Highest posterior mass prediction intervals for binomial and poisson distributions

K. Krishnamoorthy & Shanshan Lv

Metrika
International Journal for Theoretical and Applied Statistics

ISSN 0026-1335
Volume 81
Number 7

Metrika (2018) 81:775-796
DOI 10.1007/s00184-018-0658-z

Springer

Metrika (2018) 81:775–796
https://doi.org/10.1007/s00184-018-0658-z

Highest posterior mass prediction intervals for binomial and poisson distributions

K. Krishnamoorthy[1] · Shanshan Lv[1]

Received: 1 June 2017 / Published online: 11 April 2018

Abstract The problems of constructing prediction intervals (PIs) for the binomial and Poisson distributions are considered. New highest posterior mass (HPM) PIs based on fiducial approach are proposed. Other fiducial PIs, an exact PI and approximate PIs are reviewed and compared with the HPM-PIs. Exact coverage studies and expected widths of prediction intervals show that the new prediction intervals are less conservative than other fiducial PIs and comparable with the approximate one based on the joint sampling approach for the binomial case. For the Poisson case, the HPM-PIs are better than the other PIs in terms of coverage probabilities and precision. The methods are illustrated using some practical examples.

Keywords Coverage probability · Fiducial method · Highest probability mass function · Precision · Predicting distribution

1 Introduction

The problem of predicting a future outcome based on the past and currently available samples arises in many applications. Applications of prediction intervals (PIs) based on continuous distributions are well-known. The prediction intervals based on gamma or lognormal distributions are often used in environmental monitoring, and normal based PIs are used in monitoring and control problems (Hahn and Meeker 1991). Compared

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00184-018-0658-z) contains supplementary material, which is available to authorized users.

✉ K. Krishnamoorthy
krishna@louisiana.edu

1 University of Louisiana at Lafayette, Lafayette, LA, USA

with continuous distributions, there have been only limited investigations for discrete distributions. Prediction intervals based on a discrete distribution are needed to predict the number of events that may occur in the future. For example, if the number of breakdowns of a system in a year follows a Poisson distribution, then the objective is to predict the number of breakdowns in a future year based on available samples (Faulkenberry 1973). Bain and Patel (1993) have noted a situation where it is desired to construct binomial prediction intervals. Wang (2008) illustrated an example, where one needs to predict the number of defective chips in a wafer based on recorded number of defective chips from a sample of wafers.

The binomial prediction problem that we shall address can be described as follows. Given that X successes are observed in a sequence of n independent Bernoulli trials each with "success" probability p, we would like to predict the number of successes Y in another m independent Bernoulli trials each with the same success probability p. Formally, for a given confidence level $1-\alpha$, the problem is to find two integer valued functions $L(X; n, m, \alpha)$ and $U(X; m, n, \alpha)$ so that

$$P_{X,Y}\left(L(X; n, m, \alpha) \le Y \le U(X; m, n, \alpha)\right) \ge 1-\alpha. \tag{1}$$

Any interval that satisfies the above probability requirement for all p is an exact PI. It is now well-known that the exact statistical intervals for discrete distributions only guarantee coverage probabilities, but they are often too conservative and unnecessarily wide (e.g., Agresti and Coull 1988; Brown et al. 2001). For advantages/disadvantages of using an exact inferential result for a binomial distribution, see the recent articles by Thulin (2014) and Thulin and Zwanzig (2017). For our present prediction problem, Thatcher (1964) has proposed a conditional PI (see Sect. 2.2.1) which satisfies the coverage probability requirement (1). However, coverage studies by various authors have indicated that this exact PI is very conservative and unnecessarily wide (see Krishnamoorthy and Peng 2011 and Sect. 2.3 in the sequel). Alternative approximate closed-form PIs based on Wald's result were proposed in Nelson (1982), and they are valid only for large samples. However, coverage studies by Wang (2008) and Krishnamoorthy and Peng (2011) indicate that Nelson's PIs have poor coverage probabilities even for large samples. Wang (2008) has proposed a modification to the Nelson PI, and Krishnamoorthy and Peng (2011) coverage studies indicated that Wang's modified PI is also conservative. Among all frequentist PIs for the binomial and Poisson distributions, the approximate ones based on the joint sampling approach, proposed in Krishnamoorthy and Peng (2011), are reasonably accurate and comparable with other efficient PIs in terms of coverage probability and precision.

The fiducial approach that we shall employ in this paper to find PIs, is a useful tool to obtain accurate interval estimates for various problems involving discrete distributions (e.g., Krishnamoorthy and Lee 2010; Wang et al. 2012; Krishnamoorthy et al. 2017). Fisher (1930, 1935) has introduced the fiducial distribution as a "probable" distribution for the parameter of interest given the data. Fiducial distribution is also interpreted as the posterior distribution without assuming a prior on the parameter (Efron 1998). In general, a fiducial distribution for a parameter can be found in different ways and it is not unique. Dawid and Stone (1982) have proposed functional-model approach to find fiducial distributions. By extending Fisher's fiducial argument, Hannig (2009) and

