Making Babies by the Flip of a Coin?

Matthew A. CARLTON and William D. STANSFIELD

Many probability and genetics textbooks pose standard questions about eye color, birth defects, sexes of children, and so on. Solutions to these questions, specifically about sexes, generally make two assumptions: first, that a randomly selected embryo is equally likely to be male or female; second, that the sexes of successive children from the same parents are independent. In other words, probabilists (and some geneticists) treat sexes of children like flips of a fair coin: two possible outcomes, each equally likely, with outcomes independent from trial to trial. But are these assumptions realistic? Demographic data suggest that neither a balance of sexes nor true independence exist in nature. Yet most textbooks, both in genetics and probability theory, continue to use the binomial distribution as an acceptable approximation for solving genetics problems involving live-birth sex ratios in species where sex is determined by an XX versus XY chromosome mechanism. We look at a widely circulated article in *Parade* magazine regarding the gender distribution in human families with *two* children and analyze comparable data from federal sources to show that such families do not conform to any binomial distribution. The sequence of investigations we take here could be followed in an introductory or intermediate probability and statistics course.

KEY WORDS: Binomial model; Chi-square test; Genetics; Gender; Probability.

1. THE ORIGINAL PROBLEM AND THE MATHEMATICIAN'S SOLUTION

Popular magazine columnist Marilyn vos Savant (1997) received the following problem from one of her readers. "A woman and a man (unrelated) each have two children. At least one of the woman's children is a boy, and the man's older child is a boy. Do the chances that the woman has two boys equal the chances that the man has two boys?" Vos Savant replied that the chances the woman has two boys are about 1 in 3, while the chances the man has two boys are about 1 in 2. To substantiate her answer, which many readers found counter-intuitive, she conducted a volunteer survey (by E-mail and letters) from respondents with exactly two children, *at least one of which is a boy*, with the following results. In a total of 17,946 responses, 35.9% (about 1 in 3) reported two boys.

In what follows, we will let B and G denote the events a child is a boy or girl, respectively. Also, we will let BG denote the event a two-child family has an older boy and a younger girl (similarly for GB, BB, and GG).

The question posed to vos Savant is a famous one in conditional probability. The solution hinges on the assumption a parent with exactly two children is equally likely to have any of the four possible gender sequences BB, BG, GB, or GG. Suppose this were indeed the case. Given that at least one of the woman's children is a boy, three of the original gender sequences are possible (all but GG); because we are assuming these are equally likely, the conditional probability of BB is now 1 in 3. In contrast, knowledge that the man's *first* child is a boy leaves only BB and BG as potential gender sequences for his children, and thus the conditional probability of BB equals 1/2.

Hence, from the mathematical perspective (which might or might not have been the poser's intent), vos Savant's solution is correct. But what underlies the assumption that all four gender sequences are equally likely? In the language of probability, this assumption asserts that (1) a live birth is equally likely to be a boy or a girl, that is, $\Pr(\mathrm{B}) = \Pr(\mathrm{G}) = 1/2$, and (2) the gender of a second child is independent of the gender of the first child. Conditions (1) and (2) can be stated equivalently by saying the number of boys (or girls) in a two-child family is a binomial random variable with probability parameter 1/2.

Of these two conditions, (2) is more fundamental, and we will address this later. Suppose we grant the independence assumption for now, but we generalize to an arbitrary boy-girl ratio. Let $\Pr(\mathrm{B}) = b$ and $\Pr(\mathrm{G}) = g$, with $g = 1 - b$. Then the probabilities of BB, BG, GB, GG are bb, bg, gb, and gg, respectively. The probability the woman has two boys, given she has at least one boy, now equals

$$\frac{\Pr(\mathrm{BB})}{\Pr(\mathrm{BB,\ BG,\ GB})} = \frac{bb}{bb + bg + gb} = \frac{b^2}{b^2 + 2bg} = \frac{b}{b + 2g}.$$

Similarly, the probability the man has two boys, given that his *older* child is a boy, equals $b/(b + g)$. Notice that when the boy-girl ratio is 1:1, that is, $b = g = .5$, these reduce to the original 1/3 and 1/2, respectively. We can also rephrase these probabilities in terms of the *boy-girl ratio*: if we define $r = b/g$, the two aforementioned probabilities are $r/(r+2)$ and $r/(r+1)$, respectively.

