The n-Children Problem

John Engbers[*]
Department of Mathematics, Statistics, and Computer Science
Marquette University
john.engbers@marquette.edu

Adam Hammett
Department of Science and Mathematics
Cedarville University
ahammett@cedarville.edu

September 29, 2019

Abstract

In 1959, Martin Gardner posed the Two-Children Problem and provided its solution. He later corrected his solution due to the ambiguity involved in the procedure of obtaining the given information. In 2010, Gary Foshee posed a generalization of Gardner's problem by involving a second condition beyond gender, namely the birth-day of the week. This article generalizes the Two-Children Problem to n children with m conditions, and gives a formula to compute this probability under one explicit procedure of obtaining the given information.

1 Introduction

Martin Gardner, one of the greatest mathematical puzzlers ever to live, posed the *Two-Children Problem* in a 1959 publication of *Scientific American* (republished in Chapter 14 of [2]):

G1 Mr. Jones has two children. The older child is a girl. What is the probability that both children are girls?

G2 Mr. Smith has two children. At least one of them is a boy. What is the probability that both children are boys?

Our goal in this note is to extend this famous problem, but not before taking in the historical, and sometimes controversial, landscape surrounding it. At

[*]Research supported by the Simons Foundation grant 524418.

the conclusion of our exploration, we will be able to state and prove a robust generalization to any number of children and any number of conditions we might require of them.

2 A brief history lesson

Let's get back to Gardner's problems. Unsurprisingly, the widely accepted answer to G1 is $1/2$, and we will largely ignore this problem. But Gardner caused quite a ruckus by claiming that the answer to G2 is $1/3$. Here's his solution: "If Smith has two children, at least one of which is a boy, we have three equally probable cases: boy-boy, boy-girl, girl-boy. In only one case are both children boys, so the probability that both are boys is $1/3$."

However, there is an issue with this solution that many have since pointed out. In fact, Gardner himself wrote a follow-up column titled "Probability and Ambiguity" in which he admitted as much. According to Gardner (see Chapter 19 of [2]), "...the answer depends on the procedure by which the information 'at least one is a boy' is obtained." In a 2012 article in *The College Mathematics Journal* ("Martin Gardner's Mistake" [3]), Tanya Khovanova delineates the procedure suggested by Gardner through which we might arrive at his original answer to G2: "Pick all the families with two children, at least one of which is a boy. If Mr. Smith is chosen randomly from this list, then the answer is $1/3$."

Still, in the original problem, the procedure that determined the information to be given, or equivalently the underlying randomness inherent in the problem, is not explicitly stated. Khovanova drives this point home by providing a litany of examples whereby the resulting answer to G2 is different depending upon the way in which Smith reports the information to us.

As a new example, consider all two-child families with at least one son. Suppose we know that the random person selected (Smith) will report to us in the following way: if he has exactly one son, he rolls a die and says "I have one boy and one girl" if the outcome of the roll is 1, and says "I have at least one boy" otherwise. If he has two sons, he rolls a die and says "I have two boys" if the outcome of the roll is 1 or 2, and says "I have at least one boy" otherwise. So here, we have the following 18 equally likely outcomes:

(boy-boy, 1)	(boy-boy, 2)	(boy-boy, 3)	(boy-boy, 4)	(boy-boy, 5)	(boy-boy, 6)
(boy-girl, 1)	(boy-girl, 2)	(boy-girl, 3)	(boy-girl, 4)	(boy-girl, 5)	(boy-girl, 6)
(girl-boy, 1)	(girl-boy, 2)	(girl-boy, 3)	(girl-boy, 4)	(girl-boy, 5)	(girl-boy, 6)

Note that the 14 boxed outcomes result in Smith saying "I have at least one boy," and the 4 of those in the first row occur when Smith has two sons, and thus the answer to G2 in this scenario is $4/14 = 2/7$. The curious reader will note that modifications of the procedure just described give rise to many other possible answers to G2. So the original problem without a specified reporting procedure is ambiguous.

As Khovanova points out in her article, this was not the end of the controversy. At the ninth annual *Gathering 4 Gardner* in 2010, puzzler Gary Foshee posed this version of Gardner's *Two-Children Problem*:

F1 I have two children. One is a boy born on a Tuesday. What is the probability that I have two boys?

