NOTES

Edited by **Sergei Tabachnikov**

Probabilistic Proofs of a Binomial Identity, Its Inverse, and Generalizations

Michael Z. Spivey

Abstract. We give elementary probabilistic proofs of a binomial identity, its binomial inverse, and generalizations of both of these. The proofs are obtained by interpreting the sides of each identity as the probability of an event in two different ways. Each proof uses a classic balls-and-jars scenario.

1. INTRODUCTION. In a recent note [**3**] in this MONTHLY, Peterson uses properties of the exponential distribution to give a probabilistic proof of the identity

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{m+k} = \frac{n!}{m(m+1)\cdots(m+n)} \tag{1}$$

and a generalization of it, valid for $m > 0$.

In this note, we give a more classic probabilistic proof of Identity (1) and its generalization, both in the case where m is a positive integer. We use a standard balls-and-jars interpretation. These proofs are obtained by first finding probabilistic proofs for the binomial inverses of these identities, after which the proofs for (1) and its generalization are straightforward. Also, we prove (1) in a slightly different form, valid for positive integers m:

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{m+k} = \frac{(m-1)!\,n!}{(m+n)!}. \tag{1a}$$

Finally, we use $\bar{A}$ to denote the complement of the set A.

2. INCLUSION/EXCLUSION. Since (1a) is an alternating binomial coefficient identity, at first glance, it seems as if it should have an interpretation that uses the principle of inclusion/exclusion.

Theorem 1 (Principle of Inclusion/Exclusion, Probabilistic Version). *Given events* $A_1, A_2, \ldots, A_n$,

$$P\left(\bigcap_{i=1}^{n} A_i\right) = 1 - \sum_{i=1}^{n} P(\bar{A}_i) + \sum_{1 \le i < j \le n} P(\bar{A}_i \cap \bar{A}_j)$$

http://dx.doi.org/10.4169/amer.math.monthly.123.2.175
MSC: Primary 05A19, Secondary 60C05

$$- \sum_{1\le i<j<k\le n} P(\bar{A}_i \cap \bar{A}_j \cap \bar{A}_k) + \cdots + (-1)^n P(\bar{A}_1 \cap \cdots \cap \bar{A}_n).$$

However, when inclusion/exclusion is used to give a probabilistic proof of a binomial coefficient identity, each term on the right-hand side of Theorem 1 generally corresponds to one of the terms in the sum expression of the identity to be proved. When this happens, the $k = 0$ term evaluates to 1, the $k = 1$ term evaluates to $-\sum_{i=1}^{n} P(\bar{A}_i)$, and so forth. That is not the case with (1a), though; the $k = 0$ term evaluates to $1/m$.

We can still use inclusion/exclusion, though, by generalizing Theorem 1 in the following manner.

Theorem 2 (Principle of Inclusion/Exclusion, Generalized Probabilistic Version). *Given events $A_1, A_2, \ldots, A_n$ and B,*

$$\begin{aligned} P\left(B \cap \bigcap_{i=1}^{n} A_i\right) = & P(B) - \sum_{i=1}^{n} P(B \cap \bar{A}_i) + \sum_{1\le i<j\le n} P(B \cap \bar{A}_i \cap \bar{A}_j) \\ & - \sum_{1\le i<j<k\le n} P(B \cap \bar{A}_i \cap \bar{A}_j \cap \bar{A}_k) \\ & + \cdots + (-1)^n P(B \cap \bar{A}_1 \cap \cdots \cap \bar{A}_n). \end{aligned}$$

Theorem 2 can be easily proved from Theorem 1. First, note that if $P(B) = 0$, then Theorem 2 reduces to $0 = 0$. If $P(B) > 0$, then divide both sides of Theorem 2 by $P(B)$. By the definition of conditional probability and the fact that $P(*|B)$ is itself a probability measure, the resulting expression follows from Theorem 1. Of course, Theorem 1 is just Theorem 2 with B taken to be the entire sample space.

To prove Identity (1a) using Theorem 2, we will (among other things) need to find an event B that has probability $1/m$.

3. THE BINOMIAL INVERSE IDENTITY.

To understand the origin of our balls-and-jars proof of (1a), it is helpful to begin with the proof of its binomial inverse. The binomial inversion property is the following.

