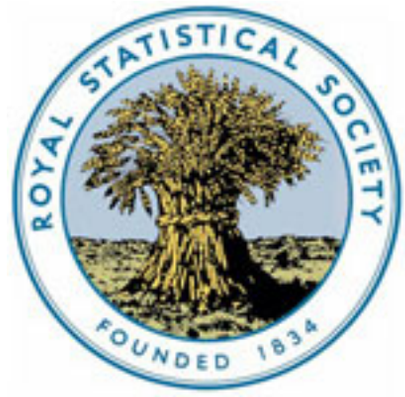

An Inquiry into the Nature of Frequency Distributions Representative of Multiple Happenings with Particular Reference to the Occurrence of Multiple Attacks of Disease or of Repeated Accidents
Author(s): Major Greenwood and G. Udny Yule
Source: *Journal of the Royal Statistical Society*, Vol. 83, No. 2 (Mar., 1920), pp. 255-279
Published by: Wiley for the Royal Statistical Society
Stable URL: http://www.jstor.org/stable/2341080
Accessed: 04/06/2014 15:18

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Wiley and *Royal Statistical Society* are collaborating with JSTOR to digitize, preserve and extend access to *Journal of the Royal Statistical Society.*

http://www.jstor.org

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

MISCELLANEA.

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.

By Major Greenwood and G. Udny Yule.

Section I.—*Introductory; the pigeon-hole schema; the Poisson series.*

In the practical applications of statistics such problems as the following often present themselves:—Of n households exposed to risk, m_0 returned 0 cases of disease, m_1 returned each a single case, m_2 each two cases . . . m_n each n cases. Might such a distribution have arisen from sampling a "population," each member of which was subject to a constant chance of infection throughout the period of exposure, or is the form of the distribution valid evidence that particular households were especially prone to take the disease in question?

A precisely similar problem is to be solved when we desire to ascertain whether the frequency of multiple accidents sustained by individual operatives in a factory is the product of uniform or of variable cause groups.

But little thought is required to make it appear that such problems are of first-rate importance; we merely give an instance. The degree of infectivity of various diseases has been assigned by many epidemiologists from an examination of the relative frequencies of multiple cases in households and this without consideration of the probabilities involved; in similar fashion the immunising value of a successful passage through an attack of disease is inferred from the absolute rarity of second or third attacks: again without recourse to any criterion of the general frequency type. This paper describes an attempt to provide more exact criteria for use in investigations of the kind.

A first step in the direction of reducing such inquiries to a uniform scientific rule was taken by Professor Karl Pearson in his study of the distribution of multiple cancer deaths in houses.* Professor Pearson looked at the problem as one of distributing n balls in N

* *Biometrika*, ix, 1913, p. 28.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

pigeon holes; he supposed that the chance of a case of cancer occurring in any one house was 1/N and devised a test for the probability that the observed discrepancy between the actual distribution and that of the binomial $\left(\frac{1}{N}+\frac{N-1}{N}\right)^n$ might have arisen by chance.

The sufficiency of this comparison needs justification. Thus, suppose we have the case of a platoon of N soldiers amongst whom n wounds have been distributed, m_1 soldiers having been wounded once, m_2 wounded twice, &c.; it might appear that the proportion n/N should be the characteristic parameter and that if the same proportional distribution of wounds amongst 2N soldiers, receiving $2n$ wounds, has been observed, the distribution should be *in pari materiâ*. In effect if the distribution over N individuals had happened to be written out upon s sheets each relating to N/s soldiers, different conclusions might be reached, by the pigeon-hole schema, on combining the s separate binomials from those that would emerge by fitting a binomial to the pooled data. It is also to be remarked that any modification of the strict pigeon-hole schema to agree with such a condition as that the receipt of one wound renders the victim more likely to be (or less likely to be) wounded again, involves the absurdity that the others will *ipso facto* become less liable to injury. The total number of balls and the total volume of the pigeon holes are constants; if a pigeon hole that has received a ball expands and becomes more likely to secure a second, it must squeeze its neighbours and diminish their capacities.

The problem of randomness or non-randomness of distribution in such cases as that of cancer houses we think must be formulated in the following terms.

1. By an uncomplicated chance distribution, we suppose that the distributing factors are independent of (*a*) the previous history of the houses in respect of the occurrence or non-occurrence of a case within any particular house during the period of the distribution and also independent of (*b*) the character of the inmates and their individual predisposition (we are supposing of course that the numbers, ages and sexes of the inmates are alike in all houses). Under these conditions the distribution should depend upon a single parameter. In the pigeon-hole schema for N large and n/N finite, the condition is fulfilled, but we do not think that for N small the employment of this binomial can be justified.

2. By a modified chance distribution we understand (*a*) that either the happening of the event at random, as under 1, differentiates the population into sub-groups within each of which the above

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

conditions are fulfilled, but the numbers in the sub-groups undergo continuous modification from the beginning to the end of the period of observation. Or (*b*) we suppose that the population is *ab initio* divisible into sub-groups for each of which the chance is different but constant throughout. The first case corresponds to that of a population of equally susceptible units exposed to a disease which is infectious or the sustaining of which affects after-liability to take it by the individual or within the household. The second case is that of varying original susceptibility. No doubt in reality both factors would be involved.

For the purpose of recognizing a case partaking of the character 1, we believe that the ordinary Poisson limit to the binomial is the correct standard to employ. It in fact provides a lower limit. A distribution which, judged by this process, does not appear to be one of uncomplicated chance is not likely to be in conformity with our postulate 1.

For consider the matter thus. Let the N persons be exposed to risk during a small interval of time—an interval so small that the chance of any one person meeting with two accidents may be regarded as negligible. Let the chance of a person meeting with an accident during this small interval be p, and the chance of his not meeting with an accident be q. Then at the end of the interval the numbers with 0 and with 1 accident respectively will be given by N $(q + p)$. At the end of 2 intervals the numbers with 0, 1 and 2 accidents will be given by the expansion of N $(q + p)^2$, and so on. At the end of T intervals, therefore, the distribution will be given by the binomial expansion of N $(q + p)^T$. But here p is very small and q near unity, while T we must suppose very large in order to give us a finite number of accidents. Hence if λ is the ratio of the number of accidents n (*i.e.*, pTN) to the number N of persons at risk the true distribution is given by the Poisson series

$$e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2!} + \quad . \quad . \quad .\right)$$

This, as it seems to us, and not the pigeon-hole binomial, is the form to which "uncomplicated distributions" should be expected to be assignable. But when the number of persons at risk, N, is large the distinction breaks down, for then as we have already pointed out the "pigeon-hole binomial" is itself merged in Poisson's series.

But we think it right to add that the *à posteriori* test of fit to the Poisson series by the χ^2 method is likely to exaggerate the evidence in favour of the distribution being random, owing to the fact that we determine λ from the observations themselves.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Thus, suppose we are given *à priori* that λ is 0.1. Then the distribution to be expected is, for 10,000 sets

0	9048
1	905
2	45
3	2

Suppose the observed data are :—

0	8997
1	950
2	50
3	3

The concordance of the observed with the expected series is not particularly close, and the probability P of a fit as bad or worse arising on random sampling would be only 0.32. We might therefore reasonably suspect some source of divergence. But fitting a Poisson series from the observations we have for the argument 0.1059 and for the "expected" values calculated from this argument :—

0	8995
1	953
2	50
3	2

giving an excellent fit and a value of P well over 0.9. It should always be remembered, in using the χ^2 method, that the results in strictness apply only to testing the fit of an observed distribution to a distribution given *à priori*. The figures obtained for P must be regarded with a little caution when the constants of the fitted curve or polygon are obtained from the observations themselves.

Section II.—*The generalized pigeon-hole schema.*

For the reasons just set out, we consider the range of experience within which the pigeon-hole schema is applicable to be narrow and dissent from the opinion that such a problem as that of random or not random distribution of cases of disease in houses can be solved by an appeal to the method. Nevertheless the schema admits of various modifications not devoid of interest and we shall therefore set out the results we have obtained.

The simple case of a pure binomial distribution, viz. :—

$$N\left(\frac{N-1}{N}+\frac{1}{N}\right)^n$$

N being the number of pigeon holes and n the number of balls to be distributed, together with the conditions under which the binomial

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

approximates to the exponential form of Poisson, are too familiar to need any comment beyond what is contained in other sections of this memoir. We pass at once to a modification of this form, making the assumption that the chance of any pigeon hole receiving a ball is not independent of the number of balls already lodged in it, *i.e.*, is not constantly equal to $\frac{1}{N}$.