Again, the procedure by which the information is given is not explicitly stated, and so the problem is ambiguous, as Khovanova illustrates beautifully with several different procedures which produce different answers. We add that a similar discussion related to the famous Monty Hall problem has Persi Diaconis saying: "The strict argument would be that the (Monty Hall probability) question cannot be answered without knowing the motivation of the host." [6] At the core of these issues lies the understanding of what the inherent probability distribution is, and whether a reader chooses to read the statements literally or with any possible ambiguity. As such, for the sake of investigation we clearly state our assumption that we will use for the rest of the article.

Assumption: All families are listed, and any satisfying the given conditions are marked. One such family is chosen uniformly at random from this marked group.

For F1, one way to explicitly enact the procedure we will use is the following, given by Khovanova [4]: "You pick a random father of two children and ask him, 'Yes or no, do you have a son born on a Tuesday?' Let's make a leap and assume that all fathers know the days of the births of their children and that they answer truthfully. If the answer is yes, what is the probability that the father has two sons?" Notice that the marked group, and therefore the notion of what is uniformly random, changes when moving from G2 to F1. We also make explicit the following assumptions.

Assumption: Sons and daughters are equally probable, and the sex of one child is independent of any other children.

Assumption: Twins do not occur.

Assumption: Date and time of birth are equally likely and have no effect on the sex of the child.

See [3, 5] for discussions on the accuracy of these assumptions. Under the above assumptions, the answer to question G1 is 1/2 and to G2 is 1/3.

What about F1? First, among two-child families with at least one son, there is a 1/3 probability that there are two sons. Thus the proportion of these that have a son born on a Tuesday is

$$\frac{1}{3}\left(\frac{1}{7}\right)^2 + 2 \cdot \frac{1}{3}\left(\frac{1}{7}\right)\left(1-\frac{1}{7}\right) = \frac{1}{3}\left(\frac{1}{7}\right)\left(2-\frac{1}{7}\right);$$

here, the first summand comes from those families with two sons born on Tuesday, and the second from those with two sons where exactly one was born on

Tuesday. Next, the proportion of two-child families with a daughter and a son born on Tuesday is simply given by $(2/3) \cdot (1/7)$, since here the daughter could be born on any day. Consequently, the answer to F1 is given by

$$\frac{\frac{1}{3}\left(\frac{1}{7}\right)\left(2-\frac{1}{7}\right)}{\frac{1}{3}\left(\frac{1}{7}\right)\left(2-\frac{1}{7}\right)+\frac{2}{3}\left(\frac{1}{7}\right)} = \frac{\frac{1}{3}\left(\frac{1}{7}\right)\left(2-\frac{1}{7}\right)}{\frac{1}{3}\left(\frac{1}{7}\right)\left(2-\frac{1}{7}+2\right)} = \frac{2-\frac{1}{7}}{4-\frac{1}{7}} = \frac{13}{27}.$$

This is the surprising answer that Foshee favored. Look again at the problems: the "Tuesday" information, when compared with Gardner's original problem G2, has *significantly increased* the probability that both children are boys! But why is it that this additional information has pushed the probability higher, in fact nearer to the answer to G1? The answer to this question requires that we mine the specifics of our solution to F1 a bit deeper, and recall that the marked group changed from G2 to F1.

3 A generalization

Suppose that each child has likelihood $p > 0$ of satisfying a given condition C, and that each child's satisfaction of C is independent of gender and of any other child's satisfaction of C. For example, C might be the condition "born on a Tuesday," in which case $p = 1/7$. We may now answer the following question (which, as a reminder, follows the assumptions that have been laid out above):

Q1 I have two children. One is a boy satisfying condition C. What is the probability that I have two boys?