Hannig et al. (2016) have proposed a general approach to find a fiducial distribution for a parameter, and they refer to it as generalized fiducial distribution. For our present problem, the paper by Wang et al. (2012) appears to be the first one that proposed a fiducial PI for a binomial distribution. These authors have proposed a general fiducial approach for finding PIs, and illustrated the method for a few distributions including the binomial distribution. The fiducial distribution for the binomial parameter p proposed in Wang et al. (2012) involves a linear combination of order statistics from a uniform (0, 1) distribution, and it can not be expressed in closed-form (see Sect. 2.1). The fiducial distributions for a binomial parameter p proposed in the literature are members of the family of beta distributions, and they can be readily used to find PIs. Indeed, two fiducial distributions for a binomial parameter were proposed in the literature (Clopper and Pearson 1934) to find the so-called exact confidence interval (CI). On the basis of Clopper and Pearson fiducial distributions, Stevens (1950) has proposed a family of approximate fiducial distributions (see Sect. 2.1). The confidence intervals based on an alternative fiducial distribution suggested by Steven are shown to be accurate for estimating functions of binomial parameters such as the difference, relative risk and odds ratio (Krishnamoorthy and Lee 2010; Krishnamoorthy et al. 2017). Furthermore, as the alternative fiducial distribution is in explicit forms, one could readily obtain a closed-form expression for predicting fiducial distribution of a future random variable. On the basis of closed-form predicting distributions, PIs can be obtained in two different ways. The methods for the binomial case can be extended to a Poisson distribution in a similar manner. In fact, Cox (1953) has used such fiducial distribution to develop an approximate CI for the ratio of two Poisson means.

The rest of the article is organized as follows. In the following section, we consider the binomial case, and describe two fiducial distributions for a binomial parameter. We then describe the exact PI by Thatcher (1964), the PI based on the join sampling approach (Krishnamoorthy and Peng 2011), fiducial PIs by Wang et al. (2012) and the highest posterior mass (HPM) PIs. These PIs are evaluated in terms of coverage probability and expected width in Sect. 2.3. In Sect. 3, we address the problem of finding PIs for a Poisson distribution by describing the fiducial approach and the joint sampling approach, and compare them in terms of coverage probability and precision. In Sect. 4, the methods are illustrated using a few practical examples. Some concluding remarks are given in Sect. 5.

2 Binomial distribution

2.1 Fiducial distributions for p

Beta fiducial distribution

Let $X \sim$ binomial (n, p), and let x be an observed value of X. Let $B_{a,b}$ denote the beta random variable with shape parameters a and b. Consider testing $H_0 : p \le p_0$ versus $H_a : p > p_0$. Using the well-known relation between the beta and binomial distributions, the p value can be expressed as

$$P(X \ge x|n, p_0) = P\left(B_{x,n-x+1} \le p_0\right).$$

Notice that, for a given x and a level of significance α, any value of p_0 for which the above p value is greater than α is a "probable" value of p, and such probable values are determined by the $B_{x,n-x+1}$ distribution. Similarly, by considering the test for $H_0 : p \geq p_0$ versus $H_a : p < p_0$, we will arrive at the $B_{x+1,n-x}$ distribution as a probable distribution for p. This pair of probable or fiducial distributions led to Clopper and Pearson's (1934) confidence interval for p, the one $B_{x,n-x+1}$ is for setting up lower confidence limit for p, and the other $B_{x+1,n-x}$ is for setting upper confidence limit for p. Specifically, the Clopper–Pearson CI based on this pair of fiducial variables is given by $\left(B_{x,n-x+1;\alpha/2}, B_{x+1,n-x;1-\alpha/2}\right)$, and is commonly referred to as the exact CI. Instead of having two fiducial variables, Stevens (1950) has suggested that a random quantity that is "stochastically between" $B_{x,n-x+1}$ and $B_{x+1,n-x}$ can be used as a single approximate fiducial variable for p. A simple choice is

$$B_{x+1/2,n-x+1/2}, \tag{2}$$

which is also the posterior distribution with the Jeffreys prior $B_{1/2,1/2}$ (e.g., Cai 2005). In the sequel, we shall use the fiducial distribution (2) to find PIs for a binomial distribution.

The generalized fiducial distribution

Wang et al. (2012) have obtained a generalized fiducial distribution for p by extending Fisher's fiducial argument. For more details on extended fiducial argument, see Sect. 2 of Hannig et al. (2016). To describe the generalized fiducial distribution for p, let $U_1, \ldots, U_n$ be independent uniform(0,1) random variables, and let $X \sim \text{Binomial}(n, p)$. Notice that

$$X \stackrel{d}{=} \sum_{i=1}^{n} I_{[0,p]}(U_i),$$

where $I_A(x)$ is the indicator function, and the notation $\stackrel{d}{=}$ means "distributed as". Following Hannig (2009), a fiducial quantity for p is expressed as

$$\widetilde{p}_{x,n} = \begin{cases} U_{1:n}D, & \text{if } x = 0, \\ (1-D)U_{x:n} + DU_{x+1:n}, & \text{if } x = 1, 2, \ldots, n-1, \\ (1-D)U_{n:n} + D, & \text{if } x = n, \end{cases} \tag{3}$$

where $D \sim \text{uniform}(0, 1)$ independently of $U_1, \ldots, U_n$. Noting that the rth order statistic for a sample of size n from a uniform(0, 1) distribution is distributed like $B_{r,n-r+1}$, we see that the fiducial distribution for p is determined by a random linear combination of $B_{x,n-x+1}$ and $B_{x+1,n-x}$ distributions that were suggested by Clopper and Pearson (1934).

