

M. MAJSNEROWSKA (Wrocław)

A NOTE ON POISSON APPROXIMATION BY
 w -FUNCTIONS

Abstract. One more method of Poisson approximation is presented and illustrated with examples concerning binomial, negative binomial and hypergeometric distributions.

1. Introduction. Let X be a non-negative integer-valued random variable with distribution \mathcal{F} and let \mathcal{P}_λ denote the Poisson distribution with mean λ . It is well known that the distributions of some types of X 's can be approximated by \mathcal{P}_λ provided natural conditions concerning their parameters are satisfied. To measure how close the distributions of X and \mathcal{P}_λ are, the total variation distance is usually applied. It is defined by

$$(1) \quad d(\mathcal{F}, \mathcal{P}_\lambda) = \sup_A |\mathcal{F}(A) - \mathcal{P}_\lambda(A)|,$$

where A runs over subsets of non-negative integers.

The investigation of Poisson convergence and approximation has a long history and enormous bibliography. Numerous methods have been developed or adapted to deal with the problem. The aim of this paper motivated by Cacoullos *et al.* (1994) studying the normal case is to present still another way of considering Poisson approximation which allows checking, in a unified and simple manner, Poisson convergence in the class of discrete distributions as well as provides upper bounds for the total variation distance.

The main tools are the so-called w -functions which we introduce and describe in Section 2. In Section 3 we use the w -functions to formulate and prove a Poisson characterization and a convergence theorem. An upper bound for the total variation distance in terms of the w -functions is given in Section 4. To obtain it we apply the Stein–Chen identity (see Barbour

1991 *Mathematics Subject Classification*: Primary 60F05.

Key words and phrases: total variation distance, w -functions, Stein–Chen identity, Poisson, binomial, negative binomial, hypergeometric distributions.

et al. (1992), pp. 5ff) according to which for every positive constant λ , every subset A of non-negative integers and some function $h = h_{\lambda,A}$,

$$(2) \quad \mathcal{F}(A) - \mathcal{P}_\lambda(A) = \mathbb{E}[\lambda h(X+1) - Xh(X)].$$

The explicit formula for the function h can be found e.g. in Barbour *et al.* (1992), p. 7, but what we really need are the following estimates valid uniformly for all A :

$$(3) \quad \sup_k |h(k)| \leq \min(1, \lambda^{-1/2}), \quad |\Delta h| = \sup_k |\Delta h(k)| \leq \lambda^{-1}(1 - e^{-\lambda}),$$

where $\Delta h(k) = h(k+1) - h(k)$, given by Barbour and Eagleson (1983).

Finally, in Section 5 we present some examples concerning basic discrete distributions.

2. The w -functions. Let a non-negative integer-valued random variable X with distribution $\mathcal{F} = \{p(k), k = 0, 1, \dots\}$ have mean μ and variance σ^2 . Define a function w associated with the random variable X (the distribution \mathcal{F}) by the relation

$$\sigma^2 w(k)p(k) = \sum_{i=0}^k (\mu - i)p(i), \quad k = 0, 1, \dots$$

Immediately from the above we have

$$(4) \quad w(0) = \frac{\mu}{\sigma^2}, \quad w(k+1) = \frac{p(k)}{p(k+1)}w(k) + \frac{\mu - (k+1)}{\sigma^2}, \quad k = 0, 1, \dots,$$

and

$$(5) \quad w(k) \geq 0, \quad k = 0, 1, \dots$$

The next relation was stated by Cacoullos and Papathanasiou (1989): if a function g satisfies

$$\mathbb{E}[w(X)\Delta g(X)] < \infty, \quad \mathbb{E}[(X - \mu)g(X)] < \infty,$$

then

$$(6) \quad \text{Cov}(X, g(X)) = \sigma^2 \mathbb{E}[w(X)\Delta g(X)].$$

Note that putting $g(x) = x$ we obtain at once

$$(7) \quad \mathbb{E}[w(X)] = 1.$$

3. Poisson characterization and convergence theorem. It is easy to see that the relation (4) and elementary properties of the Poisson distribution yield the following characterization of the Poisson distribution in terms of the w -functions:

$$\mathcal{F} \text{ is Poisson} \quad \text{if and only if} \quad w(k) = 1, \quad k = 0, 1, \dots$$

It turns out that the characterization remains valid also in, roughly speaking, “limit situations”, providing thus another necessary and sufficient condition for Poisson convergence. More precisely, we have

THEOREM 1. *Let $\{X_n, n = 1, 2, \dots\}$ be a sequence of non-negative integer-valued random variables, each X_n having a distribution \mathcal{F}_n , mean μ_n , variance σ_n^2 and the associated function $w_n, n = 1, 2, \dots$, such that for some positive constant λ ,*

$$\lim_{n \rightarrow \infty} \mu_n = \lim_{n \rightarrow \infty} \sigma_n^2 = \lambda.$$

