Statistics in the Twenty-First Century: Special Volume
In Honour of Distinguished Professor Dr. Mir Masoom Ali
On the Occasion of his 75th Birthday Anniversary
PJSOR, Vol. 8, No. 3, pages 669-684, July 2012

Some Instructional Issues in Hypergeometric Distribution

Anwar H. Joarder
Department of Mathematics and Statistics
King Fahd University of Petroleum and Minerals
Dhahran 31261, Saudi Arabia
anwarj@ kfupm.edu.sa

Abstract

A brief introduction to sampling without replacement is presented. We represent the probability of a sample outcome in sampling without replacement from a finite population by three equivalent forms involving permutation and combination. Then it is used to calculate the probability of any number of successes in a given sample. The resulting forms are equivalent to the well known mass function of the hypergeometric distribution. Vandermonde's identity readily justifies different forms of the mass function. One of the new form of the mass function embodies binomial coefficient showing much resemblance to that of binomial distribution. It also yields some interesting identities. Some other related issues are discussed.

1. Introduction

Usually hypergeometric probability distribution is introduced without really introducing sampling scheme without replacement. In this paper, we want to introduce the related issues of sampling without replacement to provide clarity in the understanding of the hypergeometric probability distribution. We represent the probability of a sample outcome in sampling without replacement from a finite population by three equivalent forms involving permutation and combination. Then it is used to calculate the probability of any number of successes in a given sample. The resulting forms are equivalent to the well known mass function of the hypergeometric distribution. Some related instructional issues are presented.

1.1 Sampling Without Replacement

Consider a population of three doctors and two nurses denoted by A, B, C and D, E respectively. Notice that the individuals are distinctly identified. The sample space of a sample of 3 persons selected without replacement is given by

$$\{ABC,\ ABD,\ ABE,\ ACD,\ ACE,\ ADE,\ BCD,\ BCE,\ BDE,\ CDE\}.$$

i) ***Probability That a Person Is Included in a Particular Draw***

Let $A_i (i=1,2,3)$ be the event that Doctor A is included in the i-th selection. Then the probability that A is included in the 1st selection is $= P(A_1) = 1/5$. Since the sampling is without replacement, the probability that A is included in the 2nd selection is given by $P(A_1 A_2') = P(A_1')P(A_2 \mid A_1')$ which equals $\left(1-(1/5)\right)(1/4) = 1/5$.

Also the probability that Doctor A is included in the 3rd selection is given by

$P(A_1 A_2' A_3') = P(A_1')P(A_2' \mid A_1')P(A_3 \mid A_1' A_2')$ which is $\left(1-(1/5)\right)\left(1-(1/4)\right)(1/3)=1/5.$

Obviously, the probability that unit j of the population of N units is included in the i-th selection is given by

$$\left(1-N^{-1}\right)\left(1-(N-1)^{-1}\right)\cdots\left(1-(N-i+2)^{-1}\right)(N-i+1)^{-1}=N^{-1}. \tag{1.1}$$

ii) ***Probability That a Person Is Included in a Sample***

The probability that Doctor A is included in the sample is

$$P(A_1)+P(A_1'A_2)+P(A_1A_2'A_3')=(1/5)+(1/5)+(1/5)=3/5.$$

Thus each of the 5 persons have the same chance $(3/5)$ of being selected in a without replacement sample of size 3.

Let $M(j)=\binom{N-j}{n-j},\ j=1,2,\cdots,n.$ The number of samples of size n that contains unit j of the population of N units is $M(1)=\binom{N-1}{n-1}$. Since the total number of samples of size n is given by $M(0)=\binom{N}{n}$, the probability that unit j of the population of N units is included in the sample is given by

$$\frac{M(1)}{M(0)}=\binom{N-1}{n-1}\div\binom{N}{n}=\frac{n}{N}. \tag{1.2}$$

1.2 ***Probability of Selecting a Sample (Equiprobable Case)***

The probability of selecting a sampleof size $n=3$ members from a population of 5 people discussed in Section 1.1 is given by

$$\frac{3}{5}\times\frac{3-1}{5-1}\times\frac{3-2}{5-2}=\left(\binom{5}{3}\right)^{-1}.$$

