© Journal of Applied Mathematics & Decision Sciences, 1(2), 151–157 (1997)
Reprints Available directly from the Editor. Printed in New Zealand.

Classroom Note

An Inductive Derivation of Stirling Numbers of the Second Kind and their Applications in Statistics

ANWAR H. JOARDER†
Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia.

MUNIR MAHMOOD
Department of Econometrics, Monash University, Clayton, Victoria, Australia 3168.

Abstract. An inductive method has been presented for finding Stirling numbers of the second kind. Applications to some discrete probability distributions for finding higher order moments have been discussed.

Keywords: Stirling numbers of the second kind, raw moments, factorial moments.

1. Introduction

Stirling numbers of the second kind are used to express x^n, where n is a non-negative integer, as sums of powers of factorial polynomials not higher than the n-th. For every non-negative integer n, the function x^n can be expressed as

$$x^n = S(n,1)x_{(1)} + S(n,2)x_{(2)} + \ldots + S(n,n)x_{(n)} \tag{1.1}$$

where $x_{(r)} = x(x-1)\ldots(x-r+1)$ is the factorial polynomial of order $r(r = 1, 2, \ldots, n)$, and $S(n,1), S(n,2), \ldots, S(n,n)$ are called Stirling numbers of the second kind. It may be mentioned that $S(n,r)$ is the number of ways to distribute n distinguishable balls into r indistinguishable urns with no urn empty. These numbers are called Stirling numbers of the second kind after the British mathematician James Stirling (1692 - 1770).

In this note we present an inductive proof to derive Stirling numbers of the second kind. For a combinatorial proof the reader is referred to Roberts (1984, pp. 182-183). We also demonstrate the application of Stirling numbers in calculating moments of some discrete distributions. The usual method of deriving raw moments of higher order of an integer-valued random variable is to derive the moment

† The work was done in part while the first author was at the University of Sydney, Australia.

generating function and then differentiate as many times as the order of the moment required. In this paper we take an advantage of factorial moments which are easily derived for integer-valued random variables. These moments can be combined with the help of Stirling numbers of the second kind for deriving raw moments and hence central moments of integer-valued random variables. The advantage of the method presented here for calculating raw moments or central moments of a distribution is that it avoids using derivatives of higher order.

2. The Description of the Method

Tables showing Stirling numbers up to $n = 8$ is well-known (see e.g. Beyer (1981, p.450)). These tables can be prepared by simple algebraic manipulations. Suppose that we want to calculate Stirling numbers in (1.1) for $n = 3$ so that

$$\begin{aligned} x^3 &= S(3,1)x_{(1)} + S(3,2)x_{(2)} + S(3,3)x_{(3)} \qquad (2.1) \\ &= S(3,1)x + S(3,2)x(x-1) + S(3,3)x(x-1)(x-2). \end{aligned}$$

Equating the coefficients of x, x^2 and x^3 we have the following triangular system of equations:

$$\begin{aligned} S(3,1) &- S(3,2) + 2\,S(3,3) = 0 \\ S(3,2) &- 3\,S(3,3) = 0 \\ S(3,3) &= 1. \end{aligned}$$

Solving the equations we have $S(3,1) = 1, S(3,2) = 3$ and $S(3,3) = 1$. If the above method of calculating Stirling numbers is repeated for $n = 2, 3$ etc., then it is easy to prepare a table showing Stirling numbers for different n and r. The following well-known recurrence relation (see e.g. Beyer, 1981, p. 450) is similar to the relation among binomial coefficients in Pascal triangle except a multiple to the second term

$$S(n,r) = S(n-1, r-1) + r\,S(n-1, r),\ r = 1, 2, \ldots, n \qquad (2.2)$$

with $S(n,1) = S(n,n) = 1$ and $S(n,r) = 0$, $r > n$. For a simple algebraic proof see Berman and Fryer (1972, pp. 217-218).

In this section we describe that the well-known formula for a Stirling number of the second kind given by

$$\begin{aligned} S(n,r) &= \frac{1}{r!}\sum_{i=0}^{r}(-1)^i \binom{r}{i} (r-i)^n \\ &= \frac{1}{r!}\sum_{i=0}^{r}(-1)^{r-i} \binom{r}{i} i^n \qquad (2.3) \end{aligned}$$

(Roberts, 1984, pp.182-183) where $0 \leq r \leq n$, and n is any non-negative integer can be derived from (2.2) by induction. The last line in the identity (2.3) follows by changing the index of summation i to j where $j = r - i$.

