Statistical Science
1994, Vol. 9, No. 2, 279–284

A Lewis Carroll Pillow Problem: Probability of an Obtuse Triangle

Stephen Portnoy

Abstract. On the 100th anniversary (1993) of Lewis Carroll's *Pillow Problems*, Eugene Seneta presented a selection of the problems the author, Charles Dodgson, claims to have solved while in bed. The selection omits the one problem in continuous probability: "Three points are taken at random on an infinite plane. Find the chance of their being the vertices of an obtuse-angled triangle." Charles Dodgson presents a solution that involves a clear error in conditioning. An alternative solution is suggested here. This solution seems rather natural and should be especially appealing to statisticians. The nature of the solution suggests a method for using transformation groups to give meaning to the phrase "at random" in somewhat general situations.

Key words and phrases: Random triangle, sampling at random, invariant measure, transformation group, homogeneous space.

1. INTRODUCTION

In 1893 Charles Dodgson presented a collection of *Pillow Problems* as Part II of *Curiosa Mathematica* (1893), written under the pen name Lewis Carroll. One of the problems asked for the probability that a triangle formed by choosing three points at random on an infinite plane would have an obtuse angle. Dodgson purports to solve this problem as follows. Let AB denote the longest side of the triangle. Then the third point, C, must lie in the lune-shaped intersection of the two discs of radius AB with centers at A and B (see Figure 1). Furthermore, the triangle has an obtuse angle if and only if the third point lies inside the circle with diameter AB. This holds since if C is on the circle, the largest angle will be a right angle exactly. As a solution, Dodgson simply takes the ratio of the area of the disc with diameter AB to the area of the lune-shaped region and obtains

$$p = \frac{\pi/8}{\pi/3 - \sqrt{3}/4} \approx 0.64.$$

Clearly, the problem is not well-posed since the notion of "at random on an infinite plane" is not defined precisely. Nonetheless, there is a rather compelling reason that Dodgson's solution cannot be correct. The following argument was described to me by Frank Wattenberg (Wattenberg, 1973). Let AB now denote the *second-longest* side. Then the third point must lie in the symmetric difference of the two discs of radius AB centered at A and B; that is, C must lie in the complement of the lune relative to the union of the discs (see Figure 1). Furthermore, the triangle is obtuse if and only if C lies outside the perpendiculars to AB through A and B; that is, the heavily shaded region in Figure 1 indicates the triangles with no obtuse angle. The ratio of the areas of these two regions is

$$p = \frac{\pi/2}{\pi/3 + \sqrt{3}/2} \approx 0.82.$$

Since both of these solutions seem equally plausible (or equally implausible), something must be wrong. The error clearly involves conditioning on a specific side. Whatever meaning can be given to "at random on the plane," there is no reason to expect that the third point has a uniform distribution *conditional* on the other two points determining the longest (or second-longest) side. The fact that the two solutions above differ shows that the conditional distribution cannot be uniform in both cases and suggests that it is uniform in neither. Clearly, a well-formulated problem can be produced by specifying some particular probabilistic mechanism for drawing a random triangle; for example, taking the vertices to have a uniform distribution over some fixed compact set in the plane. There are a number of papers giving such solutions. Unfortunately, since the solution depends

Stephen Portnoy is Professor of Statistics, Department of Statistics, University of Illinois, 725 S. Wright Street, Champaign, Illinois 61820.

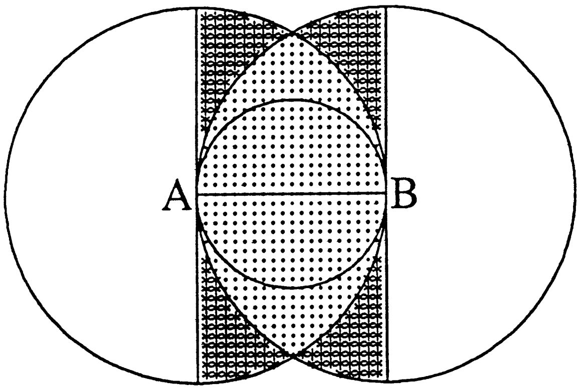

FIG. 1. *Region for third point given side AB.*

on the shape of the set (e.g., the probability of obtuseness is nearly 1 for a thin rectangle), this approach cannot provide reasonably unique answers, even by taking limits. Here we consider the question of whether any reasonable or natural answer can be provided for the original Pillow Problem.

