.$$

$$P_h = [\arccos(-r/2h) - \arccos(r/2h)]/\pi.$$

Now let us evaluate some specific cases. Suppose we choose $h = r$. Then

$$P_h = [\arccos(-1/2) - \arccos(1/2)]/\pi.$$

$$P_h = 1/3.$$

The angular probability at the circumference corresponds to Bertrand's B_1. Now let $h = r/2$.

$$P_h = [\arccos(-1) - \arccos(1)]/\pi.$$

$$P_h = 1.$$

In this case, the angular probability is unity. In other words, when a chord intersecting $r/2$ is rotated through π radians, all resultant chords meet Bertrand's condition.

Now consider Bertrand's solution B_2. Rather than fixing a distance h from the center (as does B_1, where $h = r$), solution B_2 fixes a value of $\alpha = 0$, for all h. Applying this value to inequality (1), we find

$$\cos(0) \leq |r/2h|$$

or

$$-r/2 \leq h \leq r/2.$$

This reconfirms, by analytical geometry, a Euclidean result of which we are already aware; namely, that when chords are drawn perpendicular to a diameter, those lying in the above subinterval of the diameter meet Bertrand's condition.

In general, then,

$$P_h = \begin{cases} 1, & |h| \leq r/2 \\ (1/\pi)[\arccos(-r/2h) - \arccos(r/2h)], & r/2 < |h| \leq r. \end{cases}$$

Taking the unit circle for convenience, figure 6.1 illustrates the shape of this probability function. The average value of P_h (call it $\bar{P}$) is the area under the curve. Thus $\bar{P}$ is the sum of the two areas A_1 and A_2. By inspection, $A_1 = 1/2$, while A_2 is found by integration:

$$A_2 = 1/\pi \int_{1/2}^{1} [\arccos(-1/2h) - \arccos(1/2h)]dh = 1/4.$$

So the average probability is 3/4. This theoretical result can be corroborated empirically by the Monte Carlo method. Selecting values of h at random, ten thousand random computer trials yield an average probability of 0.7515, a close approximation indeed.

Having derived a general probability function that solves Q_4, we can also show that the set of chords specified by P_h is not isomorphic with the set of chords in the circle. The probability function is clearly isomorphic with respect to Q_4: With each random chord that intersects a fixed diameter of a circle we associate exactly one ordered pair (h,α);

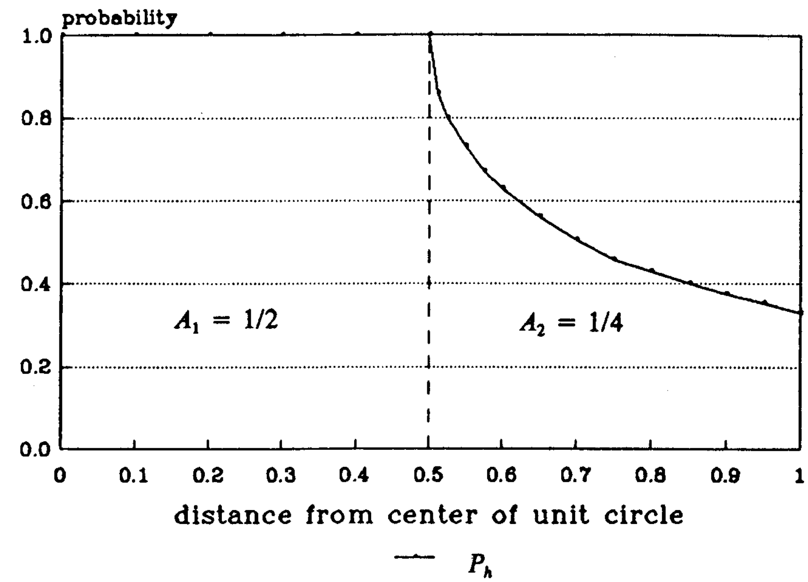

Figure 6.1. Continuum of Solutions to Q_4: $P_h = [\arccos(-r/2h) - \arccos(r/2h)]/\pi$.

and conversely, with each ordered pair (h,α) we associate exactly one chord that intersects a fixed diameter of a circle.

To show that P_h cannot embody both uniqueness and completeness when applied to Q_1 or Q_2, we distinguish three cases of P_h and show that, in each case, P_h either fails to specify a random chord uniquely, or if it specifies a random chord uniquely, then it fails to specify every random chord.

