The Negative Binomial and distributions, based on similar underlying mathematical models, may be useful in fitting data obtained in biological and medical research. This article discusses these distributions and some of the mathematical models, and the application to some medical data is presented.

SOME APPLICATIONS OF THE NEGATIVE BINOMIAL AND OTHER CONTAGIOUS DISTRIBUTIONS

John Gurland, Ph.D.

1. Introduction

MANY OF THE statistical technics for the analysis of experimental data based on samples from Normal populations are sufficiently well developed to be regarded as standard tools. In fact, it is not uncommon practice for data from non-Normal populations to be transformed in order that the resulting distribution be sufficiently close to Normal for the standard technics to be applicable.

Much of the experimentation in the biological and medical sciences pertains to statistical distributions which are far from Normal. In fact, the distributions may be discrete. For example, the data may be numbers of insects on plants, or numbers of defective teeth in people's mouths, or numbers of sicknesses suffered by industrial workers. Even if we should try to approximate the distribution of numbers of defective teeth, say, by a continuous curve, its shape would in general be far from that of the Normal frequency curve.

There are, of course, continuous distributions which are also distinctly non-Normal. A typical example in the medical sciences is the distribution of survival times of patients cured by cancer therapy (cf. Boag[1]). Although continuous non-Normal populations are widely applicable and are of interest in themselves, they will not be included within the scope of the present paper. Our attention will be focused upon such discrete distributions as the Negative Binomial and other contagious frequency distributions. The concept of contagion with respect to frequency distributions will be described below.

2. The Negative Binomial Distribution (Pascal Distribution)

2.1 *Preliminary Remarks and Background*

Among the most well known discrete distributions are the Binomial, the Poisson, and the Negative Binomial. The theoretical connections between these distributions are so close that it is hardly convenient to discuss any one of them without referring to the others.

The Binomial distribution is ascribed to Jakob Bernoulli[2] who investigated it in some detail. It is called the Binomial distribution because its probabilities may be obtained from the terms in the expansion

$$(q+p)^n = q^n + nq^{n-1}p + \frac{n(n-1)}{2!}q^{n-2}p^2 + \dots + \frac{n(n-1)(n-2)\dots(n-r+1)}{r!}q^{n-r}p^r + \dots + p^n \quad (1)$$

Here $p(0<p<1)$ is the probability of a "success" in each trial, $q=1-p$ is the probability of a "failure," and n is the number of independent trials. The notation r! represents (r) (r-1) (r-2) .... (3)(2)(1) and the Binomial coefficients appearing in (1) are usually written as

$$\binom{n}{r} = \frac{n(n-1)\dots(n-r+1)}{r!} \quad (2)$$

Thus, the probability that a Binomial variate X, say, assumes the value r is given by

$$P\left\{X=r\right\} = \binom{n}{r}p^r q^{n-r} \quad r=0, 1, 2, \dots, n \quad (3)$$

An example of a fit of the Binomial to an observed distribution is shown in Table 1. The data are from Eisenhart and Wilson[3] and give the number of monocytes in 100 blood cells of a cow in 113 successive weeks. The observed frequencies appear in the second column and the expected frequencies based on a Binomial distribution appear in the third column. Some remarks on the methods of fitting a theoretical to an observed distribution will be made in Section 3. For the present it suffices to note there is at least some correspondence between the observed and the theoretical distributions in the second and third columns, respectively.

The Negative Binomial distribution is so called because the probabilities can be obtained from the terms in the expansion

$$(q-p)^{-k} = q^{-k}\left[1 + k\frac{p}{q} + \frac{k(k+1)}{2!}\left(\frac{p}{q}\right)^2 + \dots + \frac{k(k+1)\dots(k+r-1)}{r!}\left(\frac{p}{q}\right)^r + \dots\right] \quad (4)$$

where $p>0$, $k>0$, and $q=1+p$. Thus, the probability that a Negative Binomial variate X assumes the value r is given by

$$P\left\{X=r\right\} = \left(\frac{1}{q}\right)^k \frac{k(k+1)\dots(k+r-1)}{r!}\left(\frac{p}{q}\right)^r \quad r=0, 1, 2, \dots \quad (5)$$

The following interpretation of this probability term is possible. If we imagine independent trials with $\frac{1}{q}$ as the probability of a "success" and $\frac{p}{q}$ as the probability of a "failure," then the expression in (5) is the probability that $r+k$ trials will be required to obtain k successes. Although this interpretation requires that k be a positive integer, there are interpretations arising from other mathematical models underlying this distribution which require merely that k be positive.