By the repeated use of (2.2), it is easy to check that the Stirling numbers of the second kind $S(n,r), r = 2,3,4$ can be written as

$$\begin{aligned}
S(n,2) &= \sum_{i=0}^{n-2} 2^i\, S(n-1-i,1) = 2^{n-1}-1, \;\; n \geq 2 \\
S(n,3) &= \sum_{i=0}^{n-3} 3^i\, S(n-1-i,2), \;\; n \geq 3 \\
&= \sum_{i=0}^{n-3} 3^i(2^{n-2-i}-1) = \frac{1}{6}\,[3^n - 3(2^n)+3] \\
S(n,4) &= \sum_{i=0}^{n-4} 4^i\, S(n-1-i,3), \;\; n \geq 4 \\
&= \sum_{i=0}^{n-4} 4^i \left[\frac{1}{6}\{3^{n-1-i} - 3(2^{n-1-i})+3\}\right] \\
&= \frac{1}{24}\;[4^n - 4(3^n) + 6(2^n) - 4] \\
&= \frac{1}{4!}\;\sum_{i=0}^{4}(-1)^i \binom{4}{i} (4-i)^n
\end{aligned}$$

where $\binom{r}{i}$ is the combination of r things taking i at a time.

We assume that the identity (2.2) is true and try to prove the identity in (2.3). The identity (2.3) holds for $S(n,0) = 0$ and $S(n,1) = 1$. It may be proved by induction that for all real numbers r

$$\sum_{i=0}^{n}(-1)^i \binom{n}{i} (r-i)^n = n!$$

(Ruiz, 1996) and consequently $S(n,n) = 1$. Applying induction hypothesis for $n-1$ and so $1 < r < n$, we have

$$\begin{aligned}
S(n-1,r-1) &= \frac{1}{(r-1)!}\;\sum_{i=0}^{r-1}(-1)^{r-1-i}\binom{r-1}{i}\, i^{n-1} \\
&= \frac{1}{(r-1)!}\;\sum_{i=0}^{r-1}(-1)(-1)^{r-i}(1-i/r)\binom{r}{i}\, i^{n-1}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{-1}{(r-1)!} \sum_{i=0}^{r-1} (-1)^{r-i} \binom{r}{i} i^{n-1} \\
&+ \frac{1}{(r-1)!} \sum_{i=0}^{r-1} (-1)^{r-i} \frac{i}{r} \binom{r}{i} i^{n-1}
\end{aligned}
$$

and

$$
\begin{aligned}
r\,S(n-1,r) &= r\,\frac{1}{r!} \sum_{i=0}^{r} (-1)^{r-i} \binom{r}{i} i^{n-1} \\
&= \frac{1}{(r-1)!} \sum_{i=0}^{r-1} (-1)^{r-i} \binom{r}{i} i^{n-1} + \frac{r^{n-1}}{(r-1)!}
\end{aligned}
$$

so that

$$
\begin{aligned}
S(n,r) &= S(n-1,\; r-1) + r\; S(n-1,r) \\
&= \frac{1}{(r-1)!} \sum_{i=0}^{r-1} (-1)^{r-i} \frac{i}{r} \binom{r}{i} i^{n-1} + \frac{r^{n-1}}{(r-1)!}
\end{aligned}
$$

The summand in the second term in the above expression is exactly the same as the first term if $i = r$ and consequently the two terms add to the expression in (2.3). The proof is thus complete.

The above proof has pedagogic values especially to readers who want to avoid combinatorial proof.

3. Applications

Stirling numbers connect the factorial moments to raw moments and vice versa. Taking expectations in the both sides of (1.1) with n and r replaced by r and i respectively, we have the general relationship between raw moments and factorial moments as

$$
\mu_r^{'} = \sum_{i=1}^{r} S(r,i)\;\mu_{(i)}^{'} \tag{3.1}
$$

where $\mu_r^{'} = E(X^r)$ is the r-th $(r = 1, 2, \ldots, n)$ raw moment and $\mu_{(r)}^{'} = E(X_{(r)})$ is the r-th $(r = 1, 2, \ldots, n)$ factorial moment. Harris (1966, p. 110) first used this formula clearly in the context of statistics though it has not been popular.

Stuart and Ord (1994, p.84) presented the formula (3.1) for $r \leq 4$, and also mentioned that Frisch gave general formulae showing factorial moments about one point in terms of ordinary moments about another, and vice versa. However, formulae given by Frisch involve Bernoulli polynomials and are not easy to use. The central moments are connected to the raw moments by the following relation:

$$\mu_r = \sum_{i=0}^{r} (-1)^i \binom{r}{i} \mu'_{r-i} (\mu'_1)^i. \tag{3.2}$$

The factorial moments of integer valued distributions e.g. binomial and negative binomial distributions can be easily derived. Then these moments can be combined by (3.1) and (3.2) to derive raw and central moments without having to resort to any differentiation. Some examples are presented below to illustrate the application.

