A Correction for an Underdispersed Event Count Probability Distribution

Rainer Winkelmann, Curtis S. Signorino, and Gary King

We demonstrate that the expected value and variance commonly given for a well-known probability distribution are incorrect. We also provide corrected versions and report changes in a computer program to account for the known practical uses of this distribution.

The probability distribution in question, named the continuous parameter binomial (CPB) by King (1989a), has been known for at least six decades. The publications that reported the moments incorrectly, or were at least unclear about them, include a dated article (Guldberg 1931), a dissertation and several resulting published articles (Katz 1945, 1965), a popular reference book (Johnson and Kotz 1969),[1] work in political science on event count regression models (King 1989a), and extensions of these event count models in econometrics (Winkelmann and Zimmermann 1991). Event count regression models have become increasingly common in empirical political science research; some recent examples include Wang et al. (1993) and Krause (1994).

We first prove that the commonly given expressions for the expectation and variance of the CPB are incorrect. Then, we derive the correct expressions. Finally, we evaluate the impact on the existing empirical applications of event count regression models (in political science and elsewhere), which rely in part on the CPB but the wrong expectation or variance. We show that the

1. Johnson and Kotz (1969) did not give the CPB expected value explicitly. However, they do so implicitly by giving the factorial moments in their equation 42 (p. 41) and saying that it applies to the CPB; these are wrong because they imply the wrong expression for the expected value.

corrections will not materially affect any existing substantive results. We also provide information in this section about a computer program to run event count regression models and for which the moments have been corrected.

The Problem

As is well known, a random variable Y has a *binomial distribution* with parameters n, $n \in \mathbf{N}$, and p, $p \in (0,1)$ (writing $Y \sim B(y|n,p)$) if

$$p_k \equiv P(Y = k) = \binom{n}{k} p^k(1 - p)^{n-k} \tag{1}$$

where $k = 0,1,\ldots,n$. The expectation and variance of Y are given by $E(Y) = np$ and $\mathrm{Var}(Y) = np(1 - p)$, respectively.

The CPB arises by replacing the nonnegative integer n in equation 1 by a continuous $\alpha \in \mathbf{R}^+$ where $k = 0,1,\ldots,\bar{n}$ and

$$\bar{n} = \begin{cases} \mathrm{int}(\alpha) + 1 & \text{if } \alpha \text{ non-integer} \\ \alpha & \text{if } \alpha \text{ integer} \end{cases}$$

When α is not an integer, the p_k terms do not sum to one as is required for probability distributions. Thus, the probabilities $\bar{p}_k$ are obtained by applying the following normalization:

$$\bar{p}_k = \frac{p_k}{\sum_{i=0}^{\bar{n}} p_i}, \qquad k = 0,1,\ldots,\bar{n}. \tag{2}$$

For continuous α, we rewrite equation 1, using either the generalized binomial coefficient $\binom{a}{k}$ or the Gamma function:

$$p_k = \frac{\Gamma(\alpha + 1)}{\Gamma(k + 1)\Gamma(\alpha - k + 1)} p^k(1 - p)^{\alpha-k}, \tag{3}$$

since $n! = \Gamma(n + 1)$. The claim we disprove is that the expected value of the probability distribution defined by equation 2 is equal to αp for all α and p.

Proposition 1

Let the random variable Y have a continuous parameter binomial distribution $CPB(y|\alpha,p)$, where α, $p \in (0,1)$. Then

$$E(Y) > \alpha p.$$

Proof.

