

## A Recursive Formula for Moments of a Binomial Distribution

Árpád Bényi (benyi@math.umass.edu), University of Massachusetts, Amherst, MA 01003 and Saverio M. Manago (smmanago@nps.navy.mil) Naval Postgraduate School, Monterey, CA 93943

While teaching a course in probability and statistics, one of the authors came across an apparently simple question about the computation of higher order moments of a random variable. The topic of moments of higher order is rarely emphasized when teaching a statistics course. The textbooks we came across in our classes, for example, treat this subject rather scarcely; see [3, pp. 265–267], [4, pp. 184–187], also [2, p. 206]. Most of the examples given in these books stop at the second moment, which of course suffices if one is only interested in finding, say, the dispersion (or variance) of a random variable  $X$ ,  $D^2(X) = M_2(X) - M(X)^2$ . Nevertheless, moments of order higher than 2 are relevant in many classical statistical tests when one assumes conditions of normality. These assumptions may be checked by examining the skewness or kurtosis of a probability distribution function. The skewness, or the first shape parameter, corresponds to the the third moment about the mean. It describes the symmetry of the tails of a probability distribution. The kurtosis, also known as the second shape parameter, corresponds to the fourth moment about the mean and measures the relative peakedness or flatness of a distribution. Significant skewness or kurtosis indicates that the data is not normal.

However, we arrived at higher order moments unintentionally. The discussion that follows is the outcome of a question raised by our students without knowing much about the motivation for studying moments of higher order. It has to do with discrete random variables

$$X \sim \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ p_1 & p_2 & \dots & p_m \end{pmatrix}$$

and their  $p$ -moments

$$M_p(X) = M(X^p) = \sum_{i=1}^m x_i^p p_i. \quad (1)$$

Here,  $p$  denotes a positive integer; on the first line of the table of  $X$  we have the values of the random variable in increasing order, that is  $x_1 < x_2 < \dots < x_m$ , while on the second line we read the probabilities with which  $X$  takes the values  $x_i$ , namely

$$p_i = \mathbf{P}(\{X = x_i\}), \quad 1 \leq i \leq m.$$

We note also that for  $1 \leq i \leq m$ ,  $p_i \in [0, 1]$  and  $\sum_{i=1}^m p_i = 1$ .

Let  $n \geq 1$  be an integer. Given now the discrete random variable

$$X_n \sim B\left(n, \frac{1}{2}\right) \quad \text{such that for all } k = 1, \dots, n, \quad \mathbf{P}(\{X_n = k\}) = \frac{1}{2^n} \binom{n}{k},$$

the students were asked to compute the first moment

$$M(X_n) = \frac{1}{2^n} \sum_{k=0}^n k \binom{n}{k}. \quad (2)$$

The  $k$ th column of the variable  $X_n$  refers to the event of drawing  $k$  white balls out of a bowl containing a total of  $n$  (equally numbered) white and black balls. In this drawing scheme—a particular case of the binomial or Bernoulli scheme—we require that an extracted ball is reintroduced in the bowl after each drawing. Going back to the computation of the first moment (2), it is easy to see that

$$M(X_n) = \frac{1}{2^n} \sum_{k=1}^n n \binom{n-1}{k-1} = \frac{n}{2^n} 2^{n-1} = \frac{n}{2}. \quad (3)$$

The transition from the first to the second equality follows from the well-known Newton binomial identity

$$2^q = (1+1)^q = \sum_{r=0}^q \binom{q}{r}, \quad q \in \mathbf{N}.$$

After explaining the calculation of (2), as a counter-question from the students, it was asked whether the same kind of argument would give a simplified answer to  $M_2(X_n)$  or even, say,  $M_{2003}(X_n)$ ? And how about an answer in closed form for  $M_p(X_n)$ ,  $p \geq 1$ , in general?

