CHAINS OF RARE EVENTS

By Felix Cernuschi[1] and Louis Castagnetto

Harvard University

1. Summary. The negative binomial distribution of Greenwood and Yule is generalized and modified in order to obtain distribution curves which could be used in many concrete cases of chains of rare events. Assuming that the numbers of single, double, triple, and so on, events are distributed according to Poisson's law with parameters λ_1, λ_2, $\lambda_3 \cdots$ respectively, and that λ_s is given by $\lambda_s = \lambda_1 \frac{a^{s-1}}{s!}$, the probability of obtaining M successful events is studied. In the considered relation λ_s, for convenient values of a, first increases with s and after a certain saturation value of s starts to decrease. A relation of this type is very suitable for studying the distribution of score in a match between two first class billiard players, the probability of accidents on a highway of dense traffic, etc. The general methods of finding the distribution curves for arbitrary relations between the λ's are indicated. The method of steepest descent is applied to find an acceptable approximation of the distribution function; and the advantage of this method is pointed out for other similar cases, in addition to the concrete one which was developed, in which the method of direct expansion into power series becomes inapplicable.

2. Introduction. M. Greenwood and G. U. Yule [1] have deduced the negative binomial distribution from a compound Poisson law:

$$P(m, \lambda) = \frac{\lambda^m}{m!} e^{-\lambda},$$

where λ itself is a random variable distributed according to Pearson's law of type III:

$$P(\lambda)\, d\lambda = \beta^{\alpha+1} \frac{\lambda^\alpha}{\alpha!} e^{-\beta\lambda}\, d\lambda.$$

They obtained the distribution

$$P(m) = (1 - a)^{\alpha+1} \frac{(\alpha + m)!}{\alpha!\, m!} a^m,$$

where $1 - a = \frac{\beta}{\beta + 1}$. As is easily seen, $P(m)$ is given by the coefficient of x^m in the expansion of:

$$(1 - a)^{\alpha+1}\left(1 - \frac{x}{\beta + 1}\right)^{-(\alpha+1)}.$$

[1] Research done at Harvard Astronomical Observatory as Guggenheim Fellow.

R. Lüders [2] has arrived at a negative binomial law by the following considerations. Certain events, like automobile accidents, can be classified as simple or multiple according to the number of units involved. Assume that the numbers of single, double, triple, and so on, events are distributed according to Poisson's law with the parameters $\lambda_1, \lambda_2, \lambda_3, \cdots$, respectively. The probability of obtaining n_1 single, n_2 double, n_3 triple, $\cdots$ successful events is (assuming mutual independence)

$$P(n_1, n_2, n_3, \cdots; \lambda_1, \lambda_2, \lambda_3, \cdots) = \frac{\lambda_1^{n_1}}{n_1!}\frac{\lambda_2^{n_2}}{n_2!} \cdots e^{-(\lambda_1+\lambda_2+\cdots)}. \tag{1}$$

The total number of successful events is

$$n = n_1 + 2n_2 + 3n_3 + \cdots + in_i + \cdots. \tag{2}$$

The probability of obtaining n successful events is given by the sum of all expressions (1) subject to the condition (2). This sum is given by the coefficient of x^n in the expansion

$$f(x) = e^{-(\lambda_1+\lambda_2+\cdots)} e^{\lambda_1 x + \lambda^2 x 2 + \cdots}. \tag{3}$$

Now if the parameters λ_s satisfy

$$\lambda_s = \lambda_1 \frac{a^{s-1}}{s} \tag{4}$$

one finds

$$f(x) = \left(\frac{1-a}{1-ax}\right)^{\frac{\lambda_1}{a}} \tag{5}$$

and

$$P(n) = (1-a)^{\frac{\lambda_1}{a}} \frac{\lambda_1}{a}\left(\frac{\lambda_1}{a}+1\right)\cdots\left(\frac{\lambda_1}{a}+n-1\right)\frac{a^n}{n!}. \tag{6}$$