The Negative Binomial was formulated by Montmort[4] in 1714. Feller[5] calls the distribution given by (5) with k a positive integer a Pascal distribution in ref-

Table 1—Number of Monocytes in 100 Blood Cells of a Cow in 113 Successive Weeks (Data from Eisenhart and Wilson[3])

Monocytes in 100 Blood Cells r	Observed Frequency f	Expected Frequency (Binomial)	Expected Frequency (Poisson)
0	0	0.2	0.3
1	3	1.5	1.7
2	5	4.8	5.2
3	13	10.0	10.3
4	19	15.3	15.4
5	13	18.7	18.3
6	15	18.7	18.1
7	12	15.9	15.4
8	10	11.7	11.5
9	11	7.6	7.6
10	7	4.4	4.5
11	3	2.3	2.5
12	2	1.1	1.2
13+	0	0.8	1.0

Table 2—Distribution of Dental Caries in 12-Year-Old Children. Number of Smooth Surface Cavitation Counts (Data from Grainger and Reid[8])

Count Per Child r	Observed Frequency f	Expected Frequency (Pascal)*	Expected Frequency (Pascal)†	A_r	Expected Frequency (Poisson v Pascal)
0	63	52.6	61.2	100	57.9
1	29	31.6	29.2	71	26.7
2	12	21.5	18.8	59	19.6
3	15	15.2	13.1	44	14.8
4	8	11.0	9.6	36	11.1
5	9	8.0	7.1	27	8.4
6	5	5.9	5.4	22	6.3
7	4	4.4	4.1	18	4.7
8	6	3.2	3.2	12	3.5
9	2	2.4	2.4	10	2.6
10	3	1.8	1.9	7	1.9
11	3	1.3	1.6	4	1.4
12	2	1.0	1.2	2	1.0
13+	2	3.1	4.2	0	2.9
χ^2		14.2	10.0		10.1
$P(\chi^2)$		0.22	0.53		0.43

* Determined by method of moments.
† Determined by maximum likelihood.

erence to the mathematician Blaise Pascal.[7] By synecdoche we shall often refer to the Negative Binomial as a Pascal distribution (cf. Gurland[6]).

One of the important properties of this distribution is that its variance ($kp + kp^2$) exceeds its mean (kp). This property is sometimes referred to as overdispersion.

An example of fitting the Negative Binomial to an observed frequency distribution is shown in Table 2. The data are taken from Grainger and Reid[8] on the distribution of dental caries among 12-year-old school children. The second column in the table gives the observed distribution of smooth surface cavitation counts and the third column the expected frequencies on the basis of a Negative Binomial determined by the method of moments. The fourth column exhibits the expected frequencies calculated from a Negative Binomial determined by the method of maximum likelihood. (Remarks on methods of estimating the parameters of a distribution in attempting to fit observed data will be made in Section 3.) It will be noted there is some correspondence between the observed frequencies in the second column and the expected frequencies in the third and fourth columns. Further, the correspondence between the observed distribution and the Pascal distribution determined by moments does not appear as close as the correspondence between the observed distribution and the Pascal distribution determined by maximum likelihood.

The relative mathematical simplicity in computing probabilities based on the Negative Binomial is probably a contributing factor to its widespread use as a contagious distribution. Further aspects of this distribution will be considered below. We shall now turn our attention briefly to the Poisson distribution.

The distribution known as the Poisson was considered by Simeon Poisson[9] in 1837. This distribution may be regarded as a limiting case of the Binomial distribution in (3) by letting $n \to \infty$, $p \to 0$, while keeping $np = \lambda$, a constant. The expression in (3) then becomes

$$P\left\{X = r\right\} = e^{-\lambda} \frac{\lambda^r}{r!} \qquad r = 0, 1, 2, \ldots . \qquad (6)$$

It may be remarked that for a Poisson variate the variance (λ) is equal to the mean (λ) whereas for a Pascal variate the variance exceeds the mean.

An example illustrating the fitting of a Poisson distribution to observed frequencies is given in Table 1. These data were considered above in fitting the Binomial distribution. If the same data are fitted by a Poisson distribution, we obtain the expected frequencies shown in the fourth column. It will be observed the expected frequencies corresponding to the Binomial and Poisson distributions are very close. This results from the large value of n(100) and the small value of p (0.0596) for the Binomial. The mean of the Poisson is $\lambda = np = 5.96$.

It may also be remarked that the Poisson may be regarded as a limiting case of the Negative Binomial distribution by letting $k \to \infty$, $p \to 0$ in (5) while keeping the mean constant and equal to λ, say.

2.2 *Apparent Contagion (Compound Poisson)*

It was pointed out by Feller[12] there are two kinds of contagion, which he refers to as "true contagion" and "apparent contagion."

Apparent contagion is the result of heterogeneity arising from distributions on the parameters involved in a population. Thus, in the case of the Poisson distribution (6), if the values of the parameter λ are obtained randomly through some statistical population, the resulting "mixture" of Poisson distributions would afford an example of apparent contagion. Such a distribution obtained in this manner is called a compound Poisson distribution (cf. [12]). This notion was developed by Greenwood and Yule[13] and has been widely applied in studies on accident proneness (cf. Arbous and Kerrich[14]).

