NON-EXISTENCE OF CERTAIN POLYNOMIAL REGRESSIONS IN RANDOM SUM THEORY

NORMAN L. JOHNSON
University of North Carolina
Chapel Hill, NC

SAMUEL KOTZ
University of Maryland
College Park, MD

Abstract

It is shown that the regression of a counting variable (X) on the value of a random sum of X independent and identically distributed, non-negative integer-valued random variables cannot, in general, be a polynomial of degree greater than 1.

[Key Words: Bayes' Theorem, Counting Variables, Polynomial Regression, Random Sums]

Kyriakoussis (1988) considered recently the following model. Let X, Z_1, Z_2,... be independent random variables taking nonnegative integer values, and

$$Y = \begin{cases} Z_1+Z_2+\dots+Z_X & \text{if } X > 0 \\ 0 & \text{if } X = 0 \end{cases}$$

The Z's are identically distributed, with $\Pr[Z_j=0] = p_0 \neq 1$, and the regression of X and Y is a polynomial of degree r,

$$E[X|Y=y] = \sum_{i=0}^{r} \beta_i y^i, \quad \text{with } \beta_r \neq 0. \tag{1}$$

(Kyriakoussis shows that if $E[Y^{r-1}]$ is finite, then the joint distribution of (X,Y) is uniquely determined by the common distribution of the Z's, the regression function (1), and the first (r-1) moments of Y.)

The main purpose of this paper is to show that, under the stated conditions, the value of r in (1) cannot exceed 1 - that is, polynomial regression, other than linear, is impossible - unless Y is bounded. In what follows we shall assume that Y is unbounded.

We first note that we must have $\beta_r > 0$ because if β_r is negative, then $E[X|Y=y]$ would be negative for sufficiently large y. Also, the condition that Y is unbounded implies that X is unbounded, too. This is because, if $\Pr[X < \xi] = 1$ for some finite ξ,

$$E[X|Y=y] < \xi$$

and (1) could not be satisfied for sufficiently large y. We also note that if

$p_o=0$, (so $Z_j \geq 1$, a.s.) then $X \leq Y$ a.s. and so

$$E[X|Y=y] \leq y$$

which again implies that (1) could not be satisfied for sufficiently large y. So without loss of generality, we take $0 < p_o < 1$.

Now write

$$X = X_+ + X_0$$

where

$$X_+ = \text{number of Z's which are nonzero}$$

and $$X_0 = \text{number of Z's which are zero.}$$

Since the minimum possible value for each of the X_+ Z's is 1, it follows (as above) that

$$X_+ \leq y.$$

Using Bayes' theorem

$$\Pr[X=x|X_+=x_+] = \frac{P_x \binom{x}{x_+} p_0^{x-x_+}(1-p_0)^{x_+}}{\sum_{w=x_+}^{\infty} P_w \binom{w}{x_+} p_0^{w-x_+}(1-p_0)^{x_+}}$$

$$= \frac{P_x x^{(x_+)} p_0^{x}}{\sum_{w=0}^{\infty} P_w w^{(x_+)} p_0^{w}} ,$$

where $P_x = \Pr[X=x]$ and $a^{(b)} = a(a-1)\ldots(a-b+1)$, as usual. (Note that $w^{(x_+)} = 0$ for $w < x_+$.)

Hence

$$E[X|X_+=x_+] = \frac{\sum_{x=0}^{\infty} xP_x x^{(x_+)} p_0^{x}}{\sum_{x=0}^{\infty} P_x x^{(x_+)} p_0^{x}} = g(x_+), \text{ say.}$$

We now show that (as is to be expected) $g(x_+)$ is a nondecreasing function of x_+. Using the relation $x^{(x_+ +1)} = (x-x_+)x^{(x_+)}$ we have

$$g(x_+ +1) - g(x_+) = \frac{\Sigma_{x=0}^{\infty} x(x-x_+)\alpha_x}{\Sigma_{n=0}^{\infty} (x-x_+)\alpha_x} - \frac{\Sigma_{x=0}^{\infty} x\,\alpha_x}{\Sigma_{x=0}^{\infty} \alpha_x}$$

where $\alpha_x = P_x\, x^{(x_+)}\, p_0^x$.

Continuing,

$$g(x_+ +1)-g(x_+) = \frac{\left[\Sigma_{x=0}^{\infty} \alpha_x\right]\left\{\Sigma_{x=0}^{\infty} x^2\alpha_x - \left[\Sigma_{x=0}^{\infty} x\,\alpha_x\right]^2\right\}}{\left\{\Sigma_{x=0}^{\infty} (x-x_+)\alpha_x\right\}\left[\Sigma_{x=0}^{\infty} \alpha_x\right]}$$

$$= \frac{\Sigma_{x=0}^{\infty} \alpha_x(x-\bar{x})^2}{\Sigma_{x=0}^{\infty} (x-x_+)\alpha_x} \geq 0$$

where $\bar{x} = \left[\Sigma_{x=0}^{\infty} x\,\alpha_x\right]/\left[\Sigma_{x=0}^{\infty} \alpha_x\right]$. (Note that $\alpha_x = 0$ for $x < x_+$.)

So $g(x_+)$ is a nondecreasing function of the integer-valued variable x_+.