Taking $\frac{\lambda_1}{a}$ equal to $\alpha + 1$ one gets Greenwood and Yule's distribution in the form given above [3]. The negative binomial law has useful applications, for instance in some cases of accidents of workers in factories. It is proved that with values of a near 1, the most probable value for n is $n = 0$ and the average value is a finite number different from zero. Therefore the distribution will be in some way similar to the distribution of the scores in a match between two first class billiard players whose most frequent scores are zero and their average may be, say, 50. In the case of the Poisson distribution the most frequent score and the average score should be nearly the same. The relation (4) does not provide an adequate description of many practical distributions. For instance, in a match between two first class billiard players, the probability of making a second, third, $\cdots$, point will be considerably greater than the probability of making the first. With the relation (4) λ_s is a decreasing function of s, while we shall investigate cases in which λ_s first increases with s and after a certain value of s starts to decrease. As other examples of distributions of similar types we shall mention the following: On a highway with dense traffic at high speeds the probability of only one car being involved in an accident may be smaller than the probability of having several cars involved. Something similar may be said for the cases of work accidents in factories where the work of one is interconnected with the work of others. In many cases of telephone calls (business transactions, organization of meetings, etc.) the sample Poisson law is not suitable to interpret the distribution of calls, since one call may increase the probability that the called person makes one or more calls.

The purpose of this paper is to treat the problem when, instead of (4), we take other expressions which may in a better way describe some processes such as the ones which we have referred to.

3. Modification and generalization of the scheme of Greenwood-Yule and Lüders. According to the relation (4) λ_s is a decreasing function of s and the parameter a must be in the interval $0 \leq a < 1$. Instead of (4) we shall use

$$\lambda_s = \lambda_1 \frac{a^{s-1}}{s!}, \tag{7}$$

where a may have any positive value. In particular for $a = 0$ our case reduces to the Poisson case.

From (7) it follows that

$$\frac{\lambda_{s+1}}{\lambda_s} = \frac{a}{s+1} \tag{8}$$

and we see that λ_s increases with s for $1 \leq s < a$ and decreases for $s + 1 > a$. Substituting from (7) in (3) we get

$$f(x) = e^{-(\lambda_1/a)e^a} e^{(\lambda_1/a)e^{ax}}. \tag{9}$$

As the probability of obtaining n successful events is given by the coefficient of x^n in (9), we shall expand $e^{\alpha e^{\beta x}}$ in power series (α, β being two arbitrary constants). We have

$$e^{\alpha e^{\beta x}} = 1 + \sum_{n=1}^{\infty} \frac{\alpha^n}{n!} e^{n\beta x} = e^{\alpha} + \sum_{n=1}^{\infty} \frac{\beta^n}{n!} x^n \sum_{m=1}^{\infty} \frac{m^n \alpha^m}{m!}. \tag{10}$$

Now

$$\sum_{m=1}^{\infty} \frac{m^n}{m!} \alpha^m = y_n(\alpha) e^{\alpha} \tag{11}$$

where [4]

$$
\begin{aligned}
y_1(\alpha) &= \alpha \\
y_2(\alpha) &= \alpha^2 + \alpha \\
y_3(\alpha) &= \alpha^3 + 3\alpha^2 + \alpha \qquad (12)\\
&\cdots\cdots\cdots\cdots \\
y_n(\alpha) &= \sum_{i=1}^{n} \frac{\Delta^i 0^n}{i!} \alpha^i .
\end{aligned}
$$

Here we use the notation of differences of zero: $\Delta^i 0^n$. We have

$$e^{\alpha e^{\beta x}} = e^{\alpha}\left[1 + \sum_{n=1}^{\infty} \frac{y_n(\alpha)\beta^n}{n!} x^n\right] \tag{13}$$

or

$$e^{\alpha e^{\beta x}} = e^{\alpha}\left[1 + \sum_{n=1}^{\infty} \frac{\beta^n}{n!} x^n \sum_{i=1}^{n} \frac{\Delta^i 0^n}{i!} \alpha^i\right]. \tag{14}$$