In what follows, we plan to answer—satisfactorily, we hope—the former question by displaying a recurrence relation for the general  $p$ -moments. The reader should note that the recursive formula is useful for calculations using pencil and paper as long as  $p$  is in a relatively small range. Observe also that, even for the particular case of  $X_n$  in discussion, the recursion does not fall into a very nice shape. Most of the students would rarely think of using pencil and paper for such computations considering the significant use of software such as STS when teaching a statistics class. However, our decision of writing this short note was made by a need of showing our students what we as mathematicians do in certain situations: we try to prove rigorously and in most generality the observations we make by looking at several particular cases displayed on a computer screen.

Our main goal is to show the following result:

**Proposition.** *Let  $p \geq 1$ . Then*

$$(a) \quad M_p(X_n) = n \left( M_{p-1}(X_n) - \frac{1}{2} M_{p-1}(X_{n-1}) \right); \quad (4)$$

$$(b) \quad M_p(X_n) = \frac{n}{2^p} Q_{p-1}(n), \text{ where } Q_{p-1}(n) \text{ is the monic polynomial of degree } p-1 \text{ in the variable } n \text{ which solves the recursion } Q_p(n) = 2nQ_{p-1}(n) - (n-1)Q_{p-1}(n-1).$$

*Proof.* Let us re-denote

$$S(n, p) = 2^n M_p(X_n) = \sum_{k=0}^n k^p \binom{n}{k}.$$

Since

$$k^p C_n^k = k^{p-1} k \binom{n}{k} = nk^{p-1} \binom{n-1}{k-1},$$

we can write

$$\begin{aligned}
 S(n, p) &= n \sum_{k=1}^n k^{p-1} \binom{n-1}{k-1} = n \sum_{k=1}^n k^{p-1} \left( \binom{n}{k} - \binom{n-1}{k} \right) \\
 &= n \left( \sum_{k=0}^n k^{p-1} \binom{n}{k} - \sum_{k=0}^n k^{p-1} \binom{n-1}{k} \right)
 \end{aligned}$$

or

$$S(n, p) = n(S(n, p-1) - S(n-1, p-1)); \tag{5}$$

in the second sum above we used the convention  $\binom{n-1}{n} \equiv 0$ . It is clear that (5) translates to the  $p$ -moments exactly as the recursion (a).

We now prove part (b). Newton's binomial formula gives  $S(n, 0) = 2^n$ . From here and (5), we get

$$S(n, 1) = n(S(n, 0) - S(n-1, 0)) = n(2^n - 2^{n-1}) = n2^{n-1}.$$

Note that, in particular, we recover the computation of the first moment (3). Similarly, one gets  $S(n, 2) = n(n+1)2^{n-2}$ ,  $S(n, 3) = n^2(n+3)2^{n-3}$ , etc. The formulas obtained for the first four sums  $S(n, p)$ ,  $0 \leq p \leq 3$ , suggest that

$$S(n, p) = n2^{n-p} Q_{p-1}(n), \tag{6}$$

where  $Q_{p-1}(n)$  is some polynomial of degree  $p-1$  in the variable  $n$  with unitary leading coefficient. Indeed, one can prove this using (5) and an induction argument on  $p$ . Assuming that (6) holds true, we have

$$\begin{aligned}
 S(n, p+1) &= n(n2^{n-p} Q_{p-1}(n) - (n-1)2^{n-1-p} Q_{p-1}(n-1)) \\
 &= n2^{n-p-1} (2n Q_{p-1}(n) - (n-1) Q_{p-1}(n-1)) = n2^{n-(p+1)} Q_p(n),
 \end{aligned}$$

where we denoted

$$Q_p(n) = 2n Q_{p-1}(n) - (n-1) Q_{p-1}(n-1).$$

By our inductive assumption,  $Q_{p-1}(n) = n^{p-1} +$  lower order terms. To complete the induction step, we only need to show that  $Q_p(n)$  is a polynomial of degree  $p$  in the variable  $n$  with unitary leading coefficient. To see this, note that the leading term of  $Q_p(n)$  comes from the difference

