Available online at www.sciencedirect.com
SCIENCE DIRECT®
Statistics & Probability Letters 71 (2005) 249–256
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

On the Poisson–binomial relative error

Sabrina Antonelli[1], Giuliana Regoli*

Dipartimento di Matematica e Informatica, University of Perugia, Via Vanvitelli 1, 06100 Perugia, Italy

Received 29 December 2003; received in revised form 30 September 2004
Available online 24 December 2004

Abstract

This paper analyses the goodness of Poisson approximation to the binomial distribution with parameters n and p.

In order to stress the importance of the smallness of p to obtain a good approximation, the relative error is analysed when p is fixed and n tends to infinity. We find the relative error pointwise limit and its asymptotic distribution; we also show that the asymptotic mean relative error gives an approximation of the mean relative error.

Finally we give numerical estimates of all these approximations.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Approximation of distribution; Binomial distribution; Poisson distribution; Relative error; Asymptotic distribution

1. Introduction

The binomial distribution with parameters n and p is usually approximated by the Poisson distribution with parameter $\lambda = np$, when n is large and p is small.

How small must p be and how large must n be in order to have a good approximation?

In Johnson et al. (1992) it is noted that the maximum error is practically independent of n and approaches zero as p approaches zero. This observation makes us focus our attention on the

*Corresponding author. Tel.: +39 075 5855022.
E-mail addresses: antonell@stat.unipg.it (S. Antonelli), regoli@dipmat.unipg.it (G. Regoli).
[1]Ple Caduti dei Lager 13, 06024 Gubbio (PG), Italy.

0167-7152/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.spl.2004.11.009

importance of the smallness of p in order to have a good approximation. We point out this concept, showing that, for every fixed p, the goodness of the approximation cannot be improved by increasing n and that the error strictly depends on p. This result is obtained by studying the value

$$f_n(x) = \left|\frac{\text{poi}(x;\lambda) - \text{bin}(x;n,p)}{\text{bin}(x;n,p)}\right| = \left|\frac{\text{poi}(x;\lambda)}{\text{bin}(x;n,p)} - 1\right| \quad \text{with } \lambda = np,$$

which is the relative error at x if $\text{bin}(x;n,p)$ is replaced by $\text{poi}(x;\lambda)$, where obviously $\text{bin}(x;n,p) = \binom{n}{x}p^x(1-p)^{n-x}$ and $\text{poi}(x;\lambda) = (\lambda^x/x!)e^{-\lambda}$.

In this paper, we analyse the relative error for a fixed p as n tends to infinity. Given p, the pointwise limit of the relative error is studied in Section 2. In Section 3 we find the relative error asymptotic distribution and we show that its mean gives an approximation of the mean relative error. In the last section numerical examples show that the pointwise approximation and the asymptotic mean relative error provide good estimates for finite n.

Burr (1973) gave another approximation of $f_n(x)$, which is accurate as p tends to zero. Inequalities for the ratio $\text{poi}(x;\lambda)/\text{bin}(x;n,p)$ are in Johnson et al. and Feller (1950). On the other hand, Anderson and Samuels (1965) analysed the absolute error between cumulative distribution functions. Other approaches relating to the problem of the Poisson approximation to the binomial distribution can be found in the literature. For example, bounds on the tails of both distributions can be found in Shorack and Wellner (1986) and Johnson et al.

2. Relative error

The relative error $f_n(x)$ depends on the ratio

$$a(x) = \frac{\text{poi}(x;np)}{\text{bin}(x;n,p)} \quad \text{with } x \in \{0, 1, \ldots, n\}.$$

Proposition 1. *As x runs from 0 to n, the ratio $a(x)$ first decreases, then increases, reaching the minimum at $\bar{x} = \lfloor np + 1 \rfloor$.*

Proof. Since

$$\frac{a(x)}{a(x-1)} = \frac{\text{poi}(x;np)}{\text{poi}(x-1;np)} \frac{\text{bin}(x-1;n,p)}{\text{bin}(x;n,p)} = \frac{np}{x} \frac{x(1-p)}{(n-x+1)p} = \frac{n(1-p)}{(n-x+1)},$$

the term $a(x)$ is smaller than the preceding one for $x < np + 1$ and is greater for $x > np + 1$. If $\bar{x} = np + 1$ happens to be an integer, then $a(\bar{x}) = a(\bar{x} - 1)$. Therefore, the minimum of $a(x)$ is achieved at the largest integer, $\bar{x}$, less than or equal to $np + 1$, that is $\bar{x} = \lfloor np + 1 \rfloor$. $\square$

As x increases, the terms $\text{bin}(x;n,p)$ are first smaller, then larger, and then again smaller than $\text{poi}(x;np)$ and the maximum of $f_n(x)$, is reached for $x = 0$, for $x = \lfloor np + 1 \rfloor$ or for $x = n$. When $n \geqslant 3$, this maximum is reached for the extreme values 0 or n. For n sufficiently large, the maximum of $f_n(x)$ is reached for $x = 0$, when $p \geqslant 1 - e^{-1}$, and for $x = n$, when $p < 1 - e^{-1}$.

We focus our attention on the values around the mean np, so that, for a given $n \in N$, every $x \in \{0, 1, 2, \ldots, n\}$ can be written as $x(n,c) = \lfloor np + c\sqrt{np(1-p)} \rfloor$, for suitable real numbers c.

Download English Version:

https://daneshyari.com/en/article/10526020

Download Persian Version:

https://daneshyari.com/article/10526020

Daneshyari.com