Let us write ${}_m{}^n u$ for the number of pigeon holes which after the nth ball has been distributed contain each m balls. Thus, after $n-1$ balls have been distributed, on the average ${}^{n-1}{}_0 u$ contain no balls, ${}^{n-1}{}_1 u$ contain 1 ball . . . ${}^{n-1}_{n-1} u$ contain $n-1$ balls.

Then on the hypothesis of constant chance the distribution of the nth ball will be governed by the relation

$$ {}^{n-1}{}_0 u + {}^{n-1}{}_1 u + \ldots\ldots {}^{n-1}_{n-1} u = N \qquad (1) $$

But if the previous disposition of balls affects the fall of the nth then the governing condition is not (1) but

$$ \alpha\, {}^{n-1}{}_0 u + \beta\, {}^{n-1}{}_1 u + \gamma\, {}^{n-1}{}_2 u + \ldots\ldots = N \qquad (2) $$

The empty pigeon holes receiving $1 - \dfrac{\beta\, {}^{n-1}{}_1 u + \gamma\, {}^{n-1}{}_2 u + \ldots\ldots}{N}$

or $${}^n{}_0 u = {}^{n-1}{}_0 u - 1 + \frac{\beta\, {}^{n-1} u_1 + \gamma\, {}^{n-1}{}_2 u + \ldots\ldots}{N} \qquad (3)$$

$\alpha, \beta, \gamma \ldots\ldots$ being positive constants.

As will hereafter appear, there is no loss of generality in writing $\beta = \gamma = \ldots\ldots s$.

so that (3) becomes—

$$ {}^n{}_0 u - \frac{N-s}{N}\, {}^{n-1}{}_0 u = s - 1 \qquad (4) $$

where s is a constant. Equation (4) is the fundamental equation of what we shall term the " biassed pigeon-hole schema."

When there is no bias, $s = 1$ and (4) reduces to

$$ {}^n{}_0 u - \frac{N-1}{N}\, {}^{n-1}{}_0 u = 0 \qquad (5) $$

which is a simple homogeneous equation of differences, its solution being ${}^n{}_0 u = C\left(\frac{N-1}{N}\right)^{n-1}$.

Since ${}^1{}_0 u = N - 1$, $C = N - 1$; and

this becomes $N\left(\frac{N-1}{N}\right)^n$, the appropriate term of the binomial.

If $s \neq 1$, we have a first order difference equation with constant coefficients, the solution being—

$$ {}^n{}_0 u = \frac{N}{s}\left(\frac{N-s}{N}\right)^n + \frac{N(s-1)}{s} \qquad (5A) $$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

and it appears that when all pigeon holes into which balls have fallen are equally favoured as against empty pigeon holes, the number which will remain empty after n balls have been distributed is :—

$$\left(\frac{N-s}{N}\right)^n \left\{ N + (s-1)\left[\frac{N}{s}\left(\left\{ \frac{N}{N-s} \right\} - 1 \right) \right] \right\} \qquad (6)$$

while ${}^n{}_1u$, ${}^n{}_2u$, etc., are given by the terms of

$$\frac{N}{s}\left(\frac{N-s}{N} + \frac{s}{N}\right)^n . \qquad (7)$$

omitting the first term.

The constant s can be deduced from the second moment coefficient of the statistics.

For the second moment coefficient of $\left(\frac{N-s}{N} + \frac{s}{N}\right)^n$ about the centre of the zero group is $\frac{s\,n\,(N-s)}{N^2} + \frac{n^2\,s^2}{N^2}$, and consequently the value of μ_2 for the complete frequency provided by (6) and (7) is :—

$$\mu_2 = \frac{n\{N - n + s\,(n-1)\}}{N^2} \qquad (8)$$

from which s can at once be calculated.

The expressions (6) and (7) assume that all pigeon holes which have received one or more balls prior to the throw considered are equally favoured at the expense of the empty pigeon holes. But (4) can be interpreted in a more general sense. In effect the position merely is that $1 - ({}^{n-1}{}_0u - {}^n{}_0u)$ balls are to be distributed between $N - {}^{n-1}{}_0u$ pigeon holes and we can make any condition we please as to how.

Let us accordingly contrast the pigeon holes containing one ball apiece with those holding two or more.

We have—

$$\begin{cases} {}^n{}_1u = {}^{n-1}{}_0u - {}^n{}_0u + \dfrac{N-\alpha}{N}\,{}^{n-1}{}_1u \\ \dfrac{\alpha}{N}\,{}^{n-1}{}_1u + \dfrac{\beta}{N}\,(N - {}^{n-1}{}_1u - {}^{n-1}{}_0u) = 1 - ({}^{n-1}{}_0u - {}^n{}_0u) \end{cases}$$

Hence—

$${}^n{}_1u - \frac{N-\beta}{N}\,{}^{n-1}{}_1u = \left(2 - \frac{\beta}{N}\right){}^{n-1}{}_0u - 2\,{}^n{}_0u + \beta - 1 \qquad (9)$$

${}^{n-1}{}_0u$ and ${}^n{}_0u$ are given by (5A) as functions of n and s.

Substituting, we have—

$${}^{n-1}{}_1u - \frac{N-\beta}{N}\,{}^n{}_1u = \frac{\beta}{s} - 1 + \frac{2s-\beta}{s}\left(\frac{N-s}{N}\right)^n \qquad (10)$$

The solution of this difference equation being :—

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

$$ {}^{n}{}_{1}u = \left(\frac{N-\beta}{N}\right)^{n}\frac{N}{s}\left\{\frac{2s-\beta}{s-\beta}\left[1-\left(\frac{N-s}{N-\beta}\right)^{n}\right]-\frac{(\beta-s)N}{\beta(N-\beta)}\left[\left(\frac{N-\beta}{N}\right)^{n}-1\right]\right\} \quad (11) $$

As tests of the correctness of this solution, we shall determine whether (α) it gives ${}^{1}{}_{1}u = 1$ for all values of β and s and (β) for $s = \beta$ it reduces to (7). We find :—

(α) If $n = 1$. $(11) = \dfrac{N-\beta}{N}\cdot\dfrac{N}{s}\left\{\dfrac{2s-\beta}{s-\beta}\left(\dfrac{s-\beta}{N-\beta}\right)+\dfrac{\beta-s}{\beta}\cdot\dfrac{N}{N-\beta}\cdot\dfrac{\beta}{N}\right\}$

$$ = \frac{N-\beta}{s}\left\{\frac{2s-\beta}{N-\beta}+\frac{\beta-s}{N-\beta}\right\} = 1 $$

(β) The second term in the large bracket obviously vanishes so that we need only evaluate :—

$\left(\dfrac{N-\beta}{N}\right)^{n}\dfrac{N}{s}\left\{\dfrac{2s-\beta}{s-\beta}\left[1-\left(\dfrac{N-s}{N-\beta}\right)^{n}\right]\right\}$. Put $\beta = rs$ and re-arrange : We have :—

$$ \frac{1}{sN^{n-1}}\cdot\frac{s(2-r)}{s(1-r)}\left\{\left(\overline{N-s}+\overline{1-r}.s\right)^{n}-(N-s)^{n}\right\} $$

$$ = \frac{1}{sN^{n-1}}\cdot\frac{s(2-r)}{s(1-r)}\left\{n(1-r)s.(N-s)^{n-1}\right. $$

$$ \left.+\frac{n(n-1)}{2!}(1-r)^{2}s^{2}(N-s)^{n-2}+\ldots.\right\} $$

$$ = \frac{1}{sN^{n-1}}(2-r)\left\{ns\,(N-s)^{n-1}\right. $$

$$ \left.+\frac{n(n-1)}{2!}(1-r)\,s^{2}\,(N-s)^{n-2}+\ldots.\right\} \quad (12) $$

Now putting $r = 1$, (12) becomes $n\left(\dfrac{N-s}{N}\right)^{n-1}$ the value required by (7).

To obtain the second distributing constant β, we must equate (11) to the number of pigeon holes found to contain one ball apiece.

This process can plainly be carried further, a difference equation for ${}^{n}{}_{2}u$ being formed and solved precisely on the lines of (9) to (11), as many constants, s, β, γ . . . being obtainable as there are sensible frequencies.