Example 1. Let us calculate the fourth raw moment of the binomial distribution $B(n,p)$ by the use of (3.1). Here we have

$$\mu'_4 = S(4,1)\mu'_{(1)} + S(4,2)\mu'_{(2)} + S(4,3)\mu'_{(3)} + S(4,4)\mu'_{(4)}.$$

It follows from (2.3) that $S(4,1) = 1, S(4,2) = 7, S(4,3) = 6$ and $S(4,4) = 1$. The i-th factorial moment of the binomial distribution is easily shown to be $\mu'_{(i)} = n_{(i)}p^i$, $i = 1, 2, \ldots, n$. Therefore the fourth raw moment of the binomial distribution is given by

$$\mu'_4 = np + 7n_{(2)}p^2 + 6n_{(3)}p^3 + n_{(4)}p^4.$$

Example 2. Consider deriving the moments of the geometric distribution. Let X be the number of trials prior to the first success with probability density function (p.d.f.)

$$P(X = x) = q^x p, \;\; 0 < p < 1, \;\; q + p = 1, \;\; x = 0, 1, \ldots.$$

It can be easily verified that $\mu'_{(i)} = i!(q/p)^i$, $i = 1, 2, \ldots$. By the use of the formula (3.1) we have

$$\begin{aligned}
\mu'_1 &= \mu'_{(1)} = q/p, \\
\mu'_2 &= S(2,1)\mu'_{(1)} + S(2,2)\mu'_{(2)} = (q/p) + 2(q/p)^2, \\
\mu'_3 &= (q/p) + 6(q/p)^2 + 6(q/p)^3, \\
\mu'_4 &= (q/p) + 14(q/p)^2 + 36(q/p)^3 + 24(q/p)^4.
\end{aligned}$$

Now by the use of the moment relationship in (3.2) followed by some algebraic manipulations, we find that the second, third and fourth order central moments of the geometric distribution are

$$\begin{aligned}
\mu_2 &= qp^{-2}, \\
\mu_3 &= (q + q^2)p^{-3}, \\
\mu_4 &= (q + 7q^2 + q^3)p^{-4}.
\end{aligned}$$

Example 3. Let us derive moments of the negative binomial distribution. Let X be the number of trials prior to s-th ($s \geq 1$) success with p.d.f.

$$P(X = x) = \binom{s+x-1}{x} q^x p^s,\ q+p=1,\ 0<p<1,\ x=0,\ 1, \ldots$$

It can be easily verified that $\mu'_{(i)} = (s+i-1)_{(i)}(q/p)^i, i = 1, 2, \ldots$. By the use of the formula (3.1) we have

$$\begin{aligned}
\mu'_1 &= \mu'_{(1)} = sq/p, \\
\mu'_2 &= sq(1+sq)\,p^{-2}, \\
\mu'_3 &= s(q/p) + 3(s+1)_{(2)}\ (q/p)^2 + (s+2)_{(3)}\ (q/p)^3, \\
\mu'_4 &= s(q/p) + 7(s+1)_{(2)}\ (q/p)^2 + 6(s+2)_{(3)}\ (q/p)^3 + (s+3)_{(4)}(q/p)^4.
\end{aligned}$$

Then by the use of (3.2) we find that the second, third and fourth order central moments of the negative binomial distribution are

$$\begin{aligned}
\mu_2 &= sqp^{-2}, \\
\mu_3 &= s(q/p) + 3s(q/p)^2 + 2s(q/p)^3 \\
&= s(q+q^2)p^{-3}, \\
\mu_4 &= s(q/p) + s(7+3s)(q/p)^2 + 6s(2+s)(q/p)^3 + 3s(2+s)(q/p)^4 \\
&= s[q + (3s+4)q^2 + q^3)]p^{-4}
\end{aligned}$$

(cf. Evans, M. et al., 1993 p. 110).

Acknowledgments

The authors thank Dr. David Easdown, the University of Sydney and Dr. M. Iqbal, King Fahd University of Petroleum and Minerals for their constructive suggestions on an earlier draft of this paper. The authors are also grateful to the editor Dr. Mahyar Amouzegar and an anonymous referee for their constructive suggestions that considerably improved the presentation of this paper.

References

1. Berman, G. and Fryer, K.D. *Introduction to Combinatorics.* Academic Press New York, 1972.
2. Beyer, W.H. *Handbook of Mathematical Sciences.* CRC Press, Florida, 1981.
3. Evans, M., Hastings, N., and Peacock B. *Statistical Distributions.* John Wiley, New York, 1993.
4. Harris, B. *Theory of Probability.* Addison-Wesley Pub. Co., Massachusetts, 1966.

5. Roberts, F.S. *Applied Combinatorics.* Prentice Hall, New Jersey, 1984.
6. Ruiz, S.M. An algebraic identity leading to Wilson's theorem. *Mathematical Gazette*, 80, 579-582, 1996.
7. Stuart, A. and Ord, J.K. *Kendal's Advanced Theory of Statistics*, Vol. 1. Charles Griffin and Co. Ltd., London, 1994.