Matthew A. Carlton is Assistant Professor, Statistics Department, Cal Poly State University, San Luis Obispo, CA 93407 (E-mail: mcarlton@calpoly.edu). William D. Stansfield is Professor Emeritus, Biology Department, Cal Poly State University, San Luis Obispo, CA 93407 (E-mail: wstansfi@calpoly.edu). The authors thank Debra Blackwell at the National Center for Health Statistics for providing data and Ulric Lund and the referees for providing useful comments.

Table 1. All Biological Children 10 and Younger Years of Age in Two-Child Families. Data from 1998–2002 NHIS, courtesy of Debra Blackwell, NCHS.

Gender sequence	*Number of families*	*Percent of sample*
girl-girl (GG)	5,844	22.95%
boy-girl, boy older (BG)	6,628	26.02%
girl-boy, girl older (GB)	6,451	25.33%
boy-boy (BB)	6,545	25.70%
Total	25,468	100.00%

We mentioned before that vos Savant's survey found 35.9% of women with at least one boy (in families of exactly two children) had two boys, which is close to 1 in 3. She commented, "Given that there are about 106 boys born for every 100 girls, the actual percentage in the population would be closer to 33.9%." This is, in fact, incorrect. Presumably, vos Savant simply multiplied 35.9% by the ratio 100/106, since this is indeed 33.9% to one decimal place. The correct "adjustment," still operating under the assumption of gender independence, would be to substitute $r = 1.06$ into the first fraction above, yielding $1.06/3.06 = 34.6\%$.

In any case, the voluntary survey at least validates vos Savant's correct assertion that the "chances" posed in the original question, though similar-sounding, are different, and that the first probability is certainly nearer to 1 in 3 than to 1 in 2.

2. FEDERAL DATA ON TWO-CHILD FAMILIES

Since vos Savant's survey was purely voluntary, it would obviously be incorrect to use her data for inferential purposes. We turn instead to data from the National Health Interview Survey (NHIS), conducted annually by the Center for Disease Control's National Center for Health Statistics (NCHS) under contract with the U.S. Census Bureau. The NHIS employs a probability sampling plan; for details, consult their Web site http://www.cdc.gov/nchs. The data in Table 1 show the gender sequence for all biological children ages 10 and younger in two-child families from the NHIS five-year cohort 1998 to 2002. We discuss the consequences of using this particular dataset in the following.

We realize that this is not the best dataset for our purposes, because these data do not account for monozygotic twins, still births, abortions, and child fatalities up to age 10 (i.e., a two-child family in our data may have had older or intervening children who died). Ideally, we would have data from all parents who have given birth to exactly two children, but such data are not readily available. That said, we believe the aforementioned categories are relatively small; for example, monozygotic twins comprise less than .5% of all live births in the United States (Stern 1960, pp. 533–534).

3. GENDER (IM)BALANCE AND (IN)DEPENDENCE

We wish to see whether the NHIS data substantiates the 1:1 boy-girl ratio and/or the assumption of independent sequential sexes. We shall rely on standard statistical techniques for the analyses; see Samuels and Witmer (2003) or DeGroot and Schervish (2002) for reference. In what follows, we present three partly redundant tests for illustrative purposes; we caution against performing numerous, overlapping hypothesis tests on the same dataset in practice. Instructors who elect to use this data can follow our sequence or choose to show only the test(s) appropriate to their students' level.

Table 2. Observed Counts, Expected Counts (in parentheses), and Chi-Square Contributions (in italics) for a Binomial Goodness-of-Fit Test, using the NHIS Dataset

Number of boys	*0*	*1*	*2*
Observed	5,844	13,079	6,545
Expected	(6,021.32)	(12,724.35)	(6,722.32)
Chi-square	*5.222*	*9.885*	*4.677*

First, we consider the boy-girl ratio among all families with exactly two children. Most statisticians and biologists are aware that boy babies and girl babies are *not* equally likely. Curiously, the boy/girl birth ratio is greater than one; this contrasts with the well-known fact that women outnumber men overall. (According to the 2000 Census, there were 96.3 males of all ages for every 100 females in the United States.) Although it is impossible at present to determine the actual numbers of fertilized human eggs bearing XX versus XY chromosomes, data from the 2000 U.S. Census show that the number of males per 100 females under five years of age is about 104.8. Let us compare that figure to our NHIS data.