As before, the proportion of two-child families with two sons, at least one of which satisfies C is given by

$$\tfrac{1}{3}p^2 + 2 \cdot \tfrac{1}{3}p(1-p) = \tfrac{1}{3}p(2-p);$$

here, the first summand is the probability that both children are boys satisfying C, and the second is from the two cases where both children are boys and exactly one satisfies C. On the other hand, the proportion of two-child families with a daughter and a son satisfying C is simply given by $(2/3) \cdot p$ (since the daughter may or may not satisfy C), and so the answer to Q1 is

$$\frac{\frac{1}{3}p(2-p)}{\frac{1}{3}p(2-p)+\frac{2}{3}p} = \frac{\frac{1}{3}p(2-p)}{\frac{1}{3}p(2-p+2)} = \frac{2-p}{4-p}. \tag{1}$$

We note that a geometric derivation of (1) can be found in [5]. Taking $p = 1/7$ into (1) we once again obtain the answer to F1, and taking $p = 1$ we get the answer to G2! We also see that this probability always lies in the interval $[1/3, 1/2)$, and gets closer to $2/4 = 1/2$ for conditions C that are more rare. For example, if C is the condition "born between 12:00pm and 12:01pm," then $p = 1/1440$ since there are 1440 minutes in a day, and (1) becomes $2879/5759 \approx 0.4999$. Of course here we are using our assumption of a uniform distribution of birth-times, and while this is not strictly true due to high proportions of births by scheduled caesarean section (see the 2018 National Vital Statistics Reports [1]), it is a reasonable approximation.

4 A further generalization

We are now ready to give a generalization that unifies and extends both the problems of Gardner and Foshee. For this, we shall need three fundamental tools:

1. The formula for conditional probability; namely that if A and B are events with $\mathbb{P}(B) > 0$, then $\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. Here, $\mathbb{P}(A \mid B)$ is understood to mean "the probability that event A happens, conditioned on B having occurred." B is called the *conditioning event*.
2. If A and B are independent events, then $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.
3. If A' denotes the complement of the event A, then $\mathbb{P}(A) + \mathbb{P}(A') = 1$.

Now on to the generalization: suppose we have m (enumerated) conditions, n children, and that C_{ij} denotes the event "child i (by birth order) satisfies the jth condition." Moreover, we assume that the events C_{ij}, $1 \le i \le n$, $1 \le j \le m$, are mutually independent, with $\mathbb{P}(C_{ij}) = p_j > 0$. In other words, each child has the same likelihood of satisfying condition j, independent of any other child, and child i's satisfying condition j does not impact child i's likelihood of satisfying condition $j' \neq j$. Writing $[m] := \{1, 2, \ldots, m\}$, is it possible to compute the following?

$$\begin{aligned}\mathbb{P}\big(&\text{"all } n \text{ children satisfy the conditions in } X\text{"} \\ &\quad \mid \text{"at least } k \text{ children satisfy all conditions in } Y\text{"}\big) \\ &\qquad := \mathbb{P}\left(\mathcal{X}_n \mid \mathcal{Y}_k\right), \quad X, Y \subseteq [m], \quad 0 \le k \le n\end{aligned}$$

Surprisingly, we can!

Notice that if $X = Y = \varnothing$, then $\mathbb{P}(\mathcal{X}_n) = \mathbb{P}(\mathcal{Y}_k) = 1$, since there are no requirements on the children. Letting $\Pi_S := \prod_{s \in S} p_s$ for a subset $S \subseteq [m]$ (with the convention $\Pi_\varnothing = 1$), the probability of the conditioning event $\mathcal{Y}_k$ may be computed as follows. Each child has probability Π_Y of satisfying the conditions in Y, and $1 - \Pi_Y$ of failing to meet them. Thus, since the n children are independent, the probability that some j children ($0 \le j \le n$) meet the conditions in Y is given by $\binom{n}{j}(\Pi_Y)^j(1 - \Pi_Y)^{n-j}$ (binomial probability distribution). Consequently

$$\Sigma_{Y,k} := \sum_{j \ge k} \binom{n}{j} (\Pi_Y)^j \left(1 - \Pi_Y\right)^{n-j} = 1 - \sum_{j < k} \binom{n}{j} (\Pi_Y)^j \left(1 - \Pi_Y\right)^{n-j} \tag{2}$$

is the likelihood that *at least* k children meet all conditions in Y. Of course both versions of this formula (2) may be applied, depending upon which sum is less arduous to utilize. And for $\mathcal{X}_n$, since each child has probability Π_X of meeting all conditions in X, independence implies that $(\Pi_X)^n$ is the probability that all n children satisfy X. Thus we have proved

$$\mathbb{P}(\mathcal{X}_n) = (\Pi_X)^n \quad \text{and} \quad \mathbb{P}(\mathcal{Y}_k) = \Sigma_{Y,k}. \tag{3}$$

