Some Combinatorial Identities on the Stirling Number Based on a Probabilistic Model

Tan Mingshu †

(Department of Mathematics, Chongqing Three-Gorges University, Chongqing 404000, P.R.China)

Abstract: An urn contains m distinguishable balls with m distinguishable colors. Balls are drawn for n times successively at random and with replacement from the urn. The mathematical expectation of the number of drawn colors is investigated. Some combinatorial identities on the Stirling number of the second kind $S(n, m)$ are derived by using probabilistic method.

Key words: Mathematical expectation; Stirling number of the second kind; Harmonic number; Combinatorial identity

AMS Classification(2000):05A19; 60C05

1 INTRODUCTION

It is an important method to find and prove some combinatorial identities on Stirling numbers based on a probabilistic model[1−7]. In

†Received date:2009-12-26, Tan Mingshu(Tan_mingshu@163.com)

Supported by Chongqing Education Committee of China (KJ081107)

this paper, We build a model as follows:

An urn contains m distinguishable balls with m distinguishable colors. Draw balls for n times successively at random and with replacement from the urn, which is noted as experiment T. We investigate the the number of drawn different colors or balls.

2 DISTRIBUTION

Let X be the number of drawn different balls(or colors) in experiment T($X = 1, 2, \cdots, m$), A_i represent the event that the ith kind ball(or color) is drawn($i = 1, 2, \cdots, m$), and $P(A_i)$ denote the probability of A_i.

$$
\begin{aligned}
&P(A_1A_2\cdots A_k) = 1 - P(\bar{A}_1 \cup \bar{A}_2 \cup \cdots \cup \bar{A}_k) \\
= &\ 1 - \binom{k}{1}P(\bar{A}_1) + \binom{k}{2}P(\bar{A}_1\bar{A}_2) - \cdots + (-1)^k\binom{k}{k}P(\bar{A}_1\bar{A}_2\cdots\bar{A}_k) \\
= &\ 1 - \binom{k}{1}\Big(\frac{m-1}{m}\Big)^n + \binom{k}{2}\Big(\frac{m-2}{m}\Big)^n - \cdots + (-1)^k\binom{k}{k}\Big(\frac{m-k}{m}\Big)^n.
\end{aligned}
$$

thus,

$$
P(A_1A_2\cdots A_k) = \sum_{k=0}^{k}(-1)^i\binom{k}{i}\Big(\frac{m-i}{m}\Big)^n.
$$

Note that

$$
\begin{aligned}
&P\Big(\bigcup_{l=k+1}^{m}(A_1A_2\cdots A_kA_l)\Big) \\
= &\sum_{i=1}^{m-k}(-1)^{i-1}\binom{m-k}{i}P(A_1A_2\cdots A_kA_{k+1}\cdots A_{k+i}) \\
= &\sum_{i=1}^{m-k}(-1)^{i-1}\binom{m-k}{i}P(A_1A_2\cdots A_{k+i}).
\end{aligned}
$$

so

$$\begin{aligned}P\{X=k\} &= \binom{m}{k}P(A_1A_2\cdots A_k\bar{A}_{k+1}\bar{A}_{k+2}\cdots\bar{A}_m)\\ &= \binom{m}{k}\Big[P(A_1A_2\cdots A_k)-P\Big(\bigcup_{l=k+1}^{m}(A_1A_2\cdots A_kA_l)\Big)\Big]\\ &= \binom{m}{k}\sum_{i=0}^{m-k}(-1)^i\binom{m-k}{i}P(A_1A_2\cdots A_{k+i})\\ &= \binom{m}{k}\sum_{i=0}^{m-k}(-1)^i\binom{m-k}{i}\sum_{j=0}^{k+i}(-1)^j\binom{k+i}{j}\Big(\frac{m-j}{m}\Big)^n.\end{aligned}$$

Especially, the probability of the all kinds of colors are drawn in n trials is

$$P\{X=m\}=\sum_{i=0}^{m}(-1)^i\binom{m}{i}\Big(\frac{m-i}{m}\Big)^n=\frac{m!S(n,m)}{m^n}. \tag{1}$$

where $n\geq m$, and $S(n,m)$ is the Stirling number of the second kind.

For (1), it is equivalent to n distinguishable balls are set into $m(n\geq m)$ distinguishable urns with no empty urn, then the number of ways is $m!S(n,m)$. It is also equivalent to a set with n elements is divided into m no empty subsets.