Case (1) involves the endpoint of the angular probability at which $h = r$. Let the inscribed equilateral triangle have apex a, and let all possible chords be drawn from apex a. The circle is now densely packed with chords, each of which is uniquely defined by some (r,α). But we can readily find other chords that are not yet defined in this coordinate system. Figure 6.2a illustrates two such chords, c and c'. In order to define chord c in this system, we are obliged to admit another coordinate. Call it β, where β is an angle of rotation. We rotate the endpoint p of chord c through an angle β so that p coincides with vertex a. Now chord c is defined by the coordinates (r,α,β). Thus it is specified within the distribution P_h, and meets Bertrand's condition with probability 1/3 in this case.

Note that while the rotational coordinate is necessary in order to define the chord c, this coordinate has no effect on the probability distribution

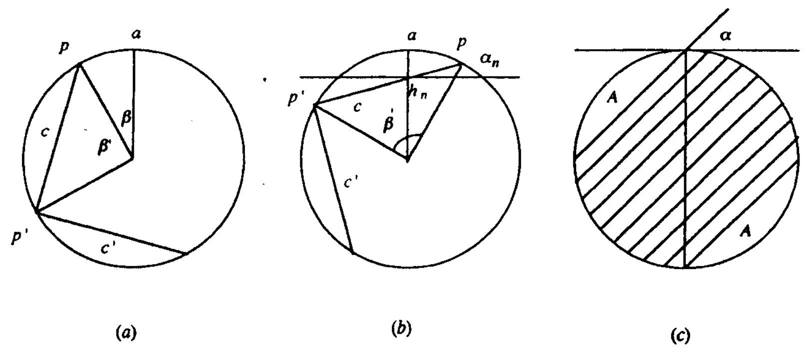

Figure 6.2. Nonisomorphism of P_h with Q_1 and Q_2.

itself owing to circular symmetry. This rotational coordinate, which must make one complete revolution (through 2π radians) in order to specify all possible chords, is introduced implicitly in Bertrand's solution B_1.

Refer again to figure 6.2*a*. Suppose we wish to define the chord c'. We similarly rotate the endpoint p' of chord c' through an angle β' so that p' coincides with apex a. Now this chord is uniquely defined by the coordinates $(r,\alpha',\beta + \beta')$. Thus it too is specified within the distribution P_h, and also meets Bertrand's condition with probability 1/3 in this case. However, look again at chord c, whose other endpoint p' now coincides with apex a. Chord c is now defined by the coordinates $(r,\alpha'',\beta + \beta')$.

Chord c was previously defined by the coordinates (r,α,β). Thus, chord c is not uniquely defined. We incur both definitions of chord c in order to accommodate chord c' in P_h. Hence, if we demand that all possible chords be specifiable within P_h, then P_h lacks the property of uniqueness. If we relax this demand, then P_h maps one-to-one, but not onto the space. If we enforce it, then P_h maps onto the space, but one-to-two. Thus, P_h must lack either uniqueness or completeness.

Case (2) involves the angular probability generated over the domain of h inside the circle, that is, in the region $-r < h < r$. Now the circle is even more densely packed with chords, each of which is uniquely defined by some (h,α). We can readily find chords c and c', which are not defined by any (h,α). Again, we introduce a rotational coordinate and rotate c through an angle β such that c is now defined by (r,α,β). Similarly, we rotate chord c' through β' such that c' is defined by (r,α',β').

However, consider the path of chord c as it rotates through angle β' (see figure 6.2*b*). When the endpoint of the chord coincides with the vertical diameter, P_h specifies this chord as (r,α_0,β_0). At the next rotational increment, P_h specifies this chord as (h_1,α_1,β_1). At the nth rotational increment, P_h specifies this chord as (h_n,α_n,β_n). An infinite number of rotational increments are specified by P_h in the angular interval $[\beta,\beta + \beta']$, and thus P_h lacks the property of uniqueness. But if we constrain P_h to defining uniquely only a subset of chords, then P_h lacks the property of completeness. In other words, P_h maps one-to-one but not onto the space. If we introduce the rotational coordinate, then P_h maps onto the space, but one-to-c, where c is the power of the continuum. Again, P_h must lack either uniqueness or completeness.