Theorem 3 (Binomial Inversion). *If*

$$g(n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k f(k),$$

then

$$f(n) = \sum_{k=0}^{n} \binom{n}{k} (-1)^k g(k).$$

For a proof of Theorem 3, see Graham, Knuth, and Patashnik [**2**, p. 193]. We cite binomial inversion here not to prove (1a) and its generalizations via their binomial inverses, although this would of course be logically valid. Instead, our intention is to motivate strictly probabilistic balls-and-jars proofs of (1a) and its generalizations via similar proofs for their binomial inverses.

The binomial inverse of (1a) is

$$\sum_{k=0}^{n}\binom{n}{k}\frac{(m-1)!\,k!\,(-1)^k}{(m+k)!}=\frac{1}{m+n}. \tag{2}$$

Proof. Thinking about (2) probabilistically, the right-hand side looks like the probability of selecting one special item out of $m+n$ choices. Working with balls and jars, let's suppose we have a black ball, n blue balls, and $m-1$ red balls. The right-hand side of (2) is the probability that if we choose the balls one-by-one from the jar, the black ball is the first ball chosen.

To use Theorem 2 for the left-hand side, we have to interpret "the black ball is the first ball chosen" as the intersection of a collection of n specific events and a special event B. The straightforward choice is to let A_i, for $1 \le i \le n$, be the event that the black ball is drawn before blue ball i and B be the event that the black ball is chosen before any of the red balls. This interpretation yields $P(B) = 1/m$, the probability that the black ball is chosen first of m specific balls. The event $B \cap \bar{A}_i$ is the event that blue ball i is chosen first and the black ball is chosen second of $m+1$ specific balls and thus has probability $1/(m(m+1))$. In general, the intersection of any k of the $\bar{A}_i$'s with B is the event that k specific blue balls are chosen before the black ball, which in turn is chosen before all of the red balls. There are $k!(1)(m-1)!$ ways to place the k blue balls, followed by the black ball, followed by the $m-1$ red balls, out of $(m+k)!$ total ways to place $m+k$ balls. Thus, this event has probability $(m-1)!\,k!/(m+k)!$. Since there are $\binom{n}{k}$ ways to select k specific blue balls from the n total blue balls, the probability that the black ball is chosen first is, by Theorem 2, also given by

$$\frac{(m-1)!}{m!}-\frac{n(m-1)!}{(m+1)!}+\cdots+(-1)^k\binom{n}{k}\frac{(m-1)!\,k!}{(m+k)!}+\cdots+(-1)^n\frac{(m-1)!\,n!}{(m+k)!}$$
$$=\sum_{k=0}^{n}\binom{n}{k}\frac{(m-1)!\,k!\,(-1)^k}{(m+k)!},$$

the left-hand side of (2). ∎

4. THE BINOMIAL IDENTITY. With the proof of (2) in mind, it is now easy to see how to construct a balls-and-jars proof of (1a):

$$\sum_{k=0}^{n}\binom{n}{k}\frac{(-1)^k}{m+k}=\frac{(m-1)!\,n!}{(m+n)!}. \tag{1a}$$

Proof. Using our interpretation for the nth term on the left-hand side of (2), the right-hand side of (1a) is the probability that, given a jar containing a black ball, n blue balls, and $m-1$ red balls, the black ball is chosen after all of the blue balls but before any of the red balls. We take A_i to be the event that the black ball is chosen after blue ball i. (This makes A_i the complement of the event we labeled "A_i" in the proof of (2). We shall see the same phenomenon in our proofs of (3) and (4); it is a consequence of the relationships between the identities via binomial inversion.) We take B to be the event that the black ball is chosen before all the red balls. Thus,

$$P\left(B\cap\bigcap_{i=1}^{n}A_i\right)=\frac{(m-1)!\,n!}{(m+n)!}.$$

The event $\bar{A}_i$ is the event that the black ball is chosen before blue ball i. Thus, the intersection of any k of the $\bar{A}_i$'s with B is the event that the black ball is chosen first of $m + k$ specific balls. This, of course, has probability $1/(m + k)$. There are $\binom{n}{k}$ ways to select k of the $\bar{A}_i$'s, and thus, by Theorem 2, the probability that the black ball is chosen after the blue balls but before any of the red balls can also be expressed as

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{m + k},$$

the left-hand side of (1a). ■

5. A GENERALIZATION. Peterson also outlines how to use his approach with exponential random variables to prove the following generalization of (1a), valid for nonnegative integers r:

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{m + k}\right)^r (-1)^k$$
$$= \frac{(m-1)!\,n!}{(m+n)!} \sum_{0 \le k_1 \le \cdots \le k_{r-1} \le n} \frac{1}{(m + k_1)(m + k_2)\cdots(m + k_{r-1})}. \tag{3}$$

Our balls-and-jars argument can be extended to prove this as well.