If $\alpha \in (0,1)$ then $k \in \{0,1\}$, and

$$p_0 = \frac{\alpha\Gamma(\alpha)}{\Gamma(1)\Gamma(\alpha + 1)} p^0(1 - p)^{\alpha-0} = (1 - p)^\alpha$$

$$p_1 = \frac{\alpha\Gamma(\alpha)}{\Gamma(2)\Gamma(\alpha)} p^1(1 - p)^{\alpha-1} = \alpha p(1 - p)^{\alpha-1}$$

Moreover,

$$E(Y) = \bar{p}_1 = \frac{\alpha p(1 - p)^{\alpha-1}}{(1 - p)^\alpha + \alpha p(1 - p)^{\alpha-1}}$$

$$= \alpha p\left[\frac{(1 - p)^{\alpha-1}}{(1 - p)^{\alpha-1}(1 - p) + \alpha p(1 - p)^{\alpha-1}}\right]$$

$$= \alpha p\left[\frac{(1 - p)^{\alpha-1}}{(1 - p)^{\alpha-1}(1 - p + \alpha p)}\right]$$

$$= \alpha p\left[\frac{1}{1 + (\alpha - 1)p}\right] > \alpha p \qquad \square$$

The Solution

The correct expected value can be calculated by the usual formula:

$$E(Y) = \sum_{i=0}^{\bar{n}} y_i CPB(y_i|\alpha,p). \tag{4}$$

Using the same method as above, $E(Y)$ can be calculated for successive intervals of α and generalized. For the first three unit regions, the expected value is as follows.

$$E(Y|0 < \alpha \le 1,p) = \alpha p\left[\frac{1}{1 + (\alpha - 1)p}\right]$$

$$E(Y|1 < \alpha \le 2,p) = \alpha p\left[\frac{1 + (\alpha - 2)p}{1 + (\alpha - 2)p + \frac{(\alpha - 2)(\alpha - 1)}{2}p^2}\right]$$

$$E(Y|2 < \alpha \le 3, p) = \alpha p$$

$$\times \left[\frac{1 + (\alpha - 3)p + \dfrac{(\alpha - 3)(\alpha - 2)}{2} p^2}{1 + (\alpha - 3)p + \dfrac{(\alpha - 3)(\alpha - 2)}{2} p^2 + \dfrac{(\alpha - 3)(\alpha - 2)(\alpha - 1)}{6} p^3} \right]$$

The same pattern continues for subsequent intervals of α. Generalizing over all α and p yields

$$E(Y) = \alpha p \left[\frac{\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)}{\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)} \right] \qquad (5)$$

where

$$\Theta^{\kappa}(\gamma, \rho) = \sum_{i=0}^{\kappa} \frac{(\gamma)_i}{i!} \rho^i \qquad (6)$$

and $(\gamma)_i$ is Pochhammer's symbol, given by

$$(\gamma)_i = \gamma(\gamma + 1)(\gamma + 2) \cdots (\gamma + i - 1)$$

$$(\gamma)_0 = 1.$$

Equation 6 may alternatively be written as

$$\Theta^{\kappa}(\gamma, \rho) = \sum_{i=0}^{\kappa} \binom{\gamma + i - 1}{i} \rho^i \qquad (7)$$

where $\binom{n}{k}$ is the *generalized binomial coefficient*, defined for real n and integer k as

$$\binom{n}{k} = n(n - 1) \cdots (n - k + 1)/k!$$

$$\binom{n}{0} = 1.$$

Appendix A gives a more formal proof of this same result.

The difference between the expected value αp assumed in the literature and the correct expression in equation 5 is not large, but it is not zero, and it varies with the two parameters of the CPB. Fortunately, it can be easily shown that the Θ-ratio in equation 5 converges for both α and p. There are a number of interesting characteristics to the error's limits centering around this ratio. First, notice that in the Θ-ratio all but the first term in both the numerator and denominator polynomials contain a p^i term, where $i \ge 1$. Thus,

$$\lim_{p \to 0} \left[\frac{\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)}{\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)} \right] = 1.$$

Similarly, all but the first term in both the numerator and denominator polynomials contain an $(\alpha - \tilde{n})$ term. Therefore, for the intervals $\tilde{n} - 1 < \alpha \le \tilde{n}$,

$$\lim_{\alpha \to \tilde{n}} \left[\frac{\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)}{\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)} \right] = 1.$$

