Bernoulli Trials and Family Planning

EUGENE F. KRAUSE
University of Michigan
Ann Arbor, MI 48104

Introduction A sports-minded young couple resolves to continue having babies until they have enough boys for a touch football team (7) and enough girls for a basketball team (5). How many babies should they plan on having? Assuming boys and girls are equally probable, we shall see that they had better plan on having $15\frac{81}{512} \approx 15.158$ babies. If we use the statistically more realistic probabilities of 0.514 for a boy and 0.486 for a girl, then (we shall see) they should expect to have about 14.745 babies.

We have two goals in this paper. One, of course, is to pose and solve the general problem of which the family-planning question above is a special case. The other is to provide some insight into the process of problem-solving by describing two quite different strategies that we employed. Before we can look at the strategies, though, we need to formulate the general problem.

We are concerned with Bernoulli trials: repeated performances of a single random experiment such that on each performance there are just two possible outcomes, success with probability p or failure with probability $q = 1 - p$. We suppose that these Bernoulli trials continue until the (first) moment when the number of successes has reached s and the number of failures has reached r. The question is this: On average, how many trials (performances) are required? We shall refer to this expected number of trials as the waiting time for s successes and r failures, and we shall denote it by $w(s, r)$.

A front-end, recursive-inductive strategy On the *first* trial, one of two things will happen. Either we will have a success, with probability p, and then expect to wait an additional $w(s-1, r)$ trials; or we will have a failure, with probability $q = 1 - p$, and then expect to wait an additional $w(s, r-1)$ trials. That is,

$$w(s,r) = p[1 + w(s-1,r)] + q[1 + w(s,r-1)]$$

which simplifies to a fundamental recurrence relation

$$w(s,r) = 1 + qw(s,r-1) + pw(s-1,r) \qquad (r, s = 1, 2, \ldots) \tag{1}$$

Our plan is to use this recurrence relation to generate a list of formulas for $w(s,1)$, $w(s,2), \ldots$ in which we will look for a pattern. To get started we need two initial conditions. The initial condition

$$w(s,0) = s/p \qquad (s = 1, 2, \ldots) \tag{2}$$

follows from the simple observation that $w(s,0) = s \cdot w(1,0)$ and the well-known fact [**1**, p. 189] that the waiting time for *one* success in Bernoulli trials is the reciprocal of the probability of success on a single trial. (For example, the waiting time for a "4" to turn up when one rolls a die repeatedly is 6.) The other initial condition, of course, is

$$w(0,r) = r/q \qquad (r = 1, 2, \ldots) \tag{3}$$

Fixing $r = 1$ in (1) yields a recurrence relation for the function of one variable, $w(s,1)$

$$w(s,1) = pw(s-1,1) + 1 + qs/p$$

Iterating this recurrence relation produces a formula for $w(s,1)$

$$w(s,1)=\sum_{i=0}^{s-1}p^i+q\sum_{i=0}^{s-1}p^{i-1}(s-i)+p^s/q$$

which, upon use of the easily derived identities (4) and (5),

$$\sum_{i=0}^{s-1}p^i=\frac{1-p^s}{1-p}=\frac{1-p^s}{q} \tag{4}$$

$$\sum_{i=0}^{s-1}p^{i-1}(s-i)=\frac{1}{q}\left(\frac{s}{p}-p^{s-1}\right)+\frac{1}{q^2}\left(p^{s-1}-1\right) \tag{5}$$

becomes the first formula in (6) below.

Next we fix $r=2$ in (1) and use the formula we just found for $w(s,1)$ to arrive at a recurrence relation for the function of one variable, $w(s,2)$:

$$w(s,2)=pw(s-1,2)+1+qs/p+p^s.$$

Iterating this recurrence relation produces a formula for $w(s,2)$:

$$w(s,2)=\sum_{i=0}^{s-1}p^i+sp^s+q\sum_{i=0}^{s-1}p^{i-1}(s-i)+2p^s/q,$$

which, on using identities (4) and (5), simplifies to the second formula in (6) below.

Fixing r at 3, then 4 in (1), and repeating twice the process that we carried out in cases $r=1$ and $r=2$, leads us to the final two formulas in (6):

$$\begin{aligned}
w(s,1)&=\frac{s}{p}+p^s\left[\frac{1}{q}\binom{s-1}{0}\right]\\
w(s,2)&=\frac{s}{p}+p^s\left[\frac{2}{q}\binom{s-1}{0}+\binom{s}{1}\right]\\
w(s,3)&=\frac{s}{p}+p^s\left[\frac{3}{q}\binom{s-1}{0}+2\binom{s}{1}+q\binom{s+1}{2}\right]\\
w(s,4)&=\frac{s}{p}+p^s\left[\frac{4}{q}\binom{s-1}{0}+3\binom{s}{1}+2q\binom{s+1}{2}+q^2\binom{s+2}{3}\right]
\end{aligned} \tag{6}$$

The decision to express the polynomials in s (the expressions inside the brackets in (6)) in terms of the basis $\binom{s-1}{0}$, $\binom{s}{1}$, $\binom{s+1}{2}$,... instead of the standard basis $s^0, s^1, s^2,\ldots$ was based on exploratory work with the special case $p=q=\frac{1}{2}$ where the standard basis gave a chaotic picture while the basis of binomial coefficients revealed a pattern.

