The Probability of Heads
Author(s): Joseph B. Keller
Source: *The American Mathematical Monthly*, Vol. 93, No. 3 (Mar., 1986), pp. 191-197
Published by: Taylor & Francis, Ltd. on behalf of the Mathematical Association of America
Stable URL: https://www.jstor.org/stable/2323340
Accessed: 27-09-2024 12:58 UTC

TIETZE EXTENSION THEOREM. *If M is a closed subset of a normal space X, then any bounded continuous real-valued function on M may be extended to a continuous function on X with the same bound.*

Proof. Let T be the restriction map from the space of bounded continuous functions on X to the space of bounded continuous functions on M. When these function spaces are given the sup norms, T becomes a bounded operator. We will verify that T satisfies the hypothesis of the Approximation Lemma with $m = 1/3$ and $r = 2/3$ by repeating the first step of the standard proof of the Tietze Extension Theorem [**1**, Theorem 3.2, p. 212]. Suppose that g is a continuous function on M with sup norm 1. Let $A = g^{-1}[-1, -1/3]$ and $B = g^{-1}[1/3, 1]$. By Urysohn's lemma, there is a continuous function f from X to $[-1/3, 1/3]$ with f identically equal to $-1/3$ on A and to $1/3$ on B. Then $\|f\| = 1/3$ and $\|Tf - g\| \leqslant 2/3$. This completes the proof.

Another form of the Approximation Lemma [**2**, Lemma 4.13(a), p. 94–95], which is useful in operator theory, states that if B_1 and B_2 are open balls about the origins in E and F, respectively, and if $\overline{T(B_1)} \supseteq B_2$, then $T(B_1) \supseteq B_2$. This follows easily by applying the form of the Approximation Lemma proved above to a sequence $\{r_n\}$ with limit 0.

References

1. J. R. Munkres, Topology: A First Course, Prentice-Hall, Englewood Cliffs, NJ, 1975.
2. Walter Rudin, Functional Analysis, McGraw-Hill, New York, 1973.

THE PROBABILITY OF HEADS*

JOSEPH B. KELLER
Departments of Mathematics and Mechanical Engineering, Stanford University, Stanford, CA 94305

1. Introduction. Why is the outcome of a coin toss considered to be random, even though it is uniquely determined by the laws of physics and the initial conditions? If it is random, why is there a definite probability associated with each outcome, regardless of how the coin is tossed? Finally, if there is a definite probability for each outcome, how can it be calculated? We shall try to answer these questions by analyzing the motion of a tossed coin. Then we shall extend our considerations to a wheel and other "chance" devices.

2. Mechanics of a tossed coin. Let us consider a circular coin of radius a and negligible thickness, one side of which is marked heads and the other side tails. We assume that the center of gravity of the coin is at its geometrical center, the height of which we denote $y(t)$ at time t. Then Newton's equation for the vertical motion of the center of gravity of the coin is

$$\frac{d^2y(t)}{dt^2} = -g. \tag{1}$$

Here the positive constant g is the acceleration of gravity. To supplement (1) we suppose that initially, at time $t = 0$, the center of the coin is at height a and that it has an upward velocity u. Thus we have the initial conditions

$$y(0) = a, \quad \frac{dy(0)}{dt} = u. \tag{2}$$

The differential equation (1) and the initial conditions (2) determine $y(t)$. Instead of $y(0) = a$ we could have prescribed any other initial value. This particular choice will simplify some of the subsequent calculations.

*Research supported by the Office of Naval Research, the Air Force Office of Scientific Research, the Army Research Office, and the National Science Foundation.