Case (3) involves the fixing of α in order to determine the corresponding domain of h that meets Bertrand's condition. This case is equivalent to considering all α's, as in case (2), only one at a time. Let α assume a given value. Then, as figure 6.2*c* illustrates, the circle is not densely packed with chords for the given α. There are two empty sectors of combined area $r^2(2\alpha + \sin\alpha\cos\alpha)$. The neglected area amounts to πr^2 when $\alpha = \pi/2$, and amounts to zero when $\alpha = 0$. For all values of α other than zero, infinite numbers of chords with slope α can be drawn in the neglected areas. These sectors are empty because all such chords have h-coordinates that lie outside the given domain, $-r \leq h \leq r$.

(Recall that Q_2 has a fundamentally different meaning than Q_4: It asks about random chords generated by means external to the circle. This question imposes no finite constraint upon h. Thus, in solving Q_2, we find no empty sectors for any fixed α.)

If, by the introduction of a rotational coordinate β, we attempt to bring any such chord into specification under P_h, then we will encounter the same problem as in case (2). That is, P_h maps either one-to-one but not onto the space, or else maps onto the space but one-to-c, where c is the power of the continuum. So P_h lacks either completeness or uniqueness when applied to the set of all chords in the circle (note that this failure is intrinsic to P_h; i.e., that P_h fails independently of the coordinate system in which P_h is expressed). Thus P_h solves Q_4, but not Q_1 or Q_2. Moreover, we have seen that B_1 and B_2 are derived as extrema of P_h. Thus B_1 and B_2 provide correct answers but incorrect solutions to Q_1 and Q_2, respectively.

7. Paradox Lost. There exists a multiplicity, if not an infinite number, of procedures for generating random chords of a circle. The answers that one finds to Bertrand's generic question Q vary according to the way in which the question is interpreted, and depend explicitly upon which geometric entity or entities are assumed to be uniformly distributed. A final example should prove instructive.

Suppose we adopt the following randomizing procedure: Our highly unskilled dart thrower repeatedly lets fly at a rectangle of length π and height 2. Each dart selects an ordered pair (x,y). The set of ordered pairs

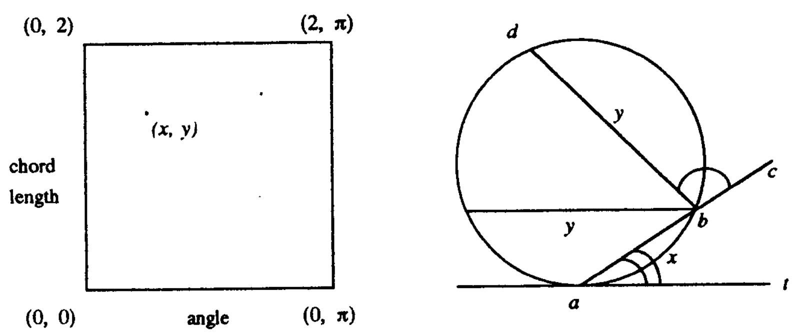

Figure 7.1. Randomizer Yielding Uniformly Distributed Chord Lengths.

on this rectangle maps isomorphically to the set of chords in the unit circle as follows (see figure 7.1). Fix any point a on the circumference. Locate point b on the circumference by drawing the chord ab at angle x to the tangent through a. Now, exactly two chords of length y can be drawn from b (except for the unique diameter when $y = 2$). Produce ab to any point c. Draw the chord bd of length y such that, in every case, angle cbd is the first of the two possible angles so formed in a counterclockwise rotation from bc. (Naturally, this procedure also maps isomorphically if, in every case, we choose instead the second of the two possible angles.) With what probability does this randomly chosen chord bd meet Bertrand's condition?

To solve this problem, look again at the randomizer. Given a uniform distribution of darts, the probability that a random chord meets Bertrand's condition is the probability that the y-coordinate of a random dart equals or exceeds $\sqrt{3}$; namely, $1 - \sqrt{3}/2$ (approximately 0.134). Note that this result is independent of the angular variable (the x-coordinate); that is, it is generated by the theory of a single uniformly distributed random variable.