Equation (11) can be written in a less unsuitable form for computation provided the conditions justify Poisson's approximation. If so, writing $\beta_1 = \dfrac{\beta n}{N}$, $s_1 = \dfrac{sn}{N}$ (11) reduces to :—

$$ {}^{n}{}_{1}u = \frac{n}{s_1}\left[\frac{2s_1-\beta_1}{s_1-\beta_1}\left(e^{-\beta_1}-e^{-s_1}\right)-\frac{n(\beta_1-s_1)}{\beta_1(n-\beta_1)}\left(e^{-2\beta_1}-e^{-\beta_1}\right)\right] \quad (13) $$

But even with this simplification the calculation of β is troublesome and, *a fortiori*, that of further constants.

Nor does it seem practicable to reduce to a relation between the moments of a form similar to that of (8).

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Since, as remarked above, we do not believe that the theoretical basis of this schema either in the modified or original form is appropriate to the class of problems with which we are dealing, the resultant expressions are at best mere smoothing formulæ, and when their application involves a large amount of arithmetic do not have any value even from that point of view. An exception must, however, be made in favour of the single bias case which in virtue of equation (8) can be readily employed and, as will appear in our arithmetical examples, frequently graduates the statistics with considerable success. It will therefore be of some interest to determine the probable error of s calculated by (8).

Assuming n and N to be large we can write (8)

$$\frac{s}{N} = \frac{N\mu_2}{n^2} - \frac{N-n}{Nn} \tag{14}$$

Hence, remembering that $\frac{n}{N} = \mu'_1$

$$s = \frac{\mu_2}{\mu'^2_1} - \frac{1-\mu'_1}{\mu'_1} \tag{15}$$

Taking differentials.

$$\delta s = \frac{\mu'_1\delta\mu_2 - 2\mu_2\delta\mu'_1}{\mu'^3_1} + \frac{\delta\mu'_1}{\mu'^2_1}$$

Squaring, summing and dividing by the number of samples

$$\left.\begin{aligned}\sigma_s^2 = \frac{1}{\mu'^6_1}\Big(\mu'^2_1\,\sigma^2_{\mu_2} + 4\mu^2_2\,\sigma^2_{\mu'_1} - 4\mu'_1\,\mu_2\,\sigma_{\mu'_1}\,\sigma_{\mu_2}r_{\mu'_1\mu_2}\Big)\\ + \frac{2}{\mu'^5_1}\Big(\mu'_1\,\sigma_{\mu'_1}\sigma_{\mu_2}r_{\mu'_1\mu_2} - 2\,\mu_2\,\sigma^2_{\mu_1}\Big) + \frac{\sigma^2_{\mu_1}}{\mu'^4_1}\end{aligned}\right\} \tag{16}$$

But, $\sigma^2_{\mu'_1} = \frac{\mu_2}{N}$, $\sigma^2_{\mu_2} = \frac{\mu_4 - \mu^2_2}{N}$, $\sigma_{\mu_2}.\sigma'_{\mu_1}.r_{\mu_2\mu'_1} = \frac{\mu_3}{N}$

Substituting in (16)

$$\sigma_s^2 = \frac{1}{N}\left[\frac{4\mu^3_2}{\mu'^6_1} - \frac{4\mu_2}{\mu'^5_1}(\mu_3+\mu_2) + \frac{1}{\mu'^4_1}(\mu_4 - \mu^2_2 + \mu_2 + 2\mu_3)\right] \tag{17}$$

Now if conditions (6) and (7) hold, all the moments can be expressed in terms of N, n and s.

The moments of the binomial $\left(\frac{s}{N} + \frac{N-s}{N}\right)^n$ about the zero successes are (the Poisson approximation being assumed to be valid) $M_2 = m + m^2$, $M_3 = m + 3m^2 + m^3$, $M_4 = m^4 + 6m^3 + 7m^2 + m$, where $m = \frac{sn}{N}$ and these divided by s and referred to the mean of the whole distribution, $\frac{n}{N}$, will be the required moments.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

We thus obtain :—

$$\left.\begin{aligned}
\mu'_1 &= \frac{n}{N} \\
\mu_2 &= \frac{n}{N} + \frac{n^2}{N^2}(s-1) \\
\mu_3 &= \frac{n}{N} + \frac{n^2}{N^2}3(s-1) + \frac{n^3}{N^3}(s-1)(s-2) \\
\mu_4 &= \frac{n}{N} + \frac{n^2}{N^2}(7s-4) + \frac{n^3}{N^3}6(s-1)^2 \\
&\qquad + \frac{n^4}{N^4}(s^3 - 4s^2 + 6s - 3)
\end{aligned}\right\} \quad (18)$$

and

$$\mu_4 - \mu^2_2 = \frac{n}{N} + \frac{n^2}{N^2}(7s-5) + \frac{n^3}{N^3}(3s-4)(s-1) + \frac{n^4}{N^4}(s-2)^2(s-1)$$

As a check, note that when $s = 1$, the complete distribution becoming the ordinary binomial,

$N\left(\frac{1}{N} + \frac{N-1}{N}\right)^n$, (18) gives $\mu'_1 = \mu_2 = \mu_3 = \frac{n}{N}$ and $\mu_4 = \frac{n}{N} + \frac{3n^2}{N^2}$, which are the correct values for a Poisson binomial.

Substituting from (18) in (17), we obtain after lengthy but straightforward reduction :—

$$\sigma_s^2 = \frac{s^2(s-1)}{N} + \frac{(s-1)(4-3s)}{n} + \frac{2sN}{n^2} \qquad (19)$$

If s is so small in comparison with either n or N that the first two terms on the right may be neglected, the standard deviation of s is effectively equal to :—

$$\frac{1}{n}\sqrt{2sN}$$

It is, however, distinctly to be noted that (19) is only valid on the assumption stated, while (17) is exact.

Cases frequently occur when the biassed schema based on (6) to (8) reproduces the observations well, but the momental relations of (18) are not even approximately fulfilled.

This statement will now be illustrated.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

B. Shop. L. Factory. [High-explosive shell manufacture.]

Accidents.	Frequency.	Calculated from the unmodified exponential.	Calculated from the biassed schema.
0	397	379	403
1	133	163	125
2	47	36	44
3	5	5	10
4	1	1	2
5	0	·1	·3
6	0		
7	1		
	584		

Here N = 584, $n = 253$.

The exponential limit of the binomial fails completely to reproduce these data.

The mean is .4332, $\mu_2 = .5504$.

Hence $s = 1.6303$ and the resultant distribution agrees satisfactorily with the original data.

The remaining moments of the observations up to the 4th are $\mu_3 = 1.003$ and $\mu_4 = 4.194$, while the theoretical values from equation (18) are .769 and 2.020. Hence σ_s^2 from (17) is .030 if the theoretical and .103 if the observed moments are substituted in the equation.

The inference is that although the method provides a reasonable fit (P = .51 when the usual goodness of fit test is used*) we should not regard the result as any substantiation of the theoretical basis. On this account it seems unprofitable to perform the heavy arithmetical work involved in the fitting of further constants.

In other words, equation (8) is merely a useful basis for an interpolation formula.

Section III.—*The generalised Poisson series.*

As we have seen, the method just examined involves the assumption that the happening of the event not only improves the prospects of the successful candidates but militates against the chances of those who had hitherto failed; this assumption cannot be entertained and we proceed to develop a frequency system not involving it.