In addition to its vertical motion, the coin is assumed to be rotating about a horizontal axis that lies along a diameter of the coin. We shall choose the z-axis to be parallel to this rotation axis. Then we can describe the angular position of the coin at time t by the angle $\theta(t)$ between the positive y-axis and the normal to the side of the coin marked heads, both of which lie in the x, y plane. (See Fig. 1.) Now the equation governing the rotational motion of the coin is simply

$$\frac{d^2\theta(t)}{dt^2} = 0. \tag{3}$$

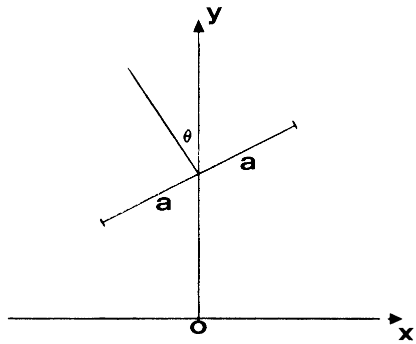

FIG. 1. The x, y plane intersects the coin along a diameter of length $2a$. The normal to the side of the coin marked heads makes the angle θ with the positive y-axis.

We assume that initially the coin is horizontal with side heads up, so that $\theta(0) = 0$, and that it has the positive angular velocity ω. Thus we specify the initial conditions

$$\theta(0) = 0, \quad \frac{d\theta(0)}{dt} = \omega. \tag{4}$$

Here also we could have replaced $\theta(0) = 0$ by any other initial value, but this choice will simplify our calculations.

The solutions of (1) and (2) for $y(t)$ and of (3) and (4) for $\theta(t)$ are

$$y(t) = ut - \frac{gt^2}{2} + a, \tag{5}$$

$$\theta(t) = \omega t. \tag{6}$$

These equations hold from $t = 0$ until the first time $t_0 > 0$ at which the coin lands on a surface, which we take to be the plane $y = 0$. We shall assume that whichever side of the coin is up at t_0 remains up. This will be the case if the coin lands in sand or mud, but not if it lands on a hard surface where it will bounce, roll, etc. Thus the coin will end its motion with heads up if

$$2n\pi - \frac{\pi}{2} < \theta(t_0) < 2n\pi + \frac{\pi}{2}, \quad n = 0, 1, 2, \ldots. \tag{7}$$

To find t_0, we consider the lowest point of the coin at time t, which is at $y(t) - a|\sin\theta(t)|$. Then t_0 is the smallest positive root of the equation

$$y(t_0) - a|\sin\theta(t_0)| = 0. \tag{8}$$

3. The pre-image of heads. We shall now analyze (7) and (8) to find out if the coin ends up heads for given values of the initial velocity u and the initial angular velocity ω. The set of all pairs u, ω of nonnegative values for which it ends up heads, we shall call the pre-image of heads in the u, ω plane, and we shall denote it H. First we consider the end points of the intervals in (7), which are given by $\theta(t_0) = (2n \pm \frac{1}{2})\pi$. At them (6) yields

$$\omega t_0 = \left(2n \pm \frac{1}{2}\right)\pi. \tag{9}$$

Now since $\sin\theta(t_0) = \pm 1$, (8) becomes $y(t_0) - a = 0$. When (5) is used in this equation it becomes

$$(10) \qquad ut_0 - gt_0^2/2 = 0.$$

The positive solution of (10) is $t_0 = 2u/g$. Then we use this result in (9) to obtain

$$(11) \qquad \omega = \left(2n \pm \frac{1}{2}\right)\frac{\pi g}{2u}, \quad n = 0,1,2\dots.$$

The relation (11) corresponds to the endpoints of the intervals in (7), and therefore it determines the boundaries of the region H in the u, ω plane. This relation is graphed in Fig. 2 for many values of n. Each curve is a hyperbola. On the axis $\omega = 0$ heads remains up throughout the toss, so this axis and the adjacent strip lie in H. The next strip lies in T, the pre-image of tails, and the successive strips alternate between H and T, as we can see by examining (7).

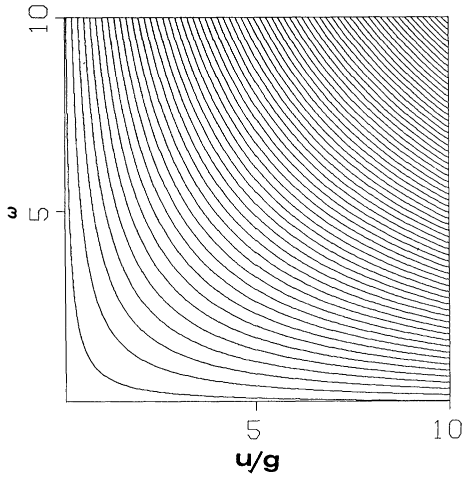

FIG. 2. The curves which separate the sets H and T, the preimages of heads and tails in the u, ω plane of initial conditions, are shown for various values of n. These are based upon (11), with the abscissa being u/g. The lowest strip, adjacent to the axes, belongs to H, the next to T and so on alternately.