In general, at the first draw the probability that one of the n specified units is selected is n/N . At the second draw the probability that one of the remaining $(n-1)$ specified units is drawn is $(n-1)/(N-1)$, and so on. Hence the probability that all n specified units are selected in n draws is

$$\frac{n}{N}\cdot\frac{n-1}{N-1}\cdot\frac{n-2}{N-2}\cdots\frac{n-(n+1)}{N-(n+1)}=\frac{n!(N-n)!}{N!}=\left(\binom{N}{n}\right)^{-1}=\frac{1}{M(0)} \tag{1.3}$$

(Cochran, 1977). An alternative argument is provided now. Since the number of samples that includes unit j and k of the population in the sample is given by $M(2)=\binom{N-2}{n-2}$, the probability that unit j and k of the population will be included in the sample is given by

$$\frac{M(2)}{M(0)}=\binom{N-2}{n-2}\div\binom{N}{n}=\frac{n(n-1)}{N(N-1)}.$$

Again the number of samples that includes unit j, k and l of the population is given by $M(3)=\binom{N-3}{n-3}$, and hence, the probability that unit j, k and l of the population will be included in the sample is given by

$$\frac{M(3)}{M(0)}=\binom{N-3}{n-3}\div\binom{N}{n}=\frac{n(n-1)(n-3)}{N(N-1)(N-2)}.$$

Since the number of samples that include specific n units of the population is $M(n)=\binom{N-n}{n-n}$, the probability that specified n units of the population of N units is included in a sample of size n is given by

$$\frac{M(n)}{M(0)}=\binom{N-n}{n-n}\div\binom{N}{n}$$

which is the same as (1.3).

1.3 Hypergeometric Probabilities

There are two major ways of calculating hypergeometric probabilities. One assumes that the items in the population are distinguishable, or can be labeled to make distinguishable. This will be discussed in Section 3.

In the other case, it is immaterial whether the items in the population are distinguishable or indistinguishable. This will yield a sample space $\mathbb{S}$ where outcomes are based on dichotomous nature of the population. This will be discussed in Section 2.

Suppose that a population containing K items are of one kind (say defective) and $N-K$ items are of different kind (say non-defective). Let n items be drawn at random in succession, without replacement, and X denote the number of defective items selected. The quantity $D_1D_2\cdots D_xD'_{x+1}\cdots D'_n$ denotes x successive defectives items and $n-x$ successive non-defective items. The probability of $D_1D_2\cdots D_xD'_{x+1}\cdots D'_n$ is expressed by truncated factorial by Joarder and Al-Sabah (2007). In this paper, we show that it can be represented by permutation or combination. We have used it for the probability of any number of successes which results in equivalent but insightful forms of the mass function of the hypergeometric distribution. Since the combinatorial function is available in almost all calculators, these forms are preferred to that presented by Joarder and Al-Sabah (2007) and Kendal and Stuart (1969, 133). Vandermonde's identity readily justifies the equivalence of the different forms of the mass function. On the other hand, any of the mass functions can also be used to prove Vandermonde's identity.

One of the new form of the mass function embodies binomial coefficient $\binom{n}{x}$ showing much resemblance to that of binomial distribution. It will be more transparent that hypergeometric mass function converges to that of the binomial distribution.

The paper is organized as follows. In Section 2, we will clearly demonstrate unequally likely sample space $\mathbb{S}$ and related representations of the hypergeometric probabilities. In Section 3, we will discuss the equally likely sample space and the wellknow hypergeometric mass function. We compare them by putting the two sample spaces based on Example 2.1 and Example 3.1 side by side. In Section 4, we show by an example how exact hypergeometric probabilities can be calculated. In Section 6, we present Vandermonde's identities related to hypergeometric distribution. Some other related issues are discussed. In Section 7, we prove that the probability of a particular sequence of outcomes with exactly x defective items and $n-x$ nondefective items in sampling without replacement, with unequally likely sample space, converges to that of a similar sequence in binomial distribution.

2. Conditional Probability Method (Unequally Likely Sample Space)

Let the population items be divided into units of two exhaustive kinds and $\mathbb{S}$ denote the unequally likely sample space of at most 2^n outcomes. It is usually done by a tree diagram for the case of indistinguishable items. But in this method, it is really immaterial whether the items are distinguishable or not. Then we have the following lemma.

