JOURNAL OF COMBINATORIAL THEORY, Series A **68**, 212–214 (1994)

Note

A Probabilistic Proof of Gauss' $_2F_1$ Identity

ROBERTA S. WENOCUR

University of Pennsylvania, Department of Statistics, Philadelphia, Pennsylvania 19104

Communicated by the Managing Editors

Received December 4, 1992

Using order statistics, we prove Gauss' $_2F_1$ identity probabilistically. As a consequence, we show that Gauss' $_2F_1$ summation formula is related to an inverse Pólya distribution. We observe that a relationship exists between *WZ*-pairs and our probabilistic approach.

INTRODUCTION

Gauss' $_2F_1$ summation theorem often appears as

$$ {}_2F_1[a,b;c;1] = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \tag{1} $$

for $c > a + b$, where $_2F_1[a, b; c; 1]$ is defined by

$$ {}_2F_1[a,b;c;1] = \sum_{k=0}^{\infty} \frac{a \cdots (a+k-1) b \cdots (b+k-1)}{c(c+1) \cdots (c+k-1) k!}; $$

see, for example, Slater [4, pp. 1, 27–28]. The primary purpose of this paper is to present a probabilistic proof of (1) where a, b, and c are positive integers.

PROBABILISTIC PROOF

Fix positive integers N, j, m, k, with $1 \le j \le N$. The probabilistic approach involves identifying $\{p_k\}_{k=0}^{\infty}$ as the probabilities of pairwise mutually exclusive exhaustive events $\{\mathscr{E}_k\}_{k=0}^{\infty}$, respectively, and proceeding to prove Gauss' identity in the form $\sum_{k \ge 0} p_k = 1$.

Draw N real numbers independently and uniformly from the interval $(0, 1)$ and let $Y_{(j)}$ denote the jth smallest of them. Then, continue

0097-3165/94 $6.00

sampling such random numbers until $Y_{(j)}$ is exceeded by m of the newly drawn values, for the first time.

Let $\mathscr{E}_k$ denote the event that exactly $m + k$ trials beyond the first N are needed, with p_k equal to the probability of event $\mathscr{E}_k$. It is apparent that p_k is the probability that a permutation of $N + m + k$ letters has exactly m of its last $m + k$ values, including its final value, exceeding the jth smallest of its first N values. To determine p_k, proceed as follows:

Obtain the number of permutations σ, of $N + m + k$ letters, that are of the following form:

(a) Among the values $\sigma_1, \ldots, \sigma_N$, exactly $j - 1$ values are less than r, r is some value, and $N - j$ values are greater than r.

(b) Among the values $\sigma_{N+1}, \ldots, \sigma_{N+m+k-1}$, exactly $m - 1$ are greater than r and k are less than r.

(c) The value σ_{N+m+k} is greater than r.

The total number of values less than r must be equal to $r - 1$, but it is also equal to $j - 1 + k$, so $r = j + k$. Hence restate (a) through (c) as follows:

(a) Among the (*past*) values $\sigma_1, \ldots, \sigma_N$, exactly $j - 1$ values are less than $j + k$, $j + k$ is a value, and $N - j$ values are greater than $j + k$.

(b) Among the (*future*) values $\sigma_{N+1}, \ldots, \sigma_{N+m+k-1}$, exactly $m - 1$ are greater than $j + k$, and k are less than $j + k$.

(c) the value σ_{N+m+k} is greater than $j + k$.

Now construct all such permutations. Choose the $j - 1$ *past* values that are less than $j + k$ (in $\binom{k+j-1}{j-1}$ ways), then choose the $N - j$ *past* values that are greater than $k + j$ (in $\binom{N+m-j}{N-j}$ ways); arrange the chosen *past* values in some sequence (in $N!$ ways), then choose which one of the m remaining values that are greater than $k + j$ shall be the final value σ_{N+m+k} (in m ways), and finally, arrange the *future* values in sequence (in $(m + k - 1)!$ ways). Multiply all of these counts together and divide by $(N + m + k)!$ to obtain the probability of such a permutation, namely

$$p_k = \frac{(m + k - 1)!(N + 1 - j) \cdots (N + m - j)(j) \cdots (j + k - 1)}{k!(m - 1)!(N + 1) \cdots (N + m)(N + m + 1) \cdots (N + m + k)}$$

$$= \frac{(N + m - j)!N!}{(N + m)!(N - j)!} \cdot \frac{(m)(m + 1) \cdots (m + k - 1)(j) \cdots (j + k - 1)}{k!(N + m + 1) \cdots (N + m + k)}.$$

The p_k's are equal to the probabilities for an inverse Pólya distribution; see, for example, [1; 2, pp. 194–200; 6]. In particular, the p_k's sum to 1. Since

$$\sum_{k=0}^{\infty} p_k = 1,$$

(1) follows. If we set $a = m$, $b = j$, and $c = N + m + 1$, then $N + m + 1 > j + m$ is equivalent to Gauss' condition $c > a + b$.

Remarks

• For earlier probabilistic proofs of identities in the literature see [3, 5]. These both use order statistics, as does the present proof.

• Many combinatorial identities can now be verified by means of the powerful Wilf–Zeilberger certification theorems [7], which provide an elegant method for certifying couples of identities via *WZ*-pairs. A probabilistic approach provides a counterpoint to the method of *WZ*-pairs; begin that method also, when appropriate, by dividing the identity to be proved by its right hand side, to obtain a sum of terms that is equal to 1.

Acknowledgments

Many thanks to Herbert S. Wilf and to the referee for a number of comments that have improved the presentation.

References

1. K. G. Janardan and G. P. Patil, On multivariate modified Pólya and inverse Pólya distributions and their properties, *Ann. Inst. Statist. Math.* **26** (1974), 271–276.
2. N. L. Johnson and S. Kotz, "Urn Models and Their Application," Wiley, New York, 1977.
3. K. W. J. Kadell, A probabilistic proof of Ramanujan's ${}_1\psi_1$ sum, *SIAM J. Math. Anal.* **18** (1987), 1539–1548.
4. L. J. Slater, "Generalized Hypergeometric Functions," Cambridge Univ. Press, Cambridge, 1966.
5. R. S. Wenocur, Rediscovery and alternate proof of Gauss's identity, *Ann. Discrete Math.* **9** (1980), 79–82.
6. R. S. Wenocur, Waiting times and return periods related to order statistics: An application of urn models, *Statist. Distrib. Sci. Work* **6** (1981), 419–434.
7. H. S. Wilf and D. Zeilberger, Rational functions certify combinatorial identities, *J. Amer. Math. Soc.* **3** (1990), 147–158.