Finally, there is an overall convergence in the Θ-ratio as α increases. Note that the numerator and denominator polynomials of the Θ-ratio are exactly the same, except for the last term in the denominator. As α increases, this additional term decreases in its effect on the ratio, giving

$$\lim_{\alpha \to \infty} \left[\frac{\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)}{\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)} \right] = 1.$$

To get a feel for this convergence we present figure 1. This figure is a three-dimensional plot with values of α and p along the two axes on the floor (with ranges $1 < \alpha \le 10$ and $0 \le p \le 1$, respectively) and with the height of the surface representing the effect of $\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)/\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)$. As the above limits indicate, the surface is on or near the floor (i.e., a ratio of one) for most combinations of α and p. The effect of the Θ-ratio on αp decreases exponentially as p decreases, decreases in each interval $\tilde{n} - 1 < \alpha \le \tilde{n}$ as $\alpha \to \tilde{n}$, and decreases exponentially overall as $\alpha \to \infty$.

To get a feel for the difference between the expected value αp assumed in the literature and the correct expression in equation 5, we present two additional figures. Figure 2 is, again, a three-dimensional plot with values of α and p along the two axes on the floor (with ranges $0 < \alpha \le 5$ and $0 \le p \le 1$, respectively), and with the height of the surface representing the error $\epsilon(\alpha, p) = E(Y) - \alpha p$. As can be seen, the surface is on or near the floor (i.e., zero error) for most combinations of α and p. The error decreases as p decreases, decreases in each interval $\tilde{n} - 1 < \alpha \le \tilde{n}$ as $\alpha \to \tilde{n}$, and decreases overall as $\alpha \to \infty$. The largest errors are for values of p near 1.0 and α near zero.
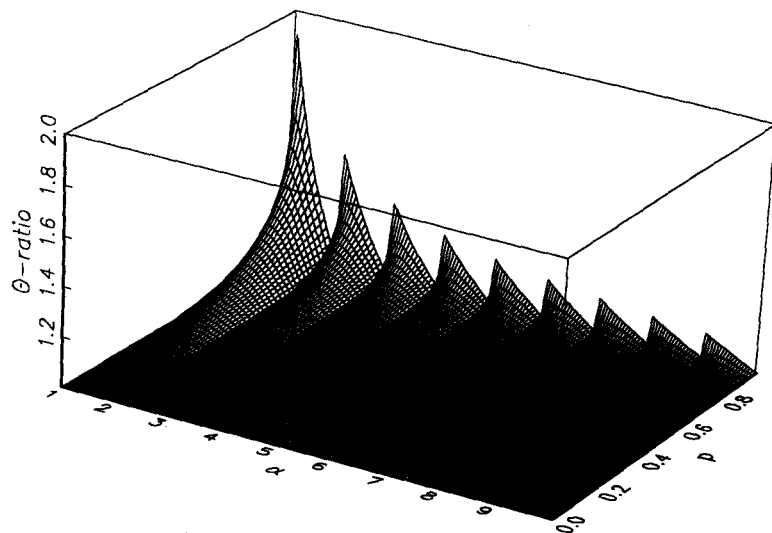
Fig. 1. The effect of $\Theta^{\bar{n}-1}(\alpha - \bar{n},p)/\Theta^{\bar{n}}(\alpha - \bar{n},p)$ for $1 \leq \alpha \leq 10$ and $0 \leq p \leq 1$. The effect of $\Theta^{\bar{n}-1}(\alpha - \bar{n},p)/\Theta^{\bar{n}}(\alpha - \bar{n},p)$ on αp decreases exponentially as p decreases, decreases in each interval $\bar{n} - 1 < \alpha \leq \bar{n}$ as $\alpha \to \bar{n}$, and decreases exponentially overall as $\alpha \to \infty$.

However, the maximum error is never as large as 1.0. The discrete jumps in the graph are due to the fact that, when α is an integer, the CPB reduces to exactly the binomial and the error drops to zero.

