THE NEGATIVE BINOMIAL AND POISSON DISTRIBUTIONS COMPARED

BY

LEROY J. SIMON

Section I — Preliminaries

For much statistical work the binomial distribution is the most suitable mathematical model. It involves n independent trials, each having a probability of success equal to p. In automobile and other branches of casualty insurance, we are not concerned with a limited number of independent trials, but with an exposure to accident such that n becomes very great while $n \times p$ remains finite and is the number of "successes". In this case, the Poisson distribution is the correct statistical model to produce the probability that 0, 1, 2 . . . . . "successes" will be experienced by a given observational unit (one car, a fleet of cars, all the cars in one territory, etc.).

In computing the probability distribution of the number of experimental units which will have 0, 1, 2 . . . . . "successes", the Poisson distribution is also a good representation if the loss frequency is the same for each element in the group; or in other words the group is isohazardous.[1] In many cases, particularly in automobile insurance, we know that even classified experience is not isohazardous and in such circumstances the negative binomial distribution is the most appropriate model. Considerable interest has been recently stimulated[2] in the negative binomial distribution as it applies to the emergence of claims in automobile insurance. This note shows how the negative binomial distribution compares with the better known Poisson distribution. The method employed is to develop the first four moments of each distribution and then to compare their kurtosis and skewness.

Moment-generating functions (mgf) will be used to develop the moments needed. Let f(x) denote the frequency function being studied. Recall that for discrete data the mgf, designated $M(\theta)$, is defined as follows:

$$M(\theta) = \sum_{x=0}^{\infty} e^{\theta x} f(x) \tag{1}$$

By substituting the power series for $e^{\theta x}$, multiplying through by f(x) and applying the summation to each individual term of the expanded series, it develops that

$$M(\theta) = 1 + \theta \mu_1' + \frac{\theta^2}{2!}\mu_2' + \frac{\theta^3}{3!}\mu_3' + \ldots\ldots + \frac{\theta^k}{k!}\mu_k' + \ldots\ldots$$

[1] "Isohazardous" is a coined word. *Adj.* [Gr. *isos* equal + O.F. *hasard,* fr. An. *alzahr* the die.] *1.* Having the same or equal inherent hazard. *2.* Homogeneous in propensity for accident involvement. *3. Ins.* Having the same loss frequency potential. The nominative form is "isohazard".

[2] Dropkin, Lester, "Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records", CAS XLVI, p. 165 and discussion thereof in this volume.

where μ'_k is the k^{th} moment about the origin. From this latter form it can be seen that if one takes the k^{th} derivative of $M(\theta)$ with respect to θ, all terms with powers of θ less than k will differentiate to zero. If we then set $\theta=0$, all terms originally having powers of θ greater than k will also go to zero. This will leave the k^{th} moment about the origin with the factorial in the denominator being exactly cancelled by the factorial produced by k differentiations of θ^k. Hence, we may state that

$$\mu'_k = \left.\frac{d^k M}{d\theta^k}\right|_{\theta=0} \tag{2}$$

It is further necessary to recall that we reduce moments about the origin to moments about the mean through use of the following equations:

$$\left.\begin{aligned} \mu_2 &= \mu'_2 - (\mu'_1)^2 \\ \mu_3 &= \mu'_3 - 3\mu'_2\,\mu'_1 + 2(\mu'_1)^2 \\ \mu_4 &= \mu'_4 - 4\mu'_3\,\mu'_1 + 6\mu'_2\,(\mu'_1)^2 - 3(\mu'_1)^4 \end{aligned}\right\} \tag{3}$$

Finally, the customary measurement of skewness is $\alpha_3 = \mu_3/\mu_2^{3/2}$ (4)

Kurtosis is measured by $\alpha_4 = \mu_4/\mu_2^2$ (5)

Section II – The Poisson Distribution

The Poisson distribution is described by the following function:

$$f(x) = \frac{m^x e^{-m}}{x!}$$

The moment-generating function for the Poisson would be given by

$$\begin{aligned} M(\theta) &= \sum_{x=0}^{\infty} \frac{m^x e^{-m+\theta x}}{x!} \\ &= e^{-m}\left[1 + me^{\theta} + \frac{(me^{\theta})^2}{2\,!} + \ldots\right] \\ &= e^{-m} e^{me^{\theta}} \end{aligned}$$

Hence $M(\theta) = e^{m(e^{\theta}-1)}$

The first four derivatives of $M(\theta)$ are:

$M^{i}\ (\theta) = M(\theta)me^{\theta}$

$M^{ii}\ (\theta) = M(\theta)me^{\theta}(1 + me^{\theta})$

$M^{iii}(\theta) = M(\theta)me^{\theta}[me^{\theta} + (1 + me^{\theta})^2]$

$M^{iv}\ (\theta) = M(\theta)me^{\theta}[me^{\theta}(3 + 2me^{\theta}) + (1 + me^{\theta})(1 + 3me^{\theta} + m^2e^{2\theta})]$

Evaluating the above four equations at $\theta=0$, we will produce the first four moments about the origin, that is

$\mu'_1 = m$

$\mu'_2 = m^2 + m$

$\mu'_3 = m^3 + 3m^2 + m$

$\mu'_4 = m^4 + 6m^3 + 7m^2 + m$

Substituting these values back in equations (3), we produce the following moments about the mean:

$\mu_2 = m$

$\mu_3 = m$

$\mu_4 = 3m^2 + m$

The measurements of kurtosis and skewness thus become:

$$\alpha_3 = \frac{1}{\sqrt{m}}$$

$$\alpha_4 = 3 + \frac{1}{m}$$

Section III – Negative Binomial Distribution

The negative binomial distribution is described by the following function:

$$f(x) = \left(\frac{a}{1+a}\right)^r \binom{-r}{x} \left(\frac{-1}{1+a}\right)^x$$