2.2 Binomial prediction intervals

Let x be an observed value of $X \sim \text{binomial}(n, p)$. Let $Y \sim \text{binomial}\ (m, p)$ independently of X. The problem is to find a prediction interval for Y based on X. In the following, we shall describe four PIs for Y based on (X, n).

2.2.1 An exact prediction interval

We shall now describe the exact PI by Thatcher (1964) as given in Krishnamoorthy and Peng (2011). The conditional distribution of X given the sum $X + Y = s$ is hypergeometric with the sample size s, the number of "white balls" n, and the number of "black balls" m. The conditional probability mass function is given by

$$P(X = x|X + Y = s, n, m) = \frac{\binom{n}{x}\binom{m}{y}}{\binom{m+n}{s}}, \quad \max\{0, s - m\} \leq x \leq \min\{n, s\}.$$

The cumulative distribution function (cdf) of X given $X + Y = s$ is

$$F_X(t|n, m, s) = P(X \leq t|n, m, s) = \sum_{i=0}^{t} \frac{\binom{n}{i}\binom{m}{s-i}}{\binom{m+n}{s}}. \tag{4}$$

Thatcher (1964) developed the following exact PI on the basis of the conditional distribution of X given $X + Y$. Let x be an observed value of X. The $1-\alpha$ lower prediction limit L is the smallest integer for which

$$P(X \geq x|n, m, x + L) = 1 - F_X(x - 1|n, m, x + L) > \alpha. \tag{5}$$

The $1-\alpha$ upper prediction limit U is the largest integer for which

$$F_X(x|n, m, x + U) > \alpha. \tag{6}$$

Furthermore, $[L, U]$ is a $1-2\alpha$ two-sided PI for Y. The exact PIs for extreme values of X are defined as follows. When $X = 0$, the lower prediction limit for Y is 0, and the upper one is determined by (6); when $x = n$, the upper prediction limit is m, and the lower prediction limit is determined by (5).

2.2.2 The prediction interval based on the joint sampling approach

Krishnamoorthy and Peng (2011) have proposed a PI that is based on the complete sufficient statistic $X + Y$ for the binomial$(n + m, p)$ distribution. Let $c = z_{1-\alpha/2}$ denote the $100(1-\alpha/2)$ percentile of the standard normal distribution. To handle the extreme cases, these authors have defined the point predictor as

$$\widehat{Y} = \begin{cases} .5\frac{m}{n} & \text{if } X = 0, \\ X\frac{m}{n} & \text{if } X = 1, \ldots, n-1, \\ (n - .5)\frac{m}{n} & \text{if } X = n. \end{cases}$$

The $100(1-\alpha)\%$ PI is expressed as

$$\frac{\left[\widehat{Y}\left(1-\frac{c^2}{m+n}\right)+\frac{mc^2}{2n}\right] \pm c\sqrt{\widehat{Y}(m-\widehat{Y})\left(\frac{1}{m}+\frac{1}{n}\right)+\frac{m^2c^2}{4n^2}}}{\left(1+\frac{mc^2}{n(m+n)}\right)} = (L, U), \text{ say.} \tag{7}$$

The $1-2\alpha$ PI, based on the roots in (7), is given by $[\lceil L \rceil, \lfloor U \rfloor]$, where $\lceil x \rceil$ is the ceiling function and $\lfloor x \rfloor$ is the floor function. We refer to this PI as the "joint sampling–prediction interval" (JS-PI).

2.2.3 *Fiducial prediction interval by Wang et al.*

On the basis of the fiducial distribution by Wang et al. (2012) described in Sect. 2.1, the fiducial prediction distribution is determined by the conditional distribution $\widetilde{Y}|\widetilde{p}_{x,n} \sim$ binomial$(m, \widetilde{p}_{x,n})$ and the distribution of $\widetilde{p}_{x,n}$ in (3). The fiducial PI based on this predicting distribution can be obtained using Monte Carlo simulation as shown in the following algorithm.

Algorithm 1
Let x be an observed value of $X \sim$ binomial(n, p).
1. Generate $U_1, \ldots, U_n$ from uniform (0, 1) distribution
2. Sort U_j's generated in step 1, generate a D from uniform (0, 1), and find $\widetilde{p}_{x,n}$ using (3)
3. Generate $\widetilde{Y}$ from binomial $(m, \widetilde{p})$
4. Repeat steps 1–3 for a large number of times, say, 10,000
5. The lower and upper α quantiles of 10,000 $\widetilde{Y}$'s form a $100(1-2\alpha)\%$ PI for a future observation from binomial (m, p) distribution.

2.2.4 *Equal-tailed and HPM prediction interval*

Let x be an observed value of $X \sim$ binomial (n, p). We shall now describe two PIs for $Y \sim$ binomial (m, p) based on the beta fiducial distribution $B_{x+.5,n-x+.5}$. For a given x, the predicting fiducial distribution is determined by

$$Y^*|W \sim \text{binomial}(m, W) \quad \text{and} \quad W \sim B_{x+.5,n-x+.5}. \tag{8}$$

The probability mass function (pmf) of Y^* (see "Appendix A") is given by

$$P(Y^* = y) = \frac{\Gamma(x+y+.5)}{\Gamma(y+1)\Gamma(x+.5)} \frac{\Gamma(m+n-x-y+.5)}{\Gamma(n-x+.5)\Gamma(m-y+1)} \bigg/ \binom{m+n}{m}, \quad y = 0, 1, \ldots, m. \tag{9}$$

The above pmf resembles the pmf of a hypergeometric distribution.

