Statistics and Probability Letters 83 (2013) 2422–2431

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Improved lower bounds on the total variation distance for the Poisson approximation

Igal Sason *

Department of Electrical Engineering, Technion, Haifa 32000, Israel

ARTICLE INFO

Article history:
Received 28 April 2013
Received in revised form 3 July 2013
Accepted 3 July 2013
Available online 11 July 2013

Keywords:
Chen–Stein method
Poisson approximation
Total variation distance

ABSTRACT

New lower bounds on the total variation distance between the distribution of a sum of independent Bernoulli random variables and the Poisson random variable (with the same mean) are derived via the Chen–Stein method. The new bounds rely on a non-trivial modification of the analysis by Barbour and Hall (1984) which surprisingly gives a significant improvement. A use of the new lower bounds is addressed.

1. Introduction

Convergence to the Poisson distribution, for the number of occurrences of possibly dependent events, naturally arises in various applications. Following the work of Poisson, there has been considerable interest in how well the Poisson distribution approximates the binomial distribution.

The basic idea which serves as a starting point for the so called *Chen–Stein method for the Poisson approximation* is the following (see Chen (1975)). Let $\{X_i\}_{i=1}^n$ be independent Bernoulli random variables with $\mathbb{E}(X_i) = p_i$. Let $W \triangleq \sum_{i=1}^n X_i$ and $V_i \triangleq \sum_{j \neq i} X_j$ for every $i \in \{1, \ldots, n\}$, and $Z \sim \text{Po}(\lambda)$ with mean $\lambda \triangleq \sum_{i=1}^n p_i$. It is easy to show that

$$\mathbb{E}[\lambda f(Z+1) - Zf(Z)] = 0 \quad (1)$$

holds for an arbitrary bounded function $f : \mathbb{N}_0 \to \mathbb{R}$ where $\mathbb{N}_0 \triangleq \{0, 1, \ldots\}$. Furthermore (see, e.g., Chapter 2 in Ross and Peköz (2007))

$$\mathbb{E}\left[\lambda f(W+1) - Wf(W)\right] = \sum_{j=1}^n p_j^2 \, \mathbb{E}\left[f(V_j+2) - f(V_j+1)\right] \quad (2)$$

which then serves to provide rigorous bounds on the difference between the distributions of W and Z, by the Chen–Stein method for Poisson approximations. This method, and more generally the so called *Stein method*, serves as a powerful tool for the derivation of rigorous bounds for various distributional approximations. Nice expositions of this method are provided by, e.g., Arratia et al. (1990), Ross and Peköz (2007) and Ross (2011). Furthermore, some interesting links between the Chen–Stein method and information-theoretic functionals in the context of Poisson and compound Poisson approximations are provided by Barbour et al. (2010).

Throughout this letter, the term 'distribution' refers to a discrete probability mass function of an integer-valued random variable. In the following, we introduce some known results that are related to the presentation of the new results.

* Tel.: +972 4 8294699.
E-mail address: sason@ee.technion.ac.il.

0167-7152/$ – see front matter
http://dx.doi.org/10.1016/j.spl.2013.07.002

Definition 1. Let P and Q be two probability measures defined on a set $\mathcal{X}$. Then, the total variation distance between P and Q is defined by

$$d_{\text{TV}}(P, Q) \triangleq \sup_{\text{Borel}\, A \subseteq \mathcal{X}} (P(A) - Q(A)) \tag{3}$$

where the supremum is taken w.r.t. all the Borel subsets A of $\mathcal{X}$. If $\mathcal{X}$ is a countable set then (3) is simplified to

$$d_{\text{TV}}(P, Q) = \frac{1}{2} \sum_{x \in \mathcal{X}} |P(x) - Q(x)| = \frac{\|P - Q\|_1}{2} \tag{4}$$

so the total variation distance is equal to half of the L_1-distance between the two probability distributions.

