The Truth about Models: How Well Do Mechanical Models Mimic the Observed Gender Distributions in Two-Child Families?

WILLIAM D. STANSFIELD, MATTHEW A. CARLTON

ABSTRACT

We question the use of mechanical models, such as coin flipping, to represent the probabilities of gender distributions in sibship families consisting of two children. Both the assumptions of the models and the reliability of the data should be evaluated. Using models without these critical evaluations may tend to perpetuate myths rather than elucidate biological realities.

Key Words: *Binomial distribution; birth order; chi-square test; conditional probability; data quality; Lexis variation; model assumptions; Poisson variation.*

Biology and genetics textbooks sometimes use the mechanical model of coin flipping to represent the process whereby the gender (sex) of babies is determined. Can this process be simulated by any other mechanical model, and, if so, which model best reflects biological reality? What deeper understandings about scientific models, in general, can be transmitted to our students by investigation of this subject?

The mechanical models presented … offer hands-on experience that may help students better understand probability theory.

Coin-Flipping Model

Honeycutt & Pierce (2007) recommend "that students collect at least some data using the manipulative activities (e.g., coin tossing, rolling dice, pulling beans from bag), and not rely entirely on simulations for data generation." They supply URLs for electronic tools that illustrate probability using these mechanical models.

In the coin-tossing model, we assume that the coin is well balanced (not biased). Let heads represent boys (B) and tails represents girls (G). In a simple model, the probabilities of B and G are equally frequent in a population at birth (Pr[B] = Pr[G] = 0.5), and the probability distribution of the second child's gender in the same family is identical to, and independent of, the first child's gender. To simulate families consisting of two children, we toss the coin twice. Four gender combinations (birth orders) are possible (BB, BG, GB, GG). Over an infinite number of such trials, each combination is expected to be equally frequent (1/4 each); equivalently, the number of boys (or girls) follows a binomial distribution with n = 2 and p = 0.5.

Dice-Rolling Model

Throwing a die twice or a single roll of two dice can also simulate a family of two children. Let even-numbered sides (2, 4, 6) represent boys and odd-numbered sides (1, 3, 5) represent girls, so again Pr(B) = Pr(G) = 0.5. The outcomes would resemble the coin-tossing model, with equally frequent probabilities of 1/4 each for BB, BG, GB, GG.

Card-Game Model

Marilyn vos Savant authors a weekly column (titled "Ask Marilyn") in *Parade* magazine. One of her readers recently posed the following problem:

During a card game, I said that the probability of getting dealt two aces is the same as getting an ace and a deuce – that the chance of getting any two cards is the same, whatever they are. My friends say I'm wrong. Who's right?

Vos Savant responded as follows. "Your friends are right. To illustrate, let's narrow the question to the red or black half of the deck with only two aces and deuces. Lay the four cards in a grid – aces on the left and deuces on the right, like this:

    ace   deuce
    ace   deuce

To get two aces, you must be dealt the left column. (And to get two deuces, you must be dealt the right column.) But to get one ace and one deuce, you can be dealt either the top row or the bottom row. So getting one of each is more likely" (vos Savant, 2010).

Some readers may assume from vos Savant's presentation that being dealt "one of each" is twice as likely as being dealt two aces. In fact, the probability of drawing two aces from a standard 52-card deck is $Pr(AA) = (4/52)(3/51) = 12/2652$; the fraction 3/51 is the conditional chance of the second card also being an ace. By contrast,

the chance of drawing an ace and a deuce – in either order – is Pr(A2 or 2A) = (4/52)(4/51) + (4/52)(4/51) = 32/2652. The ratio of the probabilities is 32/12 ≈ 2.66, so the chance of drawing an ace and a deuce is 2.66 times greater than the chance of drawing two aces.

Others may misinterpret the question asked of vos Savant, thinking that, if an ace is the first card dealt, what is the probability that the second card is an ace (or a deuce)? In this case, the conditional probability that the second card is an ace = 3/51; for the deuce it is 4/51. The ratio of the probabilities is 4/3 ≈ 1.33. Thus, a deuce on the second card is about one and a third times more likely than an ace on the second draw, given that the first draw was an ace.

