Encore We cannot resist one more picture. Let

$$f(x) = x^{1/5}(x-1)^{2/3}/(x-2)^{1/7},$$

let $n = 76/105$, and let $h = 10/19$. Below are the graphs of $y = f(x)$ (solid) and $y = (x-h)^n$ (dashed).

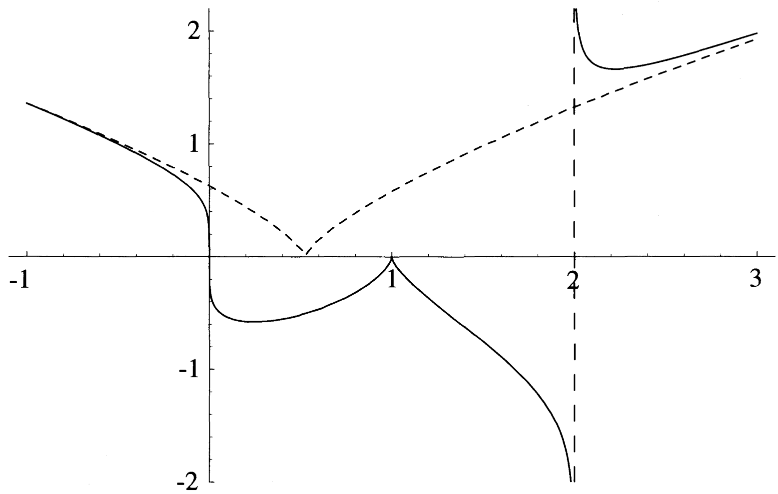

Figure 4 A more general example

Duality and Symmetry in the Hypergeometric Distribution

JAMES JANTOSCIAK
Brooklyn College (CUNY)
Brooklyn, NY 11210
jamesj@brooklyn.cuny.edu

WILLIAM BARNIER
Sonoma State University
Rohnert Park, CA 94928
bill.barnier@sonoma.edu

A CNN posting on the internet **[1]** reports that the 1996 trial of a rap star on manslaughter charges resulted in a hung jury. The posting says that the jury was composed of 7 men and 5 women and was hung at 9 to 3. Not reported is how many men or women voted with the majority. Several interesting probability problems come to mind. For example, what is the probability that exactly 5 jurors among the majority are men?

The usual solution for such a problem utilizes methods that are associated with the hypergeometric probability distribution and involves designating successes and choosing a sample from a population. For the problem stated, it is clear that the population is the 12 jurors. But for the men and the majority, it is not so clear which should be considered the successes and which the sample; either way seems artificial and contrived. We solve this problem later, in a more natural way, using a symmetric form of the hypergeometric distribution derived in this note.

Using the classical terminology, as we have, for describing the hypergeometric distribution, we consider a population of size n, a number i of which are specified as successes. Our probability experiment is the random choice of a sample of size j from the population. We would like to know the probability that the random variable X, which is the number of successes among the sample, has value x. This is given by the hypergeometric distribution, specifically,

$$P(X=x)=\binom{i}{x}\binom{n-i}{j-x}\Big/\binom{n}{j}$$

when $0, i+j-n \leq x \leq i, j$.

The hypergeometric distribution admits what may be called *dual* and *symmetric* forms. In the dual formulation, we can interchange the sample with the successes and arrive at the equally valid

$$P(X=x)=\binom{j}{x}\binom{n-j}{i-x}\Big/\binom{n}{i}.$$

This duality has been brought out in these Notes by Davidson and Johnson [**2**], is mentioned by Feller [**3**, page 44], and is easily verified by expanding all the binomial coefficients into factorials. Common examples where duality emerges naturally include lottery games [**4**, Exercises 9 and 10, page 172] and selections by lot. Here is an instance of the latter.

Three persons are chosen by lot from a group of 10. What is the probability that a specific person from the group is chosen?

Here, a sample of 3 is chosen from a population of 10 that includes 1 success, the specific person. Hence, the probability is $P(X=1)=\binom{1}{1}\binom{9}{2}/\binom{10}{3}=\frac{3}{10}$. Dually, the specific person, the sample, has 3 chances, the successes, out of 10 of being chosen. Thus, $P(X=1)=\frac{3}{10}=\binom{3}{1}\binom{7}{0}/\binom{10}{1}$.

Returning to our beginning example, we see that the way the men and majority are regarded as the successes and sample does not matter, and once a decision is made, the probability is easily determined. However, having no grounds to decide, we prefer to respect the symmetry in the problem.