2. SOLUTION TO THE PILLOW PROBLEM

Consider the following approach to choosing three points at random on an infinite plane. First, introduce coordinates: $A = (X_1, Y_1)$, $B = (X_2, Y_2)$ and $C = (X_3, Y_3)$. Then the set of triangles can be identified with six-dimensional Euclidean space R^6, and the set of obtuse triangles T_O is a subset of R^6. Now, multiplying each coordinate by the same constant produces a similar triangle. Thus, if a point $u \in R^6$ corresponds to an obtuse triangle, any point on the ray through u from the origin is in T_O. Therefore, T_O is a double cone, and hence it intersects each sphere about the origin in R^6 in a similar set. As a consequence, $P\{T_O\}$ is the same value for every distribution P that is spherically symmetric in R^6. This holds since, for any spherically symmetric probability distribution, the conditional distribution given the distance from the origin is uniform on the appropriate sphere; thus the conditional probability of T_O on any sphere is the same value (whatever the radius).

Next, it seems reasonable that one consequence of taking points "at random in the plane" is that the induced distribution in R^6 be spherically symmetric. A general approach to justifying this statement will be presented in the next section. If this is accepted, the Pillow Problem can be solved by using any spherically symmetric distribution in R^6. One such distribution, especially appealing to statisticians and probabilists, is the unit multivariate normal distribution in R^6, having each of the six coordinates independent and identically distribution as $\mathcal{N}(0, 1)$.

PROPOSITION 1. *Suppose the six coordinates $\{X_i, Y_i : i = 1, 2, 3\}$ are independent and identically distributed as $\mathcal{N}(0, 1)$. Then $P\{T_O\} = \frac{3}{4}$. Hence, $P\{T_O\} = \frac{3}{4}$ for every distribution that is spherically symmetric in R^6.*

The proof will require the following four well-known facts about normal sampling:

1. If U and V are independent $\mathcal{N}(0, 1)$, then $\frac{1}{2}(U + V)$ and $\frac{1}{2}(U - V)$ are independent $\mathcal{N}(0, \frac{1}{2})$.
2. If U and V are independent $\mathcal{N}(0, 1)$, then $aU + bV$ is $\mathcal{N}(0, a^2 + b^2)$.
3. If U and V are independent $\mathcal{N}(0, a^2)$, then $U^2 + V^2$ has the same distribution as $a^2 S$, where $S \sim \chi_2^2$ (a chi-square distribution with two degrees of freedom).
4. If R and S are independent χ_2^2 random variables, then $R/(R + S)$ is uniformly distributed on the interval $[0, 1]$.

PROOF. Since at most one angle of a triangle can be obtuse and the distribution of coordinates is invariant under permutation (of coordinates in R^6),

$$\begin{aligned} (1) \quad & P\{\triangle ABC \text{ obtuse}\} \\ &= P\{\angle ABC \geq 90^\circ \text{or} \\ &\qquad \angle CAB \geq 90^\circ \text{or} \angle BCA \geq 90^\circ\} \\ &= 3P\{\angle ABC \geq 90^\circ\}. \end{aligned}$$

Now, $\angle ABC$ is obtuse if and only if the median from B to the midpoint of AC is smaller than $\frac{1}{2}AC$. This follows (as in Dodgson's solution), since if these lengths are equal, B will lie on a circle with diameter AC, and $\angle ABC$ would be a right angle. Calculating these lengths in terms of coordinates,

$$\begin{aligned} L^2(\text{median}) &= \left(X_2 - \tfrac{1}{2}(X_1 + X_3)\right)^2 \\ &\quad + \left(Y_2 - \tfrac{1}{2}(Y_1 + Y_3)\right)^2, \\ L^2\left(\tfrac{1}{2} AC\right) &= \left(\tfrac{1}{2}(X_1 - X_3)\right)^2 + \left(\tfrac{1}{2}(Y_1 - Y_3)\right)^2. \end{aligned}$$