We often hear the metaphor that the meiotic mechanism "shuffles" the genetic deck of cards and deals out a new hand with the production of every egg or sperm. We have never seen a deck of cards used to demonstrate this principle as a teaching aid, but cards can be used to illustrate some aspects of biological reality. Can a deck of cards be used to represent the gender distributions in two-child families? For example, let the 26 black cards represent boys, and the 26 red cards represent girls. The probability that two children are both boys under this mechanical model is (26/52)(25/51) = 0.2451. Reshuffle the deck of 52 cards after each two-card trial (two-child family) has been dealt. The probability of the first child being a boy and the second a girl is (26/52)(26/51) = 0.2549. The same probability exists for the first child being a girl and the second a boy; combined = 2(0.2549) = 0.5098. The ratio of Pr(BG or GB) to Pr(BB) is 0.5098/0.2451 = 2.079. Thus, one of each gender is expected to be about 2.08 times more frequent than two boys. For comparison, under the coin-flip model the chance that a two-child family includes one child of each sex is 2 times the chance that the children are two boys.

○ Discussion

Models are tentative schemes or structures that correspond to real objects, events, or classes of events, and that have explanatory power. Models help scientists and engineers understand how things work. Models take many forms, including physical objects, plans, mental constructs, mathematical equations, and computer simulations. Scientific explanations incorporate existing scientific knowledge and new evidence from observations, experiments, or models into internally consistent, logical statements. Different terms, such as "hypothesis," "model," "law," "principle," "theory," and "paradigm" are used to describe various types of scientific explanations. As students develop and as they understand more science concepts and processes, their explanations should become more sophisticated. That is, their scientific explanations should more frequently include a rich scientific knowledge base, evidence of logic, higher levels of analysis, greater tolerance of criticism and uncertainty, and a clearer demonstration of the relationship between logic, evidence, and current knowledge. (National Research Council, 1996: p. 117)

In an article titled "Making babies by the flip of a coin" (Carlton & Stansfield, 2005), we provided statistical evidence that the probability of the gender of the second child in a two-child family is not independent of the probability of the gender of the first child. The coin-tossing experiment draws upon an unlimited source of independent events (birth of each child), and in this regard the simulation resembles the theoretically unlimited source of producing two-child families. However, the probability of the second child's gender in a real family is not independent of the gender of the first child. This is one "Achilles heel" of the coin-tossing model. Another problem is that traditional "unbiased" coin-tossing is only good for modeling populations in which boys and girls are equally frequent. In the U.S. white population, the gender ratio at birth is approximately 105 boys for every 100 girls. The reasons for this discrepancy, which is even greater at earlier periods of gestation, are not well established. The sex ratio at birth also varies between different populations.

In the card game, the probability of the second child being a boy is different from the probability of the first child being a boy, and in this respect the card game more accurately reflects the reality of biological families. Also, the number of black or red cards can be roughly adjusted to resemble other sex ratios at birth. This principle cannot be demonstrated with balanced coin-tosses. For example, if we remove one red card from a standard deck of 52 cards, we are left with a black/red (boy/girl) ratio of 26/25 = 1.04. This is the minimal boy-biased B/G ratio possible with a deck of 52 cards; boy-biased B/G ratios between 1.04 and 1.00 cannot be simulated with only 52 cards but can be produced by adding more cards to a standard deck. For example, a deck containing 102 black and 100 red cards simulates a population with boy/girl ratio = 1.02.

At this point, we might just as well replace cards with a jar of black and white beans whose bean numbers can be easily adjusted as needed. For example, to simulate a population with boy/girl ratio = 105/100, we could place 105 black and 100 white beans in the jar. The probability that a black bean will be the first draw is 105/205; the conditional probability of black on the second draw is similar (104/204), but not identical, to the first draw. Although the probability of the second bean being black is different and dependent on the color of the first bean, that probability would be very unlikely to reflect the biological population the mechanical model is intended to represent (as suggested by a mathematical formula akin to that of Malinvaud, 1955). Edmond Malinvaud studied almost 4 million births in France for years 1946–1950 and concluded that, if boys represent 51.45% of first-born children in a population, the probability estimate p (measured as a percentage) of a pregnancy producing a boy in subsequent pregnancies is fairly well fitted by the linear relationship p = 51.45 + 0.3b – 0.5f, where b is the number of preexisting boys and f is the number of preexisting girls in a sibship. Thus, if the first child is a boy, then the probability of the second being a boy is 51.45 + 0.3(1) – 0.5(0) = 51.75%. Similarly, if the first child is a girl, the probability that the second child will be a boy is 50.95%. Malinvaud's formula predicts that the third child in a family with two preexisting boys has probability p = 51.45 + 0.3(2) – 0.5(0) = 52.05% of also being a boy. The card or bean model predicts that the number of male births decline as family size increases, whereas Malinvaud's model suggests the opposite.