Equal-tailed prediction intervals

The $100(1-2\alpha)\%$ equal-tailed PI for a future outcome from a binomial (m, p) distribution is formed by the lower 100α percentile and the upper 100α percentile

Table 1 95% PIs by various methods for a binomial (m, p) distribution

(n, k, m)	Exact PI (1)	Other methods*	HPM PI
(5, 3, 15)	(1, 15)	(2, 14)	(2, 14)
(20, 10, 15)	(2, 13)	(3, 12)	(3, 12)
(10, 4, 12)	(0, 10)	(1, 10)	(1, 9)
(5, 1, 9)	(0, 7)	(0, 7)	(0, 6)
(50, 43, 33)	(22, 32)	(22, 32)	(23, 32)
(11, 3, 15)	(0, 11)	(0, 10)	(0, 9)
(343, 30, 200)	(9, 29)	(9, 28)	(8, 27)
(5000, 200, 1000)	(27, 54)	(27, 54)	(27, 53)

*PI based on Algorithm 1 by Wang et al. (2012), equal-tailed PI by simulation using 10,000 runs and the equal-tailed PI by the numerical method

of Y^*. Such percentiles can be readily obtained using Monte Carlo simulation. The lower and upper percentiles of the fiducial predicting distribution can also be obtained in an exact manner using the pmf in (9). Let L^* be the smallest integer for which $\sum_{y=0}^{L^*} P(Y^* = y) > \alpha$ and let U^* be the largest integer for which $\sum_{y=U^*}^{m} P(Y^* = y) > \alpha$, where $P(Y^* = y)$ is defined in (9). Then (L^*, U^*) is the $100(1-2\alpha)\%$ equal-tailed fiducial PI. The PIs based on simulation with $N =$10,000 and the one based on the exact numerical approach are practically the same; see Table 1 for some numerical evidence.

Highest posterior mass prediction intervals

Instead of equal-tailed PIs, one could also use the highest posterior mass (HPM) fiducial PIs, which are expected to be shorter than equal-tailed PIs in the preceding section (see Fig. 1). A HPM-PI can be obtained by calculating $P(Y^* = y)$ at the mode or the mean of Y^*. Let y_{m} denote the mean $E(Y^*) = m(x + .5)/(n + 1)$, where Y^* is defined in (8). The HPM-PI can be obtained by calculating probability $P(Y^* = y_{\mathrm{m}})$, and then by adding integers $y_{\mathrm{m}} + 1$ and $y_{\mathrm{m}} - 1$ to the prediction set until the probabilities

$$P(Y^* = y_{\mathrm{m}}),\ P(Y^* = y_{\mathrm{m}} \pm 1),\ P(Y^* = y_{\mathrm{m}} \pm 2), \ldots$$

add up to at least $1-\alpha$. Even though there is not much difference between the mode and the mean, the PI calculated using the mean as a starting point has better frequentist coverage properties than the one using the mode as a starting point. The following algorithm can be used to find the HPM-PI for a future observation from a binomial (m, p) distribution.

Algorithm 2
For a given (x, n, m) and a confidence level $1-\alpha$,

1. Set the mean

$$y_{\mathrm{m}} = \begin{cases} 0 & \text{if } x = 0, \\ m & \text{if } x = n \\ \lceil m(x + .5)/(n + 1) \rceil, & \text{for } x = 1, \ldots, n - 1. \end{cases}$$

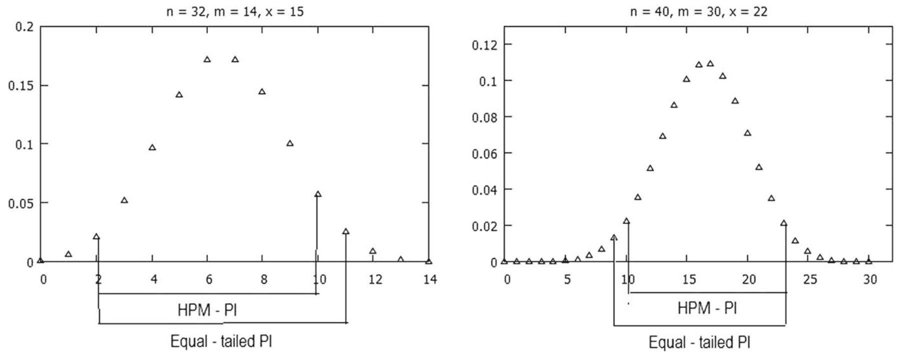

Fig. 1 Binomial predicting probability mass function and 95% PIs

2. If $y_m = 0$, find the smallest integer U so that $\sum_{y=0}^{U} P(Y^* = y) \geq 1-\alpha$. In this case $[0, U]$ is the desired PI.
3. If $y_m = m$, find the largest integer L so that $\sum_{y=L}^{m} P(Y^* = y) \geq 1-\alpha$. In this case $[L, m]$ is the desired PI.
4. For other values of y_m, determine the smallest integer j so that