Therefore,

$$\begin{aligned} (2) \quad & P\{\angle ABC \geq 90^\circ\} \\ &= P\left\{L(\text{median}) \leq L\left(\tfrac{1}{2} AC\right)\right\} \\ &= P\left\{L^2(\text{median}) \leq L^2\left(\tfrac{1}{2} AC\right)\right\} \\ &= P\Big\{\left(X_2 - \tfrac{1}{2}(X_1 + X_3)\right)^2 + \left(Y_2 - \tfrac{1}{2}(Y_1 + Y_3)\right)^2 \\ &\qquad \leq \left(\tfrac{1}{2}(X_1 - X_3)\right)^2 + \left(\tfrac{1}{2}(Y_1 - Y_3)\right)^2\Big\} \\ &\equiv P\{U_1^2 + U_2^2 \leq V_1^2 + V_2^2\}. \end{aligned}$$

Now from fact 1, $X_1 + X_3$ and $X_1 - X_3$ are independent, and similarly for $Y_1 + Y_3$ and $Y_1 - Y_3$. It follows that U_1, U_2, V_1 and V_2 are independent. Furthermore, from fact 2, U_1 and U_2 are $\mathcal{N}(0, \frac{3}{2})$, and V_1

and V_2 are $\mathcal{N}(0, \frac{1}{2})$. Thus, from fact 3, we can define R and S as independent χ_2^2 random variables, and equation (2) becomes

$$\begin{aligned} &P\{\angle ABC \geq 90°\} \\ &= P\left\{\frac{3}{2}R \leq \frac{1}{2}S\right\} = P\left\{\frac{R}{S} \leq \frac{1}{3}\right\} \\ &= P\left\{\frac{R}{R+S} \leq \frac{1}{4}\right\}. \end{aligned} \tag{3}$$

However, from fact 4, $R/(R+S)$ is uniform on $[0, 1]$; so the probability above is just $\frac{1}{4}$. Therefore, from equation (1),

$$P\{\triangle ABC \text{ obtuse}\} = \tfrac{3}{4}. \qquad \square$$

3. AN APPROACH TO DEFINING "AT RANDOM"

The solution given in the previous section suggests a general approach to giving meaning to the phrase "at random." First it seems clear that some notion of indistinguishability among objects is necessary for one of the objects to be chosen at random. In particular, any such "random" mechanism should not depend on how objects are labeled. There are several mathematical ways to define concepts concerning equivalence among a set of objects. One of the most fruitful, especially in statistics, is the use of transformation groups: a transformation group is a set $\mathcal{X}$, together with a group $\mathcal{G}$ of invertible transformations from $\mathcal{X}$ to $\mathcal{X}$. The transformation group, $(\mathcal{X}, \mathcal{G})$ is said to be *transitive* or is called a *homogeneous space* if, for each pair of objects in $\mathcal{X}$, there is a transformation taking one object into the other. Such mathematical structures are familiar in statistical decision theory; see Wijsman (1990) for a rather thorough treatment. A homogeneous space would seem to provide an ideal candidate for a set of indistinguishable objects. If the group is locally compact, there is an invariant ("Haar") measure on $\mathcal{G}$, and this measure induces an invariant measure on $\mathcal{X}$ [at least, under appropriate smoothness conditions; see Wijsman (1990)]. The measure is unique up to multiplication by a scalar. If $\mathcal{X}$ is compact, then the measure is finite and may be taken to be a probability measure. Otherwise, further development is necessary.

Some simple examples should clarify this idea. If $\mathcal{X}$ is finite, we may wish to take $\mathcal{G}$ to be the group of all permutations on $\mathcal{X}$. The invariant measure on $\mathcal{G}$ is just counting measure, and this induces counting measure on $\mathcal{X}$. Normalizing this induced measure makes each object equally likely; that is, it gives the usual model for random sampling from a finite population. Note that $\mathcal{G}$ is larger than $\mathcal{X}$ in this case. $\mathcal{X}$ is also invariant under the smaller group of cyclic permutations, and this smaller group would still induce the same probability measure on $\mathcal{X}$. This exemplifies a general phenomenon: if $\mathcal{G}' \supset \mathcal{G}$ and $(\mathcal{G}', \mathcal{X})$ is also a homogeneous space, then the measures induced by $\mathcal{G}'$ and by $\mathcal{G}$ are generally the same (up to multiplication by a scalar) under appropriate conditions.

