Journal of Statistical Research *ISSN 0256 - 422 X*
2009, Vol. 43, No. 2, pp. 29-35
Bangladesh

ASYMPTOTIC DISTRIBUTION OF THE NUMBER OF RECTANGLES ARISING IN BINOMIAL TRIALS

Zafar Iqbal
National College of Business Administration and Economics, Lahore, Pakistan
Email: zafariqbal75@yahoo.com

Ahmed Zogo Memon
National College of Business Administration and Economics, Lahore, Pakistan
Email: azogomemon@hotmail.com

summary

The simultaneous occurrence of independent and identical Binomial trials at n^2 locations of an $n \times n$ lattice may randomly produce a specified configuration comprising the same event. One configuration being a rectangle, this paper finds factorial moments of the number of these rectangles. Asymptotic behavior of the random variable defined is also investigated.

Keywords and phrases: rectangle, factorial moments, lattice.

1 Introduction

Natural disasters, ecological disorders and epidemiological patterns of a disease that randomly inflict human dwellings, forests and plant nurseries over space and time often engage a scientist's curiosity to understand how these events are statistically distributed. For events that occur with same probability the concept of join count statistics is generally found useful. The number of these counts becomes a random variable of interest in situations specially where the locations define a lattice. For instance, fruit trees are often grown in this fashion for the purpose of improving efficiency of production but the risk of disease is there. Experiments with artificially generated lattices are conducted to examine the effect of soil infection and the infection patterns that it creates in trees. A disease has binary nature, its presence or absence, as it is the case of a single binomial trial where an event may or may not occur. A join occurs when two adjacent trees are simultaneously infected by disease in a horizontal, vertical or some other manner.

Single binominal trials find their application in scientific inquiries where a trial results in some specified event with a constant probability. A researcher may be interested in the number of these events that happen randomly when the trial is independently repeated.

Suppose that we have a set of n^2 locations of a square lattice, arranged in n rows and n columns, and that binomial trials occur simultaneously at all these locations. If each trial results in a specified event E with some probability, various configurations are likely to emerge. A configuration of these events and so its probability, often becomes a subject of interest. Moran (1948), Fuchs and David (1965) and Memon and David (1968) for instance, study the distribution of various patterns that arise in a similar situation based on single binomial trials. We investigate here the probability distribution of the number of particular configuration which we call here a rectangle of events and define it below with reference to the above mentioned situation.

Let $\phi_{i,j}$ denote j^{th} link in i^{th} row with a value = 1 (if the event E occurs at locations j and $j+1$ in i^{th} row) and zero otherwise. Thus the condition $\phi_{i,j} = \phi_{i+1,j} = 1$ entails a rectangle with the event E that occurs at the locations (i, j) , $(i, j + 1)$, $(i + 1, j)$, $(i + 1, j + 1)$ and zero otherwise. Let X be the number of rectangles that arise when independent Bernoulli trials materialize simultaneously at n^2 locations of the lattice. The use of Memon and David (1968) Theorem in Appendix will be made to find the factorial moments of X in this paper.

2 Factorial Moments of the Random Variable X

In this section we find the first three factorial moments.

2.1 First Factorial Moment

A particular rectangle occurs when $\phi_{i,j} = \phi_{i+1,j} = 1$, for $i, j = 1, \ldots, n-1$; and zero otherwise. For $r = 1$, in Memon and David Theorem we have the following first factorial moment

$$\mu_{[1]} = \sum_k p_k,$$

where p_k is the probability of a particular rectangle. The probability p_k of a particular rectangle is p^4 and the total number of such possible rectangles are

$$\sum_{i,j,r,s=1}^{n-1} \phi_{ij}\phi_{rs}$$

which simplifies to

$$\sum_{i,j=1}^{n-1} \phi_{ij}\phi_{i+1j} = (n-1)^2$$

that is

$$\mu_{[1]} = (n-1)^2 p^4. \tag{2.1}$$

2.2 Second Factorial Moment

The second factorial moment follows from $r = 2$, in the Memon and David Theorem, that is,

$$\mu_{[2]} = 2! \sum prob(\text{two particular rectangles})$$

where $\sum$ is carried over all possible sets of two rectangles in the $n \times n$ lattice. Since two rectangles may have a common side, a common corner, or non-contiguous with probabilities p^6 , p^7 and p^8 respectively. It follows that the above moment is expressed as

$$2!(a_{2,6}p^6 + a_{2,7}p^7 + a_{2,8}p^8). \tag{2.2}$$

We find below the coefficients of probabilities p^6, p^7 and p^8.