Now in our case

$$\alpha = \frac{\lambda_1}{a}, \qquad \beta = a, \tag{15}$$

whence

$$P(n) = e^{-(\lambda_1/a)(e^a - 1)} \frac{a^n}{n!} \sum_{i=1}^{n} \frac{\Delta^i 0^n}{i!} \left(\frac{\lambda_1}{a}\right)^i, \qquad n > 0. \tag{16}$$

$$P(0) = e^{-(\lambda_1/a)(e^a - 1)}, \qquad \text{for } n = 0. \tag{17}$$

We have in particular

$$
\begin{aligned}
P(1) &= \lambda_1 P(0) \\
P(2) &= \frac{1}{2!}(\lambda_1^2 + a\lambda_1)P(0) \\
P(3) &= \frac{1}{3!}(\lambda_1^3 + 3\lambda_1^2 a + \lambda_1 a^2)P(0) \\
P(4) &= \frac{1}{4!}(\lambda_1^4 + 6\lambda_1^3 a + 7\lambda_1^2 a^2 + \lambda_1 a^3)P(0) \\
P(5) &= \frac{1}{5!}(\lambda_1^5 + 10\lambda_1^4 a + 25\lambda_1^3 a^2 + 15\lambda_1^2 a^3 + \lambda_1 a^4)P(0) \\
P(6) &= \frac{1}{6!}(\lambda_1^6 + 15\lambda_1^5 a + 65\lambda_1^4 a^2 + 90\lambda_1^3 a^3 + 31\lambda_1^2 a^4 + \lambda_1 a^5)P(0) \\
(18)\quad P(7) &= \frac{1}{7!}(\lambda_1^7 + 21\lambda_1^6 a + 140\lambda_1^5 a^2 + 350\lambda_1^4 a^3 + 301\lambda_1^3 a^4 + 63\lambda_1^2 a^5 \\
&\qquad + \lambda_1 a^6)P(0) \\
P(8) &= \frac{1}{8!}(\lambda_1^8 + 28\lambda_1^7 a + 266\lambda_1^6 a^2 + 1050\lambda_1^5 a^3 + 1701\lambda_1^4 a^4 + 966\lambda_1^3 a^5 \\
&\qquad + 127\lambda_1^2 a^6 + \lambda_1 a^7)P(0)
\end{aligned}
$$

$$P(9) = \frac{1}{9!}(\lambda_1^9 + 36\lambda_1^8 a + 462\,\lambda_1^7 a^2 + 2646\lambda_1^6 a^3 + 6951\lambda_1^5 a^4 + 7770\lambda_1^4 a^5 + 3025\lambda_1^3 a^6 + 255\lambda_1^2 a^7 + \lambda_1 a^8)P(0)$$

$$P(10) = \frac{1}{10!}(\lambda_1^{10} + 45\lambda_1^9 a + 750\lambda_1^8 a^2 + 5880\lambda_1^7 a^3 + 28827\lambda_1^6 a^4 + 42525\lambda_5^1 a^5 + 34105\lambda_1^4 a^6 + 9330\lambda_1^3 a^7 + 511\lambda_1^2 a^8 + \lambda_1 a^9)P(0).$$

For $\lambda_1 = a$ it follows that

(19) $$P(0) = e^{-e^a+1}$$

(20) $$P(n) = e^{-e^a+1}\frac{a^n}{n!}\,y_n(1)$$

$$\sum_{n=0}^{\infty} P(n) = ee^{-e^a}\left[1 + \sum_{n=1}^{\infty}\frac{a^n}{n!}\,y_n(1)\right].$$