Thus, by the definition of $S(n,k)$, we have

$$P\{X=k\}=\frac{A_m^kS(n,k)}{m^n},\quad k=1,2,\cdots,\min(n,m).$$

Then, we obtain a combinatorial identity

$$\sum_{k=1}^{\min(n,m)}A_m^kS(n,k)=m^n. \tag{2}$$

where $A_m^k = m(m-1)\cdots(m-k+1)$.

Because $\sum_{k=1}^{m} P\{X=k\} = 1$, we have

$$\sum_{k=1}^{m}\sum_{i=0}^{m-k}\sum_{j=0}^{k+i}(-1)^{i+j}\binom{m}{k}\binom{m-k}{i}\binom{k+i}{j}\left(\frac{m-j}{m}\right)^n = 1.$$

that is,

$$m^n = \sum_{k=1}^{m}\sum_{i=0}^{m-k}\sum_{j=0}^{k+i}(-1)^{i+j}\binom{m}{k}\binom{m-k}{i}\binom{k+i}{j}(m-j)^n. \quad (3)$$

At the same time, by the distribution of X using different methods as above, a new explicit expression on $S(n,k)$ is derived:

$$S(n,k) = \frac{1}{k!}\sum_{i=0}^{m-k}\sum_{j=0}^{k+i}(-1)^{i+j}\binom{m-k}{i}\binom{k+i}{j}(m-j)^n. \quad (4)$$

Where $n \geq m \geq k$. For example, for $n=5, k=3$, (4) holds when $m=3$ or $m=4$ showed as follows.

$$S(5,3) = \frac{1}{3!}\sum_{i=0}^{3-3}\sum_{j=0}^{3+i}(-1)^{i+j}\binom{3-3}{i}\binom{3+i}{j}(3-j)^5$$

$$= \frac{1}{3!}\sum_{j=0}^{3}(-1)^j\binom{3}{j}(3-j)^5$$

$$= \frac{1}{3!}(3^5 - 3\times 2^5 + 3) = 25.$$

$$S(5,3) = \frac{1}{3!}\sum_{i=0}^{4-3}\sum_{j=0}^{3+i}(-1)^{i+j}\binom{4-3}{i}\binom{3+i}{j}(4-j)^5$$

$$= \frac{1}{3!}\left[\sum_{j=0}^{3}(-1)^j\binom{3}{j}(4-j)^5 + \sum_{j=0}^{4}(-1)^{1+j}\binom{4}{j}(4-j)^5\right]$$

$$= \frac{1}{3!}[4^5 - 3 \times 3^5 + 3 \times 2^5 - 1 - 4^5 + 4 \times 3^5 - 6 \times 2^5 + 4] = 25.$$

3 MATHEMATICAL EXPECTATION

Set $X_i = \begin{cases} 1, & \text{the } i\text{th kind color is drawn in } n \text{ trials} \\ 0, & \text{otherwise} \end{cases}$, then

$$X = X_1 + X_2 + \cdots + X_m.$$

Because $P\{X_i = 0\} = \left(\frac{m-1}{m}\right)^n$, then the mathematical expectation $EX = m\left(1 - (\frac{m-1}{m})^n\right)$. By $EX = \sum_{k=1}^{m} kP\{X = k\}$, we obtain

$$\sum_{k=1}^{m}\sum_{i=0}^{m-k}\sum_{j=0}^{k+i}(-1)^{i+j}\binom{m}{k}\binom{m-k}{i}\binom{k+i}{j}\left(\frac{m-j}{m}\right)^n k = m\left(1-(\frac{m-1}{m})^n\right). \tag{5}$$

$$\sum_{k=1}^{\min(n,m)} kA_m^k S(n,k) = m^{n+1} - m(m-1)^n. \tag{6}$$

For example, $n = 5, m = 4$, we have

$$\sum_{k=1}^{4} kA_4^k S(5,k) = 3124 = 4^6 - 4 \times 3^5.$$

4 FURTHER INVESTIGATION

Suppose we need at least Y trials to draw the all colors in experiment T, then $Y = m, m+1, \cdots$. Set $p_n = P\{X = n\}, q_n = P\{Y = n\}$, then

$$p_n = q_m + q_{m+1} + \cdots + q_{n-1} + q_n.$$

That is

$$\begin{aligned} q_n &= p_n - p_{n-1} \\ &= \frac{m!S(n,m)}{m^n} - \frac{m!S(n-1,m)}{m^{n-1}} = \frac{(m-1)!S(n-1,m-1)}{m^{n-1}} \\ &= \sum_{k=0}^{m-1} (-1)^k \binom{m-1}{k} \Big(\frac{m-1-k}{m}\Big)^{n-1}. \end{aligned}$$