The coefficient $a_{2,6}$: To calculate this coefficient we consider the following model for two adjacent rectangles in a row $\phi_{i,j} = \phi_{i,j+1} = \phi_{i+1,j} = \phi_{i+1,j+1} = 1$, for $i = 1, \ldots, n-1$; $j = 1, \ldots, n-2$; and zero otherwise. Under these conditions the number of these possibilities can be obtained from

$$\sum_{i_1,\ldots,i_8} \phi_{i_1,i_2}\phi_{i_3,i_4}\phi_{i_5,i_6}\phi_{i_7,i_8}$$

over $i_1, \ldots, i_8 = 1, \ldots, n-1$; which simplifies to

$$n_1 = (n-1)(n-2).$$

Similarly the number of possibilities for column-wise two adjacent rectangles is

$$n_2 = (n-1)(n-2).$$

Hence,

$$a_{2,6} = n_1 + n_2 = 2(n-1)(n-2).$$

The coefficient $a_{2,7}$: To calculate this coefficient we consider the following model for the two diagonally ($> 90°$) adjacent rectangles with a common corner, i.e.,

$$\phi_{i,j} = \phi_{i,j+1} = \phi_{i+1,j+1} = \phi_{i+2,j+1} = 1,$$

for $i, j = 1, \ldots, n-2$ and zero otherwise. The number of these configurations simplifies to

$$\sum_i \sum_j \phi_{i,j}\phi_{i,j+1}\phi_{i+1,j+1}\phi_{i+2,j+1}$$

that is,

$$n_3 = (n-2)^2.$$

Similarly the number of two diagonally adjacent ($< 90°$) rectangles with a common corner is

$$n_4 = (n-2)^2.$$

so

$$a_{2,7} = n_3 + n_4 = 2(n-2)^2.$$

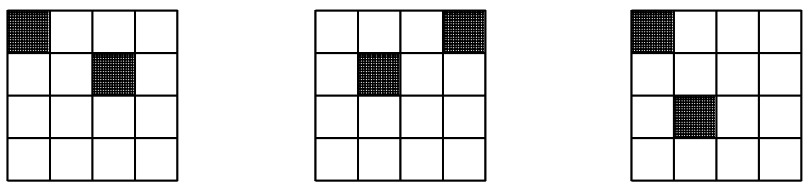
The coefficient $a_{2,8}$: For this coefficient we have the following model for the two non-contiguous rectangles in a row: $\phi_{i,j} = \phi_{i+1,j} = \phi_{i+1,m} = \phi_{i+1,m} = 1$ for $i = 1,\ldots,n-1$; $j = 1,\ldots,n-3$; $m = 3,\ldots,n-1$; $m - j \geq 2$; and zero otherwise. The number of these possibilities in the lattice can be shown as

$$n_5 = \frac{(n-1)(n-2)(n-3)}{2}.$$

Similarly the number of possible column-wise two non-contiguous rectangles in the lattice

$$n_6 = \frac{(n-1)(n-2)(n-3)}{2}.$$

The model, for horizontally appearing non-contiguous rectangles with one or more gaps in the adjacent rows (as indicated in Figure 1(a)), is $\phi_{i,j} = \phi_{i+1,j} = \phi_{i+1,k} = \phi_{i+1,k} = 1$ for $i = 1,\ldots,n-2; j = 1,\ldots,n-3; k = 3,\ldots,n-1$; $k - j \geq 2$; and zero otherwise.