Lemma 2.1 Suppose that an urn contains K items of one kind (say defective) and $N-K$ items are of a different kind (say non-defective). Let n items be drawn at random in succession, without replacement, and X denote the number of defective items selected. The probability of x successive successes in n trials is given by any of the following three mass functions:

$$(i)\ P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n \mid \mathbb{S}) = \frac{P_x^K}{P_{N-x}^N} \times \frac{P_{n-x}^{N-K}}{P_{N-n}^{N-x}}, \tag{2.1a}$$

$$(ii)\ P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n \mid \mathbb{S}) = \binom{N-n}{K-x} \div \binom{N}{K}, \tag{2.1b}$$

$$(iii)\ P_{wor}(D_1 D_2 \cdots D_x D'_{x+1} \cdots D'_n \mid \mathbb{S}) = \frac{P_x^K\ P_{n-x}^{N-K}}{P_n^N}, \tag{2.1c}$$

where $P_x^K = \dfrac{K!}{(K-x)!}$, $\dbinom{K}{x} = \dfrac{K!}{x!(K-x)!}$ and $\max\{0, n-(N-K)\} \le x \le \min\{n, K\}$.

Proof. The left hand side of (2.1a) is given by

$$P(D_1)P(D_2 \mid D_1)\cdots P(D_x \mid D_1 D_2 \cdots D_{x-1})P(D'_{x+1} \mid D_1 D_2 \cdots D_x)\cdots$$
$$\times P(D'_n \mid D_1 D_2 \cdots D_{x-1} D_x D'_{x+1} \cdots D'_{n-1})$$
$$= \frac{K+0}{K+(N-K)} \times \frac{(K-1)+0}{(K-1)+(N-K)} \times \cdots \times \frac{(K-x+1)+0}{(K-x+1)+(N-K)}$$
$$\times \frac{0+(N-K)}{(K-x)+(N-K)} \frac{0+(N-K-1)}{(K-x)+(N-K-1)} \times \cdots \times \frac{0+[N-K-(n-x)+1]}{(K-x)+[N-K-(n-x)+1]},$$

which is the same as (2.1a). The above can be written as

$$P(D_1)P(D_2 \mid D_1)\cdots P(D_x \mid D_1 D_2 \cdots D_{x-1})P(D'_{x+1} \mid D_1 D_2 \cdots D_x)\cdots$$
$$\times P(D'_n \mid D_1 D_2 \cdots D_{x-1} D_x D'_{x+1} \cdots D'_{n-1})$$
$$= \frac{K+0}{K+(N-K)} \times \frac{(K-1)+0}{(K-1)+(N-K)} \times \cdots \times \frac{(K-x+1)+0}{(K-x+1)+(N-K)}$$
$$\times \frac{0+(N-K)}{(K-x)+(N-K)} \times \frac{0+(N-K-1)}{(K-x)+(N-K-1)} \times \cdots \times \frac{0+[N-K-(n-x)+1]}{(K-x)+[N-K-(n-x)+1]},$$

which simplifies to $\dfrac{K!}{(K-x)!} \times \dfrac{(N-K)!}{(N-K-n+x)!} \times \dfrac{(N-n)!}{N!}$ which is equivalent to (2.1c). Moreover (2.1a) or (2.1c) simplifies to (2.1b).

The representation (2.1b) appears in a technical report of Joarder, Laradji and Omar (2009) and Joarder (2010). It is obvious that the representation (2.1c) is intuitively most appealing as it will be manifested in (2.2b).

The sample space contains a total of $\dbinom{n}{x}$ outcomes having x defectives and $(n-x)$ non-defectives out of at most 2^n outcomes. The elements in the sample space are not equally likely. The motivation that led to the following theorem is also implicit in Kendal and Stuart (1969, 133), and Joarder and Al-Sabah (2007).

Theorem 2.1 Suppose that an urn contains K items of one kind (say defective) and $N-K$ items are of a different kind (say non-defective). The items may be distinguishable or indistinguishable in each of the two categories. Let n items be drawn at random in succession, without replacement, and X denote the number of defective items selected. The probability of x successes in n trials is equivalently given by any of the two equivalent forms:

$$P(X=x)=\binom{n}{x}\left[\binom{N-n}{K-x}\div\binom{N}{K}\right], \quad \max\{0,n-(N-K)\}\le x\le \min\{n,K\}, \tag{2.2a}$$

or,

$$P(X=x)=\binom{n}{x}\times\frac{P_x^K P_{n-x}^{N-K}}{P_n^N},\ \max\{0,n-(N-K)\}\le x\le \min\{n,K\}. \tag{2.2b}$$

Proof. Any sample outcome of n items that have exactly x defectives and $n-x$ non-defective items is given by (2.1). Since there are $\binom{n}{x}$ such outcomes, out of a maximum of 2^n outcomes in the sample space, we have

$$P(X=x)=\binom{n}{x}P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n\mid \mathbb{S}). \tag{2.3}$$

Hence by (2.1b) or (2.1c), the mass function is given by (2.2a) or (2.2b).

