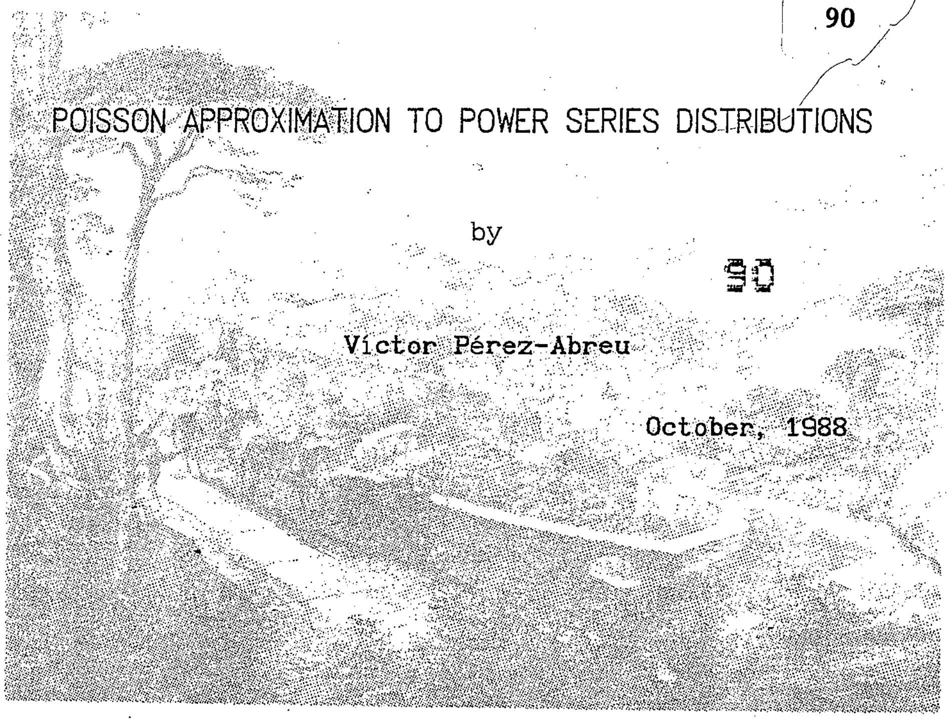
COMUNICACIONES DEL CIMAT

90



POISSON APPROXIMATION TO POWER SERIES DISTRIBUTIONS

by

90

Víctor Pérez-Abreu

October, 1988

ABSTRACT

The purpose of this article is to provide a simple proof for the Poisson approximation to power series distributions. As special cases we obtain the well known Poisson approximations to the binomial and negative binomial distributions. The proof is based on the continuity theorem for probability generating functions presented in Feller (1968). The result illustrates one justification to use Poisson models as suitable descriptions of phenomena which are cumulative effect of a large number of improbable events.

CENTRO DE INVESTIGACION EN MATEMATICAS

Apartado Postal 402
Guanajuato, Gto.
México
Tels. (473) 2-25-50
2-02-58

1. INTRODUCTION

Most introductory probability theory textbooks (see e.g. Hoel, Port and Stone 1971; Feller 1968) present the Poisson approximation to the binomial distribution. That is, if we perform n Bernoulli trials with success probability p_n at each trial and S_n is the number of sucesses in the n trials, then if $np_n \to \lambda$ as $n \to \infty$

$$\lim_{n\to\infty} P(S_n=k) = \lim_{n\to\infty} \binom{n}{k} (p_n)^k (1-p_n)^{n-k} = \exp(-\lambda)\lambda^k/k! \ . \quad (1.1)$$

Apart from the important theoretical consequences of this result, it gives a justification for the use of Poisson models for phenomena which are the cumulative effect of many improvable events.

The purpose of this article is to present a simple proof of the Poisson approximation to the distribution of the sum of random variables having power series distributions. This fact emphasizes the importance of using the Poisson distribution as a model for phenomena of the above described type. The proof of the result is based on the continuity theorem for probability generating functions presented in Feller (1968): "A sequence of discrete probability distributions converges to a limiting discrete distribution if and only if the corresponding probability generating functions converge". This theorem has had limited applicability, since the most interesting limiting forms of discrete distributions are continuous and general results of approximation type are obtained from the theory of infinite divisible distributions. However, it is also the purpose of this work to illustrate one important application of this theorem whose proof can be given in a calculus based probability course (see Feller 1968, p. 281). In Section 2 we present the main result and its proof. Section 3 contains the Poisson approximation to the binomial and negative binomial distributions as special cases of the main result. In Section 4 the result is extended to the sum of nonidentically distributed random variables with power series distributions.