From (11) we find that the vertical separation between any two adjacent boundary curves is $\pi g/2u$, except that the lowest one is only $\pi g/4u$ above the axis. Thus the strips are of equal vertical width, and this width tends to zero as u increases.

In order to find where actual tosses lie on this figure, we can consider the maximum height h to which a coin rises. From (5) we find that $dy/dt = 0$ at $t_m = u/g$, and that

$$h = y(t_m) = u^2/2g + a.$$

Thus $u = [2g(h-a)]^{1/2}$. Now $g = 32$ feet per second^2, so if $h - a = 1$ foot we find $u = 8$ ft/sec and $t = u/g = 1/4$ sec.

To find ω Professor Persi Diaconis observed a number of typical tosses of a coin under stroboscopic illumination, and found that $\omega \approx 38$ revolutions/sec $= 38(2\pi)$ radians/sec. The number of revolutions per toss is

$$n = \omega t_0 = 2u\omega/g \approx \frac{2}{4}(38) \text{ revs/toss} = 19 \text{ revs/toss}.$$

Thus the point $(u/g, \omega) = ((1/4)\text{ sec}, 76\ \pi/\text{sec})$ is way above the region shown in Fig. 2. It lies near the lines corresponding to $n = 19$ in (9).

4. The probability of heads. So far we have determined the sets H and T in the u, ω plane, which are the pre-images of heads and tails respectively. Now we suppose that the initial condition

u, ω is a random variable with a continuous probability density $p(u, \omega)$ with support in the region $u > 0$, $\omega > 0$. Then the probability of heads P_H is given by

$$P_H = \iint_H p(u, \omega)\, du\, d\omega. \tag{12}$$

Thus the outcome is random because we have assumed that the initial conditions are random. This is the answer to the first question we asked in the introduction.

For any value P_H in the interval zero to one, there are densities p for which the integral in (12) has that value. Thus (12) does not seem to place any restriction on P_H. However, we shall now show that when the support of p is shifted to sufficiently large values of u, and possibly to large values of ω also, then P_H tends to a fixed limit which is independent of the density $p(u, \omega)$. This is the content of the following theorem.

THEOREM 1. *Let $p(u, \omega)$ be a continuous probability density with support in the region $u > 0$, $\omega > 0$, and let β be a fixed constant satisfying $0 \leqslant \beta \leqslant \pi/2$. Then*

$$\lim_{U \to \infty} P_H = \lim_{U \to \infty} \iint_H p\left(u - U\cos\beta, \omega - a^{-1}U\sin\beta\right) d\omega\, du = \frac{1}{2}. \tag{13}$$

Proof. The proof is given in the appendix, although the conclusion is evident from Fig. 2.

The significance of this theorem is that there is a unique probability of heads which is approximately achieved by any continuous probability density of the initial values u, ω that is shifted to sufficiently large values of u and ω. The approximation improves as the density $p(u, \omega)$ is translated to larger values of u and ω.

The limiting value of P_H is $1/2$, despite the fact that the initial condition is not symmetric in heads and tails, since the coin always starts out with heads up. Therefore the traditional method of calculating P_H, based upon symmetry, is not applicable.

The reason why P_H has a limit as U increases is that the pre-images H and T both consist of many strips which become very narrow at infinity. Thus both H and T occupy fixed fractions of the area of any disk which is shifted to infinity. This answers the second question in the introduction. By calculating those fractions we get the limiting values of P_H and P_T, which answers the third question in the introduction.

5. Wheels. Another common gambling device, which is often used in carnivals, is a rotating wheel with numbers marked along its outer edge, and a pointer. The wheel is spun and ultimately comes to rest, with the number indicated by the pointer being the winning number. Usually there are nails between adjacent numbers which are hit by the pointer, to aid in slowing down and stopping the wheel, and to make clear which number is indicated by the pointer when the wheel stops. We shall ignore the nails, and analyze the motion of the wheel.