Let us suppose that the chance of meeting with an accident (acquiring some disease, &c., or, in general, being the subject of any happening whatever) is p_0 for those who had at the time of observation never had an accident; that for those previously credited with

* When the whole distribution is used and the freqencies from 4 accidents upwards are combined. But, as Pearson has shown (*loc. cit* p. 255 *supra*), it is proper to omit the zero group in comparisons of the present type. In that case, P becomes ·37. The P's discussed on p. 276 *infra* have been compiled by the second method.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Time interval.	Accidents. 0	1	2	3	4	5
Δ_0	1					
Δ_1	q_0	p_0				
Δ_2	q_0^2	$p_0\,(q_0 + q_1)$	p_0p_1			
Δ_3	q_0^3	$p_0\,(q_0^2 + q_0q_1 + q_1^2)$	$p_0p_1\,(q_0+q_1+q_2)$	$p_0p_1p_2$		
Δ_4	q_0^4	$p_0\,(q_0^3 + q_0^2q_1 + q_0q_1^2 + q_1^3)$	$p_0p_1\,(q_0^2 + q_1^2 + q_2^2 + q_0q_1 + q_0q_2 + q_1q_2)$	$p_0p_1p_2\,(q_0+q_1+q_2+q_3)$	$p_0p_1p_2p_3$	
Δ_5	q_0^5	$p_0\,(q_0^4 + q_0^3q_1 + q_0^2q_1^2 + q_0q_1^3 + q_1^4)$	$p_0p_1\,(q_0^3 + q_1^3 + q_2^3 + q_0^2q_1 + q_0q_1^2 + q_0^2q_2 + q_0q_2^2 + q_1^2q_2 + q_1q_2^2 + q_0q_1q_2)$	$p_0p_1p_2\,(q_0^2 + q_1^2 + q_2^2 + q_3^2 + q_0q_1 + q_0q_2 + q_1q_2 + q_0q_3 + q_1q_3 + q_2q_3)$	$p_0p_1p_2p_3\,(q_0 + q_1 + q_2 + q_3 + q_4)$	$p_0p_1p_2p_3p_4$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

an accident it is p_1; that for those who had had 2, 3 . . . r accidents the chances are p_2, p_3 . . . p_r. Then writing $p_r + q_r = 1$, the development of the series is illustrated in the table on page 265.

It will be seen that any row may be written :—

$$ {}_rA_0 + p_0\,{}_rA_1 + p_0p_1\,{}_rA_2 + p_0p_1p_2\,{}_rA_3, \text{ etc.} \tag{20}$$

where the A's are functions of q's.

We have :—

$$ {}_nA_r = {}_{n-1}A_{r-1} + q_r\,{}_{n-1}A_r \tag{21}$$

$$ {}_{n-1}A_r = {}_{n-2}A_{r-1} + q_r\,{}_{n-2}A_r \tag{22}$$

$$\vdots$$

Hence :—

$$ {}_nA_r = {}_{n-1}A_{r-1} + q_r\,{}_{n-2}A_{r-1} + q^2_r\,{}_{n-3}A_{r-1} + \ldots . q_r^{n-r}\,{}_{r-1}A_{r-1} \tag{23}$$

Now suppose the whole period of exposure to consist of very small intervals of time and the total time of exposure T to be very great in comparison with the length of such intervals, then if T—1 is sensibly equal to T, we may write (23) :—

$$ \mathrm{T}A_r = \int_0^{\mathrm{T}} {}_tA_{r-1}\, q_r^{\,\mathrm{T}-t}\, dt. \tag{24}$$

Now suppose all p's to be small but their several products with T finite and write :—

$$\begin{aligned} p_0\mathrm{T} &= \lambda_0 \\ p_1\mathrm{T} &= \lambda_1 \\ p_2\mathrm{T} &= \lambda_2 \\ &\vdots \\ p_r\mathrm{T} &= \lambda_r. \end{aligned}$$

Then, by Poisson's theorem,

$$ {}_tA_0 = e^{-p_0t}$$

and

$$ f_0 = {}_{\mathrm{T}}A_0 = e^{-\lambda_0}, \tag{25}$$

where f_0 is the relative frequency of no accidents.

Hence :—

$$ {}_{\mathrm{T}}A_1 = \int_0^{\mathrm{T}} e^{-p_0t}\, q^{\mathrm{T}-t}\, dt \tag{26}$$

$$= \int_0^{\mathrm{T}} e^{-p_0t}\, e^{-p_1(\mathrm{T}-t)}\, dt$$

$$= \frac{e^{-\lambda_1}}{p_0 - p_1}\left\{1 - e^{-(\lambda_0-\lambda_1)}\right\}$$

Then using f's to denote the relative frequencies

$$ f_1 = p_0A_1 = \frac{\lambda_0}{\lambda_0 - \lambda_1}\,(e^{-\lambda_1} - e^{-\lambda_0}). \tag{27}$$

So λ_0 is given from the zero frequency by (25) and λ_1 from (27) which may conveniently be written :—

$$ \frac{f_1}{\lambda_0} = \frac{e^{-\lambda_1} - e^{-\lambda_0}}{\lambda_0 - \lambda_1} \tag{28}$$

Passing to the next term, we must integrate :—

$$ \int_0^{\mathrm{T}} (e^{-p_1t} - e^{-p_0t})\, e^{-p_2(\mathrm{T}-t)}\, dt \tag{29}$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

$$= e^{-\lambda_2}\left[\frac{1-e^{-(\lambda_1-\lambda_2)}}{p_1-p_2} - \frac{1-e^{-(\lambda_0-\lambda_2)}}{p_0-p_2}\right] \tag{30}$$

Hence :—

$$f_2 = p_0 p_1 A_2 = \frac{\lambda_0\lambda_1}{\lambda_0-\lambda_1}\left\{\frac{e^{-\lambda_2}-e^{-\lambda_1}}{\lambda_1-\lambda_2} - \frac{e^{-\lambda_2}-e^{-\lambda_0}}{\lambda_0-\lambda_2}\right\} \tag{31}$$

By a precisely similar method successive f's can be calculated.

After simplification we have for the first six f's the following equations :—

$$f_0 = e^{-\lambda_0} \tag{32A}$$

$$f_1 = \frac{\lambda_0}{\lambda_0-\lambda_1}e^{-\lambda_1} - \frac{\lambda_0}{\lambda_0-\lambda_1}e^{-\lambda_0} \tag{32B}$$

$$f_2 = \frac{\lambda_0\lambda_1}{(\lambda_0-\lambda_2)(\lambda_1-\lambda_2)}e^{-\lambda_2} - \frac{\lambda_0\lambda_1}{(\lambda_0-\lambda_1)(\lambda_1-\lambda_2)}e^{-\lambda_1} + \frac{\lambda_0\lambda_1}{(\lambda_0-\lambda_1)(\lambda_0-\lambda_2)}e^{-\lambda_0} \tag{32C}$$

$$f_3 = \frac{\lambda_0\lambda_1\lambda_2}{(\lambda_0-\lambda_3)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)}e^{-\lambda_3} - \frac{\lambda_0\lambda_1\lambda_2}{(\lambda_0-\lambda_2)(\lambda_1-\lambda_2)(\lambda_2-\lambda_3)}e^{-\lambda_2} + \frac{\lambda_0\lambda_1\lambda_2}{(\lambda_0-\lambda_1)(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)}e^{-\lambda_1} - \frac{\lambda_0\lambda_1\lambda_2}{(\lambda_0-\lambda_1)(\lambda_0-\lambda_2)(\lambda_0-\lambda_3)}e^{-\lambda_0} \tag{32D}$$

$$f_4 = \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{(\lambda_0-\lambda_4)(\lambda_1-\lambda_4)(\lambda_2-\lambda_4)(\lambda_3-\lambda_4)}e^{-\lambda_4} - \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{(\lambda_0-\lambda_3)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)(\lambda_3-\lambda_4)}e^{-\lambda_3} + \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{(\lambda_0-\lambda_2)(\lambda_1-\lambda_2)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)}e^{-\lambda_2} - \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{(\lambda_0-\lambda_1)(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)}e^{-\lambda_1} + \frac{\lambda_0\lambda_1\lambda_2\lambda_3}{(\lambda_0-\lambda_1)(\lambda_0-\lambda_2)(\lambda_0-\lambda_3)(\lambda_0-\lambda_4)}e^{-\lambda_0} \tag{32E}$$

$$f_5 = \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_5)(\lambda_1-\lambda_5)(\lambda_2-\lambda_5)(\lambda_3-\lambda_5)(\lambda_4-\lambda_5)}e^{-\lambda_5} - \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_4)(\lambda_1-\lambda_4)(\lambda_2-\lambda_4)(\lambda_3-\lambda_4)(\lambda_4-\lambda_5)}e^{-\lambda_4} + \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_3)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)(\lambda_3-\lambda_4)(\lambda_3-\lambda_5)}e^{-\lambda_3} - \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_2)(\lambda_1-\lambda_2)(\lambda_2-\lambda_3)(\lambda_2-\lambda_4)(\lambda_2-\lambda_5)}e^{-\lambda_2} + \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_1)(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_1-\lambda_4)(\lambda_1-\lambda_5)}e^{-\lambda_1} - \frac{\lambda_0\lambda_1\lambda_2\lambda_3\lambda_4}{(\lambda_0-\lambda_1)(\lambda_0-\lambda_2)(\lambda_0-\lambda_3)(\lambda_0-\lambda_4)(\lambda_0-\lambda_5)}e^{-\lambda_0} \tag{32F}$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

The validity of equations (32) can be tested in a variety of ways. One method is to evaluate the indeterminate forms which result when λ_0 is made equal to λ_1, λ_2, ... In that case we ought to find that the series reduces to :—

$$e^{-\lambda_0} + \lambda_0 e^{-\lambda_0} + \frac{\lambda^2_0}{2!} e^{-\lambda_0} \ldots .$$

the ordinary Poisson limit.