2. POISSON APPROXIMATION

A discrete probability distribution $\{p_k\}_{k\geq 0}$, that is

$$0 \leq p_k \leq 1 \ , \quad \sum_{k=0}^{\infty} p_k = 1$$

is called a *power series distribution* (see Johnson and Kotz 1969, p.33) if each p_k can be written in the form

$$p_k = a_k\theta^k/g(\theta) \quad k = 0,1,\ldots; \ \theta > 0 \qquad (2.1)$$

where $a_k \geq 0$ and $g(\theta) = \sum_{k=0}^{\infty} a_k\theta^k$. The Bernoulli, binomial, geometric, negative binomial, Poisson and logarithmic series are among the probability distributions belonging to this class. The probability generating funcion $\phi(t) = \sum_k t^k p_k$ of the distribution (2.1) is

$$\phi(t) = g(t\theta)/g(\theta) \quad 0 \leq t \leq 1 \qquad (2.2)$$

and its mean is given by

$$\mu = \theta g^{(1)}(\theta)/g(\theta) \ . \qquad (2.3)$$

The following result is an extension of the Poisson approximation to the binomial distribution to the distribution of the sum of random variables having power series distributions.

Theorem 1. For each $n \geq 1$ let $X_1, \ldots, X_n$ be independent nonnegative integer valued random variables with common power series distribution

$$p_{k,n} = a_k \theta_n^k / g(\theta_n) \quad k = 0,1,\ldots \tag{2.4}$$

where $a_k \geq 0 \quad k = 0,1,\ldots$ are independent of n and

$$g(\theta_n) = \sum_{k=0}^{\infty} a_k \theta_n^k \quad \theta_n > 0 . \tag{2.5}$$

Let $a_0 > 0$, $\lambda > 0$ be fixed and $S_n = \sum_{i=1}^{n} X_i$. If $n\theta_n \to \lambda$ as $n \to \infty$, then

$$\lim_{n\to\infty} P(S_n=k) = \exp(-\lambda_0)\lambda_0^k/k! \quad k = 0,1,\ldots \tag{2.6}$$

where $\lambda_0 = \lambda a_1/a_0$.

The proof of the above theorem is based on the following result which is the continuity theorem for probability generating functions (see Feller 1968, p. 280).

Theorem 2. Suppose that for every fixed $n \geq 1$ the sequence $\{a_{k,n}\}_{k\geq 0}$ is a discrete probability distribution. In order that a limit

$$a_k = \lim_{n\to\infty} a_{k,n} \tag{2.7}$$

exists for every $k = 0,1,\ldots$ it is necessary and sufficient that the limit

$$A(t) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{k,n} t^k \tag{2.8}$$

exists for each t in the open interval $0 < t < 1$. In this case automatically

$$A(t) = \sum_{k=0}^{\infty} a_k t^k.$$

It is important to point out that the a_k's may not sum to one in every case, so $\{a_k\}$ may not yield a probability distribution and therefore $A(t)$ may not be a probability generating function.

Proof of Theorem 1.

Let $\phi_n(t)$ be the probability generating function of the distribution (2.4). Then using (2.2) we have that

$$\phi_{S_n}(t) = [\phi_n(t)]^n = [g(t\theta_n)/g(\theta_n)]^n \tag{2.9}$$

is the probability generating function of S_n. Now using the expansion (2.5) of $g(\theta_n)$ we have that for $t \in (0,1)$

$$\begin{aligned} [g(t\theta_n)/g(\theta_n)]^n &= [1-(g(\theta_n)-g(t\theta_n))/g(\theta_n)]^n \\ &= [1-a_1\theta_n(1-t)/g(\theta_n)-r(\theta_n)]^n \\ &= [1+a_1(1-t)n\theta_n/(ng(\theta_n))-r(\theta_n)]^n \end{aligned} \tag{2.10}$$

where $r(\theta_n) = (\sum_{k=2}^{\infty} a_k(\theta_n^k - t^k\theta_n^k))/g(\theta_n)$. By assumption we have that as $n \to \infty$ $n\theta_n \to \lambda$ and therefore $\theta_n \to 0$. Thus