While figure 2 shows the absolute error $\epsilon(\alpha,p)$ for a small range of α, it is perhaps more interesting to examine the proportion of this error to the expected value itself. In this way, we get a feel for the seriousness of the error over α and p—for example, $\epsilon(\alpha,p) = 0.9$ makes more of a difference when $E(Y) = 2.1$ than it does when $E(Y) = 20.1$. Figure 3 presents a three-dimensional plot with values of α and p along the two axes on the floor (with ranges $0 < \alpha \leq 10$ and $0 \leq p \leq 1$, respectively), and with the height of the surface representing the proportion of error to expected value, $\epsilon(\alpha,p)/E(Y)$. Again, the surface is on or near the floor (i.e., zero percent) for most of the ranges of α and p. By far the largest effect is seen in the interval $0 < \alpha \leq 1$, when the count takes on values $\{0,1\}$. Outside of that, the error decreases exponentially as p decreases, decreases in each interval $\bar{n} - 1 < \alpha \leq \bar{n}$ as $\alpha \to \bar{n}$, and decreases exponentially overall as $\alpha \to \infty$.

We also provide the correct variance here. Since $V(Y) = E(Y^2) - E(Y)^2$,

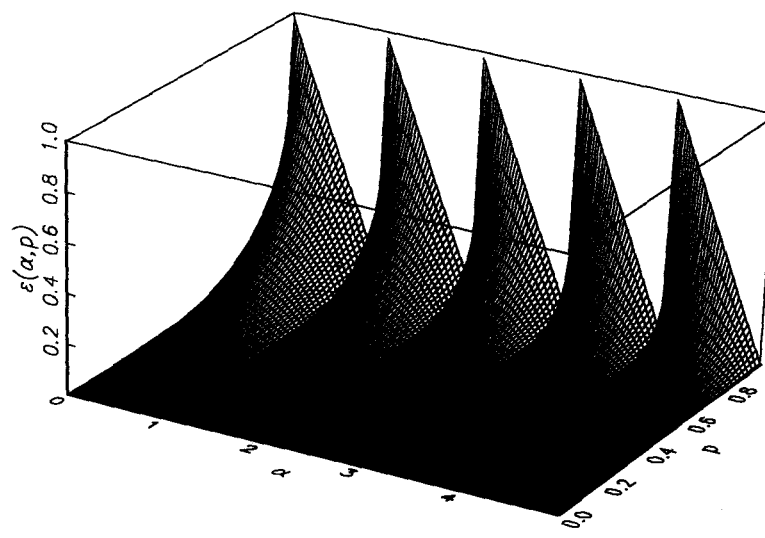**Fig. 2.** Plot of the error $\epsilon(\alpha,p) = E(Y|\alpha,p) - \alpha p$ for $0 \leq \alpha \leq 5$ and $0 \leq p \leq 1$. The error decreases exponentially as p decreases, decreases in each interval $\bar{n} - 1 < \alpha \leq \bar{n}$ as $\alpha \to \bar{n}$, and decreases overall as $\alpha \to \infty$.

and we already have the correct expression for $E(Y)$, we only need $E(Y^2)$, which can be calculated by the usual formula.

$$E(Y^2) = \sum_{i=1}^{\bar{n}} y_i^2 CPB(y_i|\alpha,p). \tag{8}$$

As with $E(Y)$, $E(Y^2)$ may be calculated for successive intervals of α and then generalized. The first three intervals of α are as follows.

$$E(Y^2|0 < \alpha \leq 1) = \alpha p \left[\frac{1}{1 + (\alpha - 1)p} \right]$$