As examples, we may reduce f_1 and f_2

$$f_1 = \frac{\lambda_0 (e^{-\lambda_1} - e^{-\lambda_0})}{\lambda_0 - \lambda_1}$$

Differentiating the numerator and denominator with respect to λ_1 we have :—

$$\frac{-\lambda_0 e^{-\lambda_1}}{-1} = \lambda_0 e^{-\lambda_0} \text{ when } \lambda_1 = \lambda_0.$$

$$f_2 = \lambda_0\lambda_1 \left(\frac{e^{-\lambda_2}}{(\lambda_0 - \lambda_2)(\lambda_1 - \lambda_2)} - \frac{e^{-\lambda_1}}{(\lambda_0 - \lambda_1)(\lambda_1 - \lambda_2)} + \frac{e^{-\lambda_0}}{(\lambda_0 - \lambda_1)(\lambda_0 - \lambda_2)} \right)$$

$$= \lambda_0\lambda_1 \left(\frac{e^{-\lambda_2}}{(\lambda_0 - \lambda_2)(\lambda_1 - \lambda_2)} + \frac{(\lambda_1 - \lambda_2) e^{-\lambda_0} - (\lambda_0 - \lambda_2) e^{-\lambda_1}}{(\lambda_0 - \lambda_1)(\lambda_1 - \lambda_2)(\lambda_0 - \lambda_2)} \right)$$

proceeding to the limit for $\lambda_1 = \lambda_0$ we have

$$\lambda^2_0 \left(\frac{e^{-\lambda_2} - (1 + \lambda_0 - \lambda_2) e^{-\lambda_0}}{(\lambda_0 - \lambda_2)^2} \right)$$

Differentiating numerator and denominator with respect to λ_2 twice we reach

$$\lambda^2_0 \frac{e^{-\lambda_2}}{2} = \frac{\lambda_0^2 e^{-\lambda_0}}{2} \text{ when } \lambda_2 = \lambda_0.$$

It may be remarked that equations (32) can be deduced by a method analogous with that employed in generalising the pigeon-hole schema.

We have (in the notation there employed)

$${}^n u_m = q_m {}^{n-1} u_m + p_{m-1} {}^{n-1} u_{m-1}$$

writing

$$p_m = \alpha_m p_0 \text{ and } q_m = \beta_m q_0$$

$${}^n u_0 = q_0 {}^{n-1} u_0 = q_0 {}^n$$

$${}^n u_1 = \beta_1 q_0 {}^{n-1} u_0 + p_0 {}^{n-1} u_0$$

Solving the difference equation we have :—

$${}^n u_1 = c\beta_1 {}^n q_0 {}^n + \frac{p_0 q_0 {}^{n-1}}{1 - \beta_1}$$

and since

$${}^0 u_1 = 0 ; \; c = -\frac{p_0}{(1 - \beta_1) q_0}$$

or

$${}^n u_1 = \frac{p_0 q_0 {}^{n-1}}{1 - \beta_1} \left(1 - \beta_1 {}^n \right) \qquad (32_G)$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Then
$$^{n+1}u_2 - \beta_2 q_0{}^n u_2 = p_1{}^n u_1 = \frac{\alpha_1 p_0{}^2 q_0{}^{n-1}}{1-\beta_1}(1-\beta_1{}^n)$$

The solution of which is

$$^nu_2 = \frac{\alpha_1 p_0{}^2 q_0{}^{n-2}}{1-\beta_1}\left\{\frac{1}{\beta_1-\beta_2}(\beta_2{}^n - \beta_1{}^n) - \frac{1}{1-\beta_2}(\beta_2{}^n - 1)\right\} \quad (32\text{H})$$

remembering that $\alpha_1 = \frac{p_1}{p_0}$, $\beta_1 = \frac{q_1}{q_0}$, $\beta_2 = \frac{q_2}{q_0}$

and writing as before $\lambda_0 = p_0 t$, $\lambda_1 = p_1 t$, and $\lambda_2 = p_2 t$ 32G and 32H reduce to 32B and 32C.

This method of solution is, however, much more laborious than that adopted above.

By means of equations (32) values of the λ's can be obtained in any particular case. The arithmetical labour is, however, very heavy, and before discussing the general problem further we shall examine in detail the particular case of $\lambda_0 \neq \lambda_1$, $\lambda_1 = \lambda_2 = \ldots = \lambda_r$.

The required expressions might be deduced from the general equations (32) but the evaluation of the indeterminate forms becomes increasing troublesome as we proceed, and it is better to take the case independently.

If we write $r = \frac{q_1}{q_0} = \text{approx. } 1 - (p - p_0)$

$$r^t = \left(1 - \frac{\lambda_1 - \lambda_0}{t}\right)^t = e^{-(\lambda_1 - \lambda_0)} \text{ and } q_0{}^t = \left(1 - \frac{\lambda_0}{t}\right)^t = e^{-\lambda_0}$$

when $t = \mathrm{T}$ and is very large.

Hence $p_0 \mathrm{A}_1 = p_0 q_0{}^{t-1}(1 + r + r^2 + \ldots . r^{t-1}) = q_0{}^{t-1}\lambda \frac{1 - e^{-\delta}}{\delta}$

where $\delta = \lambda_1 - \lambda_0$

Similarly,

$$p_0 p \mathrm{A}_2 = p p_0 q_0{}^{t-2}(1 + 2r + 3r^2 + \ldots . t-1 . r^{t-2})$$

The sum of the series in brackets is

$$\frac{1 - r^{t-1}}{(1-r)^2} - \frac{(t-1)\, r^{t-1}}{1-r}$$

Hence, assuming t so large that $t - 1$ may be put equal to t.

$$p_0 p \mathrm{A}_2 = q_0{}^{t-2}\lambda_0\lambda_1\left(\frac{1 - e^{-\delta}}{\delta^2} - \frac{e^{-\delta}}{\delta}\right)$$

Proceeding in this way and expanding $e^{-\delta}$ in powers of δ we reach :—

$$e^{-\lambda_1}\Big[1 + \lambda_0\Big(1 - \frac{\delta}{2!} + \frac{\delta^2}{3!} - \frac{\delta^3}{4!} + \frac{\delta^4}{5!} + \ldots\Big) + \lambda_0\lambda_1\Big(\frac{1}{2!} - \frac{2\delta}{3!} + \frac{3\delta^2}{4!} - \frac{4\delta^3}{5!} + \frac{5\delta^4}{6!} + \ldots\Big) + \lambda_0\lambda_1{}^2\Big(\frac{1}{3!} - \frac{3\delta}{4!} + \frac{6\delta^2}{5!} - \frac{10\delta^2}{6!} + \ldots\Big)$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

$$+\lambda_0\lambda_1{}^3\left(\frac{1}{4!}-\frac{4\delta}{5!}+\frac{10\delta^2}{6!}-\frac{20\delta^3}{7!}+\ .\ .\ .\right) \qquad (33)$$
$$+\lambda_0\lambda_1{}^4\left(\frac{1}{5!}-\frac{5\delta}{6!}+\frac{15\delta^2}{7!}\cdots\right)+\ .\ .\ .\Big]$$

The series in (33) may be treated as follows :—

Write $$f(\delta) = 1 - \frac{\delta}{2!} + \frac{\delta^2}{3!} - \frac{\delta^3}{4!} + \ .\ .\ -$$

Then
$$\left.\begin{aligned} f'(\delta) &= -\frac{1}{2}+\frac{2\delta}{3!}-\frac{3\delta^2}{4!}+\ .\ . \\ f''(\delta) &= \frac{2}{3!}-\frac{6\delta^2}{4!}+\ .\ . \\ &\ \ \vdots \end{aligned}\right\} \qquad (34)$$

So the coefficient of λ_0 is

$$\left.\begin{aligned} f(\delta) - \lambda_1 f'(\delta) + \frac{\lambda_1{}^2}{2!}f''(\delta) - \frac{\lambda_1{}^3}{3!}f'''(\delta)\ .\ .\ . \\ = f(\delta-\lambda_1) = f(-\lambda_0) \end{aligned}\right\} \qquad (35)$$

But $$f(\delta) = \frac{1-e^{-\delta}}{\delta} \quad \therefore\ f(-\lambda_0) = \frac{e^{\lambda_0}-1}{\lambda_0}$$

Hence (33) is $e^{-\lambda_0}(1 + e^{\lambda_0} - 1) = 1$, or (33) gives the complete frequency.