In effect, this is the solution to the (one-dimensional) orthogonal projection of Bertrand's problem. Consider a bob attached to one end of a string, whose other end is fixed. Suppose the bob rotates with a constant angular velocity. The bob's trajectory, then, describes a circle which lies in a plane perpendicular to the axis of rotation. Let the diameter of the circle be 2 units. Now consider the orthogonal projection of the bob's trajectory: It describes a straight line of length 2. The probability that a point "tossed" randomly onto a given interval falls within a given subinterval is the ratio of the respective lengths of the intervals. In this case, the quantity $1 - \sqrt{3}/2$ represents the probability that a point tossed randomly onto the interval [0,2] falls within the subinterval $[\sqrt{3},2]$.

The inclusion of the random angular coordinate allows this solution to be reconstituted in two dimensions, although the probability in question is independent of the angular coordinate (that is, it remains one-dimensional). An appropriate mapping then specifies random chords both uniquely and completely, thus ensuring that the set of ordered pairs (x,y) is isomorphic with the set of chords in the circle.

The question is whether this procedure solves Bertrand's problem. The answer is that it solves a variation of the problem whose precise wording can be discovered by inspecting the randomizer. Clearly, the random variable that is uniformly distributed on the interval [0,2] represents the chord length itself. The mean chord length is 1 unit. The relatively low probability obtained in this solution, compared with higher probabilities in other solutions, is a reflection of the low mean chord length generated by this randomizer, which in turn depends upon the shape of the overall distribution of random chord lengths.

In sum, the foregoing solution solves the following problem, which we label Q_5: With what probability does a random chord of a circle fulfill Bertrand's condition if we require that the chord length be uniformly distributed?

Let us inspect the randomizers used in solving Q_1, Q_2 and Q_3, reformulating those questions with similar precision. In Q_1, two random variables representing endpoints of a chord are uniformly distributed, each in an interval of $[0,2\pi]$. I conducted a computer simulation of ten thousand trials of darts thrown randomly at Douglas's square. By randomly selecting two numbers (the angular coordinates of the endpoints of the random chord) on the interval $[0,2\pi]$, I computed the length of the resultant chord (L) as a function of the angle (θ) subtended at the center: $L = 2\sin(\theta/2)$, where $\theta \leq \pi$. This procedure yields a nonuniform distribution of chord length (see the histogram in figure 7.2). The shape of the distribution appears uniform at the lower end of the spectrum, but becomes exponential at the high end. The mean chord length is 1.26 units (in the unit circle).

In Q_2, recall that a chord is formed by a straight line through a point p at a random distance h from the center where h is uniformly distributed on the projection of the diameter between the circumference and some arbitrary remote limiting point. The straight line forms a random angle $\pm\alpha/2$ to the produced diameter through p where $\pm\alpha/2$ is uniformly distributed in the angular interval $[-\beta/2,\beta/2]$. The value of β is determined by h. That is, $\tan(\beta/2) = 1/h$, and if h is sufficiently large, then $\tan(\beta/2) \approx \beta/2$. Thus, for large enough h (e.g., for $h > 100$), $\beta \approx 2/h$. The length L of a random chord is a function of the distance z of its midpoint from the center where $\tan(|\alpha/2|) = z/h$. Again, for large h, $\tan(|\alpha/2|) \approx$

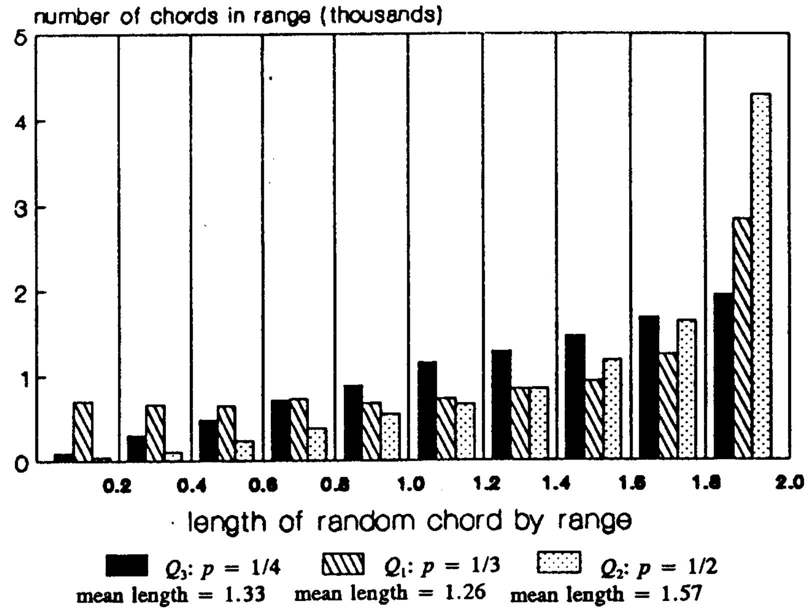

Figure 7.2. Nonuniform Distributions of Random Chords (based on 10,000 computer trials).