$$E(Y^2|1 < \alpha \leq 2) = \alpha p \left[\frac{1 + (2\alpha - 3)p}{1 + (\alpha - 2)p + \dfrac{(\alpha - 2)(\alpha - 1)}{2}p^2} \right]$$
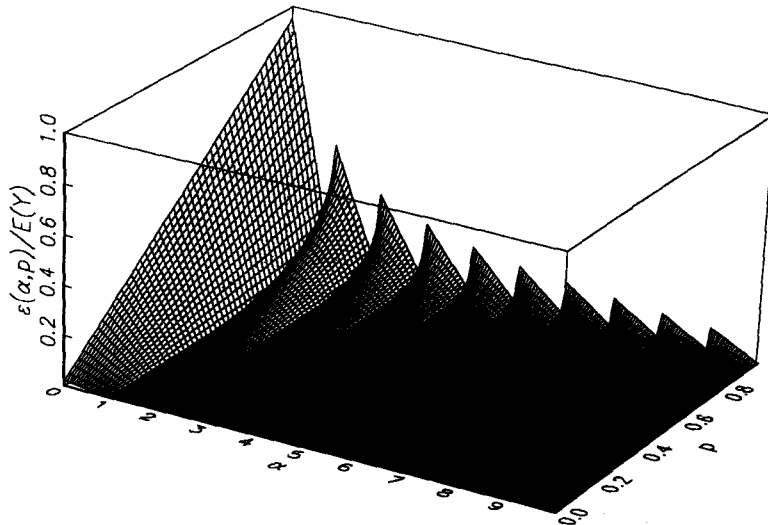
Fig. 3. Ratio of the error $\epsilon(\alpha,p)$ to the expected value $E(Y|\alpha,p)$ for $0 \le \alpha \le 10$ and $0 \le p \le 1$. The ratio of the error $\epsilon(\alpha,p)$ to the expected value $E(Y|\alpha,p)$ is greatest for $0 < \alpha < 1$. Therefore, it (1) decreases exponentially as p decreases, (2) decreases in each interval $\bar{n} - 1 < \alpha \le \bar{n}$ as $\alpha \to \bar{n}$, and (3) decreases exponentially overall as $\alpha \to \infty$.

$$E(Y^2|2 < \alpha \le 3) = \alpha p$$

$$\times \left[\frac{1 + (2\alpha - 4)p + \dfrac{(3\alpha - 4)(\alpha - 3)}{2} p^2}{1 + (\alpha - 3)p + \dfrac{(\alpha - 3)(\alpha - 2)}{2} p^2 + \dfrac{(\alpha - 3)(\alpha - 2)(\alpha - 1)}{6} p^3} \right]$$

Notice that in all cases the denominator is simply $\Theta^{\bar{n}}(\alpha - \bar{n},p)$. The pattern for the numerator becomes clearer after a few more intervals.

$$E(Y^2|3 < \alpha \le 4) = \alpha p$$

$$\times \left[\frac{1 + (2\alpha - 5)p + \dfrac{(3\alpha - 5)(\alpha - 4)}{2} p^2 + \dfrac{(4\alpha - 5)(\alpha - 4)(\alpha - 3)}{6} p^3}{\Theta^{\bar{n}}(\alpha - \bar{n},p)} \right]$$

$$E(Y^2|4 < \alpha \le 5) = \alpha p \left[1 + (2\alpha - 6)p + \frac{1}{2} (3\alpha - 6)(\alpha - 5)p^2 \right.$$

$$+ \frac{1}{6} (4\alpha - 6)(\alpha - 5)(\alpha - 4)p^3$$

$$\left. + \frac{1}{24} (5\alpha - 6)(\alpha - 5)(\alpha - 4)(\alpha - 3)p^4 \right]$$

$$/\Theta^{\bar{n}}(\alpha - \bar{n}, p)$$

Generalizing the above gives

$$E(Y^2) = \alpha p \left[\frac{1 + \Sigma_{i=2}^{\bar{n}} \left(\dfrac{(i\alpha - \bar{n} - 1)(\alpha - \bar{n})_{i-2}}{(i - 1)!} p^{i-1} \right)}{\Theta^{\bar{n}}(\alpha - \bar{n},p)} \right] \qquad (9)$$

where $\Sigma_{i=j}^{k} x_i = 0$ if $j > k$.