Using all 50,936 children in the NHIS dataset, we have 26,169 boys (51.38%) and 24,767 girls (48.62%). This yields an estimated boy-girl ratio of $\hat{r} = 1.0566$, or about 106 boys for every 100 girls (the same number offered by vos Savant without citation). However, these 50,936 children comprise a random sample *only if we believe sexes of second children are independent of the sexes of their older siblings*. Since we do not believe this to be true (in fact, we will refute this claim shortly), we will estimate the boy-girl ratio using first-born children only.

Of the 25,468 first-born children in the NHIS dataset, 13,173 (51.72%) are boys. A standard large-sample z-test of the hypothesis $\Pr(\mathrm{B}) = \Pr(\mathrm{G}) = 1/2$ yields a test statistic of $z = 5.50$ (p value ≈ 0). The NHIS data decisively refute the 1:1 human sex ratio, at least among first live births in two-child families. But, again, this is hardly news, since we already noted that U.S. Census data also shows a gender imbalance in American children.

Do two-child families follow *any* binomial model? To find out, we test conformity of the number of boys in a two-child family to a binomial distribution with $n = 2$ and an *arbitrary* probability parameter, b. This model allows for gender imbalance but still assumes that a single parameter characterizes the chance of a male baby for both the first and second child. Under this null hypothesis (which further hypothesizes sexes of siblings to be independent), our maximum likelihood estimate of b is $\hat{b} = 26{,}169/50{,}936 = .5138$. Table 2 shows the elements of the corresponding goodness-of-fit test.

The chi-square goodness-of-fit test statistic is a statistically huge 19.784 at 1 degree of freedom (p value ≈ 0); though there are three categories, we lose one additional degree of freedom to the estimation of b. We conclude that the number of boys in two-child families does not follow *any* binomial distribution, even one with a gender imbalance. Notice that the binomial model predicts considerably more same-sex sibships than actually appear in the NHIS data, and that GG families $(x = 0)$ contribute more to the chi-square statistic than BB families $(x = 2)$.

Table 3. Observed Cell Counts, Expected Cell Counts (in parentheses), and Chi-Square Contributions (in italics) for a Chi-Square Test of Independence of Successive Sexes, Using the NHIS Dataset.

		Second child G	*Second child* B
First child	G	5,844 (6,021) *5.204*	6,451 (6,274) *4.994*
	B	6,628 (6,451) *4.857*	6,545 (6,722) *4.662*

Finally, we test for independence of siblings' sexes without any other restrictions on the probabilities of boy and girl babies. In particular, we no longer require that the chances of the first child being a girl and the second child being a girl be equal. Table 3 displays the actual and expected cell counts for a standard chi-square test of independence. The results are a chi-square statistic of 19.717 at 1 degree of freedom (p value ≈ 0), with GG families making the largest contribution to the chi-square value. It is interesting to note that the NHIS data show 354 more opposite-sex siblings (and, concordantly, 354 fewer same-sex siblings) than expected. If anything, we would anticipate that monozygotic twins would increase the number of same-sex siblings over what is expected under an independence assumption.

4. WHAT COULD EXPLAIN THE DISCREPANCIES?

If we accept that the sample data accurately reflect the population, we must seek a biological mechanism to explain these results. Can a fetus sensitize its mother against the fetus' sex, making a same-sex sibling less likely? Recent research (e.g., Condon et al. 2004) shows that a mouse fetus can send specific messages to its mother. However, the research does not indicate that these messages affect the mother long enough to influence future fertilizations. Fetuses can induce sustained *immunological* responses over more than one pregnancy. For example, it is well known that an Rh^+ baby born to an Rh^- woman can stimulate her to produce a long-lasting antibody response that could affect the survival of any subsequent Rh^+ fetus. However, there is no known biological correlation between sex and Rh status.