$$S = \sum_{k=y_{\mathrm{m}}-j}^{y_{\mathrm{m}}+j} P(Y^* = k) \geq 1-\alpha.$$

Set $[L, U] = [y_{m-j}, y_{m+j}]$. If $S - P(Y^* = L) > 1-\alpha$ then the PI is $[L+1, U]$; else if $S - P(Y^* = U) > 1-\alpha$, then the PI is $[L, U-1]$.
5. In Step 4, if $y_{\mathrm{m}} - j = 0$ and $S < 1-\alpha$, add the probabilities $P(Y^* = y_{\mathrm{m}} + j + i)$, $i = 1, 2, \ldots$ to S until it becomes greater than or equal to $1-\alpha$. In this case $L = 0$ and U is the value of $y_{\mathrm{m}} + j + i$ at which $S \geq 1-\alpha$.
6. In Step 4, if $y_{\mathrm{m}} + j = m$ and $S < 1-\alpha$, add the probabilities $P(Y^* = y_{\mathrm{m}} - j - i)$, $i = 1, 2, \ldots$ to S until S becomes greater than or equal to $1-\alpha$. In this case $U = m$ and L is the value of $y_{\mathrm{m}} - j - i$ at which $S \geq 1-\alpha$.
7. The interval $[L, U]$ thus obtained is the $100(1-\alpha)\%$ HPM prediction interval.

The R code based on Algorithm 2 is available at the journal website and at the first author's homepage www.ucs.louisiana.edu/~kxk4695. In the following Table 1, we reported PIs based on some values of (n, x, m) by different approaches. For the data considered in Table 1, the PIs by Wang et al. (2012), the equal-tailed PIs and the equal-tailed PIs by simulation are the same for each case, and they are reported under the column "other methods." We observe from this table that the exact PIs are wider than other PIs in many cases. Furthermore, the HPM PIs are either the same as the others or shorter than others.

2.3 Comparison of binomial prediction intervals

All the PIs that we considered and proposed in the preceding sections are approximate except the exact PI by Thatcher (1964). So we shall evaluate the merits of all the PIs in terms of coverage probability and precision. For a given (n, m, p, α), the exact coverage probability of a PI $[L(x, n, m, \alpha), U(x, n, m, \alpha)]$ can be evaluated using the expression

$$\sum_{x=0}^{n}\sum_{y=0}^{m} f_X(x|n, p) f_Y(y|m, p) I_{[L(x,n,m,\alpha),U(x,n,m,\alpha)]}(y), \tag{10}$$

where $f_X(x|n, p)$ is the pmf of binomial (n, p) distribution and $I_A(.)$ is the indicator function. For a good PI, its coverage probabilities should be close to the nominal level. The expected width of a PI $[L(x, n, m, \alpha), U(x, n, m, \alpha)]$ can be evaluated using (10) with the indicator function replaced by the width $U(x, n, m, \alpha) - L(x, n, m, \alpha)$.

We evaluated the coverage probabilities of the exact PI, the fiducial PI by Wang et al. (2012) which is based on the fiducial distribution (3), the equal-tailed PI, the HPM-PI and the JS-PI for some values of (n, m) and for p in (0, 1). Calculated coverage probabilities at the nominal level of .95 are presented in Fig. 2. The plots on the left column of Fig. 2 are for the exact PI, fiducial PI and the equal-tailed PI while the plots in the right column are for the HPM-PI and the JS-PI. We first observe from these plots that the fiducial and equal-tailed PIs perform very similar for all values of (m, n). The coverage probabilities of these two PIs practically coincide for most cases, and they are conservative. For all the cases, the exact PI is more conservative than the fiducial and equal-tailed PIs. This is true even for sample sizes of $(n, m) = (50, 80)$. Examination of the plots in the right column of Fig. 2 indicate that the HPM-PI and the JS-PI are less conservative than other three PIs in the left column. These two PIs could be anti-conservative having coverage probabilities smaller than the nominal level .95, but no less than .94 in most cases. In general, the HPM-PI and the JS-PI maintain coverage probabilities close to the nominal level. The HPM-PI and the JS-PI are quite comparable in terms of coverage probabilities except that the coverage probabilities of HPM-PI are close to the nominal level over larger parameter space than the JS-PI in some cases; see the plots for $n = 10, m = 5$ and $n = 25, m = 25$ in Fig. 2. The JS-PI could be liberal in some cases; see the plots for $(n, m) = (20, 15)$ and $(25, 25)$.

We also calculated the expected widths of all PIs for some values (n, m) and plotted them in Fig. 3. Plots for the exact, fiducial and equal-tailed PIs are presented in the left column of Fig. 3 while the plots of the HPM-PI and JS-PI are given in the right column. The expected widths of these PIs reflect their coverage performances in Fig. 2. In particular, the exact PI is too conservative, as a result, its expected widths are larger than other two PIs. The equal-tailed PI and the fiducial PI based on (3) have similar coverage properties, and so their expected widths are practically the same. All the plots in Fig. 3 clearly indicate that the expected widths of the exact PIs are appreciably larger than those of other four PIs. The plot for the case of $(n, m) = (10, 5)$ shows that the HPM-PIs are expected to be shorter than the JS-PIs for small values of (n, m). These six plots on the right column of Fig. 3, in general, show that the HPM-PIs are shorter than the JS-PIs for small/large values of p, and the JS-PIs are narrower than the HPM-PIs for values of p in the middle of the interval (0, 1). Thus, the HPM-PI is preferable to the JS-PI if $\widehat{p} = x/n$ is not close to 0.5, and the JS-PI is preferable to the HPM-PI when $\widehat{p}$ is around 0.5.