The angular position of the wheel can be described by specifying the angular distance θ from the pointer to some mark on the wheel. Let $\theta(t)$ be this angle at time t. To determine $\theta(t)$, we assume that there is a constant frictional torque retarding the motion, and that this frictional torque vanishes when the wheel comes to rest. Then while the wheel is turning in the counterclockwise direction, so that $\theta(t)$ is increasing with t, its equation of motion is

$$\frac{d^2\theta(t)}{dt^2} = -\alpha, \quad 0 \leqslant t < t_0. \tag{14}$$

Here α is a positive constant equal to the retarding torque divided by the moment of inertia of the wheel, and t_0 is the first time at which the wheel comes to rest. Thus

$$\frac{d\theta(t_0)}{dt} = 0. \tag{15}$$

We assume that initially the mark is under the pointer, so that $\theta(0) = 0$, and that the wheel has an initial angular velocity $\omega > 0$. Thus

$$\theta(0) = 0, \quad \frac{d\theta(0)}{dt} = \omega. \tag{16}$$

The solution of (14) and (16) is

$$\theta(t) = \omega t - \alpha t^2/2, \quad 0 \leqslant t \leqslant t_0. \tag{17}$$

To find t_0 we use (17) in (15) to get $\omega - \alpha t_0 = 0$. Thus $t_0 = \omega/\alpha$ and (17) yields

$$\theta(t_0) = \frac{\omega^2}{2\alpha}. \tag{18}$$

Now (18) gives the final position of the mark, since the wheel stops moving at t_0.

Let us now solve (18) for ω to find which points on the positive ω-axis correspond to a given final position θ. To do so we set $\theta(t_0) = \theta + 2n\pi$ in (18) with $n = 0,1,2\ldots$, since all such coordinates represent the same position of the wheel. Then the solution of (18) yields

$$\omega_n = \left[2\alpha(\theta + 2n\pi)\right]^{1/2}, \quad n = 0,1,2,\ldots. \tag{19}$$

This set is the pre-image of θ on the positive ω-axis, which is the set of initial conditions. From (19) we find that for large n the spacing between successive values of ω_n is

$$\omega_{n+1} - \omega_n = \frac{2\alpha\pi}{\omega_n}\left[1 + O\left(\omega_n^{-2}\right)\right]. \tag{20}$$

Thus the ω_n become ever more closely spaced as n increases.

Now we suppose that ω is random with a continuous probability density $p(\omega)$ which vanishes for $\omega < 0$. Then the corresponding probability density $P(\theta)$ of θ is

$$P(\theta) = \sum_{n=0}^{\infty} p[\omega_n(\theta)]\frac{d\omega_n(\theta)}{d\theta} = \alpha\sum_{n=0}^{\infty} p[\omega_n(\theta)]\Big/\omega_n(\theta). \tag{21}$$

When $p(\omega)$ is shifted to larger values of ω by the amount Ω, (21) yields

$$P_\Omega(\theta) = \alpha\sum_{n=0}^{\infty} p[\omega_n(\theta) - \Omega]\Big/\omega_n(\theta). \tag{22}$$

In the appendix we obtain from (22) the following result:

THEOREM 2. *If $p(\omega)$ is a continuous probability density which vanishes for $\omega < 0$, then*

$$\lim_{\Omega\to\infty} P_\Omega(\theta) = \frac{1}{2\pi}. \tag{23}$$

Thus in this case also there is a unique probability distribution of the outcome that results approximately from any initial density $p(\omega)$ which has its support shifted to sufficiently large values of ω. Again the limit distribution is what would be given by a symmetry argument, although the initial condition is not symmetric because $\theta(0) = 0$.

6. Other chance devices. We shall now consider any mechanical device used in a game of chance. We assume that its motion is determined by the laws of mechanics and its initial condition, which we denote u. We also assume that its possible final states can be partitioned into a finite collection of subsets $S_1, S_2, \ldots S_n$, each of which we identify with one outcome or event also denoted $S_1, S_2, \ldots S_n$. The laws of mechanics determine a unique final state corresponding to each initial condition u. All those initial conditions which lead to a final state in the set S_i we call the pre-image of the outcome S_i. We denote the pre-image of S_i by H_i.