Then

$$d(\mathcal{F}_n, \mathcal{P}_\lambda) \rightarrow 0 \text{ as } n \rightarrow \infty \text{ if and only if } w_n(X_n) \xrightarrow{P.1} 1 \text{ as } n \rightarrow \infty.$$

Proof. First assume that

$$w_n(X_n) \xrightarrow{P.1} 1 \text{ as } n \rightarrow \infty.$$

From (5) and (7) we have for $n = 1, 2, \dots$,

$$w_n(k) \geq 0, \quad k = 0, 1, \dots, \quad \text{and} \quad E[w_n(X_n)] = 1.$$

Thus by Scheffe’s theorem (see e.g. Billingsley (1979))

$$E|w_n(X_n) - 1| \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Now we use the inequality (8) which will be given and independently proved in the next section, the triangle inequality and the assumptions to get

$$d(\mathcal{F}_n, \mathcal{P}_\lambda) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

To show the converse, set $p_\lambda(k) = \mathcal{P}_\lambda(\{k\})$ for $k = 0, 1, \dots$, and recall that in the discrete case

$$d(\mathcal{F}_n, \mathcal{P}_\lambda) = \sum_{k=0}^{\infty} |p_n(k) - p_\lambda(k)|.$$

So by the assumptions $p_n(k) \rightarrow p_\lambda(k)$ as $n \rightarrow \infty$, for $k = 0, 1, \dots$. This means that

$$\frac{p_n(k)}{p_n(k+1)} \rightarrow \frac{k+1}{\lambda} \text{ as } n \rightarrow \infty, \text{ for } k = 0, 1, \dots,$$

which in turn implies

$$w_n(0) = \mu_n / \sigma_n^2 \rightarrow 1 \text{ as } n \rightarrow \infty.$$

Hence $w_n(1) \rightarrow 1$ as $n \rightarrow \infty$ in view of (4), and proceeding in the same way we obtain $w_n(k) \rightarrow 1$ as $n \rightarrow \infty$, for $k = 2, 3, \dots$

The proof of the theorem is thus complete.

4. Upper bounds for the total variation distance. Using Theorem 1 we can easily test whether Poisson convergence holds for the distributions considered but from the practical point of view some quantities

measuring the distance between the distributions and a Poisson one are more desirable. In the result below we express estimates of the total variation distance in terms of the w -functions (compare Cacoullos *et al.* (1994) and Papathanasiou and Utev (1995)).

THEOREM 2. *The following inequality holds:*

$$(8) \quad d(\mathcal{F}, \mathcal{P}_\lambda) \leq \lambda^{-1}(1 - e^{-\lambda})\mathbb{E}|\sigma^2 w(X) - \lambda| + \min(1, \lambda^{-1/2})|\mu - \lambda|.$$

If additionally $\mu = \sigma^2 = \lambda$, then

$$d(\mathcal{F}, \mathcal{P}_\lambda) \leq (1 - e^{-\lambda})\mathbb{E}|w(X) - 1|.$$

PROOF. In view of (2) and (6), for every subset A of non-negative integers we get

$$\begin{aligned} |\mathcal{F}(A) - P_\lambda(A)| &= |\mathbb{E}[Xh(X)] - \lambda\mathbb{E}[h(X+1)]| \\ &= |\text{Cov}(X, h(X)) + \mu\mathbb{E}[h(X)] - \lambda\mathbb{E}[h(X+1)]| \\ &= |\sigma^2\mathbb{E}[w(X)\Delta h(X)] + (\mu - \lambda)\mathbb{E}[h(X)] - \lambda\mathbb{E}[\Delta h(X)]| \\ &\leq \mathbb{E}|\sigma^2 w(X) - \lambda| \cdot |\Delta h(X)| + |\mu - \lambda|\mathbb{E}[h(X)]. \end{aligned}$$

Now, (1) and (3) prove the result.

5. Examples. 1. Let \mathcal{F} be the binomial distribution

$$\mathcal{B}(n, p) = \left\{ p(k) = \binom{n}{k} p^k q^{n-k}, \quad q = 1 - p, \quad k = 0, 1, \dots, n \right\}.$$

Then $\mu = np$ and $\sigma^2 = npq$ and it is easily seen from (4) that

$$w(k) = \frac{1}{q} - \frac{k}{nq}.$$

If we now let $np_n \rightarrow \lambda$ as $n \rightarrow \infty$, for $\lambda > 0$, then

$$w_n(k) \rightarrow 1 \quad \text{as } n \rightarrow \infty, \quad k = 0, 1, \dots,$$

and from Theorem 1 we obtain the classical Poisson Theorem:

$$d(\mathcal{B}(n, p_n), \mathcal{P}(\lambda)) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

To find the accuracy of the above convergence we apply Theorem 2. Computing

$$\mathbb{E}|\sigma^2 w(X) - np| = np^2$$

and writing henceforth $\mathcal{P}(\lambda)$ instead of \mathcal{P}_λ , we get in view of (8),

$$d(\mathcal{B}(n, p), \mathcal{P}(np)) \leq (1 - e^{-np})p.$$

This is precisely the best known result in the literature (see e.g. Barbour *et al.* (1992), p. 8).