$|\alpha/2|$. Owing to circular symmetry, we need only consider positive α; thus, $z \approx h\alpha/2$ where $\alpha/2 \in [0,\beta/2]$.

So, for a given h, the distance of the midpoint of a random chord from the center is a linear function of the random angle. In consequence, if α is uniformly distributed in the interval $[-\beta/2,\beta/2]$, which we demand, then z is also uniformly distributed in the interval $[-1,1]$. The corresponding chord length $L = 2(1 - z^2)^{1/2}$. Taking h to be uniformly distributed on $[1,10^4]$, (that is, on the interval between the circumference and an arbitrary remote limiting point 10^4 radii distant from the center), a computer simulation of ten thousand random trials yields a nonuniform distribution of chord length (see figure 7.2). In this case the shape of the distribution appears purely exponential. The mean chord length is 1.57 units.

In Q_3, we require that randomly selected points in a circle (the midpoints of chords) be uniformly distributed with respect to area. Imagine the unit circle partitioned into n concentric rings, each of equal area π/n. The radius of the first ring is $(1/n)^{1/2}$; of the second, $(2/n)^{1/2}$; of the kth ring, $(k/n)^{1/2}$. The mean radius r is given by

$$r = 1/n \sum_{k=1}^{n} \sqrt{k/n} \approx 2/3 \text{ (for } n = 10^5).$$

However, the mean chord length is *not* a function of the mean radius since the values of the radii are *not* uniformly distributed in the interval $[0,1]$. Rather, the areas of concentric rings are uniformly distributed in the interval $[0,\pi r^2]$, where $r \in [0,1]$.

To simulate this distribution, let the computer select a random area A in the unit circle (i.e., $A \in [0,\pi]$), then find the associated radius $r = (A/\pi)^{1/2}$. The length of the associated chord is $L = 2(1 - r^2)^{1/2}$. The distribution can also be simulated by randomly selecting Cartesian coordinates (x,y) such that $x \in [0,1]$ and $y \in [0,1]$, subject to the constraint $x^2 + y^2 < 1$ (which ensures that the random point falls within the circle). Every random point represents the midpoint of a random chord at a distance $r = (x^2 + y^2)^{1/2}$ from the center. The length of such a chord is $L = 2[1 - (x^2 + y^2)]^{1/2}$. After ten thousand random trials, either method yields the same nonuniform but apparently linear distribution of chord length (see figure 7.2), with a mean chord length of 1.33 units.

In sum, Q_1, Q_2 and Q_3 can now be reformulated in the clearest possible terms. With what probability is Bertrand's condition fulfilled if we require (Q_1) that the randomly selected endpoints of chords be uniformly distributed on the circumference? (Q_2) that both (i) a randomly selected point outside the circle be uniformly distributed on the projection of the diameter between the circumference and some remote limiting point; and (ii) a randomly selected angle be uniformly distributed in the angular interval defined by the two tangents to the circle from the above randomly selected point? and (Q_3) that the randomly selected points in the circle be uniformly distributed within a series of concentric rings of equal area?

The salient findings for Q_1 through Q_5 are summarized in table 7.1. Specific requirements of different questions entail uniform distributions of random variables representing different geometric entities. In consequence, differing but respectively consistent answers ensue. It would surely be paradoxical only if the different solutions gave rise to the same answer.