The moment-generating function for the negative binomial would be given by

$$M(\theta) = \sum_{x=0}^{\infty} e^{\theta x} \left(\frac{a}{1+a}\right)^r \binom{-r}{x} \left(\frac{-1}{1+a}\right)^x$$

which can be re-written in the form

$$M(\theta) = \left(\frac{a}{1+a}\right)^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(\frac{-e^\theta}{1+a}\right)^x (1)^{-r-x}$$

The last term in this equation is inserted merely to permit an analogy to the formula for the expansion of a binomial function. Upon simplification, the mgf can be written

$$M(\theta) = a^r(1 + a - e^\theta)^{-r}$$

For simplicity of notation, $k = \frac{r+1}{r}$ and $M(\theta)$ is written as M in the formulas below. From $M(\theta)$ we get:

$$M^i\ (\theta) = Mre^\theta(1 + a - e^\theta)^{-1}$$

$$M^{ii}\ (\theta) = \frac{k(M^i)^2}{M} + M^i$$

$$M^{iii}(\theta)=\frac{2kM^{i}M^{ii}}{M}-\frac{k(M^{i})^{3}}{M^{2}}+M^{ii}$$

$$M^{iv}(\theta)=\frac{2k}{M^{2}}[MM^{i}M^{iii}+M(M^{ii})^{2}-(M^{i})^{2}M^{ii}]$$
$$-\frac{k}{M^{4}}[3M^{2}(M^{i})^{2}M^{ii}-2M(M^{i})^{4}]+M^{iii}$$

Evaluating the above four equations at $\theta=0$ we will produce the first four moments about the origin; that is

$$\mu_1'=\frac{r}{a}$$

$$\mu_2'=\frac{r}{a^2}(1+a+r)$$

$$\mu_3'=\left(\frac{2kr}{a}+1\right)\left(\frac{kr^2}{a^2}+\frac{r}{a}\right)-\frac{kr^3}{a^3}$$

$$\mu_4'=\frac{r^4}{a^4}(6k^3-7k^2+2k)+\frac{r^3}{a^3}(12k^2-6k)+\frac{r^2}{a^2}(7k)+\frac{r}{a}$$

Substituting these values back in equations (3), we produce the following moments about the mean:

$$\mu_2=\frac{r}{a^2}(a+1)$$

$$\mu_3=\frac{r}{a^3}(a+1)(a+2)$$

$$\mu_4=\frac{r(a+1)}{a^4}[(a+1)(a+3r+5)+1]$$

The measurements of kurtosis and skewness thus become:

$$\alpha_3=\frac{a+2}{\sqrt{r(a+1)}}$$

$$\alpha_4=3+\frac{a}{r}+\frac{5}{r}+\frac{1}{r(a+1)}$$

Section IV – A Comparison

In Section II the letter "m" can now be replaced by "r/a" so that a direct comparison may be made of the two distributions. From this we find that the negative binomial α_3 minus Poisson α_3 equals

$$\frac{a+2}{\sqrt{r(a+1)}} - \frac{\sqrt{a}}{\sqrt{r}} = \frac{a+2-\sqrt{a^2+a}}{\sqrt{r(a+1)}}$$

Since a and r are both positive, this latter quantity is always a real, positive number, and it follows that the negative binomial is always more skew to the right than the Poisson distribution.

If we take the negative binomial α_4 minus the Poisson α_4, we obtain after simplification, $\frac{5}{r} + \frac{1}{r(a+1)}$. Since both a and r are positive, this quantity is also positive, which means that the negative binomial is always more peaked than the Poisson.[3]

The table which follows assumes that we are given a population which is distributed as a negative binomial with $r = .8$ and $a = 8$ (thus, mean $= .10$ and variance $= .1125$). The first column of probabilities are those of this negative binomial population. The next column is the result of fitting a Poisson distribution to the first column (i.e., $m = .10$). The last two columns indicate the differences and point out how the Poisson underestimates the probability in all cases except for $x = 1$. It is also apparent that for a first or rough approximation, or for some special purpose, the Poisson distribution is still a fairly good representation even in those cases where the negative binomial is suspected or known to apply. The usefulness of this approximation diminishes as the ratio of the variance to the mean increases and also is much less valid as we move toward the higher number of accidents.

PROBABILITY DISTRIBUTION OF NUMBER OF RISKS

Number of Accidents x	Negative Binomial	Poisson	Difference	Percent Difference
0	.910076	.904837	+.005239	1%
1	.080896	.090484	−.009588	12
2	.008090	.004524	+.003566	44
3	.000839	.000151	+.000688	82
4	.000089	.000004	+.000085	96
5	.000009	.000000	+.000009	100
6	.000001	.000000	+.000001	100
Total	1.000000	1.000000		

[3]The curves are both J-shaped for the small values of r/a which we usually encounter in insurance (ordinarily much less than one accident per year) and the same conclusion would be indicated if we proved the theorem: The number of risks having zero accidents for a negative binomial distribution exceeds that for a Poisson distribution with the same mean. In our notation, this means f(o) for negative binomial $>$f(o) for Poisson, which is true if $[a/(1+a)]^r > e^{-r/a}$. Take the r^{th} root of both sides (r is positive), take the reciprocal and write the series expansion giving

$$1+\frac{1}{a} < e^{1/a} = 1+\frac{1}{a}+\frac{1}{2\,!a^2}+\frac{1}{3\,!a^3}+\cdots$$

Since a is positive, the theorem is proven, thus proving that for small r/a the negative binomial is more peaked.