Passing to the moments referred to zero, we must evaluate

$$f(\delta) - 2\lambda_1 f'(\delta) + \frac{3\lambda_1{}^2}{1\cdot 2}f''(\delta) - \frac{4\lambda_1{}^3}{3!}f'''(\delta)\ .\ .\ .$$

Expand $f(\delta-\lambda_1)$ and multiply by λ_1 ; we have :—

$$\lambda_1 f(\delta-\lambda_1) = \lambda_1 f(\delta) - \lambda_1{}^2 f'(\delta) + \frac{\lambda_1{}^3}{2!}f''(\delta)\ .\ .\ .$$

Differentiating with respect to λ, treating δ as constant, we find :—

$$\frac{d[\lambda_1 f(\delta-\lambda_1)]}{d\lambda_1} = f(\delta) - 2\lambda_1 f'(\delta) + \frac{3\lambda_1{}^2}{2!}f''(\delta)\ .\ .\ .$$

Hence the mean is given by :—

$$\mathrm{M} = \lambda_0 e^{-\lambda_0}\frac{d[\lambda_1 f(\delta-\lambda_1)]}{d\lambda_1} \qquad (36)$$

Now substituting $\dfrac{\lambda_1(e^{\lambda_1-\delta}-1)}{\lambda_1-\delta}$ for $\lambda_1 f(\delta-\lambda_1)$

(36) reduces to

$$\mathrm{M} = \frac{\lambda_1(e^{\lambda_0}-1)}{e^{\lambda_0}}\left\{\frac{1}{\lambda_1}-\frac{1}{\lambda_0}+\frac{e^{\lambda_0}}{e^{\lambda_0}-1}\right\} \qquad (37)$$

Multiplying the expansion of $f(\delta-\lambda_1)$ by $\lambda_1{}^2$, and differentiating twice with respect to λ_1, we find :—

$$\frac{d^2[\lambda_1{}^2 f(\delta_1-\lambda_1)]}{d\lambda_1{}^2} = 2f(\delta) - 2\cdot 3\lambda_1 f(\delta) + 3\cdot 4\frac{\lambda_1{}^2}{1\cdot 2}f''(\delta) - 4\cdot 5\frac{\lambda_1{}^3}{1.2.3}f'''(\delta)$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

which is $\frac{1}{\lambda_0 e^{-\lambda_0}}$ times the sum of the first and second moments about zero.

Hence
$$_0\mu_2 + \mathrm{M} = \lambda_0 e^{-\lambda_0} \frac{d^2\left[\lambda_1{}^2 f(\delta - \lambda_1)\right]}{d\lambda_1{}^2}$$

Substituting $\frac{\lambda_1{}^2 (e^{\lambda_1 - \delta} - 1)}{\lambda_1 - \delta}$ for $\lambda_1{}^2 f(\delta - \lambda_1)$

and reducing, we find :—
$$_0\mu_2 - \mathrm{M} = \lambda_1{}^2\left(\frac{2}{\lambda_1} - \frac{1}{\lambda_0} + 1\right) - \frac{\lambda_1{}^2(e^{\lambda_0} - 1)}{\lambda_0 e^{\lambda_0}} \left\{\frac{2}{\lambda_1} - \frac{2}{\lambda_0} + \frac{e^{\lambda_0}}{e^{\lambda_0} - 1}\right\} \quad (38)$$

Hence
$$\mathrm{M} = \left(1 - \frac{\lambda_1}{\lambda_2}\right)\left(1 - e^{-\lambda_1}\right) + \lambda_1 \quad (39)$$

$$_0\mu_2 - \mathrm{M} = \lambda_1(\lambda_1 + 2) - 2\frac{\lambda_1}{\lambda_0}\mathrm{M}\,. \quad (40)$$

from which λ_0 and λ_1 may be obtained by successive approximation.

(39) and (40) are fairly suitable for calculation, but we have not succeeded in obtaining a solution of the general case in terms of moments which would be of the least use in practice.

As arithmetical examples we may take the following.

Accidents in a 60-lb. shrapnel shop.

Accidents.	Observed persons.	Calculated from (39 and 40) $\lambda_0 = \cdot 6025$ $\lambda_1 = \cdot 5018$	Calculated from (8) $s = \cdot 8943$.
0	398	410·6	412
1	294	260·3	258
2	43	66·4	66
3	10	11·2	11·5
4	3	1·4	1·5
5	2	·2	1
	750		

Actually the fit is almost identical with that afforded by the one biassed pigeon-hole schema. But neither result is good.

It has been found that when λ_0 is not greatly different from λ_1, equation (8) graduates the data effectively; for $\lambda_1 = 2\lambda_0$ the fit is excellent, but when $\lambda_1 > 5\ \lambda_0$ the graduation fails. Thus we find :—

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Successes.	$\lambda_0 = \cdot 5$, $\lambda_1 = 1 \cdot 0$.		$\lambda_0 = 1$, $\lambda_1 = 2$.		$\lambda_0 = \cdot 5$, $\lambda_1 = 2 \cdot 5$.		$\lambda_0 = \cdot 5$, $\lambda_1 = 5 \cdot 0$.	
	True.	Eq. (8) $s = 1 \cdot 558$.	True.	Eq. (8) $s = 1 \cdot 351$	True.	Eq. (8) $s = 2 \cdot 320$.	True.	Eq. (8) $s = 2 \cdot 737$.
0	607	608	368	377	607	621	607	641
1	239	235	233	216	131	110	67	27
2	109	111	194	199	113	116	70	54
3	35	35	118	123	77	82	69	71
4	8	8	56	57	42	43	61	71
5	2	2	22	21	19	18	48	57
6	....	....	7	6	8	7	34	38
7	....	....	2	2	3	2	21	22
8	....	....	....	....	1	1	12	11
9	....	....	....	....	....	....	6	5
10	....	....	....	....	....	....	3	2
11	....	....	....	....	....	....	2	1
12	....	....	....	....	....	....	....	....
13	....	....	....	....	....	....	....	....

The empirical conclusion to be drawn is that data which by the method of equation (8) yield values of s below $2 \cdot 0$ and are effectively graduated may be regarded as examples of the present class. Owing to the rapidity with which this test can be made and the comparative laboriousness of (39) and (40) the point is of some practical interest.

Equations (32) were used to fit the following statistics :—

Cancer Houses (*Biometrika* viii. 431).

Cases.	Houses.	Calculated from (8).
0	2523	2530
1	315	296
2	20	36
3	6	3
4	1	
	2865	

These lead to $\lambda_0 = \cdot 1271$, $\lambda_1 = \cdot 1632$, $\lambda_2 = \cdot 95$, $\lambda_3 = \cdot 61$.

Although equations (32) provide a complete formal solution of the general problem, and by a scrutiny of the probable errors[1] of the λ's it would be possible to determine whether their differences are significant and warrant the conclusion that we are really dealing with a case of varying chance in the terms of the hypothesis, the form reached is not altogether satisfactory. For the values of the λ's are subject to high probable errors, especially in the tail of the

[1] A method was worked out for obtaining these, but we have not thought it worth while to give space to it here. As an illustration we find the probable errors of the λ's just given to be ·0046, ·0215, ·2635, ·4240.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

distribution, and in any practical case it would become very difficult to say what was the probable form of the functional relation between λ and the serial number of the accident to which it referred. For practical purposes it would be much more satisfactory to assume some fairly flexible form of functional relation between λ and the serial number of the accident—a form involving not more than a moderate number of constants such as three—and to find equations for these constants in terms of the moments of the distribution. We have made some attempts to find such a solution, but so far they have not succeeded.