Inside the brackets in the formula for $w(s,4)$ in (6), patterns finally reveal themselves in the exponents of q, in the numerical coefficients, and in the upper and lower entries in the binomial coefficients. Since these patterns persist in the formulas for $w(s,3)$, $w(s,2)$, and $w(s,1)$, it is a good bet that they hold in general. That is the content of our first theorem. And we call it a theorem rather than a conjecture because it can be proved by double induction. We omit that proof, which relies on equations (1)–(3).

THEOREM 1. *If $w(s,r)$ denotes the waiting time for s successes and r failures in Bernoulli trials, where the probability of success is p and the probability of failure is*

$q = 1 - p$, then

$$w(s,r) = \frac{s}{p} + p^s \sum_{k=1}^{r} kq^{r-1-k} \binom{s+r-1-k}{r-k} \tag{7}$$

where we agree that $\binom{-1}{0} = 1$ so that $w(0,r) = r/q$.

Note. When we apply this theorem to the case of the sports-minded couple, we have two choices. We can think of having a son as success and a daughter as failure, and calculate $w(7,5)$ with $p = q = \frac{1}{2}$ to arrive at $15\frac{81}{512}$. With $p = 0.514$ and $q = 0.486$, we get $w(7,5) \approx 14.745$. Or we can view sons as failures and daughters as successes and calculate $w(5,7)$. Again we get $15\frac{81}{512}$ when $p = q = \frac{1}{2}$, and we get (approximately) 14.745 when $p = 0.486$ and $q = 0.514$.

A back-end, deductive strategy For the front-end strategy we looked at what happens on the *first* trial, and this led us to a recurrence relation for $w(s,r)$. For the back-end strategy we look at what happens on the *last* trial; that is, on the trial that completes the project of attaining (at least) s successes and (at least) r failures. This will lead us to a probability distribution and then the expected value for an appropriate random variable.

Let K be the waiting time (random variable) until s successes and r failures have been attained. Now s successes and r failures can be attained in k trials in two ways: either $s-1$ successes occur in the first $k-1$ trials, and $(k-1)-(s-1) \geq r$, and the k^{th} trial results in a success; or $r-1$ failures occur in the first $k-1$ trials, and $(k-1)-(r-1) \geq s$, and the k^{th} trial results in a failure. Thus

$$P(K=k) = \binom{k-1}{s-1} p^{s-1} q^{k-s} p + \binom{k-1}{r-1} q^{r-1} p^{k-r} q \qquad (k \geq r+s)$$

That is, the probability distribution for K is

$$P(K=k) = \binom{k-1}{s-1} p^s q^{k-s} + \binom{k-1}{r-1} p^{k-r} q^r \qquad (k = r+s, r+s+1, \ldots) \tag{8}$$

Thus the expected value of K, which we called $w(s,r)$ before, is

$$w(s,r) = \sum_{k=r+s}^{\infty} k\left[\binom{k-1}{s-1} p^s q^{k-s} + \binom{k-1}{r-1} p^{k-r} q^r\right] \tag{9}$$

In order to deduce a finite formula for $w(r,s)$ from (9), we need the following identity, which is just our old equation (2) in different notation.

$$\sum_{k=s}^{\infty} k \binom{k-1}{s-1} p^s q^{k-s} = \frac{s}{p} \tag{10}$$

Note that the righthand side, s/p, is the waiting time for s successes, i.e., $w(s,0)$. But the lefthand side is also $w(s,0)$ because it is just (9) in the special case of $r = 0$. The deduction of a finite formula proceeds in a straightforward way:

$$\begin{aligned} w(s,r) &= \sum_{k=r+s}^{\infty} k\binom{k-1}{s-1}p^s q^{k-s} + \sum_{k=r+s}^{\infty} k\binom{k-1}{r-1}p^{k-r}q^r \\ &= \sum_{k=s}^{\infty} k\binom{k-1}{s-1}p^s q^{k-s} - \sum_{k=s}^{r+s-1} k\binom{k-1}{s-1}p^s q^{k-s} \\ &\quad + \sum_{k=r}^{\infty} k\binom{k-1}{r-1}p^{k-r}q^r - \sum_{k=r}^{r+s-1} k\binom{k-1}{r-1}p^{k-r}q^r \\ &= \frac{s}{p} - p^s \sum_{k=s}^{r+s-1} k\binom{k-1}{s-1}q^{k-s} + \frac{r}{q} - q^r \sum_{k=r}^{r+s-1} k\binom{k-1}{r-1}p^{k-r} \\ &= \frac{s}{p} - p^s \sum_{k=s}^{r+s-1} s\cdot\frac{k}{s}\binom{k-1}{s-1}q^{k-s} + \frac{r}{q} - q^r \sum_{k=r}^{r+s-1} r\cdot\frac{k}{r}\binom{k-1}{r-1}p^{k-r} \\ &= \frac{s}{p} + \frac{r}{q} - sp^s \sum_{k=s}^{r+s-1} \binom{k}{s}q^{k-s} - rq^r \sum_{k=r}^{r+s-1} \binom{k}{r}p^{k-r}. \end{aligned}$$