But what about $\mathcal{X}_n \cap \mathcal{Y}_k$? For this, we must determine what fraction of the outcomes in $\mathcal{X}_n$ also have at least k children meeting the conditions in Y. Arguing as we did above in (2), with $Y \setminus X$ replacing Y (since on the event $\mathcal{X}_n$ all of the children *already* satisfy X, we need only consider $Y \setminus X$), this fraction is given by $\Sigma_{Y\setminus X,k}$. Invoking independence once more we see that $(\Pi_X)^n \cdot \Sigma_{Y\setminus X,k}$ is the overall probability that all children satisfy X, and at least k children meet all conditions in Y. So (2) implies

$$\begin{aligned}
\mathbb{P}(\mathcal{X}_n \cap \mathcal{Y}_k) &= (\Pi_X)^n \cdot \Sigma_{Y\setminus X,k} \\
&= (\Pi_X)^n - \sum_{j<k} \binom{n}{j} (\Pi_{X\cup Y})^j (\Pi_X - \Pi_{X\cup Y})^{n-j} \\
&= \sum_{j\geq k} \binom{n}{j} (\Pi_{X\cup Y})^j (\Pi_X - \Pi_{X\cup Y})^{n-j}\,;
\end{aligned} \tag{4}$$

here, we have used $\Pi_X \cdot \Pi_{Y\setminus X} = \Pi_{X\cup Y}$. Combining (4) and (3), we obtain then

$$\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_k) = (\Pi_X)^n \cdot \left(\frac{\Sigma_{Y\setminus X,k}}{\Sigma_{Y,k}}\right). \tag{5}$$

We shall make use of all variations of (5) implicit from (4), depending upon what is most convenient in context. One special case of (5) that seems particularly pertinent is the $k = 1$ variant, and we shall make frequent reference to it below. Plugging $k = 1$ into (5) simplifies to

$$\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_1) = \frac{(\Pi_X)^n\big(1 - \big(1 - \Pi_{Y\setminus X}\big)^n\big)}{1 - (1 - \Pi_Y)^n} = \frac{(\Pi_X)^n - (\Pi_X - \Pi_{X\cup Y})^n}{1 - (1 - \Pi_Y)^n}. \tag{6}$$

5 Some concluding applications

We finish with some applications of (5) and (6).

1. Let's consider the question "Given that at least one of my n children is a boy born on a Tuesday, what is the probability that they are all born on Tuesday?" Here we have $k = 1$, $m = 2$ with $p_1 = 1/2$, $p_2 = 1/7$, $X = \{2\}$ and $Y = \{1, 2\}$. Then equation (6) delivers

$$\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_1) = \frac{\left(\frac{1}{7}\right)^n - \left(\frac{1}{7} - \frac{1}{2}\cdot\frac{1}{7}\right)^n}{1 - \left(1 - \frac{1}{2}\cdot\frac{1}{7}\right)^n} = \frac{2^n - 1}{14^n - 13^n}.$$

   Why is this answer sensible? There are 14 possible gender/birth-day combinations, and hence 14^n possible gender/birth-day sequences of n children. Moreover, 13^n of these avoid the "boy/Tuesday" combination for each child. Hence $14^n - 13^n$ equally likely n-child sequences have the property that at least one child is a boy born on Tuesday. And $2^n - 1$ of these have all children born on Tuesday, the "-1" coming from the one excluded case where each child is a girl. All of this is cleverly encapsulated in formula (5)! When $n = 2$, we compute the probability to curiously equal 1/9.

2. Suppose that $k = n = 1$. Then (6) delivers $\mathbb{P}(\mathcal{X}_1 \,|\, \mathcal{Y}_1) = \Pi_{X \cup Y}/\Pi_Y = \Pi_{X \setminus Y}$. How do we interpret this formula? Since only one child is being considered in this case, the given information that they satisfy the conditions in Y means that we only need to ensure that in addition they also satisfy the conditions in $X \setminus Y$, hence the formula above. Of course, if in addition $X \subseteq Y$ we have $X \setminus Y = \varnothing$ and so this formula gives a probability $\Pi_\varnothing = 1$ that the child satisfies the conditions in X *given* that they already satisfy those in Y. Indeed, given that the conditions in Y hold for this child, those in X are *guaranteed* to hold if $X \subseteq Y$!