The Y chromosome could, in theory, be at least partly responsible for the appearance of fewer-than-expected BB families. After all, the Y chromosome of a male embryo or fetus carries genes whose products would be foreign to its mother's immune system, and so it seems plausible that a boy fetus could sensitize his mother to make antibodies against protein products of Y-linked genes. But this would obviously not explain the scarcity of GG families, and we find no such Y-linked immunological incompatibilities reported in the literature.

On the other hand, it might be possible for an X-linked gene product in a girl fetus to stimulate her mother to make antibodies that might affect her own survival as well as that of the next girl sibling, but not of a boy sibling. Because a boy inherits his X chromosome only from his mother, a male fetus would not stimulate his mother to make antibodies against his X-linked gene products. However, if the mother lacks an X-linked gene that her female baby has received from the father, that girl fetus could sensitize the mother and cause an immune attack on any subsequent female embryos sired by the same man. The *Xg* locus (Sutton 1980) is the only example of this phenomenon we could find. Although this model immunological system might contribute to the dearth of GG sibships in families of size two, Sutton reports that mothers who form such X-linked antibodies are "exceedingly rare." And even if such a maternal-fetal incompatibility system exists, it would not explain the dearth of BB families in the NHIS data.

Of course, one obvious nonbiological explanation is that parents may be more likely to stop having children if they have had one child of each sex than if they have had two boys or two girls. This would clearly result in fewer same-sex sibships in two-child families. Unfortunately, data are not available on the reasons parents stopped having children. It must thus be acknowledged that, at present, a very plausible biological mechanism to explain the results of this study has yet to be found.

5. RECOMMENDATIONS

So, it appears that sexes of human births in two-child families do not follow a binomial model with $\Pr(\mathrm{B}) = \Pr(\mathrm{G}) = 1/2$, or with any other probability parameter. Should we stop using this example and these model assumptions to teach conditional probability and independence? Not necessarily. We use genetics examples in probability for their pedagogical merits, not because the binomial model exactly reflects biological reality. (As George Box aptly put it, "All models are wrong; some models are useful.") For an introductory probability course, the mathematical solution in Section 1 still illustrates one of the more intriguing applications of conditional probability. In fact, we can use this opportunity to discuss with our students the notion of underlying model assumptions and the consequences of violating these assumptions.

At the same time, students in a biostatistics course should find the gender imbalance and lack of independence in human sexes interesting. In particular, lack of independence among sexes of children with the same parents will surely come as a surprise to many biology students (not to mention geneticists and demographers). The simple z and chi-square hypothesis tests in Section 3, accessible to introductory biostatistics students, show these two plausible assumptions are actually not supported by this dataset.

Alternatively, one can present the data and the dilemma as an open-ended problem (in a senior-level statistics course, for example). The "dilemma" can be presented either as testing the binomial model for sexes of siblings, or as validating/refuting the probability solutions of 1/2 and 1/3 from Section 1. Students must then select for themselves which statistical test(s) to use—we have not exhausted them all here—and justify their choices. More advanced statistics students are encouraged to explore alternative models for birth sexes; for example, if you know the sex of an oldest child, what are the chances the next child will be of the same/opposite sex? And do your answers depend on the sex of the oldest child?

[Received August 2004. Revised December 2004.]

REFERENCES

Condon, J. C. et al. (2004), "Surfactant Protein Secreted by the Maturing Mouse Fetal Lung Acts as a Hormone that Signals the Initiation of Parturition," *Proceedings of the National Academy of Sciences*, 101, pp. 4978–4983.

DeGroot, M., and Schervish, M. (2002), *Probability and Statistics* (3rd ed.), Boston: Addison-Wesley.

Samuels, M., and Witmer, J. (2003), *Statistics for the Life Sciences* (3rd ed.), Upper Saddle River, NJ: Prentice Hall.

Stern, C. (1960), *Principles of Human Genetics* (2nd ed.), Berkeley, CA: W.H. Freeman and Company.

Sutton, H. E. (1980), *An Introduction to Human Genetics* (3rd ed.), Philadelphia, PA: Saunders College/Holt, Rinehart, and Winston.

Vos Savant, M. (19 October 1997), *Parade Magazine*, pp. 8–9.