The name hypergeometric is derived from a series introduced by the Swiss mathematician and physicist, Leonard Euler, in 1769. The probabilities in (3) are the successive terms of

$$\frac{(N-n)!(N-K)!}{N!(N-K-n)!}\ {}_2F_1(-n,-K;\ N-K-n+1;\ 1),$$

where ${}_2F_1(a_1,a_2;b;x)$ is the generalized hypergeometric function (Johnson, Kotz and Kemp, 1993, 237).

Example 2.1 A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that 3 out of 5 devices are defective. The inspector of the retailer randomly picks up 3 devices from a 5 devices. Determine the probability that the sample will have two defective devices.

Solution: If $D_i\ (i=1,2)$ is the event that in the i-th draw we have a defective item and f $N_i\ (i=1,2,3)$ is the event that in the i-th draw we have a non-defective item. Then the sample space (use tree diagram) is given by

$$\mathbb{S}=\{D_1D_2D_3,\ D_1D_2N_3,\ D_1N_2D_3,\ D_1N_2N_3,\ N_1D_2D_3,\ N_1D_2N_3,\ N_1N_2D_3\}.$$

Then $P_{wor}(D_1D_2N_3\mid\mathbb{S})$ is given by

$$P(D_1)P(D_2\mid D_1)P(N_3\mid D_1D_2)=\frac{3+0}{3+2}\times\frac{2+0}{2+2}\times\frac{0+2}{1+2},$$

which is 12/60. Since each of the element in the event of interest $\{D_1D_2N_3,\ D_1N_2D_3,\ N_1D_2D_3\}$ is given by (12/60), we have

$$P(X=2)=\binom{3}{2}P_{wor}(D_1D_2N_3\,|\,\mathbb{S}),\ \text{ i.e., } P(X=2)=36/60.$$

Alternatively, since $N=5,\ K=3,\ n=3$ and $x=2$, $P(X=2)$ can be directly done by (2.2a) as follows:

$$P(X=2)=\binom{3}{2}\times\left[\binom{5-3}{3-2}\div\binom{5}{3}\right]=0.60.$$

It can also be done by (2.2b) as follows:

$$P(X=2)=\binom{3}{2}\frac{P_2^3P_{3-2}^{5-3}}{P_3^5}=0.60.$$

The number 60 in the denominator of probabilities is explained at the end of theorem 3.1. The probability mass function is explicitly given by the following table:

x	1	2	3
$f(x)$	3/10	6/10	1/10

3. Equiprobable Method

In this section, it is required that the populations units be distinguishable. In case, they are indistinguishable, one may label them to make them distinguishable.

Theorem 3.1. Suppose that an urn contains K items of one kind (say defective) and $N-K$ items are of a different kind (say non-defective). The items may be distinguishable or indistinguishable in each of the two categories. Let n items be drawn at random in succession, without replacement, and X denote the number of defective items selected. The probability of x successes in n trials is given by

$$P(X=x)=\left[\binom{K}{x}\binom{N-K}{n-x}\right]\div\binom{N}{n},\ \max\{0,n-(N-K)\}\le x\le\min\{n,K\}. \qquad (3.1)$$

Proof. There will be a total of $\binom{N}{n}$ equally likely elements in the sample space. The combinatorial proof of this theorem is available in most textbooks on statistics and discrete mathematics (e.g. Barnett, 1998). There are $\binom{K}{x}$ ways of choosing x of the K items (say defective items) and $\binom{N-K}{n-x}$ ways of choosing $(n-x)$ of the $(N-K)$ non-defective items, and hence there are $\binom{K}{x}\binom{N-K}{n-x}$ ways of choosing x defectives and $(n-x)$ non-defective items. Since there are $\binom{N}{n}=M$ ways of choosing n of the N elements, assuming M sample points are equally likely, the probability of any of the M sample point is $1/M$. Hence the probability of having x defective items in the sample is given by (3.1).

Vandermonde's identity readily justifies that the two forms of the hypergeometric mass functions given by (2.2) and (3.1) are equivalent (Laradji, 2009).

If one permutes each of the $\binom{N}{n}$ sample points, the total number of elements in the inflated 'sample space' would be $\binom{N}{n}\times n!=P_n^N$ in which case (3.1) would turn into $P(X=x)=\left[\left(P_x^K/x!\right)\left(P_{n-x}^{N-K}/(n-x)!\right)\right]\div\left(P_n^N/n!\right)$, which is the same as (2.2b).