(a) (b) (c)

Figure 1:

The number of possible arrangements for Figure 1(a) is given as:

$$n_7 = \frac{(n-2)^2(n-3)}{2}.$$

By symmetry, the number of arrangements for Figure 1(b) is:

$$n_8 = \frac{(n-2)^2(n-3)}{2}.$$

Figure 1(c) represents all situations involving non-adjacent rectangles with one or more horizontal and vertical gaps. The model for this situation is $\phi_{i,j} = \phi_{i+1,j} = \phi_{k,m} = \phi_{k+1,m} = 1$, for $i = 1,\ldots,n-2; j = 1,\ldots,n-1; k = 3,\ldots,n-1; m = 2,\ldots,n-1$; $k \geq i+2$; $m \geq j+1$, and zero otherwise.

We have the number of these situations

$$n_9 = \frac{(n-1)(n-2)^2(n-3)}{2}.$$

Hence

$$\begin{aligned} a_{2,8} &= n_5 + n_6 + n_7 + n_8 + n_9 \\ &= (n-1)(n-2)(n-3) + (n-2)^2(n-3) + \frac{(n-1)(n-2)^2(n-3)}{2}. \end{aligned}$$

2.3 Third Factorial Moment

The third factorial moment follows from $r = 3$ in the theorem referred above. Here,

$$\mu_{[3]} = 3! \sum prob(\text{three particular rectangles}),$$

where the summation $\sum$ is carried over all possible sets of three rectangles with probabilities p^8, p^9, p^{10}, p^{11} and p^{12} respectively, depending on how these rectangles emerge. The three rectangles may have one or more common corners, one or more common sides, or they may be non-adjacent. The third moment can be expressed as

$$3!(a_{3,8}p^8 + a_{3,9}p^9 + a_{3,10}p^{10} + a_{3,11}p^{11} + a_{3,12}p^{12}). \tag{2.3}$$

As done in the second factorial moment, we specify their locations and formulate a model involving the variables $\phi_{i,j}$ and enumerate the possibilities using the product $\prod_i^6 \phi_{m_i,n_i}$, where m_i, n_i take values from $1, \ldots, n-1$.

The coefficient $a_{3,8}$: To obtain this coefficient we consider all three rectangles adjacently appearing in a row with probability p^8. Since in this situation the adjacent rectangles have a common vertical side we set up the following model; $\phi_{i,j} = \phi_{i,j+1} = \phi_{i,j+2} = \phi_{i+1,j} = \phi_{i+1,j+1} = \phi_{i+1,j+2} = 1$ for $i = 1, \ldots, n-1; j = 1, \ldots, n-3$; and zero otherwise. The number of these possibilities can be thus determined from the simplified expression

$$\sum_i \sum_j \phi_{i,j}\phi_{i+1,j}\phi_{i,j+1}\phi_{i+1,j+1}\phi_{i,j+2}\phi_{i+1,j+2}$$

and so

$$n_{10} = (n-1)(n-3). \tag{2.4}$$

By symmetry, the number of possible cases when three rectangles appear adjacently in columns has to be same as Equation 2.3, that is,

$$n_{11} = (n-1)(n-3). \tag{2.5}$$

Three rectangles one of which has a common vertical side with the second rectangle on the right and a common horizontal side with the lower third may be described by the model; $\phi_{i,j} = \phi_{i+1,j} = \phi_{i,j+1} = \phi_{i+1,j+1} = \phi_{i+2,j} = 1$, for $i = 1, \ldots, n-2; j = 1, \ldots, n-2$; and zero otherwise. The number of such events is found as

$$n_{12} = (n-2)^2. \tag{2.6}$$

The clock-wise rotation of three rectangles considered above produces three similar situations in addition. By symmetry the number of possibilities correspondingly is the same as 2.6, that is,

$$n_{12} = n_{13} = n_{14} = n_{15}. \tag{2.7}$$

Consequently, the coefficient of p^8 from Equations 2.4, 2.5, 2.6 and 2.7 can be simplified to

$$a_{3,8} = 6n^2 - 24n + 22.$$

The other coefficients in Equation 2.3 determined similarly by using the above approach are

$$\begin{aligned} a_{3,9} &= 8n^2 - 40n + 48. \\ a_{3,10} &= 12n^3 - 96n^2 + 246n - 198. \\ a_{3,11} &= n^4 - 2n^3 - 39n^2 + 172n - 192. \\ a_{3,12} &= \frac{1}{6}(n^6 - 6n^5 + 6n^4 - 68n^3 - 725n^2 + 172n - 192). \end{aligned}$$