$$2nn^{p-1} - (n-1)(n-1)^{p-1} = n^p + n^p - (n-1)^p = n^p + \text{lower order terms},$$

which proves our claim. Observe that (6) can be rewritten as

$$2^n M_p(X_n) = n2^{n-p} Q_{p-1}(n),$$

which is equivalent to the statement (b). ■

We remark that finding explicitly the polynomials  $Q_p(n)$  for a general positive integer  $p$  turns out to be a nontrivial task. The interested reader should check for example that already for  $M_4(X_n)$ , the polynomial  $Q_3(n)$  does not factor completely over

the integers as its predecessors. In fact,  $Q_3(n) = (n + 1)(n^2 + 5n - 2)$ . Nevertheless, the recursion (b) can be used to infer certain properties of the polynomials  $Q_p(n)$  which define the moments of  $X_n$ . For example, using again an induction argument on  $p$ , one can show that the polynomial  $Q_p$  has 0 as a root if  $p$  is even, and  $-1$  as a root if  $p$  is odd. In particular, we can infer further information about the shape of the moments, namely that for all  $p \geq 1$ ,  $n(n + 1)$  divides  $2^{2p}M_{2p}(X_n)$  and  $n^2$  divides  $2^{2p+1}M_{2p+1}(X_n)$ .

It is also worth pointing out that there exists an alternative approach to the computation of the sums  $S(n, p)$  (and thus of the moments  $M_p(X_n)$ ). If  $s_1, s_2, \dots, s_{p-1}$  denote the symmetric sums associated to the numbers  $1, 2, \dots, p - 1$ , that is,  $s_1 = 1 + 2 + \dots + (p - 1)$ ,  $s_2 = 1 \cdot 2 + 1 \cdot 3 + \dots + (p - 2) \cdot (p - 1)$ ,  $\dots$ ,  $s_{p-1} = 1 \cdot 2 \cdot \dots \cdot (p - 1)$ , then the following recursion holds:

$$M_p(X_n) - s_1 M_{p-1}(X_n) + \dots + (-1)^p s_{p-1} M_1(X_n) = n(n - 1) \dots (n - p + 1) 2^{-p}. \quad (7)$$

Start by noting that for  $p \leq k$  we have

$$k(k - 1) \dots (k - p + 1) \binom{n}{k} = n(n - 1) \dots (n - p + 1) \binom{n - p}{k - p}.$$

Hence,

$$\begin{aligned} \sum_{k=0}^n k(k - 1) \dots (k - p + 1) \binom{n}{k} &= n(n - 1) \dots (n - p + 1) \sum_{k=p}^n \binom{n - p}{k - p} \\ &= n(n - 1) \dots (n - p + 1) 2^{n-p}, \end{aligned}$$

where we used yet again Newton's binomial identity to obtain the last equality. But

$$k(k - 1) \dots (k - p + 1) = k^p - s_1 k^{p-1} + \dots + (-1)^p s_{p-1} k,$$

where  $s_1, s_2, \dots, s_{p-1}$  are the symmetric sums defined above. If we now recall the definition of the sums  $S(n, p)$ , we obtain

$$\begin{aligned} S(n, p) - s_1 S(n, p - 1) + s_2 S(n, p - 2) + \dots + (-1)^p s_{p-1} S(n, 1) \\ = n(n - 1) \dots (n - p + 1) 2^{n-p}, \end{aligned}$$

which is equivalent to the recurrence relation (7). For more on the uses of recurrence relations and difference equations, the interested reader is referred to [1].

To conclude, answering the question raised by the students about  $M_{2003}(X_n)$  in a closed form using only pencil and paper is not convenient. The recursive formulas (4) and (7) however are useful from a computational point of view. While (7) looks more elegant, formula (4) seems a bit easier to work with since it does not require the a priori calculation of symmetric sums. In a way, pencil, paper and computers all come together to solve in a satisfactory manner our problem of, and for, the moment.