To dispute which of these questions—if any—"best" represents Bertrand's generic question Q is to relinquish geometry for aesthetics. Although the randomizer in Q_1 may be deemed elegant, and that in Q_5 contrived, this distinction is arguably one of degree, not of kind. Each version of Q is so precisely because it foists a specific requirement, or set of requirements, upon the problem. While the escape from paradox lies in distinguishing between and among such requirements, it does not thereby provide a relative measure of appropriateness. It seems unlikely that a majority of mathematicians or philosophers would support the claim that any particular demand is the most "natural" one to impose on the problem. Given that a randomizer is isomorphic with chords of the circle, and that the associated random variables are uniformly distributed, then the decision to toss straws or throw darts, as well as the choice of target

TABLE 7.1. COMPARISON OF SOLUTIONS

Version of Q:	Geometric Entity(ies) Randomized:	Distribution Uniform on:	Mean Length of Chord (L)	Probability that $L \geq \sqrt{3}$
Q_1	two points on the circumference	$[0,2\pi] \times [0,2\pi]$	1.26	1/3
Q_2	point on projected diameter, & sub-angle between its tangents	$[1,d] \times [-1,1]$ where $d >> 1$	1.57	1/2
Q, B_1	angle between chord and tangent to a fixed point	$[0,\pi]$	1.28	1/3
Q, B_2	position of chord intersecting fixed diameter at right angle	$[-1,1]$	1.57	1/2
Q_3, B_3	position of point in circle	$[-1,1] \times [-1,1]$	1.33	1/4
Q_4	intersection & angle of chord with respect to fixed diameter	$[-1,1] \times [0,\pi]$	1.81	3/4
Q_5	position and length of chord	$[0,\pi] \times [0,2]$	1.00	$1 - \sqrt{3}/2$

(whether real or virtual), remains grounded in subjectivity. In this light, Poincaré's, Jaynes's and van Fraassen's appeals to invariance of probability density under the group of transformations in the Euclidean plane as the supreme arbiter of Bertrand's question merely suggest another version of Q: With what probability is Bertrand's condition fulfilled if we demand such invariance? (In that case, the answer is 1/2.) Bertrand himself made no such demand; the "paradox" arose precisely because he made no demands at all.

Significantly, the many versions of Bertrand's problem are solvable, and each solution relies upon the very procedure—namely, the consistent application of the principle of indifference to infinite sets—that Bertrand proscribed. Bertrand's former paradox of random chords is resolved by the expedient of providing what he, from the outset, withheld, namely, a "sufficient specification" of such sets. This and other apparent paradoxes arise, not out of nature, rather from want of reason (see also Marinoff 1993). To paraphrase the Bard: "There is nothing paradoxical, but thinking makes it so".

APPENDIX

It remains to locate some competing and contradictory claims, culled from the literature on the problem, within the framework of this resolution.

(a) Borel (1909, 110–113) asserted that "the majority of conceivable natural procedures" leads to the answer 1/2.

Response: Natural procedures lead to the answer 1/2 if we are answering Q_2. Equally, natural procedures lead to other answers if we are answering other questions.

(b) Poincaré (1912) proved that, with regard to representations of straight lines on a Euclidean plane, the only differential element that remains invariant under the group of translations and rotations is that which corresponds to the probability density yielding the answer 1/2 to Q_2 (cited by Kendall and Moran 1963, 16).

Response: This reinforces the argument that 1/2 is the correct answer to Q_2.

(c) Gnedenko (1962, 40–41) asserted that Bertrand's three different results "would be appropriate" in three different experiments, but Gnedenko does not describe such experiments.

Response: We have herein described three experiments to which Bertrand's three different results apply.

(d) Kendall and Moran (1963, 10) affirmed that "all three solutions are correct, but they refer to different problems". They do not articulate the problems.

Response: The three different problems are Q_1, Q_2 and Q_3.

(e) Uspensky (1937, 251) averred that "we are really dealing with two different problems". He poses Q, and argues that B_1 and B_2 answer different questions.

Response: Uspensky correctly asserts that a mechanism of random choice must be clearly specified when randomizing a given geometric entity (ibid.). But his analysis of Bertrand's problem remains superficial.

(f) Northrop (1944, 181–183) simply threw up his hands and gave out that "one guess is as good as another".

Response: One guess is at least as uninformed as another.

(g) Weaver (1963, 356–357) merely cautioned that "you have to watch your step".

Response: I have endeavored to watch not only my step, but also the steps (and missteps) of others who trod here before me.