Practical Implications

Although the continuous parameter binomial probability distribution has been studied for more than 60 years in the literature on theoretical statistics, it was not until King's (1989a) development of the generalized event count (GEC) model that the distribution was used in empirical research or in political science. The CPB distribution was used by King to model underdispersion in the study of event counts, which are dependent variables measured as the number of times an event occurs. The CPB is also one special case of his GEC distribution. The parameterization given above can be translated into that used by King by letting $\lambda = \alpha p$ and $\sigma^2 = 1 - p$. King also let λ vary over the explanatory variables by setting it to $\exp(x_i\beta)$.

The results given above demonstrate that the expected values and variances are not correct in the CPB portion of the GEC model. These quantities are correct for the cases of Poisson dispersion and overdispersion, which are more common empirically, and the maximum error possible, in the underdispersed special case, is less than one event in the predicted number of events. But despite these theoretical inaccuracies, virtually all practical empirical results based on it will be unaffected by our correction. Our reanalyses of many specific empirical examples have failed to turn up any substantial differences. In fact, the error only approaches its maximum when the count is very small. The worst case occurs when the probability of the count being greater than one is zero (i.e., $0 < \alpha \le 1$). Since, in this situation, empirical re-

searchers would, and indeed should, opt for a logit or probit-type model instead of the GEC, the worst case will not occur.

Despite the small changes in empirical results that would result from a correction, it is still important to correct the record. We have done this above for the mathematical results. In this section, we suggest a strategy for correcting future statistical analyses.

Our strategy is to estimate the GEC, as suggested by King, but to interpret the empirical results in line with our findings above. In particular, the reparameterization $E(Y_i) = \exp(x_i\beta)$ is only appropriate if it turns out that the data support Poisson dispersion or overdispersion. In the underdispersed special case, we recommend still substituting $\exp(x_i\beta)$ into the likelihood function for λ but not interpreting it as the expected value. Instead, we recommend that researchers first calculate λ and σ^2 using the log-likelihood equation given in King (1989a), then calculate α and p from the reparameterization, and then use equation 5 in this article to calculate fitted values, draw graphs, and interpret causal effects. Because of the invariance-to-reparameterization property of maximum likelihood, estimates of the parameters of the GEC model will still be consistent.

To make this procedure easy and transparent for empirical researchers, we have modified a computer program to provide two methods of parameter interpretation, following the general methods suggested in King (1989b, sec. 5.2). The computer program is called COUNT, and a new version, which incorporates these features, is now available.[2]

Each of our two methods of parameter interpretation involves the expected value. The first is calculating fitted values, that is, estimating $E(Y)$, or $\hat{y}$. The program asks the user for values of the explanatory variables and prints $\hat{y}$. For the Poisson and overdispersed cases, these results are based on the formulas given in King (1989a, 1989b), whereas for the CPB special case they are based on the results in this article.

Second, we give first differences. That is, holding constant all explanatory variables at their means, COUNT will compute the difference between $E(Y)$ for two different values of the remaining explanatory variable. This is to be interpreted literally as the effect of that explanatory variable on the dependent variable, controlling for the others.

Over- and under-dispersion has its largest effects on the Poisson regression model by biasing its standard errors; the effect on the coefficients is quite minor. Thus, an alternative strategy to using the GEC is to estimate a Poisson model and use heteroskedasticity-consistent standard errors. This is now also an option in the COUNT program.

---

2. The program COUNT is available free via gopher or anonymous FTP from hdc-gopher.harvard.edu, from the ICPSR, or by contacting the authors.

APPENDIX A

Proposition A. Let the random variable Y have a continuous parameter binomial distribution $CPB(y|\alpha,p)$, where $\alpha > 0$ and $0 \le p \le 1$. Then

$$E(Y) = \alpha p \left[\frac{\Theta^{\bar{n}-1}(\alpha - \bar{n},p)}{\Theta^{\bar{n}}(\alpha - \bar{n}, p)} \right]$$

where

$$\Theta^{\kappa}(\gamma,\rho) = \sum_{i=0}^{\kappa} \frac{(\gamma)_i}{i!} \rho^i$$

and $(\gamma)_i$ is Pochhammer's symbol, given by

$$(\gamma)_i = \gamma(\gamma + 1)(\gamma + 2) \cdots (\gamma + i - 1)$$

$$(\gamma)_0 = 1.$$

Proof.