If the parameter λ in (6) is a random variate with a Gamma frequency function given by

$$p(\lambda) = \frac{\alpha^\beta}{\Gamma(\beta)} e^{-\alpha\lambda}\lambda^{\beta-1} \qquad \lambda > 0, \alpha > 0, \beta > 0 \qquad (7)$$

the resulting compound Poisson distribution can be shown (cf. [6]) to have probabilities of the form

$$P\left\{X = r\right\} = \frac{\beta(\beta+1)\ldots(\beta+r-1)}{r!} \left(\frac{\alpha}{1+\alpha}\right)^\beta \left(\frac{1}{1+\alpha}\right)^r \qquad (8)$$

On setting $\beta = k$, $p = \frac{1}{\alpha}$, and $q = \frac{1+\alpha}{\alpha}$, it is evident that (8) becomes identical with (5), the probability corresponding to a Negative Binomial distribution. Consequently the Pascal may be regarded as a compound Poisson distribution, wherein the compounding has been effected upon the parameter λ through a Gamma distribution. As a matter of interest the shape of the Gamma frequency curve (7) has been indicated in Figure 1 for a few combinations of values of α and β.

The Negative Binomial fitted to the data on dental smooth surface cavitation counts in Table 2 affords an example of the possibility of interpreting this distribution as a compound Poisson. The distribution might be regarded as arising from an aggregate of Poisson populations with mean values following a distribution of the form (7).

2.3 *True Contagion*

In the case of true contagion the probability of a "favorable" event depends on the occurrence of previous favorable events. Feller[16] discussed the following

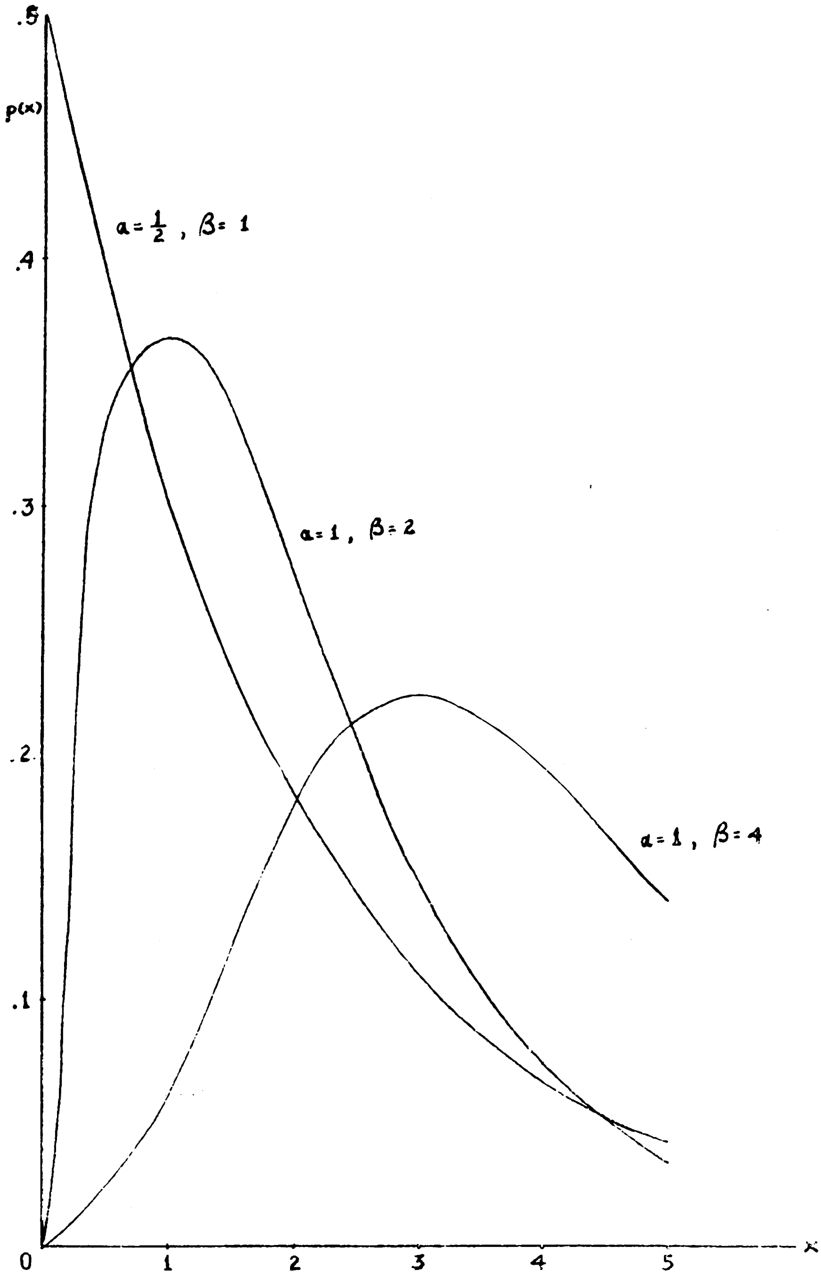

Figure 1—The Gamma Frequency Function[7] for Some Values of α and β

mathematical model due to Polya[17] which is based on true contagion and leads in a relatively simple way to the Negative Binomial distribution.