As with (1a), though, it is easier to start with the binomial inverse of (3):

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(m-1)!\,k!(-1)^k}{(m+k)!} \sum_{0 \le j_1 \le \cdots \le j_{r-1} \le k} \frac{1}{(m + j_1)(m + j_2)\cdots(m + j_{r-1})}$$
$$= \left(\frac{1}{m + n}\right)^r. \tag{4}$$

Proof. Suppose, as before, we have a jar containing a black ball, n blue balls, and $m - 1$ red balls. Generalizing our argument for (2), imagine a scenario in which we select all the balls from the jar, one-by-one, until the jar is empty. Then we place all the balls into a second jar and remove them one-by-one. Repeat this process until we obtain a total of r sequences of balls chosen from r jars. The right-hand side of (4) is the probability that the black ball is the first ball chosen from each of the r jars. Using Theorem 2 and with the proof of (2) in mind, we take B to be the event that, for all r jars, the black ball is chosen before any of the red balls and A_i to be the event that, for all r jars, the black ball is chosen before blue ball i. This gives

$$P\left(B \cap \bigcap_{i=1}^{n} A_i\right) = \left(\frac{1}{m + n}\right)^r.$$

Applying the right side of Theorem 2, we have $\bar{A}_i$ as the event that the black ball is chosen after blue ball i for at least one of the r jars. The intersection of B with any k of the $\bar{A}_i$'s is thus the event that, for all r jars, the black ball is chosen before any of the red balls but that for each of k specific blue balls, there is at least one jar such that the black ball is drawn after that blue ball from that jar. This is a rather complicated event, but we can make calculating its probability simpler by breaking it down in the following manner: Each of the k blue balls has to have a first jar for which it is drawn before the black ball. Let $i_1, i_2, \ldots, i_r$ denote the number of these k balls for which, respectively, jar $1, 2, \ldots, r$ is the first jar that the ball is drawn before the black ball. The number of ways to choose which of the k balls are to be counted by i_1, which by i_2, and so forth is given by the multinomial coefficient $\binom{k}{i_1, i_2, \ldots, i_r}$. For the first jar, we must have i_1 of the blue balls being drawn first, then the black ball, and finally the $m - 1 + k - i_1$ red balls and remaining blue balls. The probability this occurs is

$$\frac{i_1!\,(m-1+k-i_1)!}{(m+k)!}.$$

For the second jar, we do not care when the i_1 blue balls that were chosen before the black ball in jar 1 are selected, so the event we need for the second jar entails i_2 of the blue balls being drawn first, then the black ball, and finally the $m - 1 + k - i_1 - i_2$ red balls and remaining blue balls. The probability of this event is

$$\frac{i_2!\,(m-1+k-i_1-i_2)!}{(m+k-i_1)!}.$$

This pattern continues until the last jar. For the last jar, we need the i_r blue balls that remain drawn first, then the black ball, and finally the $m - 1$ red balls. The probability of this is

$$\frac{i_r!\,(m-1)!}{(m+i_r)!}.$$

Putting it all together and adding up over all possible values of $i_1, i_2, \ldots, i_r$, we have that the probability of the event associated with the intersection of B and any specific k of the $\bar{A}_i$'s is