The method in this section, also guarantees that sample outcomes are equally likely or equiprobable. Thus this method produces a Simple Random Sampling where "simple" refers to the equally likely outcomes.

Example 3.1 A random committee of size 3 is selected from 3 doctors and 2 nurses. Suppose that the doctors and members can be identified well making the individuals distinguishable. What is the probability that there will be 2 doctors in the committee?

Solution: Suppose the doctors are labeled as D^1, D^2 and D^3, while the nurses are labeled as N^1 and N^2 to make the items in the population distinguishable. The sample space of outcomes is given by

$$\{D^1D^2D^3, D^1D^2N^1, D^1D^2N^2, D^1D^3N^1, D^1D^3N^2,$$
$$D^1N^1N^2, D^2D^3N^1, D^2D^3N^2, D^2N^1N^2, D^3N^1N^2\}$$

The event of having 2 doctors in the committee is given by

$$\{D^1D^2N^1, D^1D^2N^2, D^1D^3N^1, D^1D^3N^2, D^2D^3N^1, D^2D^3N^2\}$$

which has a probability of $6/10$. This can be directly done (3.1) as the following

$$P(X=2)=\left[\binom{3}{2}\binom{5-3}{3-2}\right]\div\binom{5}{3},$$

i.e., $P(X=2)=0.6.$

That is there exactly 6 ways of selecting 2 doctors and 1 nurse in a sample of size 3. There are $\binom{N}{n} = \binom{5}{3} = 10$ ways of selecting 3 people from a population of 5 people.This isthe number of sample points in the sample space. Every sample outcome has a probability of $1/10$. The mass function is exactly the same as what we have in Section 2.

x	1	2	3
$f(x)$	$3/10$	$6/10$	$1/10$

The number 60 appearing in Example 2.1 can be explained. In fact if you permute each of the $\binom{5}{3} = 10$ sample points in Example 3.1, you will have an inflated sample space of $\binom{5}{3} \times 3! = 10(6).$ Then the probability will be

$$P(X=2) = 3!\left[\binom{3}{2}\binom{5-3}{3-2}\right] \div 3!\binom{5}{3} = \frac{3!(6)}{3!(10)} = 0.60.$$

The Example 2.1 and Example 3.1 are put side by side in the following table:

Distinguishable / Indistinguishable (Theorem 2.1)		Distinguishable (Theorem 3.1)	
Sample Space	Probability	Sample Space	Probability
$D_1D_2D_3$	$1/10$	$D^1D^2D^3$	$1/10$
$D_1D_2N_3$	$2/10$	$D^1D^2N^1$	$1/10$
$D_1N_2D_3$	$2/10$	$D^1D^2N^2$	$1/10$
$D_1N_2N_3$	$1/10$	$D^1D^3N^1$	$1/10$
$N_1D_2D_3$	$2/10$	$D^1D^3N^2$	$1/10$
$N_1D_2N_3$	$1/10$	$D^1N^1N^2$	$1/10$
$N_1N_2D_3$	$1/10$	$D^2D^3N^1$	$1/10$
		$D^2D^3N^2$	$1/10$
		$D^2N^1N^2$	$1/10$
		$D^3N^1N^2$	$1/10$

The following table will provide insight into the number of sample points:

	Unequally Likely Sample Space	Equally Likely Sample Space
Mass function	$P(X=x)=\binom{n}{x}\times\frac{P_x^K P_{n-x}^{N-K}}{P_n^N}$,	$P(X=x)=\left[\binom{K}{x}\binom{N-K}{n-x}\right]\div\binom{N}{n}$,
# of Sample points	max 2^n	$\binom{N}{n}$

We have taught both the methods in some service courses in statistics and found that students get better insight by the Conditional Probability Method. The elements of the sample space of the Conditional Probability Method are not equally likely or equiprobable. One element of the sample space in the Conditional Probability Method maps on to some elements of the sample space of the Equiprobable Method.

Example 3.2 Suppose that a shipment of 9 (N) digital voice recorders contains 4 (K) that are defective. If n voice recorders are randomly chosen without replacement for inspection, what is the probability that

a. the first two of $n=3$ checked will be defective but the third one will be non-defective?
b. 2 of the $n=3$ recorders will be defective?