Suppose an urn contains N balls of which Np are white and Nq are black $(p+q=1)$. n successive drawings of a ball are made from the urn under the procedure that after each drawing the ball is replaced and in addition $N\delta$ balls of the color last drawn are added to the urn. Let X denote the number of white balls in n successive drawings. Then the probability that X assumes the value r is given by

$$P\left\{X=r\right\}=\binom{n}{r}\frac{p(p+\delta)(p+2\delta)\ldots(p+[r-1]\delta)q(q+\delta)(q+2\delta)\ldots(q+[n-r+1]\delta)}{1(1+\delta)(1+2\delta)\ldots(1+[n-1]\delta)} \quad (9)$$

If we let $n\rightarrow\infty$, $p\rightarrow 0$, $\delta\rightarrow 0$, while keeping $np=\lambda$, $n\delta=\eta$ constant, then (9) becomes

$$P\left\{X=r\right\}=\frac{k(k+1)\ldots(k+r-1)}{r!}\left(\frac{\eta}{1+\eta}\right)^r\left(\frac{1}{1+\eta}\right)^{\frac{\lambda}{\eta}} \quad (10)$$

On setting $\frac{\lambda}{\eta}=k$, it is evident from (5) that this is a Negative Binomial distribution. Thus, two different mathematical models based on apparent and true contagion, respectively, lead to the Negative Binomial distribution.

The data in Table 2 also afford an example of the possibility of interpreting the Negative Binomial as arising from true contagion. For it is plausible to assume that the presence of a dental cavity in a child's mouth increases the probability of another cavity.

In the light of two types of contagion providing a possible explanation of the relevance of the Negative Binomial distribution it is apparently impossible from the data to distinguish between the two types. It may be remarked, however, that in the case of multidimensional forms of the Negative Binomial some progress has been made in this direction (cf. Bates and Neyman,[18] Arbous and Kerrich[14]). An interesting study on accident proneness with some applications of the bivariate Negative Binomial distribution has been made by Adelstein.[19]

2.4 *Model of Random Colonies (Generalized Poisson)*

Another widely used mathematical model which leads to the Negative Binomial distribution is based on a random distribution of colonies. As an example, in counts of soil bacteria in microscopic fields (cf. Jones, Mollison, and Quenouille[20]), if the number of bacterial colonies per field follows a Poisson distribution and the number of bacteria per colony a Logarithmic distribution, then the distribution of bacteria per field is a Negative Binomial.

An example which occurs in entomology is concerned with the counts of larvae over the plots in a field (cf. Neyman,[21] Skellam,[22] Evans[23]). The larvae are hatched from egg masses which appear at random over the field. If the number of egg masses represented on a plot under observation follows a Poisson distribution, and if the survivors from the egg masses follow a Logarithmic distribution, then the resulting distribution of larvae on plots will be a Negative Binomial.

In technical terms, on the basis of the above model, the Negative Binomial may be referred to as a generalized Poisson distribution (cf. Feller[12]) in which the "generalizer" (cf. Gurland[11]) is a Logarithmic random variable.

The Logarithmic distribution referred to above has probabilities which are the terms in the logarithmic series

$$-\alpha\log(1-\tau)=\alpha\left[\tau+\frac{\tau^2}{2}+\frac{\tau^3}{3}+\ldots\right] \quad (11)$$

where $0<\tau<1$ and $\alpha\log(1-\tau)=-1$. Thus, for a Logarithmic random variable X, say, $P\left\{X=0\right\}=0$ and

$$P\left\{X=r\right\}=\alpha\frac{\tau^r}{r};\ r=1,2,\ldots \quad (12)$$

This distribution is ascribed to Fisher[10] and is usually written in a form such as (12) in which the zero count (corresponding to $P\{X=0\}$) is excluded. It can be shown that this distribution is actually a limiting case of the Negative Binomial when $k \rightarrow 0$ and the zero count is excluded.

A Logarithmic distribution in which the zero count is not excluded has been considered by Katti and Gurland.[15]

There are other models besides those considered in 2.2, 2.3, and the present section which lead to the Negative Binomial distribution (cf. Anscombe,[24] Kendall[25]) but which are omitted here for brevity.

3. Estimation and Fitting

One of the simplest methods for estimation of the parameters in the Negative Binomial and other distributions is the method of equating the first few sample moments to the corresponding theoretical moments and solving the resulting equations for the unknown parameters. This is known as the method of moments. The number of moments used is the same as the number of unknown parameters.