We now turn to the other modification of the Poisson schema which is effected by supposing that *ab initio* the liabilities to accident, to disease, &c., are not the same for all units of the population of houses or workers.

Section IV.—*The infinitely compound Poisson distribution.*

We now suppose that the population at risk consists of persons (or other variates), the liabilities or susceptibilities of whom to accident vary, the frequencies being assigned by the ordinates of $f(\lambda)$ where λ is a variable parameter.

One naturally commences with the assumption that $f(\lambda)$ is a normal function or

$$y = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(\lambda - m)^2}{2\sigma^2}} \tag{41}$$

m being the mean of λ and σ the standard deviation. Then, supposing that within any group to which the parameter λ_s applies the frequency of multiple happenings is expressed by the Poisson exponential

$$e^{-\lambda_s}\left(1 + \lambda_s + \frac{\lambda_s^2}{2!} \cdot \cdot \cdot \cdot\right),$$

we shall have for the f_0 frequency of the complete distribution :—

$$f_0 = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(\frac{\lambda - m}{2\sigma^2}\right)^2} . e^{-\lambda} . d\lambda \tag{42}$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(m - \frac{\sigma^2}{2}\right)} e^{-\frac{(\lambda - \overline{m - \sigma^2})^2}{2\sigma^2}} d\lambda$$

$$= e^{-\left(m - \frac{\sigma^2}{2}\right)} \tag{43}$$

Similarly,

$$f_1 = \frac{1}{\sqrt{2\pi\sigma}} e^{-\left(m - \frac{\sigma^2}{2}\right)} . \int_{-\infty}^{+\infty} \lambda e^{-\frac{1}{2\sigma^2}\left(\lambda - \overline{m - \sigma^2}\right)^2} d\lambda$$

$$= e^{-\left(m - \frac{\sigma^2}{2}\right)} (m - \sigma^2) \tag{43A}$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

$$f_2 = \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\left(m-\frac{\sigma^2}{2}\right)} \int_{-\infty}^{+\infty} \frac{\lambda^2}{2!}\, e^{-\frac{1}{2\sigma^2}\left(\lambda-\overline{m-\sigma^2}\right)^2} d\lambda$$

$$= e^{-\left(m-\frac{\sigma^2}{2}\right)} \left\{ \frac{(m-\sigma^2)^2 + \sigma^2)}{2!} \right\} \tag{44}$$

$$f_3 = \frac{1}{\sqrt{2\pi\sigma}}\, e^{-\left(m-\frac{\sigma^2}{2}\right)} \int_{-\infty}^{+\infty} \frac{\lambda^3}{3!}\, e^{-\frac{1}{2\sigma^2}\left(\lambda-\overline{m-\sigma^2}\right)^2} d\lambda$$

$$= e^{-\left(m-\frac{\sigma^2}{2}\right)} \cdot \frac{(m-\sigma^2)^3 + 3\,(m-\sigma^2)\,\sigma^2}{3!} \tag{45}$$

$$f_4 = e^{-\left(m-\frac{\sigma^2}{2}\right)} \cdot \frac{(m-\sigma^2)^4 + 6\,(m-\sigma^2)^2\,\sigma^2 + \sigma^4}{4!} \tag{46}$$

&c.

But the assumption of a normal distribution cannot be justified as values of λ may be supposed to range from zero far in the positive direction so that $f(\lambda)$ should be skew. A choice of skew curves is arbitrary. The form we adopt,* since its equation is well suited to the reductions just outlined, is the binomial curve. This equation may be written, assuming the range to start from zero,

$$y = y_0 e^{-c\lambda} \lambda^{r-1} \text{ and its integral } \frac{c^r}{\Gamma(r)} \tag{46A}$$

Multiplying by $e^{-\lambda}$ is equivalent to substituting $c+1$ for c and multiplying by λ is equivalent to substituting $r+1$ for r

$$f_0 = \frac{c^r}{\Gamma(r)} \int_0^{\infty} e^{-(c+1)\lambda} \cdot \lambda^{r-1} \cdot d\lambda$$

$$= \frac{c^r}{\Gamma(r)} \cdot \frac{\Gamma(r)}{(c+1)^r} = \left(\frac{c}{c+1}\right) \tag{47}$$

Similarly :—

$$f_1 = \frac{c^r}{\Gamma(r)} \frac{\Gamma(r+1)}{(c+1)^{r+2}} = \left(\frac{c}{c+1}\right)^r \frac{r}{c+1} \tag{48}$$

$$f_2 = \left(\frac{c}{c+1}\right)^r \frac{r\,(r+1)}{2!\,(c+1)^2} \tag{49A}$$

$$f_3 = \left(\frac{c}{c+1}\right)^r \frac{r(r+1)\,(r+2)}{3!\,(c+1)^3} \tag{49B}$$

So that the successive f's are given by the terms of :—

$$\left(\frac{c}{c+1}\right)^r \left(1 + \frac{r}{c+1} + \frac{r\,(r+1)}{2!\,(c+1)^2} + \frac{r(r+1)\,(r+2)}{3!\,(c+1)^3} \ldots\ldots\right) \tag{50}$$

* Yule, *Journal Royal Statistical Society*, 1910, lxxiii, p. 26. The curve was developed from a negative binomial and might therefore seem inappropriate. It is shown in the appendix that the same curve might have been developed from a positive binomial; no definite criterion between positive and negative binomial is always possible, but see the discussion given.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

The second bracket is the expansion of $\left(1 - \frac{1}{c+1}\right)^{-r}$ so that the whole sum of (50) is unity as it must be since $f_0 + f_1 + f_2 + \ldots = 1$.

The first moment about zero of (50) is

$$\left(\frac{c}{c+1}\right)^r \frac{r}{c+1}\left\{1 + \frac{r+1}{c+1} + \frac{(r+1)(r+2)}{2!\,(c+1)^2} + \ldots\right\}$$

The second bracket is $\left(\frac{c+1}{c}\right)^{r+1}$ or

$$M = \frac{r}{c} \tag{51}$$

The second moment about zero is

$$\left(\frac{c}{c+1}\right)^r \frac{r}{c+1}\left\{1 + 2\frac{r+1}{c+1} + 3\frac{(r+1)(r+2)}{2!\,(c+1)^2} + \ldots\right\}$$

The second bracket reduces to $\left(\frac{c+1}{c}\right)^{r+1}\left(1 + \frac{r+1}{c}\right)$

Hence $$\mu_2 = \frac{r}{c}\left(\frac{c+r+1}{c}\right) - \frac{r^2}{c^2} = \frac{r(c+1)}{c^2} \tag{52}$$

Thus the distribution is fitted with great ease from the mean and second moment of the observations.

We have found this result of great service and append a few examples of its use in graduation. Several illustrations will be found in the undermentioned report by Miss Woods and one of us.*

(1) 648 *Women working on 6-inch H.E. shells for five weeks.*

No. of accidents.	Observations.	Simple Poisson series.	Method of equation (8).	Method of equations (51) and (52).
0	447	406	452	442
1	132	189	117	140
2	42	45	56	45
3	21	7	18	14
4	3	1	4	5
5	2	0·1	1	2

(2) 414 *Machinists three months' study.*

0	296	256	313	299
1	74	122	41	69
2	26	30	33	26
3	8	5	17	11
4	4	1	7	5
5	4	0·1	2	2
6	1		1	2
7	0		0·1	
8	1			

* Incidence of Industrial Accidents. Report No. 4 (Industrial Fatigue Research Board), Stationery Office, London, 1919. Some errors in the evaluations of P in that report have been corrected

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

(3) 198 *Machinists six months' study.*

No. of accidents.	Observations.	Simple Poisson series.	Method of equation (8).	Method of equations (51) and (52).
0	69	53	71	66
1	54	70	49	61
2	43	46	41	38
3	15	20	23	19
4	13	7	10	9
5	1	2	3	4
6	2	0·5	1	1
7			0·2	0·6

Using Prof. Pearson's Goodness of Fit test,* the odds against such divergences as are shown by the simple Poisson series are found to be very great. (1) and (3) are fairly well graduated by eq. (8) P = ·13 and P = ·14. Equations (51) and (52) graduate all three fairly well, the values of P being respectively 0·29, 0·64 and 0·18. In 14 sets of data published in the report above mentioned the average value of P for graduations by the method of (51) and (52) was 0·38. Of the 14 P's, 6 were under 0·25, 3 between 0·25 and 0·5, 4 between ·5 and ·75, and 1 between ·75 and 1. If the hypotheses were correctly applicable the expectation would be 3.5 P's in each case, and the value of P for the observed against the theoretical distribution is 0·30. The hypothesis comes reasonably well out of the test.†

The propriety of assuming that the Poisson approximation holds for the several distributions within the "population" depends upon the validity of the arguments adduced in the earlier pages of this memoir. The choice of the binomial curve to represent the distribution of the continuously varying liabilities throughout the "population" has been dictated by considerations of practical convenience. An infinity of skew curves fulfilling the required conditions might be imagined, but no objective evidence favouring one more than another can be produced. We think, therefore, that the proposed method should be adopted.