Particular values of (20) are

$$P(1) = aP(0)$$

$$P(2) = a^2 P(0)$$

$$P(3) = \frac{5a^3}{3!}P(0)$$

$$P(4) = \frac{15a^4}{4!}P(0)$$

$$P(5) = \frac{52a^5}{5!}P(0)$$

(21) $$P(6) = \frac{203a^6}{6!}P(0)$$

$$P(7) = \frac{877a^7}{7!}P(0)$$

$$P(8) = \frac{4140a^8}{8!}P(0)$$

$$P(9) = \frac{21147a^9}{9!}P(0)$$

$$P(10) = \frac{115975a^{10}}{10!}P(0).$$

In Figure 1 we have graphed the curves $P(n)$ for the values $\frac{\lambda_1}{a} = 1$; $\lambda_1 = 0.1$, $\lambda_1 = 1$, $\lambda_1 = 2$. We see in particular, that for $\lambda_1 = 1$ we have $P(0) = P(1)$ and for $\lambda_1 = a = 1$ we have $P(0) = P(1) = P(2)$.

4. Application of the method of steepest descent. If λ_s is not given by (7) the above method of direct expansion of $f(x)$ into a power series, usually becomes

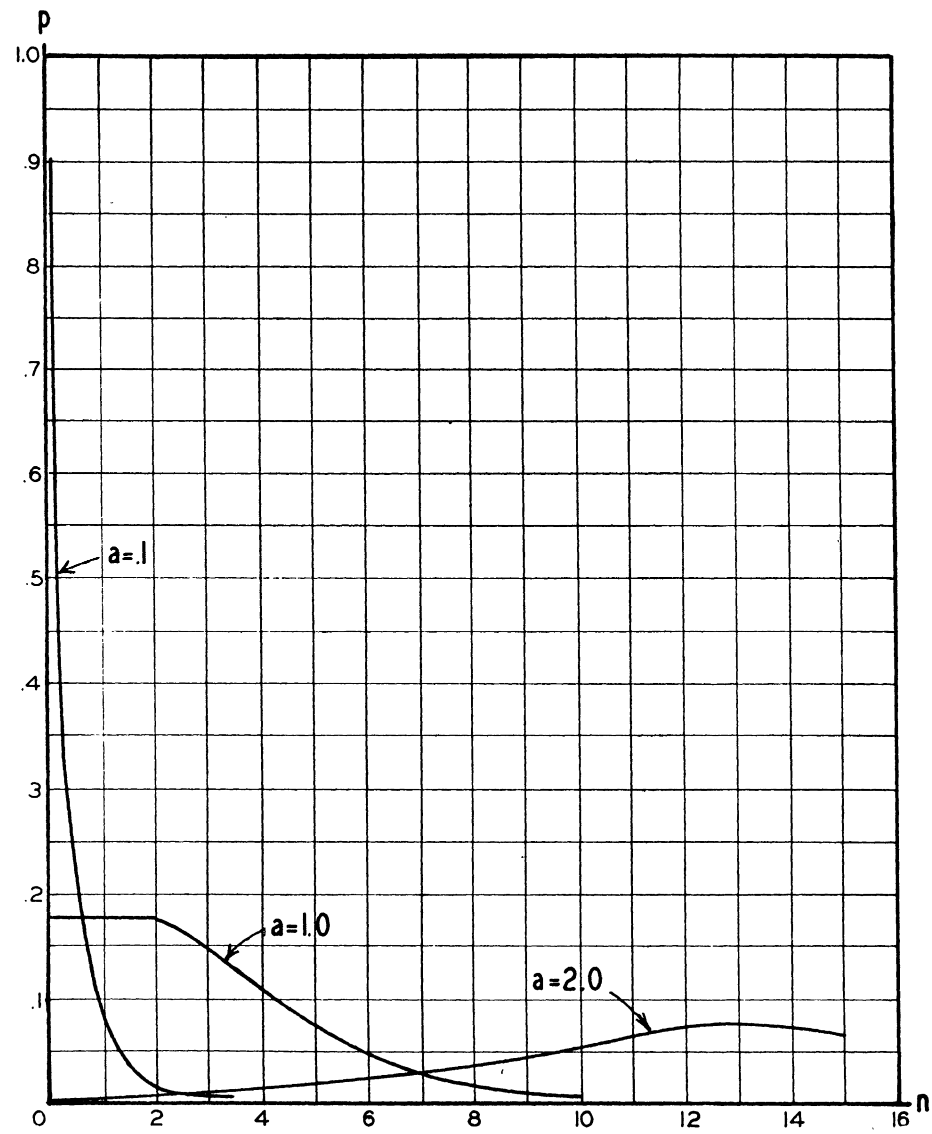

FIG. 1 DISTRIBUTION CURVES FOR $\frac{\lambda_1}{a} = 1,\ a = 0.1;\ \frac{\lambda_1}{a} = 1,\ a = 1;\ \frac{\lambda_1}{a} = 1,\ a = 2$

inapplicable. In many cases it is possible to use instead the method of steepest descent [5] in order to obtain approximate values for the coefficients of x^n in the relation (3).

As is well known, if $f(z)$ is an analytical function we have

$$\text{coeff. of } z^n = \frac{1}{2i\pi} \oint \frac{f(z)}{z^{n+1}}\, dz = \frac{1}{2i\pi} \oint e^{X(x,y)+iY(x,y)}\, dz \tag{22}$$

where $X + iY = \log \dfrac{f(z)}{z^{n+1}}$ and the integral is taken along any closed path around the origin.