In the case of the Negative Binomial distribution (5) the equations using the first two moments are

$$kp=\bar{x}; \qquad kp(1+p)=s^2 \qquad (13)$$

where $\bar{x}$, s^2 are the sample mean and variance, respectively, obtained in the usual manner

$$N\bar{x}=\sum_{i=1}^{N} x_i \ ; \ Ns^2=\sum_{i=1}^{N} (x_i-\bar{x})^2$$

where N is the total number of observations.

As an example, the observed distribution in Table 2 yields $\bar{x}=2.527$, $s^2=10.642$; and the estimates of p and k obtained by solving the two equations in (13) are $p=3.210$ and $k=0.787$. These are the values of the parameters used in calculating the expected frequencies shown in the third column of the table.

Another comparatively simple method sometimes used is to equate the first few sample frequencies to the corresponding expected frequencies and solve the resulting equations for the unknown parameters (cf. [23,24]). This method might be called the method of frequencies. The number of frequencies used is, of course, the same as the number of unknown parameters.

In the case of the Negative Binomial distribution, using the first two frequencies, the equations to solve for p and k become

$$f_0=Nq^{-k}; \ f_1=Nkpq^{-k-1} \qquad (14)$$

where f_0, f_1 are the observed frequencies of zero and one counts, respectively. From the data in Table 2, $f_0=63$, $f_1=29$, and $N=163$.

Still another relatively simple method sometimes used is to equate the first few sample moments and sample frequencies to the corresponding theoretical moments and frequencies. This is, in effect, a combination of the method of moments and the method of frequencies. Again, of course, the number of equations to be solved must be the same as the number of unknown parameters.

The efficiency of the above methods will depend on the true values of the parameters to be estimated. These methods may be reliable in some cases and unreliable in others (cf. Fisher,[39] Anscombe,[24] and Evans[23]) but are used when a simple method of estimating is desired.

A fully efficient method of estimating the parameters is the procedure known as maximum likelihood (cf. Fisher[26]). The required computations are usually not as simple as for the preceding methods, but the method of maximum likelihood is used when it is desired to extract all the possible information from the data. In the case of the Negative Binomial the equations to be solved for p and k are (cf. Fisher[27])

$$\sum_{r=0}^{\infty}\left(\frac{A_r}{k+r}\right)=N\log 1+p; \ \bar{x}=kp \qquad (15)$$

where $A_r = f_{r+1} + f_{r+2} + \ldots$ and the f_r are the observed frequencies.

The first equation in (15) requires an iterative procedure for its solution (cf. Bliss[28]). For the data in Table 2 the values of A_r are shown in the fifth column and the solution of (15) is $p = 4.299$, $k = 0.588$. The corresponding frequencies based on this Negative Binomial are shown in the fourth column of Table 2.

It may be remarked that the method of estimation known as minimum χ^2 (cf. Cramér[29], Neyman[30]) will also yield estimators efficient in the same sense as those obtained by maximum likelihood. Usually the required computations are also rather involved as in the case of maximum likelihood, but in some instances the computations are simpler. As an example of the application of the minimum χ^2 method to some other distributions the reader is referred to Berkson.[31,32]

A criterion commonly used for testing the goodness of fit to the observed frequencies by a theoretical distribution is the χ^2 statistic,

$$\chi^2 = \Sigma \frac{(\text{observed} - \text{expected})^2}{\text{expected}} \qquad (16)$$

due to Karl Pearson.[33] If the value of this statistic turns out to be not unduly large in the sense that the probability of obtaining at least such a value is above 5 per cent, the fit is not considered unsatisfactory. This does not, however, necessarily imply acceptance of the hypothesis of the given theoretical distribution (cf. Cochran[34]) especially if the data are scanty, but should rather be interpreted cautiously as some evidence in favor of the hypothesis tested.

The χ^2 statistic is also sometimes used to compare the closeness of fit to an observed frequency distribution by several different theoretical distributions (cf. Bliss,[28] Maguire, Brindley, and Bancroft,[36] Beall and Rescia[35]). If the χ^2 value using one distribution is consistently smaller than the value using another, this would suggest grounds for preferring the one distribution to the other, but again one should be cautious in making this inference.

To provide an illustrative example of the use of the χ^2 statistic, we return again to Table 2. For a fit by the Pascal distribution based on the method of moments the χ^2 statistic has the value 14.2. To obtain the probability of getting a value at least as large as this, we refer to a table of the χ^2 distribution (cf. Cramér[33]) and find the probability corresponding to 11 (i.e., $14-3$) degrees of freedom. The number of degrees of freedom is obtained by subtracting from the number of categories (14) one plus the number of parameters estimated. This leads to a probability value of approximately 22 per cent as shown at the foot of the third column in Table 2. The corresponding χ^2 value using a fit based on maximum likelihood is 10.0 and the probability is accordingly 53 per cent. The decrease in the χ^2 value for the fit based on maximum likelihood is substantial. The fits based on both the above methods are not rejected, although the fit based on maximum likelihood appears preferable.