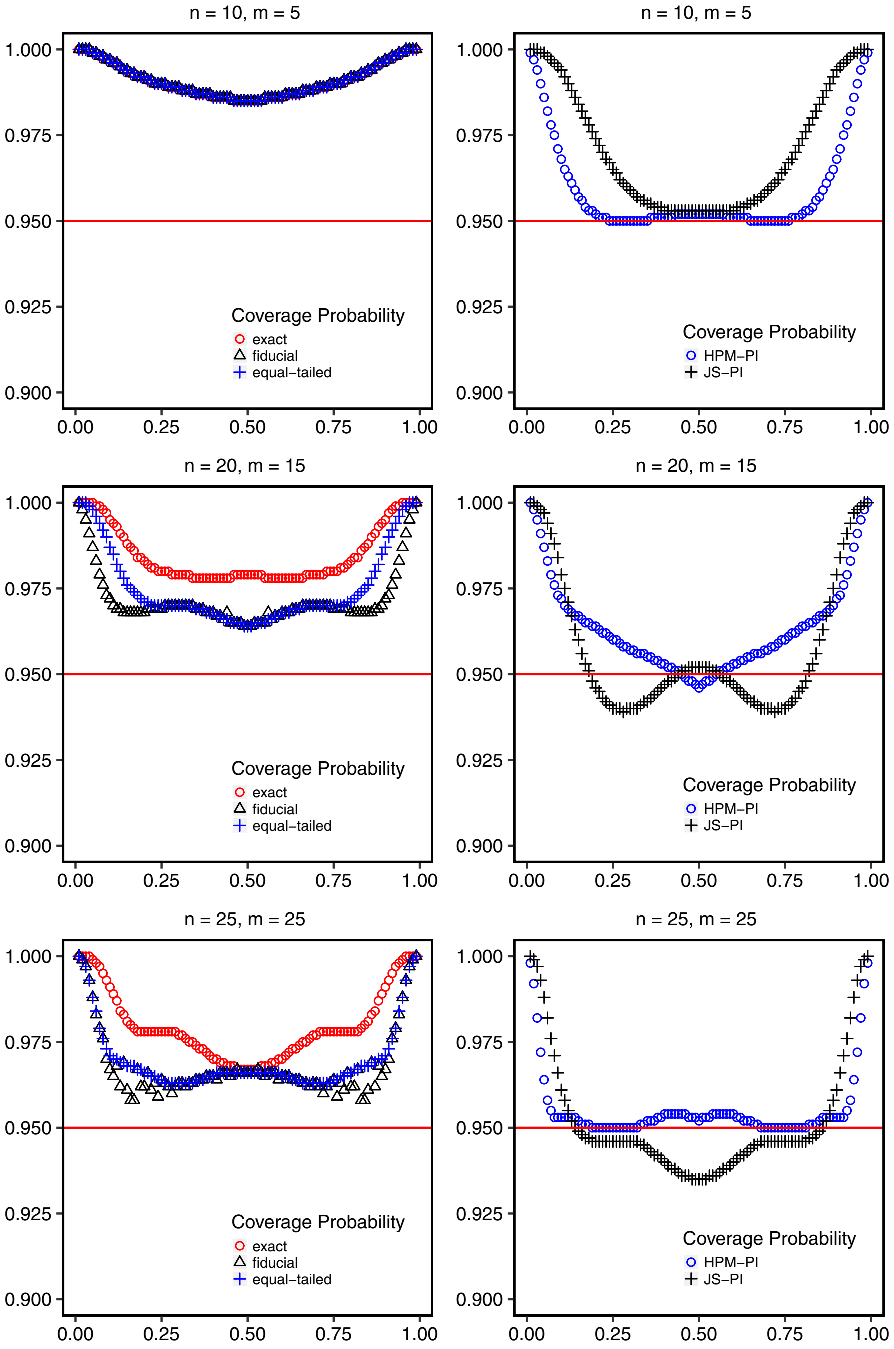

Fig. 2 Coverage probabilities of 95% prediction intervals for binomial distributions