Among old and interesting results that are related to the Poisson approximation, Le Cam's inequality (see Le Cam (1960)) provides an upper bound on the total variation distance between the distribution of the sum $W = \sum_{i=1}^n X_i$ of n independent Bernoulli random variables $\{X_i\}_{i=1}^n$, where $X_i \sim \text{Bern}(p_i)$, and a Poisson distribution $\text{Po}(\lambda)$ with mean $\lambda = \sum_{i=1}^n p_i$. This inequality states that $d_{\text{TV}}(P_W, \text{Po}(\lambda)) \le \sum_{i=1}^n p_i^2$, so if, e.g., $X_i \sim \text{Bern}\left(\frac{\lambda}{n}\right)$ for every $i \in \{1, \ldots, n\}$ (referring to the case where W is binomially distributed) then this upper bound is equal to $\frac{\lambda^2}{n}$, decaying to zero as $n \to \infty$. The following theorem combines Theorems 1 and 2 of Barbour and Hall (1984), and its proof relies on the Chen–Stein method:

Theorem 1. *Let $W = \sum_{i=1}^n X_i$ be a sum of n independent Bernoulli random variables with $\mathbb{E}(X_i) = p_i$ for $i \in \{1, \ldots, n\}$, and $\mathbb{E}(W) = \lambda$. Then, the total variation distance between the probability distribution of W and the Poisson distribution with mean λ satisfies*

$$\frac{1}{32} \left(1 \wedge \frac{1}{\lambda}\right) \sum_{i=1}^n p_i^2 \le d_{\text{TV}}(P_W, \text{Po}(\lambda)) \le \left(\frac{1 - e^{-\lambda}}{\lambda}\right) \sum_{i=1}^n p_i^2 \tag{5}$$

where $a \wedge b \triangleq \min\{a, b\}$ for every $a, b \in \mathbb{R}$.

As a consequence of Theorem 1, it follows that the ratio between the upper and lower bounds in (5) is not larger than 32, irrespectively of the values of $\{p_i\}$. The factor $\frac{1}{32}$ in the lower bound was claimed to be improvable to $\frac{1}{14}$ with no explicit proof (see Remark 3.2.2 in Barbour et al. (1992)). This shows that, for independent Bernoulli random variables, these bounds are essentially tight. Furthermore, note that the upper bound in (5) improves Le Cam's inequality; for large values of λ, this improvement is by approximately a factor of $\frac{1}{\lambda}$.

This letter presents new lower bounds on the total variation distance between the distribution of a sum of independent Bernoulli random variables and the Poisson random variable (with the same mean). The derivation of these new bounds generalizes and improves the analysis by Barbour and Hall (1984), based on the Chen–Stein method for the Poisson approximation. This letter concludes by outlining a use of the new lower bounds for the analysis in Sason (2012), followed by a comparison of the new bounds to previously reported bounds.

This work forms a continuation of the line of work in Barbour and Chen (2005)–Kontoyiannis et al. (2005) where the Chen–Stein method was studied in the context of the Poisson and compound Poisson approximations, and it was linked to an information-theoretic context by Barbour et al. (2010), Kontoyiannis et al. (2005), and Sason (2012).

2. Improved lower bounds on the total variation distance

In the following, we introduce an improved lower bound on the total variation distance and then provide a loosened version of this bound that is expressed in closed form.

Theorem 2. *In the setting of Theorem 1, the total variation distance between the probability distribution of W and the Poisson distribution with mean λ satisfies the inequality*

$$K_1(\lambda) \sum_{i=1}^n p_i^2 \le d_{\text{TV}}(P_W, \text{Po}(\lambda)) \le \left(\frac{1 - e^{-\lambda}}{\lambda}\right) \sum_{i=1}^n p_i^2 \tag{6}$$

where

$$K_1(\lambda) \triangleq \sup_{\substack{\alpha_1, \alpha_2 \in \mathbb{R}, \\ \alpha_2 \le \lambda + \frac{3}{2}, \\ \theta > 0}} \left(\frac{1 - h_\lambda(\alpha_1, \alpha_2, \theta)}{2\, g_\lambda(\alpha_1, \alpha_2, \theta)}\right) \tag{7}$$