In a 2007 article, we provided statistical support (based partly on a chi-square test of data assumed to conform to a binomial distribution) for the hypothesis that parental choice (family planning) seems likely to be responsible for more same-sex sibships than unlike-sex sibships in families of two. In our 2009 article, we

reported that the probability of a male birth declines with birth order within individual sibships (Poisson variation) and that the probability of a male birth is also affected by between-sibship variation (Lexis variation). The evidence for both types of variation is overwhelming in regard to mammalian, including human, sex ratio at birth. Because a mixture of Lexis and Poisson variation may mimic a binomial distribution (James 2000), it is invalid to infer from a seemingly binomial distribution that the probability of a male birth is equal at all trials. "In particular, the assumption that all couples have the same probability of male births (homogeneity) is invalid" (Stansfield & Carlton, 2009). To quote the National Research Council (1996: p. 23) again:

> Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires *identification of assumptions*, use of critical and logical thinking, and consideration of *alternative explanations*. Students will engage in selected aspects of inquiry as they learn the scientific way of knowing the natural world, but they also should develop the capacity to conduct complete inquiries. (Italics added)

We should question not only the assumptions of our models but also the reliability or quality of the data. In our previous three papers, the data on gender distributions in sibships of two children came from the National Health Interview Survey (NHIS) for years 1987–1993 and 1998–2002. We would have preferred data on the B/G ratio at birth, but the youngest cohort available from NHIS reports biological families 10 years of age or younger. Also, the NHIS does not obtain information on children who are not living in the household at the time the survey was taken. In addition, no information was provided on the numbers of identical twins. Monozygotic (MZ, "identical") twins represent a single fertilized egg and should not be included twice in the numbers of like-sex children. Otherwise, the numbers of like-sex families tend to be inflated.

During the years 1922–1936 in the United States, 1.129% of all white births were twins in which at least one twin was born alive. On average, about 34.2% of these twin births were identicals. Thus, (0.342)(0.0129) = 0.004412, or about 0.44% of all births in this population were estimated to be MZ. A young white mother under 20 years of age had about the same chance of having either an identical (MZ) or a nonidentical (DZ) set of twins. White mothers 35–39 years of age had about three times as many nonidentical as identical twins. Prenatal deaths of one of a pair of twins have been estimated to be as high as 20–50%, with identical twins dying 2–3 times more frequently than nonidentical twins (Stern, 1960: pp. 532–535). Some genetics textbooks offer a simple statistical model, based on W. Weinberg (1901), for estimating the numbers of MZ and DZ twins in a population:

1. Number of DZ twins = (known number of unlike-sex twins, BG and GB) + (number of DZ like-sex twins, BB and GG)
2. *Assume number of DZ like-sex twins = number of DZ unlike-sex twins, so number of DZ twins = 2(number of unlike-sex twins)
3. Number of MZ twins = (known total number of all twins) – (number of DZ twins)

This model (*) assumes that the gender probability of the second child is independent of the gender probability of the first child. It also assumes that, within a family, the probability of MZ or DZ twins does not change over the life of the mother. Neither of these assumptions exists in reality for the population cited by Stern. Nevertheless, in twinning data from Finland and Sweden, Finnish researchers compared their results with findings in the literature. "In conclusion, our findings indicate that Weinberg's differential rule is rather robust and that despite its simplicity, it gives reliable results when official birth registers are analyzed" (Fellman & Eriksson, 2006).