Now we suppose that there is a probability density $p(u)$ defined on the space of initial conditions. Then the probability P_i of outcome S_i is given by

$$P_t = \int_{H_t} p(u)\, du. \tag{24}$$

The question we consider is "When is P_t independent of the particular probability density $p(u)$; and if it is independent, what is the value of P_t?"

The preceding examples of coin tossing and wheel spinning suggest an answer. It is to consider the behavior of the H_t at infinity in the space of initial conditions. Suppose that a fixed fraction f_t of the volume of any small region at infinity is contained in H_t. Then it follows from (24) that P_t will have the limit f_t as $p(u)$ is shifted to infinity.

The meaning of this conclusion is that when the H_t have the required property, "any" probability density of initial conditions concentrated at very large initial conditions will yield nearly the same probabilities $P_t = f_t$ of the various outcomes. Whether or not the H_t do have the requisite property, and what the fractions f_t are, can be decided in principle by analyzing the mechanical behavior of the device.

7. Some related work. Poincare [**1**] treated a model of the roulette wheel which led to results like those in Section 5, although it does not involve any mechanics. His results and some related ones are mentioned in Feller [**2**]. Smoluchowski [3] presented similar ideas. Then Hopf [**4**], [**5**] introduced concepts from mechanics into the calculation of probabilities, and a point of view just like the one we have described. Recently Thorp [**6**] has used mechanical considerations in analyzing roulette.

Acknowledgement. It is a pleasure to thank Professor Persi Diaconis for raising a question about the face probabilities of noncubical dice which led me to think about this subject, and for permitting me to quote his measurements of the rotation rates of tossed coins.

Appendix. We shall now indicate the proofs of the two theorems.

Proof of Theorem 1. We first write the integral in (13) as an iterated integral, and use (11) to determine the range of $\omega' = \omega - a^{-1}U\sin\beta$:

$$P_H = \int_{U\cos\beta}^{\infty} \sum_{n=0}^{\infty} \int_{(2n-1/2)\pi g/2u - a^{-1}U\sin\beta}^{(2n+1/2)\pi g/2u - a^{-1}U\sin\beta} p(u - U\cos\beta, \omega')\, d\omega'\, du. \tag{A1}$$

The lower limit of u is $U\cos\beta$ because $p = 0$ for $u - U\cos\beta < 0$. As U tends to infinity, so does u provided that $\beta < \pi/2$. Then the range of each integral over ω' is of length $O(U^{-1})$. Therefore we can approximate each of these integrals by the length of the interval multiplied by the value of the integrand at the midpoint. The error in this approximation is $O(U^{-1})$. Thus we can rewrite (A1) as

$$P_H = \int_{U\cos\beta}^{\infty} \sum_{n=0}^{\infty} p\left(u - U\cos\beta, \frac{2n\pi g}{2u}\right) \frac{\pi g}{2u}[1 + o(1)]\, du. \tag{A2}$$

As U becomes infinite, the sum in (A2) converges to one half the Riemann integral of p with respect to ω'. This is clear if p has compact support, and if not it follows by approximating p by a sequence of functions with compact support. Thus we have

$$P_H = \int_{U\cos\beta}^{\infty} \frac{1}{2} \int_0^{\infty} p(u - U\cos\beta, \omega')\, d\omega'\, du[1 + o(1)]. \tag{A3}$$

Upon setting $u' = u - U\cos\beta$ in (A3), and remembering that the integral of p is one, we get

$$P_H = \frac{1}{2}\int_0^{\infty}\int_0^{\infty} p(u', \omega')\, d\omega'\, du'[1 + o(1)] = \frac{1}{2}[1 + o(1)]. \tag{A4}$$

When U tends to infinity, $o(1)$ tends to zero, and (A4) yields the result in (13). For $\beta = \pi/2$ we introduce $u' = u - U\cos\beta$ instead of ω' and then the proof is similar to that above.