$$\lim_{n\to\infty} g(\theta_n) = a_0 \tag{2.11}$$

$$\lim_{n\to\infty} (n\theta_n)/g(\theta_n) = \lambda/a_0 \tag{2.12}$$

and since for $k \geq 2$ $\lim_{n\to\infty} n\theta_n^k = 0$, then

$$\lim_{n\to\infty} nr(\theta_n) = 0 . \tag{2.13}$$

It is known that if $x_n \to x$ and $ny_n \to 0$ as $n \to \infty$ then

$$\lim_{n\to\infty} (1+x_n/n+y_n)^n = \exp(x) \qquad (2.14)$$

Hence using this result and (2.11)-(2.13) from (2.10) we have that for $t \in (0,1)$

$$\lim_{n\to\infty}[g(t\theta_n)/g(\theta)]^n = \exp(\lambda a_1/a_0(t-1)) \ . \qquad (2.15)$$

But the right hand side of (2.15) is the probability generating function of a Poisson distribution with parameter $\lambda_0 = a_1\lambda/a_0$. Then by the uniqueness of the probability generating function (see Hoel, Port and Stone 1971, p.74) and Theorem 2 we obtain (2.6).

Remark. An interpretation of Theorem 1 is the following: It follows from (2.3) that $\mu_n \to 0$, since $g(\theta_n) \to a_0$ and $g^{(1)}(\theta_n) \to 0$. Then when n is large S_n is the sum of random variables taking the value zero with very high probability ($p_{0,n} \to a_0/a_0=1$), hence the Poisson distribution approximates the cumulative effect of a large number of differing improbable events.

3. EXAMPLES

In this section we obtain the Poisson approximation to the binomial and negative binomial distributions. These results are obtained as easy consequences of Theorem 1.

Theorem 3. (Binomial distribution). For each $n \geq 1$ let $X_1,\dots,X_n$ be a random sample of a Bernoulli distribution with success probability p_n. Let $S_n = \sum_{i=1}^{n} X_i$ and assume that $np_n \to \lambda>0$ as $n \to \infty$. Then for each $k = 0,1,\dots$

$$\lim_{n\to\infty} P(S_n=k) = \lim_{n\to\infty} \binom{n}{k} p_n^{\,k}(1-p_n)^{n-k} = \exp(-\lambda)\lambda^k/k! \qquad (3.1)$$

Proof. The Bernoulli distribution with parameter p_n is a power series distribution with $\theta_n = p_n/(1-p_n)$ and $g(\theta_n) = 1+\theta_n$. Therefore $a_0 = a_1 = 1$, $n\theta_n \to \lambda$ as $n \to \infty$ and the result follows applying Theorem 1.

Theorem 4. (Negative binomial distribution). For each $n \geq 1$ let $X_1,\dots,X_n$ be a random sample of a geometric distribution with success probability p_n. Let $S_n = \sum_{i=1}^{n} X_i$ and assume that $(1-p_n)n \to \lambda > 0$ as $n \to \infty$. Then for each $k = 0,1,\dots$

$$\lim_{n\to\infty} P(S_n=k) = \lim_{n\to\infty} p_n^n \binom{-n}{k} (-1)^k (1-p_n)^{n-k} = \exp(-\lambda)\lambda^k/k! \quad (3.2)$$

Proof. The geometric distribution with parameter p_n is a power series distribution with $\theta_n=1-p_n$ and $g(\theta_n)=(1-\theta_n)^{-1}=\sum_{k=0}^{\infty}(-1)^{2k}\theta_n^k$. Therefore $a_0 = a_1 = 1$ and $n\theta_n \to \lambda$ as $n \to \infty$. The result then follows applying Theorem 1.

4. EXTENSION TO VARIABLE DISTRIBUTIONS

The following result is an extension of the Poisson approximation to the distribution of the sum of n Bernoulli trials with variable success probabilities.