3. What about the boundary cases $k = 0$ and $k = n$? If $k = 0$, then we certainly have $\Sigma_{Y,0} = \Sigma_{Y \setminus X,0} = 1$ and hence (5) delivers $\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_0) = (\Pi_X)^n = \mathbb{P}(\mathcal{X}_n)$ (recall (3)), as it should because $\mathcal{Y}_0 = \mathcal{U}$, the universe, in this case. On the other hand, if $k = n$ then $\Sigma_{Y,n} = (\Pi_Y)^n$ and $\Sigma_{Y \setminus X,k} = (\Pi_{Y \setminus X})^n$, and so (5) gives

$$\mathbb{P}\left(\mathcal{X}_n \,|\, \mathcal{Y}_n\right) = (\Pi_X)^n \cdot \frac{\left(\Pi_{Y \setminus X}\right)^n}{\left(\Pi_Y\right)^n} = \frac{\left(\Pi_{X \cup Y}\right)^n}{\left(\Pi_Y\right)^n} = \left(\Pi_{X \setminus Y}\right)^n .$$

This extends the last example, as we need only ensure that the n children all satisfy the remaining conditions in $X \setminus Y$, for $\mathcal{Y}_n$ guarantees they already satisfy those in Y. As above, if $X \subseteq Y$ then $\Pi_{X \setminus Y} = \Pi_\varnothing = 1$ and we get $\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_n) = 1$.

4. Suppose that $X \cap Y = \varnothing$. Then what happens? We have $Y \setminus X = Y$, hence $\Sigma_{Y \setminus X,k} = \Sigma_{Y,k}$ and so (5) reduces to $\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_k) = (\Pi_X)^n = \mathbb{P}(\mathcal{X}_n)$ (recall (3) once more). But this should be the case, because here the events $\mathcal{X}_n$ and $\mathcal{Y}_k$ are independent.

5. Let's revisit our extension Q1. In this case we have $k = 1$, $m = n = 2$, $p_1 = 1/2$, $p_2 = p$, $X = \{1\}$ and $Y = \{1, 2\}$, and so from (6) we get

$$\mathbb{P}\left(\mathcal{X}_n \,|\, \mathcal{Y}_1\right) = \frac{\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2} - \frac{1}{2}p\right)^2}{1 - \left(1 - \frac{1}{2}p\right)^2} = \frac{2 - p}{4 - p},$$

which agrees with (1). Moreover, if we extend this example to any number n of children, from (6) we obtain

$$\mathbb{P}\left(\mathcal{X}_n \,|\, \mathcal{Y}_1\right) = \frac{\left(\frac{1}{2}\right)^n - \left(\frac{1}{2} - \frac{1}{2}p\right)^n}{1 - \left(1 - \frac{1}{2}p\right)^n} = \frac{1 - (1 - p)^n}{2^n - (2 - p)^n}. \tag{7}$$

6. Let's extend Q1 further. In addition to the condition "is a boy," we consider two other mutually independent conditions C_2 and C_3 which occur with respective probabilities $p_2 = p$ and $p_3 = q$. Let's answer the question "Given that at least two of my n children are boys satisfying C_2 and C_3, what is the probability that they are all boys satisfying C_2?" Here, we take $k = 2$, $m = 3$, $p_1 = 1/2$, $p_2 = p$, $p_3 = q$, $X = \{1, 2\}$ and $Y = \{1, 2, 3\}$ into (5) and get the remarkable formula

$$\begin{aligned}\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_2) &= \frac{\left(\frac{1}{2}p\right)^n - \left(\frac{1}{2}p - \frac{1}{2}pq\right)^n - n\left(\frac{1}{2}pq\right)\left(\frac{1}{2}p - \frac{1}{2}pq\right)^{n-1}}{1 - \left(1 - \frac{1}{2}pq\right)^n - n\left(\frac{1}{2}pq\right)\left(1 - \frac{1}{2}pq\right)^{n-1}} \\ &= \frac{p^n\left(1 - (1-q)^n - nq(1-q)^{n-1}\right)}{2^n - (2-pq)^n - npq(2-pq)^{n-1}}. \end{aligned} \tag{8}$$

7. How about the direct extension of Gardner's question: "Given that at least one of my n children is a boy, what is the probability that they are all boys?" For this we can simply put $p = 1$ into (7) to obtain $\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_1) = 1/(2^n - 1)$. As above, this answer makes perfect sense in light of the fact that there are $2^n - 1$ equally likely n-child gender sequences having the property that at least one child is a boy, and exactly 1 of these consists entirely of boys. Taking $n = 2$ into this formula gives Gardner's answer to G2, $1/(2^2 - 1) = 1/3$, and taking $n = 3$ delivers an answer of $1/(2^3 - 1) = 1/7$.