$$\begin{aligned}
&\sum_{i_1+i_2+\cdots+i_r=k} \binom{k}{i_1, i_2, \ldots, i_r} \frac{i_1!\,(m-1+k-i_1)!}{(m+k)!} \cdot \frac{i_2!\,(m-1+k-i_1-i_2)!}{(m+k-i_1)!} \cdot \\
&\quad \cdots \frac{i_r!\,(m-1)!}{(m+i_r)!} \\
&= \sum_{i_1+i_2+\cdots+i_r=k} \frac{k!}{i_1!i_2!\cdots i_r!} \frac{(m-1)!}{(m+k)!} \frac{i_1!i_2!\cdots i_r!}{(m+k-i_1)(m+k-i_1-i_2)\cdots(m+i_r)} \\
&= \frac{(m-1)!\,k!}{(m+k)!} \sum_{i_1+i_2+\cdots+i_r=k} \frac{1}{(m+i_1)(m+i_1+i_2)\cdots(m+i_1+i_2+\cdots+i_{r-1})} \\
&= \frac{(m-1)!\,k!}{(m+k)!} \sum_{0\le j_1\le\cdots\le j_{r-1}\le k} \frac{1}{(m+j_1)(m+j_2)\cdots(m+j_{r-1})},
\end{aligned}$$

where we reverse the indexing of the i_x variables in the second-to-last step and let $j_x = i_1 + i_2 + \cdots + i_x$ in the final step. With $\binom{n}{k}$ ways to choose k of the n blue balls, Theorem 2 says that the left-hand side of (4) is also the probability that the black ball is chosen first for all n jars. ∎

It is now fairly straightforward to prove (3):

$$\sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{m+k}\right)^{r} (-1)^{k}$$
$$= \frac{(m-1)!\,n!}{(m+n)!} \sum_{0 \le k_1 \le \cdots \le k_{r-1} \le n} \frac{1}{(m+k_1)(m+k_2)\cdots(m+k_{r-1})}. \tag{3}$$

Proof. Suppose we have, as before, a black ball, n blue balls, and $m-1$ red balls in a jar. Suppose also, as before, that we choose the balls one-by-one from that jar. Then we place all the balls into a second jar and remove them one-by-one. Repeat this process for r jars. The right-hand side of (3) (as we can see from the proof of (4)) is the probability that, for all r jars, the black ball is drawn before any of the red balls and that, for each blue ball, there exists at least one jar for which the black ball is drawn after that blue ball. Let B be the event that, for all r jars, the black ball is drawn before any of the red balls, and let A_i be the event that the black ball is drawn after blue ball i in at least one of the jars. (As we noted would be the case in our proof of (1a), the event A_i here is the complement of the event we called "A_i" in the proof of (4).) We have that $P\left(B \cap \bigcap_{i=1}^{n} A_i\right)$ is the right-hand side of (3).

Then $\bar{A}_i$ is the event that, for all r jars, the black ball is drawn before blue ball i. The intersection of B with any k of the $\bar{A}_i$'s is thus the event that, for all r jars, the black ball is drawn as the first of $m+k$ specific balls. This probability is $\left(\frac{1}{m+k}\right)^{r}$. There are $\binom{n}{k}$ ways to select k of the $\bar{A}_i$'s, and so, by Theorem 2, the the left-hand side of (3) gives the probability of the same event as does the right-hand side. ∎

6. FINAL COMMENTS. Of course, the identities proved here can be proved by a variety of other means. For example, Peterson mentions proofs using hypergeometric functions, the Chu–Vandermonde identity, and the Rice integral formulas. In addition, (1) can be viewed from a variety of angles: It is the formula for the nth forward difference of the function $f(m) = 1/m$ (see, for example, Graham, Knuth, and Patashnik [**2**, pp. 188–189]). It also gives the partial fractions decomposition of $n!/(m(m+1)\cdots(m+n))$. Finally, it is one of the more well-known series representations of the beta function for nonnegative integers n, where the right-hand side is $B(m, n+1)$ (see, for example, Gradshteyn and Ryzhik [**1**, p. 950]).

ACKNOWLEDGMENT. The proof of (1a) presented here was first given by the author on the mathematics question-and-answer site, Mathematics Stack Exchange [**4**].

REFERENCES

1. I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, Orlando, FL, 1980.
2. R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics*. Second edition. Addison-Wesley, Reading, MA, 1994.
3. J. Peterson, A probabilistic proof of a binomial identity, *Amer. Math. Monthly* **120** no. 6 (2013) 558–562.
4. M. Spivey, Is there a combinatorial way to see the link between the beta and gamma functions? (version: 2011-10-13), Mathematics Stack Exchange, http://math.stackexchange.com/q/72135.

Department of Mathematics and Computer Science, University of Puget Sound, Tacoma, WA 98416-1043
mspivey@pugetsound.edu