$$E(Y) = \sum_{i=0}^{\bar{n}} y_i CPB(y_i|\alpha,p)$$

$$= \sum_{i=0}^{\bar{n}} i\bar{p}_i$$

$$= \frac{\sum_{i=1}^{\bar{n}} ip_i}{\sum_{i=0}^{\bar{n}} p_i}$$

$$= \frac{\sum_{i=1}^{\bar{n}} i \binom{\alpha}{i} p^i (1 - p)^{\alpha-i}}{\sum_{i=0}^{\bar{n}} \binom{\alpha}{i} p^i (1 - p)^{\alpha-i}}$$

where $\binom{n}{k}$ is the generalized binomial coefficient defined for real n and integer k as $\binom{n}{k} = n(n - 1) \cdots (n - k + 1)/k!$ and $\binom{n}{0} = 1$. Since the binomial coefficient may be rewritten in terms of Pochhammer's symbol, $\binom{n}{k} = (n - k + 1)_k/k!$, we get

$$E(Y) = \frac{\sum_{i=1}^{\bar{n}} i \frac{(\alpha - i + 1)_i}{i!} p^i (1 - p)^{\alpha-i}}{\sum_{i=0}^{\bar{n}} \frac{(\alpha - i + 1)_i}{i!} p^i (1 - p)^{\alpha-i}}$$

$$= \alpha p \left[\frac{\sum_{i=1}^{\tilde{n}} \frac{(\alpha - i + 1)_{i-1}}{(i-1)!} p^{i-1}(1-p)^{\alpha-i}}{\sum_{i=0}^{\tilde{n}} \frac{(\alpha - i + 1)_i}{i!} p^i(1-p)^{\alpha-i}} \right]$$

$$= \alpha p \left[\frac{\sum_{i=0}^{\tilde{n}-1} \frac{(\alpha - i)_i}{i!} p^i(1-p)^{\alpha-i-1}}{\sum_{i=0}^{\tilde{n}} \frac{(\alpha - i + 1)_i}{i!} p^i(1-p)^{\alpha-i}} \right]$$

Multiplying the numerator and denominator by $(1 - p)^{\tilde{n}-\alpha}$ and using the relationship $(1 - p)^n = \sum_{m=0}^{n} \binom{n}{m} (-p)^m = \sum_{m=0}^{n} [(n - m + 1)_m / m!] (-p)^m$ yields

$$E(Y) = \alpha p \left[\frac{\sum_{i=0}^{\tilde{n}-1} \left\{ \frac{(\alpha - i)_i}{i!} p^i \sum_{m=0}^{\tilde{n}-i-1} \frac{(\tilde{n} - i - m)_m}{m!} (-p)^m \right\}}{\sum_{i=0}^{\tilde{n}} \left\{ \frac{(\alpha - i + 1)_i}{i!} p^i \sum_{m=0}^{\tilde{n}-i} \frac{(\tilde{n} - i - m + 1)_m}{m!} (-p)^m \right\}} \right]$$

$$= \alpha p \left[\frac{\sum_{i=0}^{\tilde{n}-1} \sum_{m=0}^{\tilde{n}-i-1} \frac{(\alpha - i)_i}{i!} \frac{(\tilde{n} - i - m)_m}{m!} (-1)^m p^{m+i}}{\sum_{i=0}^{\tilde{n}} \sum_{m=0}^{\tilde{n}-i} \frac{(\alpha - i + 1)_i}{i!} \frac{(\tilde{n} - i - m + 1)_m}{m!} (-1)^m p^{m+i}} \right],$$

which can be respecified, using the relationship $(-1)^k (-n)_k / k! = (n - k + 1)_k / k!$, as

$$E(Y) = \alpha p \left[\frac{\sum_{i=0}^{\tilde{n}-1} \sum_{m=0}^{\tilde{n}-i-1} \frac{(\alpha - i)_i}{i!} \frac{(i - \tilde{n} + 1)_m}{m!} p^{m+i}}{\sum_{i=0}^{\tilde{n}} \sum_{m=0}^{\tilde{n}-i} \frac{(\alpha - i + 1)_i}{i!} \frac{(i - \tilde{n})_m}{m!} p^{m+i}} \right].$$