Solution:

a. The probability is $P(D_1D_2D_3')=\frac{4+0}{4+5}\times\frac{3+0}{3+5}\times\frac{0+5}{2+5}$,

which can be directly solved by (2.1c) as $\frac{P_2^4P_1^5}{P_3^9}$.

b. Since $N=9,\ K=4,\ n=3$, Conditional Probability Method (2.3), the probability that 2 of the 3 voice recorders will be defective is given by

$$P(X=2)=\binom{3}{2}P(D_1D_2D_3')=\binom{3}{2}\times\frac{4+0}{4+5}\frac{3+0}{3+5}\frac{0+5}{2+5}=\frac{5}{14},$$

which, by (2.2b), can also be written as

$$P(X=2)=\binom{3}{2}\frac{P_2^4P_1^5}{P_3^9}=\frac{180}{504}.$$

Note that the maximum number of sample points is $2^3=8$. Alternatively, by using the equiprobable method (3.1), we have

$$P(X=2)=\left[\binom{4}{2}\binom{9-4}{3-2}\right]\div\binom{9}{3},$$

which is 30/84, where the number of sample points is $\binom{9}{3}=84$. We remark that 84 sample points in formula (3.1) has been inflated to $3!(84)=504$ in the formula in (2.2b).

The students erroneously tend to use the hypergeometric probability function (3.1) based on Equally Likely Sample Space to find the probability of a simple event in part (a). Hence we recommend to use Unequally Likely Sample Space which provides probability of a simple event (say, part a by (2.1c)) or a compound event (say, part b by (2.2b)).

Example 3.3 There are N devices in a box. A total of K of them are faulty and $N-K$ are sound. One takes three devices in succession at random without replacement. The probability that first two are faulty but the third one is not faulty is $12/60$. what is the probability that in a sample of 3 devices, two are faulty and one is sound?

Solution: The probability that the first two devices selected are faulty and the third one is sound is given by

$$P(F_1F_2S_3)=\frac{K+0}{K+(N-K)}\times\frac{(K-1)+0}{(K-1)+(N-K)}\times\frac{0+(N-K)}{(K-2)+(N-K)}.$$

Hence by the given condition, we have, $\dfrac{K(K-1)(N-K)}{N(N-1)(N-2)}=0.20$ which can be solved by preparing the following table:

K	$N-K$	N	$P(F_1F_2S_3)$
2	1	3	0.3333 approx.
2	2	4	0.1667 approx
3	1	4	0.25
3	2	5	0.20
4	1	5	0.20

Notice that there are two solutions $K=3, N-K=2$, or, $K=4, N-K=1$. Hence, we have

K	$N-K$	N	$P(X=2)$
3	2	5	$\binom{3}{2}\binom{2}{1}\div\binom{5}{3}=\frac{3}{5}$
4	1	5	$\binom{4}{2}\binom{1}{1}\div\binom{5}{3}=\frac{3}{5}$

Thus $P(X=2)=\frac{3}{5}=0.60$. It will be difficult to solve the problem by (2.2a) or by (3.1), but easy by using (2.2b).

4. Computational Accuracy

Since the hypergeometric parameters are integers, it is possible to calculate the exact probability of any event. The following lemmas are in Hua (1982).

Lemma 4.1 Let p be a prime. Then the exact exponents of p that divides $n!$ is given by

$$\lfloor n/p \rfloor + \lfloor n/p^2 \rfloor + \lfloor n/p^3 \rfloor + \cdots,$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x.

Lemma 4.2 For any positive integer $n \geq 2$, the quantity $n!$ can be written as a product of prime numbers in the following manner:

$$n! = p_1^{r_1}.p_2^{r_2}.p_3^{r_3}...p_k^{r_k},$$

for some positive integer k, where p_i's are prime numbers.

Example 4.1 To simplify the probability in Example 3.2 by (2.2a)

$$P(X=2) = \binom{3}{2}\left[\binom{9-3}{4-2} \div \binom{9}{4}\right],$$

we proceed as follows:

To decompose $9!$ by Lemma 4.1, the exponent of 2 will be $\lfloor 9/2 \rfloor + \lfloor 9/2^2 \rfloor + \lfloor 9/2^3 \rfloor = 4+2+1 = 7,$ the exponent of 3 will be $\lfloor 9/3 \rfloor + \lfloor 9/3^2 \rfloor = 3+1 = 4,$ the exponent of 5 will be $\lfloor 9/5 \rfloor = 1,$ the exponent of 7 will be $\lfloor 9/7 \rfloor = 1,$ so that by Lemma 4.2, we have $9! = 2^7 \cdot 3^4 \cdot 5^1 \cdot 7^1 (= 362880).$ Then