As a second example, consider the unit circle under the group of rotations. Clearly, this structure generates the uniform distribution on the circle and provides the usual model for a random spin of a spinner. One often uses the uniform density on $[0, 1]$ as a model for choosing a point at random on the unit interval. As noted in the next section, I believe this is much more problematic. The presence of the endpoints (0 and 1) clearly complicates the issue. Since the endpoints are distinguishable, it is not clear why a uniform distribution represents randomness except by fiat. Consider the following thought experiment: imagine asking people to choose a point at random from a line segment and from a circle. I would expect that even well-trained probabilists would tend to avoid points too close to the ends of the segment, but would be much more uniform on the circle. It is of course always possible to impose the uniform distribution as a model, but I claim that this is not reasonable unless you are prepared to consider the points invariant under something like rotation—and then, both endpoints could not be present.

In the specific Pillow Problem here, the reader is asked to choose three points at random in the plane. Considering the group of translations on the plane leads to Lebesgue measure on the plane. Equivalently, letting the sample space be R^6, the invariant measure is Lebesgue measure. However, Lebesgue measure is not finite, so an additional critical ingredient is needed. Consider a general homogeneous space, and suppose the induced measure is itself invariant under some other (arbitrary) group $\mathcal{H}$. Suppose further that the events of interest are given the same probability by any $\mathcal{H}$-invariant probability measure. Then, if there is an invariant probability measure, it seems very natural to take the probability under any such measure as the appropriate probability under sampling "at random." In the case here, Lebesgue measure is invariant under rotations (in R^6), and any spherically symmetric probability gives the obtuse triangles the same probability— namely, the solution given in Section 2.

There is an alternative approach to this construction that seems to appeal to researchers well-versed in transformation groups. If the events of interest are invariant under some group $\mathcal{H}'$ (as the shape of triangles is invariant under common scale changes), it should be possible to define a "conditional" measure given the sigma field generated by the orbits. Methods of "disintegration" (see Schwartz, 1976) will provide such a measure on the space of orbits even when the original measure is not finite. The mathematics is even easier using the structure of transformation groups. General results are given in Wijsman (1990, Chapter 7), and Ambartzumian (1990) presents special cases from geometrical probability. If the space of orbits is compact, this conditional measure will be finite and can be taken to be a probability measure. Unfortunately, the resulting measure depends strongly on the choice of the group $\mathcal{H}'$. It may even depend on the representation of the space of orbits, although uniqueness does hold under certain conditions (see Wijsman, 1990, Section 7.3).

Clearly, it is possible to imagine situations where such ideas will not work. However, it does appear to be very useful to think about randomness in terms of invariance. If invariance and symmetry considerations are applicable, then a natural solution may be possible.

Of course, symmetry arguments have a long and cherished tradition in probability theory. However, as noted in the Section 4, naive application of symmetry can often be extremely misleading. Zabell (1988) provides an extensive history of the use (and misuse!) of symmetry arguments in probability. Nonetheless, the use of invariance ideas does seem to provide clarification in the Pillow Problem and in many other examples. Jaynes (1973) applies ideas similar to those above to obtain a solution of Bertrand's paradox (concerning random chords on a circle), and he argues in favor of the general usefulness of invariance. Certainly, many statisticians find these ideas valuable: statistical invariance is a well-developed area, and statisticians often suggest the use of invariance to simplify problems, to obtain optimal (minimax) procedures or to provide "informationless" priors [see Lehmann (1983), Berger (1980), Zidek (1969) and Portnoy (1971), among many others].

We conclude this section with a few more related examples of the application of these invariance ideas. First consider alternate ways of generating a "random" triangle. Suppose a side and two angles are taken "at random" (say, as angle-side-angle, to be specific). Given that the values form a triangle (i.e., that the sum of the two angles is less than π), what is the conditional probability that the triangle is obtuse? Invariance ideas suggest the following: by rotational invariance, each angle (θ_1 and θ_2) should be uniform on $[0, 2\pi)$ and, by scale invariance, the length of the side (a) should have measure da/a; that is, the element of measure is $d\theta_1\, d\theta_2\, da/a$ on the set $[0, 2\pi)^2 \times R^+$. However, the value of a is irrelevant: for any a, the desired conditional probability is the ratio of the area of the set

$$\left\{(\theta_1, \theta_2)\colon \theta_1 \ge \frac{\pi}{2} \text{ or } \theta_2 \ge \frac{\pi}{2} \text{ or } \theta_1 + \theta_2 \le \frac{\pi}{2}\right\}$$

to that of the set $\{(\theta_1, \theta_2)\colon \theta_1 + \theta_2 \le \pi\}$. Once again the probability of an obtuse triangle is $\frac{3}{4}$. An inspection of these triangular regions shows this directly.