The distribution of the χ^2 statistic (16) is really not exactly that of a χ^2 random variable; and especially for probabilities corresponding to the tail of the distribution some care is required in carrying out the test. Some writers recommend pooling frequencies of categories which have fewer than five or ten observations, but the inflexible use of this rule may be harmful (cf. Cochran[34]).

4. Other Contagious Distributions

It has been pointed out that the Negative Binomial distribution may be regarded as a compound Poisson (Section 2.2) or as a generalized Poisson distribution (Section 2.4). In the notation used by Gurland[6] the Negative Binomial can be expressed as follows:

Poisson Λ Gamma (Compound Poisson) (17)

Poisson v Logarithmic (Generalized Poisson) (18)

The symbol Λ signifies that the Poisson distribution preceding it is being compounded and that the Gamma distribution succeding it is employed as a "compounder" on the parameter λ, say, of the Poisson. As an example of a different compounder

Poisson Λ Poisson (19)

signifies that the parameter in the Poisson on the left of Λ itself follows a Poisson distribution (distribution on the right of Λ). This compound Poisson distribution given by (19) is known as the Neyman Type A distribution (cf. Feller,[12] Neyman[21]) and has been used widely in fitting entomological data (cf. Beall[37]).

Other examples of compound distributions are

Binomial Λ Poisson (20)

Pascal Λ Poisson (21)

The distribution in (20) is a compound Binomial with the Poisson distribution as a compounder, and the distribution in (21) is a compound Negative Binomial with the Poisson distribution as a compounder. To avoid ambiguity in (20) and (21) as regards which parameters are involved in the compounding process, we might write

Binomial $\underset{n}{\Lambda}$ Poisson and Pascal $\underset{k}{\Lambda}$ Poisson, respectively, to emphasize that it is the parameter n in (3) and the parameter k in (5) which follows the Poisson distrbution as the compounder in these distributions.

The symbol v in (18) signifies that the Poisson distribution which precedes it is generalized through the Logarithmic distribution which succeeds it. In terms of the descriptive language employed in Section 2.4 we can interpret (18) by saying the Poisson distribution preceding v is the distribution of colonies and the Logarithmic distribution succeding v is the distribution of individuals per colony. Some examples of other generalized Poisson distributions are

Poisson v Poisson (22)

Poisson v Binomial (23)

Poisson v Pascal (24)

As Feller[12] has pointed out, the generalized Poisson in (22) is also a Neyman Type A distribution. By a theorem in Gurland[6] it can readily be seen that the compound Binomial in (20) and the generalized Poisson in (23) are the same form of distribution; further, the compound Pascal in (21) and the generalized Poisson in (24) are the same form of distribution. Thus, the models in Sections 2.2 and 2.4 are equivalent, in a certain sense, for the class of distributions considered by Gurland.[6]

Some examples of fitting other contagious distributions than the Negative Binomial are presented in Tables 2 and 3. In Table 2, the sixth column shows the expected frequencies on the basis of a Pascal Λ Poisson distribution fitted to the data. As shown at the foot of the sixth column the χ^2 statistic has the value 10.1; and the corresponding probability obtained from the χ^2 distribution with 10 degrees of freedom is 43 per cent.

It may be remarked that the method used in estimating the three parameters in the Pascal Λ Poisson distribution was as follows: The first two sample moments were equated to the first two theoretical moments, and the ratio f_1/f_0 of the first two observed frequencies was equated to the corresponding ratio of theoretical frequencies. This is a combination of the method of moments with a modification of the method of frequencies. For this distribution the method is much simpler than the method of maximum likelihood; further, as apparent from Table 2 the fit is about as close as that accomplished by the Negative Binomial distribution based on maximum likelihood.

On assumption of a Pascal Λ Poisson distribution underlying these data two interpretations, inter alia, are possible. First, the distribution of smooth surface cavitation counts per child may be regarded as a compound Pascal with the different values of k following a Poisson distribution. A second interpretation regards the Pascal Λ Poisson as a Poisson v Pascal. Under this interpretation the number of defective teeth per child (cf. colonies per field) could be regarded as following a Poisson distribution, and the number of carious smooth surfaces per defective tooth (cf. individuals per colonies) could be regarded as following a Pascal distribution.