$\left[\binom{6}{2} \div \binom{9}{4}\right] = \frac{6!}{2!} \times \frac{5!}{9!},$ which can be expressed as $\frac{2^4 \cdot 3^2 \cdot 5}{2} \times \frac{2^3 \cdot 3^1 \cdot 5}{2^7 \cdot 3^4 \cdot 5 \cdot 7} = \frac{5}{2 \cdot 3 \cdot 7},$

so that $P(X=2) = \binom{3}{2}\frac{5}{2 \cdot 3 \cdot 7} = \frac{5}{14}.$

Trong (1993) developed an algorithm to calculate accurate Cumulative Distribution Function of Hypergeometric Distribution.

5. Acceptance Sampling Plan

Acceptance sampling is an important field of statistical quality control that was popularized by Dodge and Romig and originally applied by the U.S. military to the testing of bullets during World War II. If every bullet was tested in advance, no bullets would be left to ship. If, on the other hand, none were tested, malfunctions might occur in the field of battle, with potentially disastrous results.

Dodge and Romig reasoned that a sample should be picked at random from the lot, and on the basis of information that was yielded by the sample, a decision should be made regarding the disposition of the lot. In general, the decision is either to accept or reject the lot. This process is called *Lot Acceptance Sampling* or just *Acceptance Sampling*.

Example 5.1. Suppose that a shipment of 9 (N) digital voice recorders contains 4 (K) that are defective. If a sample of 3 $(=n)$ voice recorders contains at most one defective, the shipment is rejected. What is the probability that the shipment will be rejected?

Solution: By the Equiprobable Method (3.1), we have

x	0	1	2	3
$f(x)$	$15/126$	$60/126$	$45/126$	$6/126$

The probability that the shipment will be rejected is given by

$$P(X=0,1)=P(X=0)+P(X=1)=\frac{15}{126}+\frac{60}{126}=\frac{75}{126}\approx 0.5952.$$

6. Vandermond'e Identity

The identity $\sum_{x\ge 0}\binom{n}{x}\binom{N-n}{K-x}=\binom{N}{K}$ is proved by equating the coefficients of y^n in the following identity $(1+y)^n(1+y)^m=(1+y)^{n+m}$ with x as index of summation and $m=N-n$. Similarly, the identity $\sum_{x\ge 0}\binom{K}{x}\binom{N-K}{n-x}=\binom{N}{n}$ is proved by equating the coefficients of y^K in the following identity $(1+y)^K(1+y)^L=(1+y)^{K+L}$ with x as index of summation and $L=N-K$. It is worth noting that the above identities are forms of well knownVandermonde's identity.

Proposition 6.1 Suppose that an urn contains K items of one kind (say defective) and $N-K$ items of a different kind (say non-defective). Let n items be drawn at random, without replacement, and X denote the number of defective items selected. Then we have the following identities:

a. $$\binom{n}{x}\binom{N-n}{K-x}\binom{N}{n}=\binom{K}{x}\binom{N-K}{n-x}\binom{N}{K},$$

b. $$\binom{N}{n}\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}}\binom{n}{x}\binom{N-n}{K-x}=\binom{N}{K}\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}}\binom{K}{x}\binom{N-K}{n-x},$$

c. $$\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}}\binom{n}{x}\binom{N-n}{K-x}=\binom{N}{K},$$

d. $$\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}} \binom{K}{x}\binom{N-K}{n-x} = \binom{N}{n}.$$

e. $$\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}} \binom{n}{x} P_x^K P_{n-x}^{N-K} = P_n^N.$$

Proof. Part (a) is obvious. Summing the identity in part (a), we have

$$\sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}} \binom{n}{x}\binom{N-n}{K-x}\binom{N}{n} = \sum_{x=\max\{0,n-(N-K)\}}^{\min\{n,K\}} \binom{K}{x}\binom{N-K}{n-x}\binom{N}{K},$$

which is equivalent to part (b). Since (2.2a) is a probability mass function, part (c) follows from (2.2a). Similarly, part (d) follows from (3.1) and part (e) follows from (2.2b).