Now suppose that the three sides are drawn "at random." Again, conditioning on the event that the three sides form a triangle, find the probability that the triangle is obtuse. Here, using scale invariance, the measure for the three sides $\{a, b, c\}$ is $(da/a)(db/b)(dc/c)$ on the set $(R^+)^3$. Again using rotational invariance of this measure on $(R^+)^3$ and some rather tedious computations yields the answer, $\frac{3}{4}$. Here, there is an alternative method that avoids new computation: consider transforming (a, b, c) to (a, θ_1, θ_2) with

$$\cos\theta_1 = \frac{c^2 - a^2 - b^2}{2ab} \quad \text{and} \quad \cos\theta_2 = \frac{b^2 - a^2 - c^2}{2ac}.$$

A straightforward Jacobian calculation gives

$$\frac{da\, db\, dc}{abc} = d\theta_1\, d\theta_2 \frac{da}{a},$$

which shows that the answers must coincide (since the conditions for forming a triangle transform properly).

It is rather interesting that choosing two sides and the included angle seems to give a different answer. Here the natural measure is $(da/a)(db/b)\,d\theta$. However, there may be a problem here: θ and $2\pi - \theta$ give the same triangle. Thus, the domain of θ should be $[0, \pi]$, but this suggests that there may be an endpoint problem and perhaps indicates that this problem is somewhat less natural than the others. Ignoring this problem (and taking θ uniform on $[0, \pi]$), we can proceed as above. By rotational invariance of (a, b), we can take $a = r\cos\tau$ and $b = r\sin\tau$ with τ uniform on $[0, \pi/2)$. Letting A denote the vertex opposite the side of length a, B denote the vertex opposite the side of length b and C denote the third vertex (opposite the angle θ), elementary trigonometry shows that

$$P\left\{\angle CBA \ge \frac{\pi}{2}\right\} = P\left\{b \ge \frac{a}{\cos\theta}\right\}$$
$$= P\left\{\cos\theta \ge \frac{\cos\tau}{\sin\tau}\right\}.$$

The integral giving this probability can be calculated numerically to be approximately 0.1713, but it does not seem possible to compute the integral symbolically (at least not using *Mathematica*). The probability that the triangle is obtuse is the probability that θ exceeds $\pi/2$ plus $P\{\angle CBA \ge \pi/2\}$ plus $P\{\angle CAB \ge \pi/2\}$, which gives approximately $0.5 + 2(0.1713) = 0.8426$.

A referee suggested yet another approach. Any three points in the plane must lie on some circle. If one conditions on this circle, rotational invariance suggests taking the three points uniformly distributed. Fix the first point (A) at angle zero, and let B be the second point. Let S be the shorter arc from A to B. That is, take S to be the arc $[0, \theta]$ with $\theta < \pi$. Then it is not hard to see that the triangle is acute if and only if C lies in the arc S' exactly opposite S; that is, S' is the arc $[\pi, \pi + \theta]$. So the probability that the triangle is acute conditional on A and B is just $\theta/(2\pi)$. Since A may be fixed, the unconditional probability is

$$(4) \quad \begin{aligned} P\{\text{acute}\} &= \frac{1}{\pi}\int_0^{\pi} \frac{\theta}{2\pi}\, d\theta = \frac{1}{4}; \\ P\{\text{obtuse}\} &= \frac{3}{4}. \end{aligned}$$

None of these alternative approaches uses a measure in R^6. Thus, the connection with spherical symmetry in R^6 is unclear, and the fact that the answer $\frac{3}{4}$ arises so often seems to be fortuitous. It would be extremely interesting to find a general principle underlying all the $P = \frac{3}{4}$ answers. Note also that none of these approaches deals with choosing a random element from the set of triangles. The sample space is always strictly larger than the set of triangles. In fact, it is rather difficult (or perhaps impossible) to find a natural group operating transitively on the set of triangles (however this set is defined). As noted in the next section, this makes the notion of "random triangle" somewhat problematic.