Theorem 5. Consider the power series

$$g(\theta) = \sum_{k=0}^{\infty} a_k\theta^k \quad a_k \geq 0 \ , \ k = 0,1,\dots \ , \ \theta > 0 \ . \qquad (4.1)$$

For each $n \geq 1$ let $X_{i,n}$ $i = 1,\dots,n$ be independent random variables such that $X_{i,n}$ has a power series distribution with corresponding $g(\theta_{i,n})$ for $\theta_{i,n} > 0$. Let $S_n = \sum_{i=1}^{n} X_{i,n}$ and assume that $\theta_n^* = \max_{i=1,\dots,n} (\theta_{i,n}) \to 0$ and

$$\sum_{i=1}^{n} (\theta_{i,n})/g(\theta_{i,n}) \to \lambda \quad \text{as} \quad n \to \infty \tag{4.2}$$

for some $\lambda > 0$. Then if $\lambda_0 = a_1\lambda$, for each $k = 0,1,\dots$ we have that

$$\lim_{n\to\infty} P(S_n=k) = \exp(-\lambda_0)\lambda_0^k/k! \tag{4.3}$$

Proof. Let $\phi_{i,n}(t)$, $t \in [0,1]$ be the probability generating function of $X_{i,n}$ $i = 1,\dots,n$, $n = 1,2,\dots$. Then using (2.2) we have that

$$\phi_{S_n}(t) = g(t\theta_{1,n})\dots g(t\theta_{n,n})/(g(\theta_{1,n})\dots g(\theta_{n,n})) \quad t \in [0,1] \tag{4.4}$$

is the probability generating function of S_n. Hence for $t \in (0,1)$

$$\log \phi_{S_n}(t) = \sum_{i=1}^{n} \log(g(t\theta_{i,n})/g(\theta_{i,n})). \tag{4.5}$$

Since $\theta_n^* \to 0$ as $n \to \infty$ we have that

$$g(t\theta_{i,n})/g(\theta_{i,n}) = 1+(g(t\theta_{i,n})-g(\theta_{i,n})/g(\theta_{i,n})) \xrightarrow[n\to\infty]{} 1 \tag{4.6}$$

Then using the fact that $\log(1+x) \simeq x$, as $x \to 0$, it follows that for n large

$$\log \phi_{S_n}(t) \simeq \sum_{i=1}^{n} (g(t\theta_{i,n})-g(\theta_{i,n}))/g(\theta_{i,n})$$

$$\simeq (t-1)a_1 \sum_{i=1}^{n} \theta_{i,n}/g(\theta_{i,n}) + \sum_{i=1}^{n} \eta(\theta_{i,n})\theta_{i,n}/g(\theta_{i,n}) \tag{4.7}$$

where

$$\eta(\theta_{i,n}) = \sum_{k=2}^{n} a_k(\theta_{i,n}^{k-1}-(t\theta_{i,n})^{k-1})$$

$$\le \sum_{k=2}^{n} a_k((\theta_n^*)^{k-1})(1-t^k) \to 0 \text{ as } n \to \infty . \tag{4.8}$$

Then from (4.2), (4.7) and (4.8) we have that for $t \in (0,1)$

$$\log \phi_{S_n}(t) \to (t-1)\lambda_0 \quad \text{as } n \to \infty$$

and therefore (4.3) follows using Theorem 2 and identifying $\exp(\lambda_0(t-1))$ as the probability generating function of the Poisson distribution.

As an application of the above theorem, for each $n \geq 1$ consider n independent Bernoulli trials such that the ith trial has success probability $p_{i,n}$ $i=1,\dots,n$. Assume that $\max_{1\leq i\leq n}(p_{i,n})\to 0$ and $p_{1,n}+\dots+p_{n,n} \to \lambda > 0$ as $n \to \infty$ and let $S_n = \sum_{i=1}^{n} X_i$. Then the limit (4.3) holds. This fact follows by Theorem 5 since the Bernoulli distribution with parameter p is a power series distribution with $g(\theta) = 1+\theta$, $\theta = p/(1-p)$ and therefore

$$\sum_{i=1}^{n}(\theta_{i,n})/g(\theta_{i,n}) = \sum_{i=1}^{n} p_{i,n} \to \lambda \quad \text{as } n \to \infty .$$

This is the Poisson approximation to the distribution of the sum of n Bernoulli trials with unequal success probabilities (see Feller 1968, p. 282).

REFERENCES

Feller, W. (1968). *An Introduction to Probability Theory and its Applications* (Vol. I, 3rd Ed.), New York: John Wiley.

Hoel, P. G., Port, S. C. and Stone, C. J. (1971). *Introduction to Probability Theory*, Boston: Houghton Mifflin.

Johnson, N. L. and Kotz, S. (1969). *Discrete Distributions*, Boston: Houghton Mifflin.