Because the numerator and denominator are of the same functional form, this may be written as

$$E(Y) = \alpha p \left[\frac{\Theta^{\tilde{n}-1}(\alpha - \tilde{n}, p)}{\Theta^{\tilde{n}}(\alpha - \tilde{n}, p)} \right]$$

where

$$\Theta^{\kappa}(\gamma, \rho) = \sum_{i=0}^{\kappa} \sum_{m=0}^{\kappa-i} \frac{(\gamma - i + \kappa + 1)_i}{i!} \frac{(i - \kappa)_m}{m!} \rho^{m+i}$$

$$= \sum_{i=0}^{\kappa} \left[\sum_{m=0}^{i} \frac{(\gamma - m + \kappa + 1)_m}{m!} \frac{(m - \kappa)_{i-m}}{(i - m)!} \right] \rho^i$$

$$= \sum_{i=0}^{\kappa} \left[\sum_{m=0}^{i} \binom{\gamma + \kappa}{m} \binom{i - \kappa - 1}{i - m} \right] \rho^i$$

The bracketed term is now a summation over m of the product of two binomial coefficients. Examining the indices of the binomial coefficients, we see that the upper indices are independent of m and that the sum of the lower indices is constant for all m. Thus, the bracketed term is Vandermonde's convolution[3] and we may rewrite it as the binomial coefficient obtained by summing both the upper and lower indices, yielding

$$\Theta^{\kappa}(\gamma, \rho) = \sum_{i=0}^{\kappa} \binom{\gamma + i - 1}{i} \rho^i$$

$$= \sum_{i=0}^{\kappa} \frac{(\gamma)_i}{i!} \rho^i \qquad \square$$

REFERENCES

Abramowitz, M., and I. Stegun, eds. 1965. *Handbook of Mathematical Functions*. New York: Dover.

Guldberg, Alf. 1931. "On Discontinuous Frequency-Functions and Statistical Series." *Skandinavisk Aktvarietionskrift* 14:167–97.

Graham, Ronald L., Donald E. Knuth, and Oren Patashnik. 1989. *Concrete Mathematics: A Foundation for Computer Science*. Reading: Addison-Wesley.

Johnson, Norman L., and Samuel Kotz. 1969. *Distributions in Statistics: Discrete Distributions*. New York: Wiley.

Katz, Leo. 1965. "Unified Treatment of a Broad Class of Discrete Probability Distributions." In Ganapati P. Patil, ed., *Classical and Contagious Discrete Distributions*. Calcutta: Statistical Publishing Society.

Katz, Leo. 1945. "Characteristics of Frequency Functions Defined by First Order Difference Equations." Ph.D. diss. University of Michigan.

King, Gary. 1989a. "Variance Specification in Event Count Models: From Restrictive Assumptions to a Generalized Estimator." *American Journal of Political Science* 33:762–784.

King, Gary. 1989b. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. New York: Cambridge University Press.

Krause, George A. 1994. "Federal Reserve Policy Decision Making: Political and Bureaucratic Influences." *American Journal of Political Science* 38:124–44.

Wang, T. Y., William J. Dixon, Edward N. Muller, and Mitchell A. Seligson. 1993.

---

3. See Graham, Knuth, and Patashnik (1989, 169) for more information on Vandermonde's convolution and its variants.

"Inequality and Political Violence Revisited." *American Political Science Review* 87:979–93.

Winkelmann, Rainer, and Klaus F. Zimmermann. 1991. "A New Approach for Modeling Economic Count Data." *Economics Letters* 37:139–43.