Proof of Theorem 2. By using (20) for ω_n^{-1} we can write (22) as

$$P_\Omega(\theta) = \frac{1}{2\pi} \sum_{n=0}^{\infty} p[\omega_n(\theta) - \Omega][\omega_{n+1}(\theta) - \omega_n(\theta)][1 + O(\omega_n^{-2})]. \tag{A5}$$

Then we write $w_n(\theta) = \omega_n(\theta) - \Omega$, and note that $p(\omega_n - \Omega) = 0$ for $\omega_n < \Omega$. Thus we can write (A5) as

$$P_\Omega(\theta) = \frac{1}{2\pi} \sum_{n=0}^{\infty} p[w_n(\theta)][w_{n+1}(\theta) - w_n(\theta)][1 + O(\Omega^{-2})]. \tag{A6}$$

Now $w_{n+1} - w_n = \omega_{n+1} - \omega_n = O(\omega_n^{-1})$, and $\omega_n > \Omega$ whenever $p(\omega_n - \Omega) > 0$. Therefore $w_{n+1} - w_n = O(\Omega^{-1})$, and the sum in (A6) converges to the Riemann integral of p as Ω becomes infinite. Since this integral is one, (A6) yields (23).

References

1. H. Poincaré, Calcul des Probabilités, Gauthier-Villars, Paris, 1912.
2. W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, 2nd ed., Wiley, New York, 1971.
3. M. v. Smoluchowski, Naturwissenschaften, 6 (1918) 253.
4. E. Hopf, On causality, statistics and probability, J. Math. Phys., 13 (1934) 51–102.
5. ______, Remarks on causality and probability, J. Math. Phys., 14 (1935) 4–9.
6. E. O. Thorp, Physical prediction of roulette, (unpublished).

ANCESTORS, CARDINALS, AND REPRESENTATIVES

T. D. PARSONS
Department of Mathematics, California State University, Chico, CA 95929-0525

1. Introduction. Occasionally we come across a proof so charming that it is as memorable or more so than the theorem it proves. One such is the "parent-ancestor-descendant" proof of the Schroeder–Bernstein Theorem which appeared in Birkhoff and Mac Lane [**2**] and has ever since been a favorite of college mathematics audiences. This proof is a clever restatement of an argument of S. Banach [**1**], and it demonstrates the value of picturesque terminology in good mathematical exposition. In our expository note here, we apply the same proof to obtain a more general theorem (Theorem 1 below) and show how to derive as its corollaries not only Banach's mapping theorem and the Schroeder-Bernstein Theorem, but also some results of Mendelsohn and Dulmage [**6**], R. Rado [**7**, Theorem 4.3], and Hoffman and Kuhn [**5**] about systems of distinct representatives for families of sets.

Theorem 1 is due to Ore [**8**] and independently to Perfect and Pym [**9**], whose version we state first.

THEOREM 1. *Let $X' \subseteq X, Y' \subseteq Y$, and $E \subseteq X \times Y$ be sets. Suppose that there exist injective mappings $f: X' \to Y$ and $g: Y' \to X$ such that $(x, f(x)) \in E$ and $(g(y), y) \in E$ for all $x \in X'$ and $y \in Y'$. Then there exist sets X_0, Y_0 with $X' \subseteq X_0 \subseteq X$ and $Y' \subseteq Y_0 \subseteq Y$ and a bijection $h: X_0 \to Y_0$ such that $(x, h(x)) \in E$ for all $x \in X_0$.*

Ore's version, which is more convenient for certain applications, is stated (Theorem 1′) in the language of bipartite graphs. Here we consider X, Y to be disjoint sets of vertices forming the bipartition for the bipartite graph B with edge-set E—that is, vertex x is adjacent to vertex y if $(x, y) \in E$. A subset $S \subseteq E$ of edges is said to be *incident* to a subset X' of X (respectively, Y' of Y) if every element of X' (Y') is an end-point of some edge (x, y) of S. A set $M \subseteq E$ is *independent* if no two edges of M share a common endpoint (i.e., M is a "matching" in B).

THEOREM 1′. *Let $B = [X, Y]$ be a bipartite graph with edge-set $E \subseteq X \times Y$. Suppose that*