Since $E[X|X = x_+] = g(x_+)$ and the maximum possible value of X_+, given $Y=y$, is y, we have

$$E[X|Y=y] \leq g(y) = \frac{\Sigma_{x=0}^{\infty} x\, P_x\, x^{(y)} p_0^x}{\Sigma_{x=0}^{\infty} P_x\, x^{(y)}\, p_0^x} .$$

Since

$$x\, x^{(y)} = (y + \overline{x-y})x^{(y)} = y\, x^{(y)} + x^{(y+1)},$$

$$g(y) = y + \frac{\Sigma_{x=0}^{\infty} P_x \, x^{(y+1)} \, p_0^x}{\Sigma_{x=0}^{\infty} P_x \, x^{(y)} \, p_0^x} .$$

So

$$E[X|Y=y] \leq y + H(y) \tag{2}$$

where

$$H(y) = \frac{\Sigma_{x=0}^{\infty} P_x \, x^{(y+1)} p_0^x}{\Sigma_{x=0}^{\infty} P_x \, x^{(y)} p_0^x}$$

$$= \frac{D_t^{y+1} h(t)|_{t=1}}{D_t^y \; h(t)|_{t=1}}$$

with $D_t \equiv \frac{d}{dt}$ and $h(t) = \Sigma_{x=0}^{\infty} P_x(p_0 \; t)^x$.

Note that $h(t)$ and $D_t^s \; h(t)$ converge for $0 \leq p_0 \; t < 1$ (i.e. $0 \leq t < p_0^{-1}$) because $\Sigma_{x=0}^{\infty} P_x(=1)$ converges, and $x^{(s)} \; (p_0 \; t)^x \to 0$ as $x \to \infty$.

Hence the Taylor expansion of $h(t)$ about $t = 1$

$$h(t) = \sum_{y=0}^{\infty} \frac{(t-1)^y}{y!} D_t^y \; h(t)|_{t=1}$$

converges for $1 \leq t < p_0^{-1}$. The ratio of the (y+1)-term to the y-term in the series is

$$\frac{t-1}{y+1} H(y) .$$

By Cauchy convergence criterion the series would not converge if $H(y)/(y+1)$ tends to a limit exceeding 1 as $y \to \infty$. This would certainly be the case of $H(y)$ were of order y^2 - in fact the ratio would tend to infinity as $y \to \infty$. The same argument shows that we cannot have $H(y) = O(y^{1+\epsilon})$ for any fixed $\epsilon > 0$, while if (1) were valid we would have $E[X|Y=y] = O(y^r)$.

So (1) can only be true for r=1, (unless Y is bounded). From (2) we see that $E[X|Y=y] \leq o(y^{1+\epsilon})$.

Remarks: 1) Similar arguments can be applied when the common distribution of the Z's is not restricted to the nonnegative integers, but only satisfies the conditions

(i) Z is nonnegative.

(ii) Z can take the value zero, but no other values in an interval $[0,\zeta)$ from some $\zeta > 0$.

In this case we have $X_+ \leq [y/\zeta] + 1$, where [ ] denotes "integer part of"; this value replaces y in the subsequent analysis.

Moreover, the distribution of Z may be continuous for $Z > \zeta$.

Also, the Z_i's need not even have a common distribution, provided each satisfies (i) and (ii) above, the ζ's have a lower bound greater than zero, and there is a common value (p_0) for $Pr[Z_i=0]$. Subject to these limitations, also, the distribution of each Z_i may depend on the value of X (though the Z_i's must still be independent of one another).

2) It is, of course, possible for the regression function $E[X|Y=y]$ to be curvilinear. (Xekalaki (1980) gives examples.) Indeed, linearity of regression is a basis for a number of characterizations, just because it is not in general a feature of the joint distribution of X and Y. The seminal paper of Cacoullos and Papageorgiu (1983) contains much useful information on these matters. (It is straightforward, though tedious, to calculate the value of H(y) for the examples in that paper, which involve binomial, Pascal and Poisson distributions.)

However, as we have shown, curvilinearity of the regression cannot be expressed as a polynomial function unless, possibly, if Y is bounded.

3) If the Z's have a common Bernoulli distribution (so that Y, given X=x, has a binomial distribution with parameters $(x, 1-p_0)$) then the only nonzero value of Z_i is 1, and $X_+ = y$ if Y=y. In this case, inequality (2) becomes an equality, from which the regression function $E[X|Y=y]$ can be found. (This result does not depend on assumption that X is unbounded.)

For example, if X has a binomial (n,p) distribution, straightforward calculation shows that $H(y) = \frac{n-y}{1-p+pp_o}$, whence

$$E[X|y] = \frac{(1-p)y + npp_o}{1-p + pp_o}$$

a result which is implicit in Cacoullos and Papageorgiu (1983).

REFERENCES

Cacoullos, T. and Papageorgiu, H. (1983) Characterization of discrete distributions by a conditional distribution and a regression function, Ann. Inst. Statist. Math. (Tokyo), **35A**, 95-103.

Kyriakoussis, A. (1988) Characterization of bivariate distributions, Sankhyā, **50A**, 286-287.

Xekalaki, E. (1980) On characterizing the bivariate Poisson, binomial and negative binomial distributions, Colloq. Math. Soc. Janos Bolyai (1979), **21**, 369-379.