Table 3—Sickness Distribution of 302 Shunters for Period 1943-1947 Data from Adelstein[19]

Number of Sicknesses r	Observed Frequency f	Expected Frequency (Pascal)*	Expected Frequency (Pascal)†	Expected Frequency (Neyman Type A)	Expected Frequency (Binomial Poisson)
0	25	8.1	11.6	28.0	20.5
1	7	15.3	18.6	8.1	5.8
2	15	20.4	22.3	14.7	13.9
3	23	23.3	23.9	19.2	20.3
4	23	24.5	24.0	21.0	21.8
5	22	24.4	23.1	21.2	21.6
6	18	23.4	21.7	20.9	22.2
7	19	21.8	20.0	20.3	22.7
8	22	19.9	19.0	19.4	21.8
9	23	17.8	17.0	18.1	20.1
10	14	15.7	15.1	16.6	18.3
11	12	13.7	13.2	14.9	16.4
12	17	11.9	11.5	13.2	14.3
13	11	10.2	10.0	11.5	12.2
14	6	8.7	8.7	9.9	10.3
15	10	7.3	7.4	8.4	8.5
16	3	6.2	6.4	7.1	7.0
17	7	5.2	5.4	5.9	5.6
18	3	4.3	4.6	4.9	4.4
19	4	3.6	3.9	4.0	3.5
20	1	3.0	3.3	3.2	2.9
21	4	2.5	2.8	2.6	2.1
22	4	2.0	2.4	2.0	1.6
23	3	1.7	2.0	1.6	1.2
24	3	1.4	1.7	1.3	0.9
25+	3	5.7	2.4	4.0	2.1
χ^2		60.6	40.7	18.7	26.7
$P(\chi^2)$		0.001	0.01	0.72	0.22

* Determined by method of moments.
† Determined by maximum likelihood.

A further example is afforded by the data in Table 3 giving the frequency of sicknesses among 302 railway shunters over the period 1943-1947. These data are taken from Adelstein.[19] The second column of Table 3 shows the observed frequencies and the third column the expected frequencies based on a Negative Binomial determined by the method of moments. The fourth column shows the expected frequencies based on a Negative Binomial determined by the method of maximum likelihood and the fifth column shows the expected frequencies based on a Neyman Type A distribution determined by the method of moments. The sixth column shows the expected frequencies based on a Binomial Λ Poisson determined by the method of moments with the parameter n taken as 7.

It is apparent from these results the Negative Binomial provides a poor fit even when determined by maximum likelihood. On the basis of the χ^2 statistic the Binomial Λ Poisson provides a better fit and the Neyman Type A the best fit. A possible explanation of the poor fit by the Pascal distribution is its unimodality and the multimodality of the observed frequency distribution, with modes at $r=0$, 3, 4, 9, and so forth. The Binomial Λ Poisson and the Neyman Type A are, for certain values of the parameters, multimodal distributions and hence better adapted to situations of this kind.

Conclusion

The Negative Binomial and other distributions based on similar underlying mathematical models may be effective in fitting various types of data arising in biological and medical research. One may ask, of course, what is the purpose of trying to fit data by distributions such as those considered.

In the case of samples obtained from a population in which Normality is assumed the statistical tests performed do not usually include a test of fit because of the experience of a reasonably good fit by the Normal distribution. Furthermore, many statistical tests based on Normal theory enjoy the property of robustness (cf. Box[38]) which implies only slight distortions in the behavior of the tests when the underlying population is non-Normal.

In the case of data from discrete distributions fitting may be desirable to verify the form of the underlying population assumed. With a knowledge of the underlying population it is possible at least theoretically to construct tests or estimate parameters for the purpose of making statistical inference.

In fitting a distribution to a given set of biological data, say, it is also important for the mathematical model underlying the theoretical distribution to have a reasonable biological meaning. The contagious distributions, including the Negative Binomial, offer interesting possibilities in this direction.

Summary

The Negative Binomial distribution is discussed in its relation to the Binomial and Poisson distributions. Some mathematical models which lead to it are also discussed. These include the compound Poisson and generalized Poisson distributions, and it is shown how the Gamma and Logarithmic distributions become involved in these representations. Further, a model based on true contagion is shown to yield the Negative Binomial as a limiting case. Other compound and generalized distributions distinct from the Negative Binomial, but based on some of the models which lead to it, are presented and fitted along with the Negative Binomial to some medical data. A few remarks on estimation and fitting are also included.

Acknowledgment—The author is grateful to S. K. Katti and R. Shumway for conscientious computational aid, to L. Gerende for a useful bibliographical reference, and to Dr. W. L. Wishart for helpful comments on some dental aspects of the data considered. The further computational assistance of T. Roseberry and R. Connelly is also appreciated.