○ Conclusions

It appears that sexes of human births in two-child families do not follow a binomial statistical model with Pr(B) = Pr(G) = 0.5 or with any other probability parameter. Should we stop using this two-child family example and these model assumptions to teach probability and independence? Not necessarily. We use genetic examples in probability for their pedagogical merits, not because the binomial model exactly reflects biological reality. Nevertheless, the mechanical models presented in this article offer hands-on experience that may help students better understand probability theory. In an *ABT* editorial, William Leonard (2010) stressed the "need to use more mathematics in biology" because "our society is far too nonquantitative in general, and this only leads to misconceptions in many areas." As George Box aptly put it, "All models are wrong; some models are useful." If we must use models, it would seem irresponsible not to explain to our students the assumptions and defects of these models, and at least acknowledge the existence of any other known competitive models. Why is it important that this kind of information be transmitted to our students? If we do not question the assumptions of our models and the reliability of our data, we may be perpetuating myths rather than elucidating biological realities.

○ Teaching Aids

http://statweb.calpoly.edu/bchance/applets/CoinTossing/CoinToss.html
Binomial distributions: fair and biased coins, probability calculations, approach to normal distribution as number of trials increases.

http://www.jstor.org/pss/4448928
McKean, H.R. & Gibson, L.S. (1989). How-to-do-it: hands-on activities that relate Mendelian genetics to cell division. *American Biology Teacher*, *51*, 294–299.

http://www.hhmi.org/biointeractive/gender/pdf/shuffling_deck.pdf
Stapleton, B., Avant, R. & Avant, P. Shuffling the deck – the card game of life.

References

Carlton, M.A. & Stansfield W.D. (2005). Making babies by the flip of a coin? *American Statistician*, *59*, 180–182.
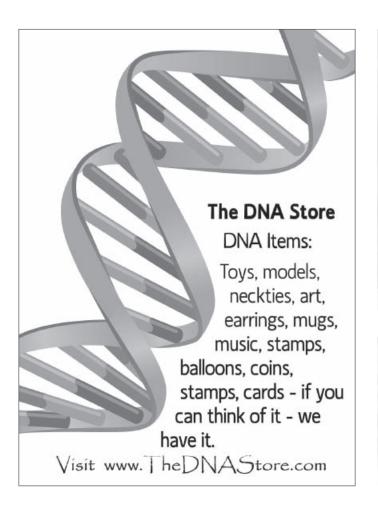
Fellman, J. & Eriksson, A.W. (2006). Weinberg's differential rule reconsidered. *Human Biology*, *78*, 253–275.

Honeycut, B.B. & Pierce, B.A. (2007). Illustrating probability in genetics with hands-on learning: making the math real. *American Biology Teacher*, *69*, 544–551.

James, W.H. (2000). The variation of the probability of a son within and across couples. *Human Reproduction*, *15*, 1184–1188.

Leonard, W. (2010). Encouraging new manuscripts in critical area of biological education. *American Biology Teacher*, *72*, 6.

Malinvaud, E. (1955). Relations entre la composition des familles et le taux de masculinité. *Journal of Social Statistics (Paris)*, *96*, 49–64.

National Research Council. (1996). *National Science Education Standards*. Washington, D.C.: National Academy Press.

Stansfield, W.D. & Carlton, M.A. (2007). Human sex ratios and sex distribution in sibships of size 2. *Human Biology*, *79*, 255–260.

Stansfield, W.D. & Carlton, M.A. (2009). The most widely publicized gender problem in human genetics. *Human Biology*, *81*, 3–11.

Stern, C. (1960). *Principles of Human Genetics, 2nd Ed.* San Francisco, CA: W.H. Freeman.

vos Savant, M. (2010). Ask Marilyn. *Parade*, 10 January, p. 19.

Weinberg, W. (1901). Beiträge zur physiologie und pathologie der mehrlingsgeburten beim menschen. *Archiv für Gesellshaft Physiologie*, *88*, 346–430.

WILLIAM D. STANSFIELD (wstansfi@calpoly.edu) is Emeritus Professor of Biological Sciences and MATTHEW A. CARLTON (mcarlton@calpoly.edu) is Professor of Statistics at California Polytechnic State University, San Luis Obispo, CA 93407.