To evaluate the integral (22) we shall follow a method similar to the one used by R. H. Fowler [6]. Putting $z = \rho e^{i\alpha}$ the relation (22) may be written:

$$\text{Coeff. of } z^n = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{f(\rho e^{i\alpha})}{\rho^n e^{in\alpha}}\, d\alpha \tag{23}$$

where the value of ρ is arbitrary. We shall put in particular $\rho = x_0$ where x_0 is the root of

$$\frac{x_0 f'(x_0)}{f(x_0)} = n. \tag{24}$$

For most functions which interest us $\dfrac{f(x)}{x^n} \to \infty$ as $x \to 0$ and as $x \to K$ (a positive number which in some cases may be infinite) and the second derivative is always positive. Consequently $f(x)/x^n$ has only one minimum between 0 and K, and (24) has therefore only one root x_0. Developing $\log \dfrac{f(x_0 e^{i\alpha})}{x_0^n e^{in\alpha}}$ into powers of α, (24) becomes

$$\text{coeff. of } x^n = \frac{1}{2\pi} \frac{f(x_0)}{x_0^n} \int_{-\pi}^{+\pi} e^{-(x_0^2\varphi''(x_0)/2)\alpha^2 + ig(x_0)\alpha^3 + h(x_0)\alpha^4 + \cdots}\, d\alpha, \tag{25}$$

where

$$\varphi(x) = \log \frac{f(x)}{x^n}.$$

In the case where $\varphi''(x_0)\, \dfrac{x_0^2}{2} \gg 1$ the first term in the exponent in (25) increases in absolute value very rapidly in the neighborhood of x_0. For small values of α we may therefore in a first approximation drop all other terms. Also, as this first term tends rapidly towards zero one does not appreciably increase the error by replacing the integral from $-\pi$ to $+\pi$ by the integral from $-\infty$ to $+\infty$.