*Acknowledgment.* The authors thank the referee for some useful suggestions.

## References

1. G. E. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis, Forecasting and Control*, 3rd ed., Prentice Hall, 1994.

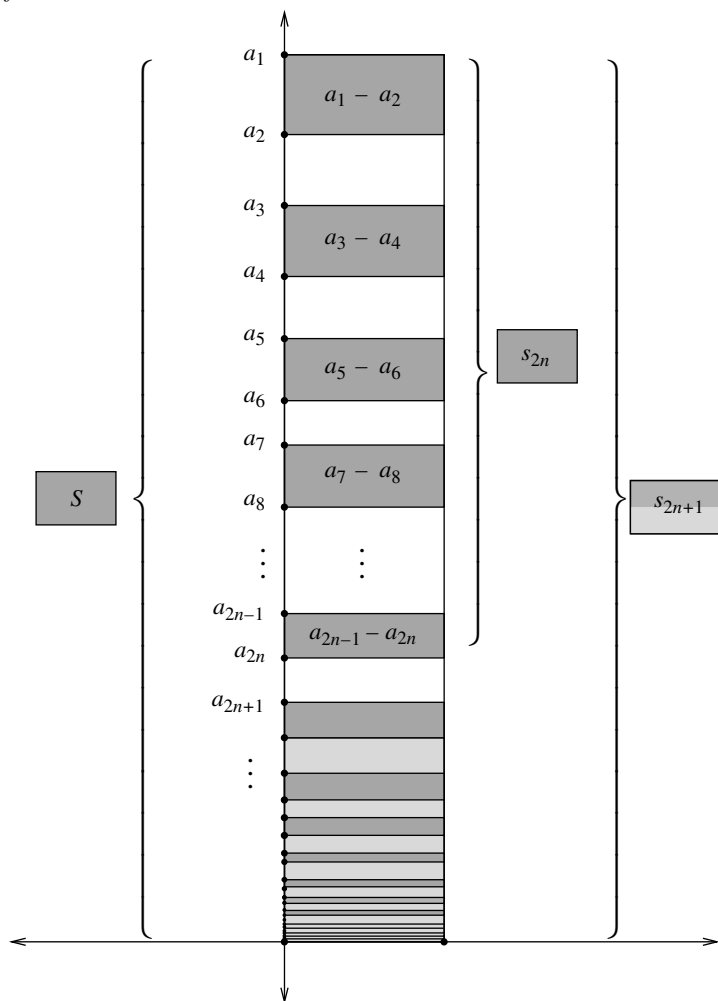
2. A. T. Craig and R. V. Hogg, *Introduction to Mathematical Statistics*, 5th ed., Macmillan, 1995.
3. J. L. Devore, *Probability and Statistics: for Engineering and the Sciences*, 4th ed., Brooks/Cole, 1995.
4. R. J. Larson and M. L. Marx, *An Introduction to Mathematical Statistics and Its Applications*, 3rd ed., Prentice Hall, 2001.

## Proof Without Words

From Richard Hammack (rhammack@rmc.edu) of Robert Madison University and David Lyons (lyons@lvc.edu) of Lebanon Valley College:

**Theorem.** An alternating series  $a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + a_7 - a_8 + \dots$  converges to a sum  $S$  if  $a_1 \geq a_2 \geq a_3 \geq a_4 \geq \dots \geq 0$  and  $a_n \rightarrow 0$ . Moreover, if  $s_n = a_1 - a_2 + a_3 - \dots \pm a_n$  is the  $n$ th partial sum then  $s_{2n} < S < s_{2n+1}$ .

*Proof.*



A different version of this result (with words) is to appear in the *Oxford Journal of Teaching Mathematics and Its Application*.