We obtains an interesting formula on $S(n,m)$

$$\sum_{n=m}^{\infty} \frac{m!S(n,m)}{(m+1)^n} = 1. \tag{7}$$

Specially,

$$\sum_{n=2}^{\infty} \frac{1}{2^{n-1}} = 1; \quad \sum_{n=3}^{\infty} \frac{2^{n-2}-1}{3^{n-1}} = \frac{1}{2}; \quad \sum_{n=5}^{\infty} \frac{S(n,5)}{6^n} = \frac{1}{5!}. \tag{8}$$

In addition,

$$EY = \sum_{n=m}^{\infty} \frac{(m-1)!nS(n-1,m-1)}{m^{n-1}}.$$

Suppose we need Z_i trials in experiment T to draw the ith kind color after the $(i-1)$th kind color appearing, $i = 1, 2, \cdots, m$, where the probability $P\{Z_1 = 1\} = 1$, random variables $Z_1, Z_2, \cdots, Z_m$ are independent each other, and

$$Y = Z_1 + Z_2 + Z_3 + \cdots + Z_m = 1 + Z_2 + Z_3 + \cdots + Z_m.$$

$$P\{Z_k = i\} = \Big(\frac{k-1}{m}\Big)^{i-1}\Big(1 - \frac{k-1}{m}\Big), \quad i = 1, 2, \cdots; \quad k = 1, 2, \cdots, m.$$

Z_k have a geometric distribution, and the mathematical expectation $E(Z_k) = \frac{m}{m-k+1}$, $k = 1, 2, \cdots, m$, thus

$$EY = m\Big(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\Big) = mh_1(m).$$

where $h_1(m) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}$ is the harmonic number.

Thinking of the above expectation EY just obtained,

$$\sum_{n=m}^{\infty} \frac{(m-1)!nS(n-1,m-1)}{m^{n-1}} = m\Big(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{m}\Big). \quad (9)$$

For example, $m = 2, 3$ respectively, we have

$$\sum_{n=2}^{\infty} \frac{n}{2^{n-1}} = 3 = 2(1 + \frac{1}{2}), \quad \sum_{n=3}^{\infty} \frac{n(2^{n-2}-1)}{3^{n-1}} = \frac{11}{4} = \frac{3}{2!}(1 + \frac{1}{2} + \frac{1}{3}).$$

By the independence of $Z_1, Z_2, \cdots, Z_m$ and the property of variance,

$$D(Z_k) = \frac{m(k-1)}{(m-k+1)^2}, k = 1, 2, \cdots, m;$$

$$E(Y^2) = DY + (EY)^2 = m^2\Big(h_1^2(m) + h_2(m)\Big) - mh_1(m).$$

then

$$\sum_{n=m}^{\infty} \frac{(m-1)!n^2S(n-1,m-1)}{m^{n-1}} = m^2\Big(h_1^2(m) + h_2(m)\Big) - mh_1(m). \quad (10)$$

where $h_2(m) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{m^2}$.

In a word, we think (4), (6), (7) and (9) are interesting results on the Stirling number of the second kind.

References

[1] Arfwedson G. A probability distribution connected with Stirling's second class numbers. Skandinavisk Aktuarietidskrift, 1951,34:121-132.

[2] Erdös P, Spencer J. Probabilistic methods in combinatorics. New

York:Academic Press, 1974.

[3] Charalambides Ch A. A review of the Stirling numbers, their generalizations and statistical applications. Communications in Statistics Theory and Methods, 1988,17:2533-2595.

[4] Yamato H. A probabilistic approach to Stirling numbers of the first kind. Communications in Statistics-Theory and Methods, 1990, 19:3915-3923.

[5] Nishimura K, Sibuya M. Extended Stirling family of discrete probability distributions. Communications in Statistics-Theory and Methods, 1997,26:1727-1744.

[6] Sun Ping, Wang Tianming. Probabilistic represents of Stirling number and its applications(In Chinese). Acta Mathematica Sinica (Chinese Series), 1998,41(2):281-290.

[7] Charalambides Ch A. Combinatorial Methods in Discrete Distribution. New Jersey: John Wiley&Sons Inc., 2005.