4. DISCUSSION

There are two common statements about choosing objects at random as discussed here that I believe are fallacious:

1. "At random" means using the uniform distribution; that is, taking a constant density over some set.
2. Since there is no unambiguous answer, all answers are equally good (or bad).

The fallacy in the first statement is perhaps clearest is some "classical" conundra. Those early gamblers who thought that there would be a $\frac{1}{3}$ probability of getting one head and one tail in tossing two coins were making this error. See also Zabell (1988). Another recent example is the "Monty Hall" problem: a contestant is asked to choose one of three doors, behind exactly one of which is the prize. After the contestant chooses one door, the host shows one of the other two doors that does not have the prize. The contestant is then permitted to switch doors. The naive (and most common!) response is to retain the original choice, arguing that the two remaining unknown doors are equally likely. In fact, if the initial chance was $\frac{1}{3}$, the chance of getting the prize by switching is $\frac{2}{3}$. These examples indicate the care with which symmetry arguments must be used—if the objects are not appropriately indistinguishable, a uniform distribution may be wrong.

Assumptions of uniformity are especially problematic in continuous probability problems like sampling from a line segment or sampling triangles. Care is clearly required whenever there are endpoint effects or other departures from symmetry. Here, symmetry arguments based on transformation groups should be especially valuable for suggesting when departures are present and when uniform assumptions may be unnatural.

The second statement is clearly fallacious as presented. It is often made more subtle by leaving the first clause implicit, but still represents the intellectual laziness of refusing to analyze a problem because its formulation is flawed in some manner. This statement is bandied about in casual argumentation all the time by persons, including myself, who should know better.

The Pillow Problem offers a good example of this fallacy. The issue of choosing a "random triangle" is indeed problematic. I believe the difficulty is explained in large measure by the fact that there seems to be no natural group of transitive transformations acting on the set of triangles. However, the pillow problem as asked in terms of random points in the plane does have a reasonable answer. I do not think it is accidental or perverse that most people with whom I have discussed this problem accept the argument using rotational invariance in R^6 as more natural than other solutions. In fact, there are informal ways of thinking about choosing a point at random in the plane. For example, one could throw darts at a large board or drop pebbles on a large floor from some height. In either case, if the target or floor were large enough so that edge effects could be ignored, one would expect the points in R^6 to satisfy rotational symmetry (at least approximately), and thus to explain why the argument does not seem so unreasonable. In any event, application of notions of symmetry seems to make some contribution to clarifying where the problems really lie. In fact, I would suggest that the value and enjoyment of mathematics is greatest when it is used to provide imperfect and incomplete solutions to ill-formulated problems.

ACKNOWLEDGMENT

This research was partially supported by NSF Grant SES-91-21776.

REFERENCES

AMBARTZUMIAN, R. V. (1990). *Factorization Calculus and Geometric Probability*. *Encyclopedia of Mathematics and Its Applications* **33**. Cambridge Univ. Press.

BERGER, J. (1980). *Statistical Decision Theory*. Springer, New York.

CARROLL, L. (1893). *Curiosa Mathematica, Part II: Pillow Problems*. MacMillan, London.

JAYNES, E. T. (1973). The well-posed problem. *Found. Phys.* **3** 477–493.

LEHMANN, E. (1983). *Theory of Point Estimation*. Wiley, New York.

PORTNOY, S. (1971). Formal Bayes estimation with application to a random effects model. *Ann. Math. Statist.* **42** 1379–1402.

SCHWARTZ, L. (1976). *Lectures on Disintegration of Measures*. Tata Inst. Fundamental Research, Bombay.

SENETA, E. (1993). Lewis Carroll's "pillow problems": on the 1993 centenary. *Statist. Sci.* **8** 180–186.

WATTENBERG, F. (1973). Personal communication.

WIJSMAN, R. A. (1990). *Invariant Measures on Groups and Their Use in Statistics*. IMS, Hayward, CA.

ZABELL, S. (1988). Symmetry and its discontents. In *Causation, Chance, and Credence* **1** (B. Skyrms and W. Harper, eds.) 155–190. Kluwer, Dordrecht.

ZIDEK, J. (1969). A representation of Bayes invariant procedures in terms of Haar measure. *Ann. Inst. Statist. Math.* **21** 291–308.