(h) Van Fraassen (1989, 298–317) argues that putative inconsistencies in the wake of the principle of indifference can often be obviated by appropriate reformulation of the given problem, careful consideration of the symmetries involved, and subsequent employment of the proper transformation group. On that basis, he accepts Q_2 (Jaynes's mistranslation of Q) and excludes other variations. Van Fraassen concludes on a pessimistic note (in a section subtitled "Pyrrhic Victory and Ultimate Defeat"), claiming that the principle of indifference is not a universal guarantor of consistent a priori predictions, "[D]ifferent models of the same situation could fairly bring us diverse answers" (p. 317).

Response: On the whole, van Fraassen's treatment of the general issue is laudable, but his acceptance of Jaynes's treatment of Bertrand's problem incorporates Jaynes's errors of translation and interpretation. This resolution shows that the principle of indifference may consistently "bring us diverse answers" to Bertrand's question.

(i) Von Mises (1964, 159–166) takes issue with the principle of indifference in similar applications, including another of Bertrand's problems as well as the Buffon needle problem. Von Mises denies that such problems can be treated coherently by the received probability calculus.

Response: The principle seems perfectly consistent with respect to Buffon's problem and Bertrand's problem (among others). We find complete correspondence between classical probabilistic predictions and empirical limiting frequencies.

(j) Keynes (1921, 63) concludes, "So long as we are careful to enunciate the alternatives in a form to which the Principle of Indifference can be applied unambiguously, we shall be prevented from confusing together distinct problems, and shall be able to reach conclusions in geometrical probability which are unambiguously valid".

Response: This study has endeavored to follow Keynes's positivistic prescription. Careful enunciations of alternatives, unambiguous applications of the principle of indifference, and clear demarcation between distinct problems together lead to conclusions in geometric probability that are self-consistent and therefore unparadoxical.

REFERENCES

Bertrand, J. (1889), *Calcul des Probabilités*. Paris: Gauthier-Villars.
Borel, E. (1909), *Éléments de la Théorie des Probabilités*. Paris: Hermann et Fils.
Bulsenko, N.; Golenko, D.; Shreider, Y.; Sobol, I.; and Sragovich, V. (1966), *The Monte Carlo Method: The Method of Statistical Trials*. s. v. G. Tee: London: Pergamon Press.
Cajori, F. (1913), *A History of Probability*. London: MacMillan & Co.
Carnap, R. ([1950] 1962), *Logical Foundations of Probability*. 2d ed. Chicago: University of Chicago Press.
Crofton, M. (1875), "Probability", *Encyclopaedia Britannica*, vol. 19. 9th ed. Edinburgh: Adam & Charles Black.
Gnedenko, B. V. (1962), *The Theory of Probability*. New York: Chelsea.
Hacking, I. (1975), *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. Cambridge, England: Cambridge University Press.
Jaynes, E. T. (1973), "The Well-Posed Problem", *Foundations of Physics 3*: 477–492.
Jeffreys, H. ([1948] 1961), *Theory of Probability*. 3d ed. Oxford: Clarendon Press.
Kendall, M. and Moran, P. (1963), *Geometrical Probability*. New York: Hafner.
Keynes, J. (1921), *A Treatise on Probability*. London: MacMillan.
Laplace, P. S. Marquis de ([1820] 1886), *Théorie Analytique des Probabilités*. 3d ed. Reprint. Paris: Gauthier-Villars.
Lucas, J. (1970), *The Concept of Probability*. Oxford: Clarendon Press.
Marinoff, L. (1993), "Three Pseudo-Paradoxes in 'Quantum' Decision Theory: Apparent Effects of Observation on Probability and Utility", *Theory and Decision 35*: 55–73.
Northrop, E. (1944), *Riddles in Mathematics: A Book of Paradoxes*. New York: Van Nostrand.
Poincaré, H. (1912), *Calcul des Probabilités*. Paris: Gauthier-Villars.
Rozanov, Y. A. (1969), *Introductory Probability Theory*. Translated and edited by R. A. Silverman. Engelwood Cliffs, NJ: Prentice Hall.
Uspensky, J. (1937), *Introduction to Mathematical Probability*. New York: McGraw Hill.
van Fraassen, B. (1989), *Laws and Symmetry*. Oxford: Clarendon Press.
von Mises, R. (1964), *Mathematical Theory of Probability and Statistics*. New York: Academic Press.
Weatherford, R. (1982), *Philosophical Foundations of Probability Theory*. London: Routledge & Kegan Paul.
Weaver, W. (1963), *Lady Luck: The Theory of Probability*. Garden City: Doubleday-Anchor.