   And for the related question "Given that at least two of my n children are boys, what is the probability that they are all boys?" we can simply take $p = q = 1$ into (8) to get $\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_2) = 1/(2^n - n - 1)$. This is reasonable, because among the 2^n equally likely n-child gender sequences, there is 1 that consists entirely of girls and n that consist of exactly one boy, and of the remaining $2^n - n - 1$ gender sequences precisely 1 consists entirely of boys.

8. Consider the following direct extension of Foshee's question: "Given that at least one of my n children is a boy born on a Tuesday, what is the probability that they are all boys?" Here we can put $p = 1/7$ into (7), obtaining then

$$\mathbb{P}(\mathcal{X}_n \,|\, \mathcal{Y}_1) = \frac{1 - \left(1 - \frac{1}{7}\right)^n}{2^n - \left(2 - \frac{1}{7}\right)^n} = \frac{7^n - 6^n}{14^n - 13^n}.$$

   We have already seen why this denominator makes sense. What about the numerator? There are $7^n - 6^n$ equally likely n-boy sequences having the property that at least one boy is born on Tuesday. Of course, taking $n = 2$ here we get Foshee's famous answer to F1, $(7^2 - 6^2)/(14^2 - 13^2) = 13/27$. When $n = 3$, we obtain a probability of $(7^3 - 6^3)/(14^3 - 13^3) = 127/547 = 0.232....$ Compare with the statement that ignores the "Tuesday" information (from the previous part) where the probability was $1/7 = 0.142....$

For the related question "Given that at least two of my n children are boys born on a Tuesday between 12:00pm and 12:01pm, what is the probability that they are all boys born on Tuesday?" we can take $p = 1/7$ and $q = 1/1440$ into (8), obtaining

$$\begin{aligned}\mathbb{P}\left(\mathcal{X}_n \,|\, \mathcal{Y}_2\right) &= \frac{\left(\frac{1}{7}\right)^n \left(1 - \left(1 - \frac{1}{1440}\right)^n - n\left(\frac{1}{1440}\right)\left(1 - \frac{1}{1440}\right)^{n-1}\right)}{2^n - \left(2 - \left(\frac{1}{7}\right)\left(\frac{1}{1440}\right)\right)^n - n\left(\frac{1}{7}\right)\left(\frac{1}{1440}\right)\left(2 - \left(\frac{1}{7}\right)\left(\frac{1}{1440}\right)\right)^{n-1}} \\ &= \frac{1440^n - 1339^n - n1339^{n-1}}{20160^n - 20159^n - n20159^{n-1}}.\end{aligned}$$

Wow! Can the reader provide the rationale for this formula?

Now it's your turn... have fun!

References

[1] Division of Vital Statistics, *National Vital Statistics Reports*, Vol. 67, No. 1, U.S. Department of Health and Human Services, Center of Disease Control and Prevention, 2018. Accessed 6 March 2018 from `https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_01_tables.pdf`

[2] Martin Gardner, *The Second Scientific American Book of Mathematical Puzzles and Diversions*, Chicago University Press, 1987.

[3] Tanya Khovanova, *Martin Gardner's Mistake*, College Mathematics Journal, Vol. 43, No. 1 (January 2012), pp. 20–24.

[4] Tanya Khovanova, *Martin Gardner's Mistake*, arXiv:1102.0173.

[5] Wendy Taylor and Kaye Stacey, *Gardner's Two Children Problems and Variations: Puzzles with Conditional Probability and Sample Spaces*, Australian Mathematics Teacher, Vol. 70, No. 2 (2014), pp. 13–19.

[6] John Tierney, Behind Monty Hall's Doors: Puzzle, Debate and Answer? *New York Times* July 21, 1991. Accessed 23 July 2018 from `https://www.nytimes.com/1991/07/21/us/behind-monty-hall-s-doors-puzzle-debate-and-answer.html`