2. Let \mathcal{F} be the negative binomial distribution

$$\mathcal{NB}(n, p) = \left\{ p(k) = \binom{n+k-1}{n-1} p^n q^k, \quad q = 1-p, \quad k = 0, 1, \dots \right\}.$$

Then $\mu = nq/p$ and $\sigma^2 = nq/p^2$ and simple calculations based on (4) lead to

$$w(k) = p(1 + k/n).$$

If $nq_n \rightarrow \lambda$ as $n \rightarrow \infty$, for $\lambda > 0$, then $w_n(k) \rightarrow 1$ as $n \rightarrow \infty$ for $k = 0, 1, \dots$, and so in view of Theorem 1 we have the Poisson Theorem for negative binomial distributions. Because

$$\mathbb{E} \left| \sigma^2 w(X) - \frac{nq}{p} \right| = \frac{nq^2}{p^2},$$

using (8) we get more, i.e.

$$d(\mathcal{NB}(n, p), \mathcal{P}(nq/p)) \leq (1 - e^{-nq/p})q/p.$$

This estimate was also given in Papathanasiou and Utev (1995). For comparison, Vervaat (1969) proved that

$$d(\mathcal{NB}(n, p), \mathcal{P}(nq/p)) \leq q/p.$$

3. Let \mathcal{F} be the hypergeometric distribution

$$\mathcal{H}(n, m, r) = \left\{ p(k) = \frac{\binom{m}{k} \binom{n-m}{r-k}}{\binom{n}{r}}, \quad \max(0, r-n+m) \leq k \leq \min(r, m) \right\}.$$

Then

$$\mu = \frac{mr}{n} \quad \text{and} \quad \sigma^2 = \frac{mr(n-r)(n-m)}{n^2(n-1)}$$

and using (4) it is not difficult to compute that

$$w(k) = \frac{n(n-1)(m-k)(r-k)}{(n-m)(n-r)mr}.$$

If we let the parameters vary according to the conditions $m_n/n \rightarrow 0$ and $m_n r_n/n \rightarrow \lambda$ as $n, m_n, r_n \rightarrow \infty$, for $\lambda > 0$, we see that $w_n(k) \rightarrow 1$ as $n \rightarrow \infty$ for $k = 0, 1, \dots$, which in view of Theorem 1 means that the Poisson Theorem holds for hypergeometric distributions. Bounds for the total variation distance are obtained by computing as before

$$\mathbb{E} \left| \sigma^2 w(X) - \frac{mr}{n} \right| = \frac{mr}{n} \left[\frac{m+r}{n} - \frac{1}{n} \left(\frac{mr}{n} + \frac{(n-r)(n-m)}{n(n-1)} \right) \right]$$

and using once more the relation (8). Then

$$d\left(\mathcal{H}(n, m, r), \mathcal{P}\left(\frac{mr}{n}\right)\right) \leq (1 - e^{-mr/n}) \left[\frac{m+r}{n} - \frac{1}{n} \left(\frac{mr}{n} + \frac{(n-r)(n-m)}{n(n-1)} \right) \right].$$

The same was proved in a different manner by Barbour *et al.* (1992), p. 112.

References

- A. D. Barbour and G. K. Eagleson (1983), *Poisson approximation for some statistics based on exchangeable trials*, Adv. Appl. Probab. 15, 585–560.
- A. D. Barbour, L. Holst and S. Janson (1992), *Poisson Approximation*, Oxford Univ. Press, Oxford.
- P. Billingsley (1979), *Probability and Measure*, Wiley, New York.
- T. Cacoullos and V. Papathanasiou (1989), *Characterizations of distributions by variance bounds*, Statist. Probab. Lett. 7, 351–356.
- T. Cacoullos, V. Papathanasiou and S. A. Utev (1994), *Variational inequalities with examples and an application to the Central Limit Theorem*, Ann. Probab. 22, 1607–1618.
- V. Papathanasiou and S. A. Utev (1995), *Integro-differential inequalities and the Poisson approximation*, Siberian Adv. Math. 5, 120–132.
- V. Veervat (1969), *Upper bounds for the distance in total variation between the binomial and negative binomial and the Poisson distribution*, Statist. Neerlandica 23, 79–86.

M. Majsnerowska
 Mathematical Institute
 University of Wrocław
 Pl. Grunwaldzki 2/4
 50-384 Wrocław, Poland
 E-mail: rmajsner@math.uni.wroc.pl

Received on 29.12.1997