In such cases we have, therefore, the approximate formula

$$\text{coeff. of } z^n \sim \frac{1}{2i\pi} \frac{f(x_0)}{x_0^m} \int_{-\infty}^{+\infty} e^{-(\varphi''(x_0)x_0^2/2)\alpha^2}\, d\alpha = \frac{f(x_0)}{x_0^{n+1}\sqrt{2\pi\varphi''(x_0)}}. \tag{26}$$

We are now in a position to deduce asymptotic values for the probabilities $P(n)$

which we have previously calculated directly. In fact, for $f(x)$ defined by (9) we obtain from (26) for large n

$$P(n) \sim \frac{e^{-(\lambda_1/a)e^a}}{\sqrt{2\pi}} \frac{e^{n/ax_0}}{x_0^n \sqrt{n(ax_0+1)}}, \tag{27}$$

where x_0 is given by

$$e^{ax_0} = \frac{n}{\lambda_1 x_0}.$$

In particular for $\lambda_1 = a$ and putting $ax_0 = y_0$ it follows that

$$P(n) \sim 0.3989\left(\frac{\lambda_1}{y_0}\right)^n e^{-e\lambda_1} \frac{e^{e^{y_0}}}{\sqrt{n(y_0+1)}}. \tag{28}$$

Comparing the numerical values given by the relation (28) with the exact values we find that even for $n = 4$ and $\lambda_1 = 1$ (28) gives an approximation with an error of about 5%.

Formula[2] (26) can also be used to evaluate the numbers $y_n(1)$ defined by (12) for $\alpha = 1$. Relation (13) gives for $\alpha = \beta = 1$

$$e^{e^x} = e\left[1 + \sum_{n=1}^{\infty} \frac{y_n(1)}{n!} x^n\right]$$

and therefore

$$\text{Coeff. of } x^n \text{ in expansion of } e^{e^x} = \frac{ey_n(1)}{n!}.$$

Putting $f(z) = e^{e^z}$ and using Stirling's formula for $n!$ we have from (26)

$$y_n(1) \sim \frac{e^{n\left[x_0 + \frac{1}{x_0} - \left(1+\frac{1}{n}\right)\right]}}{\sqrt{x_0+1}},$$

[2] Applying this relation to $f(z) = e^z$ one obtains immediately Stirling's Formula:

$$\varphi(z) = \log \frac{f(z)}{z^n} = z - n \log z$$

$$\varphi'(z) = 1 - \frac{n}{z}, \qquad x_0 = n,$$

$$\varphi''(z) = \frac{n}{z^2}, \qquad \varphi''(x_0)\frac{x_0^2}{2} = \frac{n}{2},$$

$$\frac{1}{n!} \sim \left(\frac{e}{n}\right)^n \frac{1}{\sqrt{2\pi n}}.$$

Also relation (26) is useful to find other symptotic expressions; e.g. for $f(z) = (pz + 9)^n$ one obtains for $n \to \infty$ the Laplace-Gauss formula.

where x_0 is given by

$$e^{x_0} = \frac{n}{x_0}.$$

For $n = 4$, $x_0 = 1.202$ and $y_4(1) \sim 15.56$. As the exact value of $y_4(1)$ is 15 we obtain in this case an error of less than 4%.

Repeating the calculations for $n = 6$, $x_0 = 1.432$, we find that $y_6(1)$ is given with an error of less than 3%.

REFERENCES

[1] M. GREENWOOD AND G. U. YULE, *Roy. Stat. Soc. Jour.*, Vol. 83 (1920), p. 225.
[2] R. LÜDERS, *Biometrika*, Vol. 26 (1934), p. 108.
[3] H. JEFFREYS, *Theory of Probability*, Oxford Univ. Press, 1939, pp. 67–69.
[4] H. T. DAVIS, *Tables of Higher Mathematical Functions*, Principia Press, 1935, Vol. 2, p. 212.
[5] P. DEBYE, *Mathematiche Annalen*, Vol. 67 (1909), p. 535.
L. BRILLOUIN, *Annales de l'École Normale Supérieure*, 3eme Série, Vol. 17 (1916), p. 17.
[6] R. H. FOWLER, *Statistical Mechanics*, Cambridge Univ. Press, 1936, p. 36.