Remark 1. With an $n \times n$ lattice the usual approach for finding moments of the random variable X requires the knowledge of its distribution. As n becomes larger the derivation of this distribution becomes more cumbersome. On the contrary, the use of theorem simplifies the calculation of moments whatever n may be. Even for n as small as 3 it can be seen that it is not simple to obtain the following:

$$\begin{aligned} p(0) &= q^9 + 9q^8p + 36q^7p^2 + 84q^6p^3 + 122q^5p^4 + 106q^4p^5 + 48q^3p^6 + 10q^2p^7 + qp^8 \\ p(1) &= 4q^5p^4 + 20q^4p^5 + 32q^3p^6 + 12q^2p^7 \\ p(2) &= 4q^3p^6 + 14q^2p^7 + 4qp^8 \\ p(3) &= 4qp^8 \\ p(4) &= p^9, \end{aligned}$$

where $q = 1 - p$.

The moments of X are $E(X) = 4p^4$, $E(X^2) = 4p^4 + 8p^6 + 4p^7$ and $E(X^3) = 4p^4 + 24p^6 + 12p^7 + 24p^8$. But the above results easily follow from the factorial moments given in Equations 2.1, 2.2, 2.3 for $n = 3$.

2.4 Asymptotic Distribution of X

Assuming that $n^2p^4 \to \lambda$ as $n \to \infty$, and p is small, let us find the asymptotic moments of the random variable X defined above. The moments given in Equations 2.1, 2.2, 2.3 can be expressed as

$$\mu_{[i]} = (n^2p^4)^i \phi_i\left(\frac{1}{n}\right), \tag{2.8}$$

where $\phi_i(\frac{1}{n}) = 1 + \frac{a_{i1}}{n} + \frac{a_{i2}}{n^2} + \cdots$

In fact the Expression 2.8 holds for all factorial moments $i = 1, 2, \ldots$ as the maximum power of the term n^2p^4 appearing in the *ith* factorial moment turns to be i. Since $\phi_i(\frac{1}{n}) \to 1$ when $n \to \infty$, the *ith* factorial moment of X has asymptotically the values λ^i. That is, X has asymptotically a Poisson distribution with the parameter λ.

To illustrate it, suppose that a fruit plant is grown at each of n^2 locations of an $n \times n$ lattice. Let p be the probability that a disease infects each plant independently during some time period. The number of rectangles could be an important variable of concern to a plant pathologist. For large n its distribution is approximately Poisson.

A Appendix

In their paper [2], Memon and David provide a result that facilitates a relationship between factorial moments and probabilities of specified events. They consider n possibly dependent events each of whose materialization is determined by a single binomial trial. Then the rth factorial moment of the number of materializing events is

$$\mu_{[r]} = \sum_{\binom{n}{r}} p(W)$$

where $p(W)$ denotes the probability of materialization of all events in a set of size r and the summation extends over all $\binom{n}{r}$ sets of size r.

Acknowledgement

The authors are grateful to the refree for his useful suggestions.

References

[1] Fuchs, C. E. and David, H. T. (1965). Poisson limits of multivariate run distributions. *Annals of Mathematical Statistics*, **36**(1), 215–225.

[2] Memon, A. Z. and David, H. T. (1968). The Distribution of lattice join counts. *Bulletin of the Institute of Statistical Research and Training*, **2**(2), 75–83.

[3] Moran, P. A. P. (1948). Interpretation of statistical maps. *Journal of the Royal Statistical Society, Series B*, **10**(2), 243–251.