The results obtained may now be summarised.

A general solution of the problem of the distribution arising when the chance of a happening is affected by antecedent success or failure has been obtained, although not in a form very suitable for computation. In the particular case of but two orders of

* *Biometrika* ix., 1913, p. 28.

† The distribution of P's for graduations by the method of equation (8) is 8 from 0 — ·25, 3 from ·25 — ·50, 1 from ·50 — ·75 and 2 from ·75 — 1·00, which gives P = ·04. Since P is itself subject to a large error of sampling, we do not attach much importance to this comparison.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

probability of the happening, a solution in terms of the moments of the distribution has been reached. It also appears that a first approximation to the form of the distribution can be deduced very rapidly from a modification of the ordinary pigeon-hole schema, although such modification has no correct theoretical foundation. Lastly, a very simple form of solution of the problem arising when the initial chances within the population vary has been provided.

We are alive to the inconvenience of the form in which the general solution of the first problem proposed has been expressed, but think it possible that the treatment may suggest to other investigators better lines of attack.

Appendix.

On the relation between the binomial curve and the binomial series.

In case any others may find the same difficulties as we did ourselves, we think it may be as well to give a brief note on the relations between the constants of the curve

$$y = y_0 \left(1 + \frac{x}{a}\right)^{\gamma a} e^{-\gamma x} \tag{1}$$

and the constants of the binomial series from which it may be derived. The curve was derived by Prof. Pearson in his classical memoir in the Philosophical Transactions of 1895, from the frequency-polygon given by the binomial expansion of $(q + p)^n$. In that memoir, owing to changes of notation, the relation between the p and n of the binomial series and the γ and a of the curve are not clear: following out the method of derivation in the original symbols we find, however, c being the distance apart at which the ordinates given by the binomial series are plotted :—

$$\left.\begin{aligned} \gamma &= \frac{2}{c\,(q - p)} \\ a &= \frac{2cpq\,(n + 1)}{q - p} \end{aligned}\right\} \tag{2}$$

But the curve may equally well be derived from a binomial series with negative index (Yule, *Journal of the Statistical Society*, 1910). Let this series be given by the expansion of

$$Q^N\,(1 - P)^{-N}$$

and let the polygon be plotted with an interval between the ordinates C. Then the relations between the constants of the curve and the constants of the polygon are :—

$$\left.\begin{aligned} \gamma &= \frac{2Q}{C\,(P + 1)} \\ a &= \frac{2CP\,(N - 1)}{(1 - P^2)} \end{aligned}\right\} \tag{3}$$

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

Hence the curves derived from a given positive binomial and from a given negative binomial are identical if

$$\left.\begin{aligned} \frac{1}{c\,(q-p)} &= \frac{\mathrm{Q}}{\mathrm{C}\,(\mathrm{P}+1)} \\ \frac{cpq\,(n+1)}{q-p} &= \frac{\mathrm{C.P}\,(\mathrm{N}-1)}{1-\mathrm{P}^2} \end{aligned}\right\} \qquad (4)$$

If we take the constants of the one binomial as given and wish to find constants of a binomial of the other form with the same equivalent curve, there are, in general, an infinity of solutions, for we have only two equations for the three unknowns. In one class of cases, however, solution is impossible : namely, when we are given P, C and N, and N is less than unity ; no real values of p, c, n can then satisfy the equations. In this case the initial ordinate of the curve is infinite and γ is negative. When the curve is of this form we may conclude definitely that it is derived from a negative and not from a positive binomial. In any other case the given *curve* may represent a *polygon* of either type, if we know nothing as to the position of the start of the curve, and may regard it as possible to place the binomial anywhere on the axis of abscissæ.

Supposing that the constants of the positive binomial are given, we may then assume a ratio of C to c, say

$$\mathrm{C}/c = r$$

and deduce values of P and N which will lead to the same curve-constants. We find

$$\left.\begin{aligned} \mathrm{P} &= \frac{q-p-r}{q-p+r} \\ \mathrm{N} &= \frac{1}{(q-p)^2-r^2}\left\{4pqn+1-r^2\right\} \end{aligned}\right\} \qquad (5)$$

Note that (5) shows there are limits for the admissible values of r : to give a positive value of P we must have $r < q - p$.

As an illustration, if

$$p = \tfrac{1}{8} \qquad q = \tfrac{7}{8} \qquad n = 8$$

and we assume

$$r = \tfrac{1}{2}$$

then

$$\mathrm{P} = \tfrac{1}{5} \qquad \mathrm{Q} = \tfrac{4}{5} \qquad \mathrm{N} = 13.6$$

Both sets of values give for the curve-constants

$$\gamma = 2\tfrac{2}{3} \qquad a = 2\tfrac{5}{8}$$
$$\gamma a = 7$$

The mean of the positive binomial is 1. In order that the mean of the curve may coincide with this, its origin, the mode must lie $0.375c$ to the left. The mean of the negative binomial is at C.NP/Q,

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions

that is, 3.4C (or 1.7c) to the right of its zero-point.* To make the means of the two binomials coincide, the zero of the negative binomial must therefore be taken 0.7c to the left of the zero of the positive binomial. The figure shows the points of the two binomial polygons and the fitted curve.

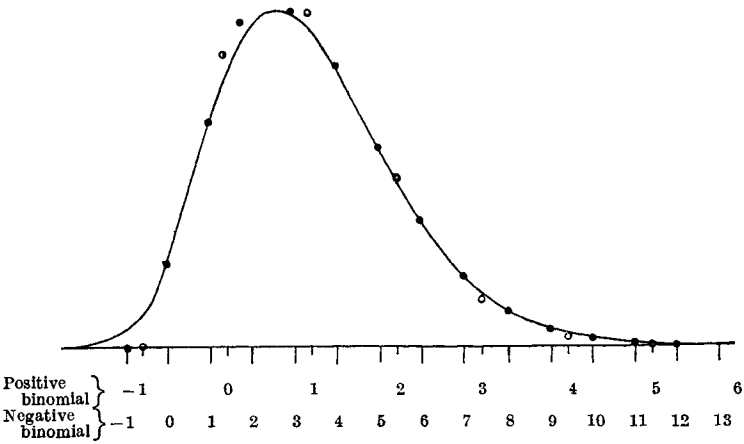

Diagram showing binomial curve fitted, by the slope relation, to both a positive binomial (hollow circles) and a negative binomial (circles blocked in).

The form of the curve then does not seem, in general, to give any criterion as between positive binomial and negative binomial, if the binomial distribution may be regarded as arbitrarily assignable to any position on the axis of abscissæ, and the distance between its ordinates as also arbitrary. But if the binomial distribution be regarded as naturally placed, and the distance between its ordinates, as usual, as unity, there is a criterion. For the second moment of the positive binomial npq is necessarily less than its mean np: while the second moment of the negative binomial NP/Q^2 is necessarily greater than its mean NP/Q. Hence for the limiting curve the same criterion will hold. In this sense the distributions given by equation (46A) of our paper for the "liabilities" of the operatives to accident are all of the negative binomial type, for the c of that equation is less than unity and hence the second moment is necessarily greater than the mean.

* The moments of the negative binomial $Q^N(1 - P)^{-N}$ may be derived from the known formulæ for the positive binomial $(q + p)^n$ by substituting $-N$ for n, $-P/Q$ for p, and $1/Q$ for q. In *Journal of the Royal Statistical Society*, vol. 73, p. 23, equation (5), for r/p in the equation for M, read rq/p.

This content downloaded from 128.118.88.48 on Wed, 4 Jun 2014 15:18:09 PM
All use subject to JSTOR Terms and Conditions