By treating the successes and the sample symmetrically, we derive another formula for the hypergeometric distribution. Consider the population as a set U in which the members are *independently* classified in two ways yielding the subsets I and J of successes and of sample points, respectively. Then X gives the size of the intersection $I \cap J$. We note that I and J are combinations of sizes i and j chosen from U. To find $P(X=x)$, we consider all such pairs (I, J) of these combinations as equally likely outcomes. The size of the sample space of all these outcomes is

$$\binom{n}{i}\binom{n}{j}.$$

The event that $I \cap J$ has size x corresponds to all the ways of partitioning U into four subsets $I \cap J$, $I \cap \bar{J}$, $\bar{I} \cap J$, and $\bar{I} \cap \bar{J}$ of sizes x, $i-x$, $j-x$, and $n-i-j+x$, respectively (FIGURE 1). Recall that the multinomial coefficient

$$\binom{n}{n_1, n_2, n_3, n_4}=\frac{n!}{n_1!\, n_2!\, n_3!\, n_4!}$$

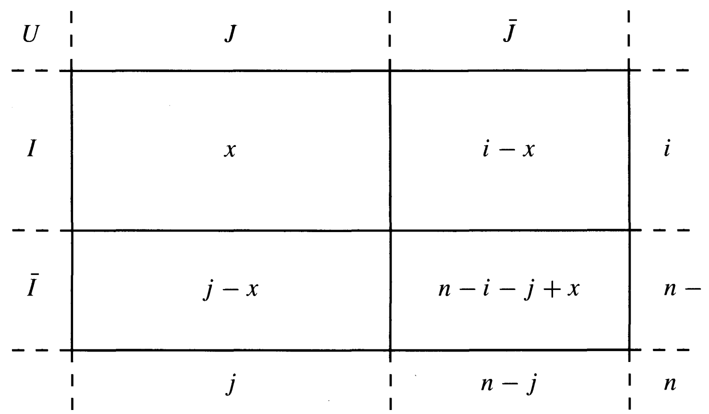

Figure 1 Partition of U

counts the number of ways to partition a set of n elements into 4 disjoint subsets with n_k elements in the kth part. It is ideal for counting the possible partitions of the kind indicated in the diagram. There are

$$\binom{n}{x,\ i-x,\ j-x,\ n-i-j+x}$$

such partitions. Therefore, the probability that $I \cap J$ has size x is equal to

$$P(X=x) = \binom{n}{x,\ i-x,\ j-x,\ n-i-j+x} \Big/ \binom{n}{i}\binom{n}{j}.$$

The use of this symmetric form of the hypergeometric distribution seems more natural than either of the dual ones in problems where no sample is chosen. Our opening example is such a problem.

We assume that being a man and being among the majority are independent events. Let U be the group of jurors, I the subgroup of men, and J the subgroup of the majority. Therefore, the probability that exactly 5 jurors among the majority are men is

$$P(X=5) = \binom{12}{5,\ 2,\ 4,\ 1} \Big/ \binom{12}{7}\binom{12}{9} = \frac{21}{44}.$$

Here is another symmetric problem in which choosing a sample and designating successes would be artificial. A coed softball team of 17 players includes 8 females and 14 players who throw right-handed. Assuming independence of these traits, what is the probability that exactly 6 players are right-handed-throwing females? The solution can be found on page 143.

Acknowledgment. We thank the anonymous referees for their useful suggestions.

REFERENCES

1. CNN, *Hung jury in rap star's manslaughter trial*, http://www.cnn.com/US/9602/snoop.
2. R. R. Davidson and B. R. Johnson, Interchanging parameters of the hypergeometric distribution, this MAGAZINE **66** (1993), 328–329.
3. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd Ed., John Wiley, New York, 1968.
4. H. J. Larson, *Introduction to Probability*, Addison-Wesley, New York, 1995.

We see that any prime divisor of $P - 2d$ that does not divide $16s$ must divide P and hence d. Thus, the number of distinct primes of $P - 2d$ is limited to those of $2s$ and d. As in the earlier cases, (8) shows that the multiplicity of such a prime is bounded. [Let $d = q^i u$ and $P - 2d = q^j v$, where q is a prime, i and j are maximal, and $j > i$. Then $P = q^i(2u - q^{j-i}v)$. Equation (8) shows that j can not exceed $2i$ + the number of factors of q in $16s$.] For a given value k, $P - 2d$ (and hence P) is bounded. This shows that $N(k)$ is finite.

REFERENCES

1. N. Altshiller-Court, *College Geometry*, 2nd ed., Barnes & Noble, New York, 1952.
2. L.E. Dickson, *History of the Theory of Numbers*, Vol. II, Chelsea, New York, 1952.
3. N. J. Fine, Solution to a problem of Phelps, *Amer. Math. Monthly* **63** (1956), 43–44.
4. I. Niven, *Maxima and Minima without Calculus*, Math. Assoc. of America, Washington, DC, 1981.
5. R. P. Phelps, Problem posed by, *Amer. Math. Monthly* **62** (1955), 365.
6. M. V. Subbarao, Perfect triangles, *Amer. Math. Monthly* **72** (1971), 384.

A Hypergeometric Problem Solved (from p. 137)

A coed softball team of 17 players includes 8 females and 14 players who throw right-handed. Assuming independence of these traits, what is the probability that exactly 6 players are right-handed-throwing females?

Solution.

$$\binom{17}{6,\ 2,\ 8,\ 1} \Big/ \binom{17}{8}\binom{17}{14} = \frac{63}{170}.$$