REFERENCES

1. Boag, J. W. Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy. J. Roy. Statist. Soc. (Supplement) 11:15–53, 1949.
2. Bernoulli, James. "Ars conjectandi," 1713.
3. Eisenhart, C., and Wilson, P. W. Statistical Methods and Control in Bacteriology. Bact. Rev. 7:57–137, 1943.
4. Montmort, P. R. "Essai d'analyse sur les jeux de hazards," 1714.
5. Feller, W. An Introduction to Probability Theory and Its Applications. New York, N. Y.: Wiley, 1957.
6. Gurland, John. Some Interrelations Among Compound and Generalized Distributions. Biometrika 44: 265–268, 1957.
7. Pascal, B. "Varia opera mathematica D. Petri de Fermat . . . Tolosae," 1679.
8. Grainger, R. M., and Reid, D. B. W. Distribution of Dental Caries in Children. J. Dent. Res. 33:613–623, 1954.
9. Poisson, S. D. "Recherches sur la probabilité des judgments en matiére criminelle et en matiére civile, précedées des régles générales du calcul des probabilitiés," 1837.
10. Fisher, R. A.; Corbett, R. S.; and Williams, C. B. The Relation Between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population. J. Animal Ecology 12:42–58, 1943.
11. Gurland, John. A Generalized Class of Contagious Distributions. Biometrics 14:229–249, 1958.
12. Feller, W. On a General Class of Contagious Distributions. Ann. Mathematical Statist. 14.389–400, 1943.
13. Greenwood, M., and Yule, G. V. An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents. J. Roy. Statist. Soc. 83:255-279, 1920.
14. Arbous, A. G., and Kerrich, J. E. Accident Statistics and the Concept of accident Proneness. Biometrics 7:340–432, 1951.
15. Katti, S. K., and Gurland, John. Some Families of Contagious Distributions. Tech. Rep., U. S. Air Force Office of Scientific Research, 1958.
16. Feller, W. "On the Theory of Stochastic Processes with Particular Reference to Applications. Proc. Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, Calif.: University of California Press, 1949.
17. Polya, G. Sur quelques points de la théorie des probabilités. An. Inst. Henri Poincaré 1:117–161, 1930.
18. Bates, Grace E., and Neyman, J. Contributions to the Theory of Accident Proneness. II: True or False Contagion. University of California Publications in Statistics 1:255–276, 1952.
19. Adelstein, A. M. Accident Proneness: A Criticism of the Concept Based Upon an Analysis of Shunters' Accidents. J. Roy. Statist. Soc., Series A 115:354–410, 1952.
20. Jones, P. C. T.; Mollison, J. E.; and Quenouille, M. H. A Technique for the Quantitative Estimation of Soil Microorganisms. J. General Microbiol. 2: 54–69, 1948.
21. Neyman, J. On a New Class of Contagious Distributions Applicable in Entomology and Bacteriology. Ann. Mathematical Statist. 10:35–57, 1939.
22. Skellam, J. G. Studies in Statistical Ecology I, Spacial Pattern. Biometrika 39:346–362, 1952.
23. Evans, D. A. Experimental Evidence Concerning Contagious Distributions in Ecology. Ibid. 40:186–211, 1953.
24. Anscombe, F. J. Sampling Theory of the Negative Binomial and Logarithmic Series Distributions. Ibid. 37:358–382, 1950.
25. Kendall, D. G. Stochastic Processes and Population Growth. J. Roy. Statist. Soc., Series B 11:230–264, 1949.
26. Fisher, R. A. On the Mathematical Foundations of Theoretical Statistics. Philosophical Trans. Roy. Soc. London, Series A 222:309–368, 1922.
27. ————. Note on the Efficient Fitting of the Negative Binomial. Biometrics 9:197–199, 1953.
28. Bliss, C. I. Fitting the Negative Binomial Distribution to Biological Data. Ibid. 9:176–196, 1953.
29. Cramér, H. Mathematical Methods of Statistics. Princeton, N. J.: Princeton University Press, 1946.
30. Neyman, J. Contribution to the Theory of the ^{2} Test. Proc. Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, 1949.
31. Berkson, Joseph. Maximum Likelihood and Minimum χ^2 Estimates of the Logistic Function. J. Am. Statist. A. 50:130-162, 1955.
32. ————. Estimate of the Integrated Normal Curve by Minimum Normit Chi-Square with Particular Reference to Bio-Assay. Ibid. 50:529–549, 1955.
33. Pearson, Karl. On the Criterion That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling. Philosophical Magazine V, 50:157–175, 1900.
34. Cochran, W. G. The χ^2 Test of Goodness of Fit. Ann. Mathematical Statist. 23:315–345, 1952.
35. Beall, G., and Rescia, R. A Generalization of Neyman's Contagious Distributions. Biometrics 9:354–386, 1953.
36. McGuire, J. V.; Brindley, T. A.; and Bancroft, T. A. The Distribution of European Corn-Borer Larvae Pyrausta Nubilalis (HBN), in Field Corn. Ibid. 13: 65–78, 1953.
37. Beall, G. The Fit and Significance of Contagious Distributions When Applied to Observations on Larval Insects. Ecology 21:460–474, 1940.
38. Box, G. E. P., and Andersen, S. L. Permutation Theory in the Derivation of Robust Criteria and the Study of Departures from Assumption. J. Roy. Statist. Soc., Series B 17:1–34, 1955.
39. Fisher, R. A. The Negative Binomial Distribution. Ann. Eugenics 11:182–187, 1941.

Dr. Gurland is professor of statistics, Iowa State University, Ames, Iowa.

This paper was presented before the Statistics Section of the American Public Health Association at the Eighty-Sixth Annual Meeting in St. Louis, Mo., October 28, 1958.