THEOREM 2. *If $w(s,r)$ denotes the waiting time for s successes and r failures in Bernoulli trials, where the probability of success is p and the probability of failure is $q = 1 - p$, then*

$$w(s,r) = \frac{s}{p} + \frac{r}{q} - sp^s \sum_{k=s}^{r+s-1} \binom{k}{s}q^{k-s} - rq^r \sum_{k=r}^{r+s-1} \binom{k}{r}p^{k-r}. \tag{11}$$

When we apply (11) in the special case of $s = 7, r = 5$ with $p = q = \frac{1}{2}$, we again get $15\frac{81}{512}$. For $p = 0.514$ and $q = 0.486$, we get (approximately) 14.745. And note that formula (11), unlike formula (7), is *obviously* symmetric in the sense that (11) is unchanged if we interchange both s with r and p with q.

Conclusion Two different strategies led to two different looking but equivalent solutions to the same problem. The front-end strategy is simple and straightforward, but soon leads one into rather heavy algebraic manipulation to establish a pattern. Induction confirms *that* the pattern holds in general, but still leaves one wondering *why*.

The back-end strategy, while requiring a stronger conceptual background in probability, leads to a simpler derivation of a qualitatively more satisfying solution: formula (11) is obviously symmetric. Quantitatively, however, formula (7) has an edge. The summation in formula (7) has r terms, while the two summations in formula (11) produce $r+s$ terms. The computational advantage of formula (7) suggests that it would be worthwhile to deduce (7) using methods similar to the ones used to deduce (11). Here is such a deduction.

The waiting time for s successes and r failures is equal to the waiting time for s successes plus the weighted average of the waiting times for k failures ($k = 0, 1, 2, \ldots, r$), assuming that $r-k$ failures occurred while we waited for the s^{th} success. The waiting time for s successes is, of course, s/p, and the waiting time for k failures is k/q. The probability (weight) that $r-k$ failures occur among the first

$s-1$ successes, and that then the s^{th} success occurs, is $\binom{s-1+r-k}{r-k}p^{s-1}q^{r-k}p$. Thus

$$w(s,r)=\frac{s}{p}+\sum_{k=0}^{r}\binom{s-1+r-k}{r-k}p^{s}q^{r-k}\frac{k}{q}$$

which reduces to (7).

Acknowledgment. I would like to acknowledge a heavy debt of gratitude to two anonymous referees and to my longtime friend and colleague, Chung-Tuo Shih, for their insights that shaped this paper.

REFERENCES

1. F. Mosteller, R. Rourke, and G. Thomas, Jr., *Probability with Statistical Applications*, 2nd edition, Addison Wesley, Reading, MA, 1970.

Sailing Around Binary Trees and Krusemeyer's Bijection

YUJIN SHEN
The Richard Stockton College of New Jersey
Pomona, NJ 08240

JIANG T. SHULT
Kansas State University
Manhattan, KS 66506

Introduction The Catalan numbers, defined numerically by the equation

$$c_n=\frac{1}{n+1}\binom{2n}{n},$$

can be used to describe many mathematical models. For example, Mark Krusemeyer presented two different types of parenthesizations in [3]. The first type of parenthesization is a sequence of n pairs of matching parentheses such that, up to any point in the sequence, the number of left parentheses is greater than or equal to the number of right parentheses. Such a sequence of n pairs of parentheses is called *well-defined.* For example, (( )) is a well-defined sequence, while ( )) (is not. The second type of parenthesization is associated with possible ways to multiply $n+1$ different things if the multiplication is non-associative and non-commutative. For example, $(ab)c$ and $a(bc)$ are two different products of three letters. We use Krusemeyer's terminology to call a product of $n+1$ different things with n pairs of parentheses, or $n-1$ pairs if you omit the outer pair, a *bracketing*. Let A_n be the set of well-defined sequences of n pairs of parentheses and B_n the set of bracketings with n multiplications. It is known (see, e.g., [4] and [5]) that $|A_n|=|B_n|=c_n$.

Richard Guy remarked in [2] that "an examination of the symmetries in the two cases makes it unlikely you'll find a direct combinatorial comparison" between the two types of parenthesizations. Interestingly, Krusemeyer [3] constructed a direct