7. Binomial and Hypergeometric Probabilities

Suppose that an urn contains K items of one kind (say defective) and $N-K$ items are of a different kind (say non-defective). Let n items be drawn at random, with replacement in succession, and X denote the number of defective items selected. The probability that any item is defective at any draw is $p = K/N$ (say). Then with arguments similar to section 2, the probability of having x successive defectives and $(n-x)$ successive non-defectives is given by

$$P_{wr}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n) = \frac{K}{N}\times\frac{K}{N}\times\cdots\frac{K}{N}\times\left(1-\frac{K}{N}\right)\times\left(1-\frac{K}{N}\right)\times\cdots\left(1-\frac{K}{N}\right),$$

which equals, $p^x q^{n-x}$, so that $P(X=x) = \binom{n}{x} P_{wr}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n),$

which equals $P(X=x) = \binom{n}{x} p^x q^{n-x}$. In case of sampling without replacement, $P(X=x) = \binom{n}{x} P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n \mid \mathbb{S})$ where $P(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n)$ is given by (2.1). Now if $N\to\infty$, and $p = K/N$, we have the following corollary.

Corollary 7.1: As $K\to\infty, N\to\infty$, but $\frac{K}{N}\to p$, the limiting the probability of x successes in n trails in case of sampling without replacement is given by $p^x q^{n-x}$**.**

Proof. The probability of x successes in n trails in case of sampling without replacement denoted by $P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n \mid \mathbb{S})$ is given by (See 2.1b)

$$P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n\,|\,\mathbb{S})=\binom{N-n}{K-x}\div\binom{N}{K}, \qquad (7.1)$$

which equals $P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n\,|\,\mathbb{S})=p_1p_2\cdots p_xq_1q_2\cdots q_{n-x}$, where

$$p_i=\frac{K-i+1}{N-i+1},\;\; q_j=\frac{N-K-j+1}{N-x-j+1},\;\; i=1,2,\cdots,x;\; j=1,2,\cdots,n-x.$$

In case $K\to\infty, N\to\infty$, the quantity $p_i\to p,\ (i=1,2,\cdots,x;\ 0<p<1)$, and $q_j\to q$. $(j=1,2,\cdots,n-x;\ 0<q<1)$. Hence we have

$$P_{wor}(D_1D_2\cdots D_xD'_{x+1}\cdots D'_n\,|\,\mathbb{S})\to p^xq^{n-x}. \qquad (7.2)$$

For any other sequence having x successes and $n-x$ failures, the probability, in the limit, will be the same as above.

This shows the equivalence of binomial and hypergeometric distribution in the limit. Though the fact is available in most textbooks on statistics, the result in (7.2) and the factor $\binom{n}{x}$ in the hypergeometric mass function (2.2) will be insightful to the students and instructors.

Acknowledgements

The author gratefully acknowledges the excellent research support provided by King Fahd University of Petroleum and Minerals, Saudi Arabia. The author is also grateful to Dr. A. Laradji, Dr. M.H. Omar, Dr. Walid S. Sabah, Dr. R.M. Latif, Dr M. Riaz and Dr. I. Rahimov for many constructive suggestions.

References

1. Barnett, S. (1998). *Discrete Mathematics: Numbers and Beyond*. Pearson Education Limited. Essex, England.
2. Cochran, W.G. (1977). *Sampling Techniques*. John Wiley.
3. Hua, L.K. (1982). Introduction to Number Theory. Springer-Verlag, New York (English Translation).
4. Joarder, A.H. (2010). Hypergeometric distribution and its application in statistics. Published in *International Encyclopaedia of Statistical Science*. Edited by MiodragLovric. Springer. (1st Edition. Edition, December 1, 2010), ISBN: 3642048978, 1852 pages)

5. Joarder, A.H. and Al-Sabah, W.S. (2007). Probability issues in without replacement sampling. *International Journal of Mathematical Education in Science and Technology*, 38(6), 823-831.

6. Joarder, A.H.; Laradji, A. and Omar, M.H. (2009). The Probability of a Sample Outcome in Sampling Without Replacement. *Technical Report No.* 405, (May 2009), Department of Mathematics and Statistics, KFUPM, Saudi Arabia.

7. Johnson, N.L.; Kotz, S. and Kemp, A.W. (1993).*Univariate Discrete Distributions.* John Wiley and Sons, New York, USA.

8. Kendal, M.G. and Stuart, A. (1969). *The Advanced Theory of Statistics*. v1: Distribution Theory. Charles Griffin, London.

9. Laradji, A. (2009). Personal communication.

10. Trong, W. (1993). An accurate computation of the hypergeometric distribution function. *ACM Transaction on Mathematical Software*, 19(1), March 1993, 33-43.