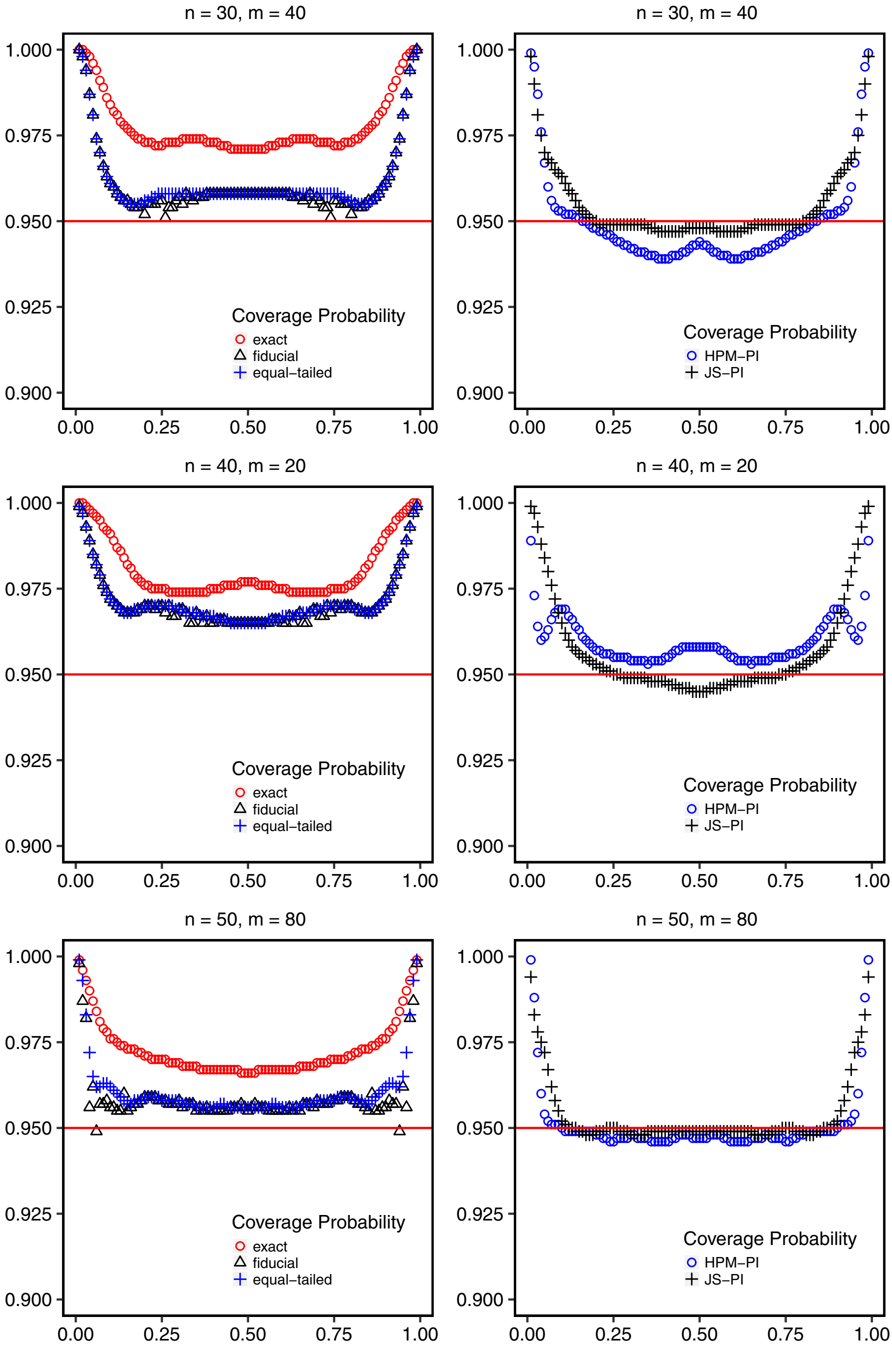

Fig. 2 continued

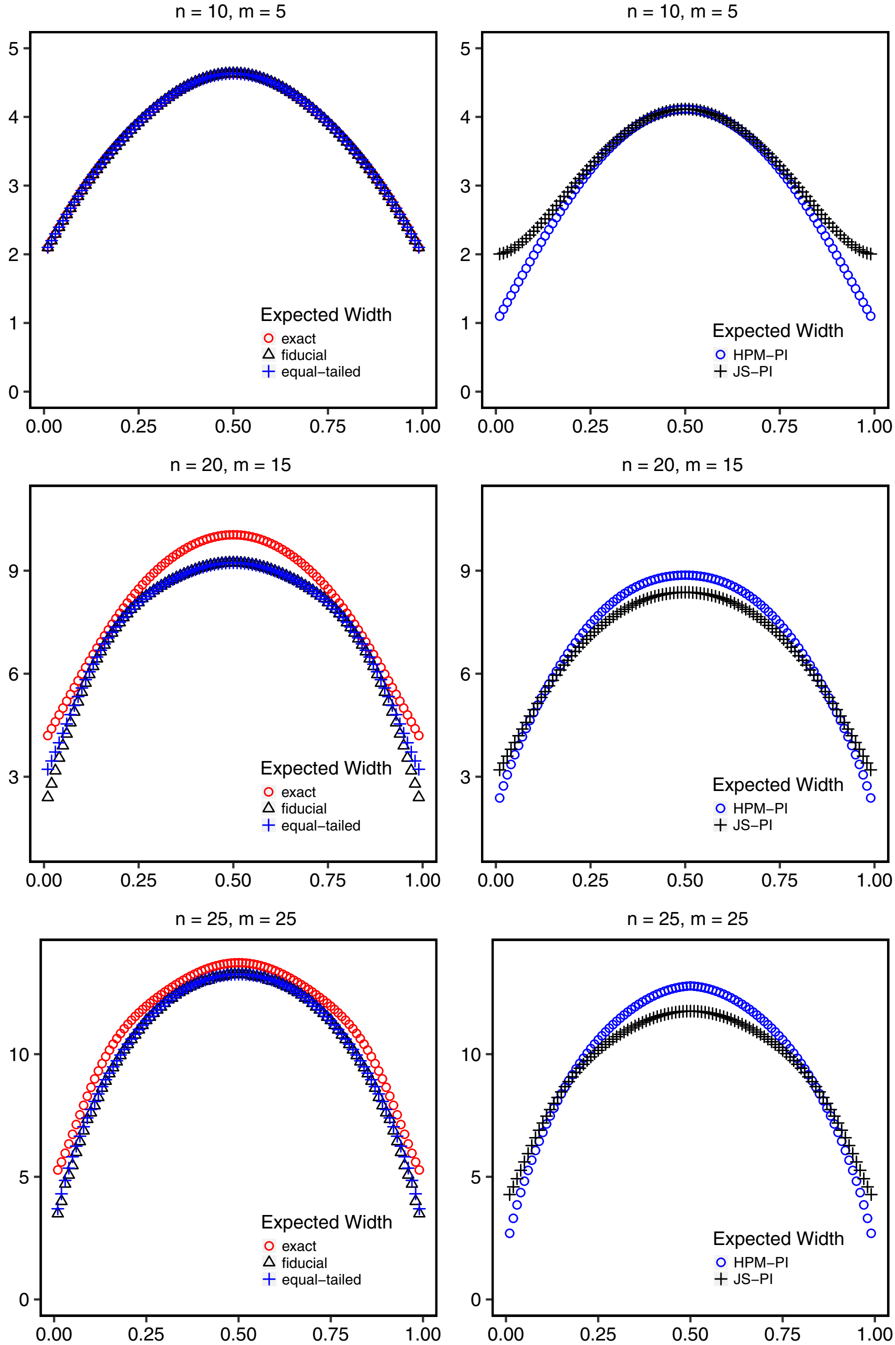

Fig. 3 Expected widths of 95% prediction intervals for binomial distributions

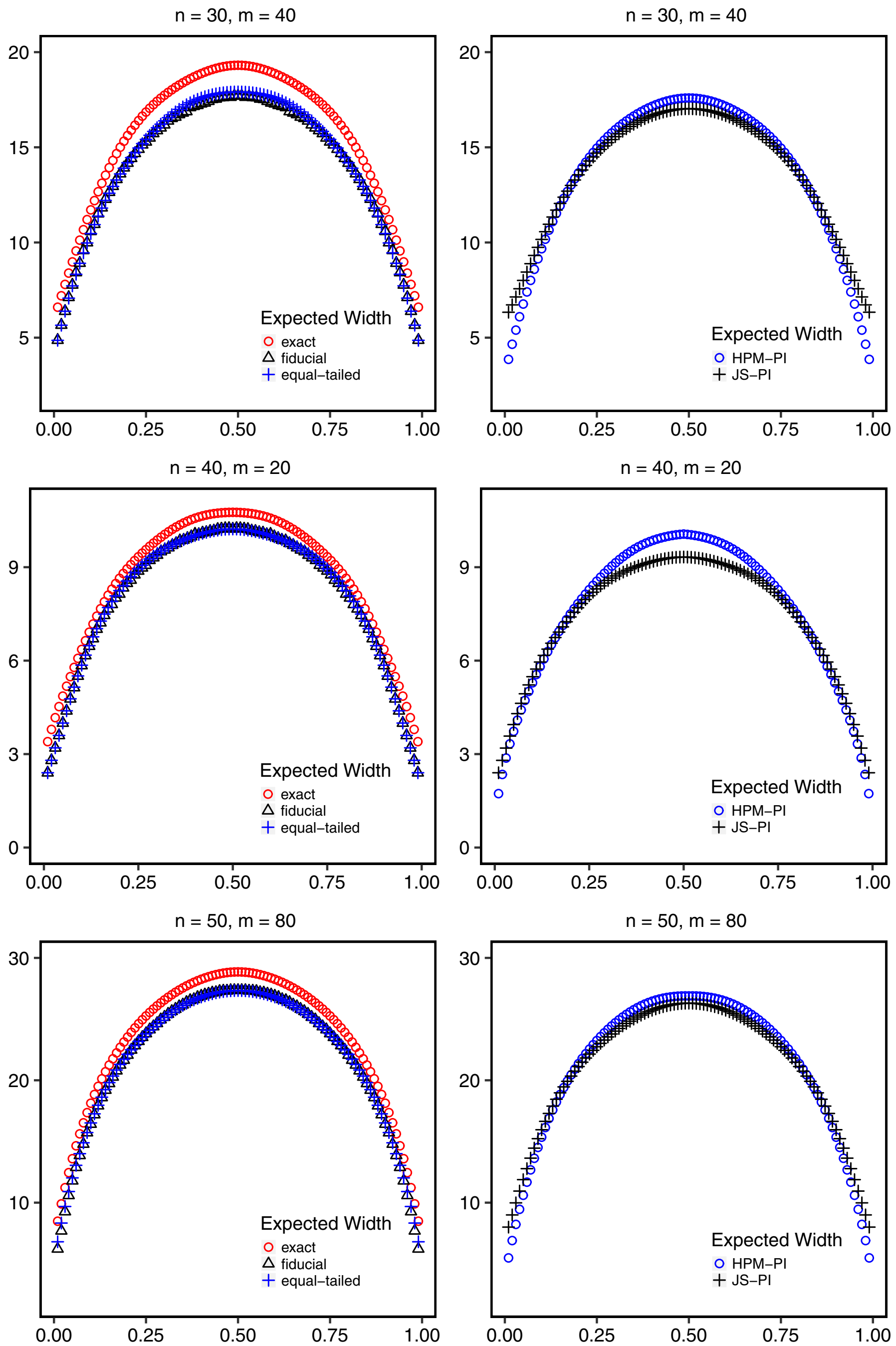

Fig. 3 continued

Overall, we see that the coverage probabilities of the HPM-PI and JS-PI are close to the nominal level despite being anti-conservative sometimes. Furthermore, these two PIs have shorter expected widths than the exact, fiducial and equal-tailed PIs.

3 Poisson distribution

Let $X_1, \ldots, X_n$ be a sample from a Poisson distribution with mean λ. Note that the total count $X = \sum_{i=1}^{n} X_i \sim$ Poisson $(n\lambda)$. Let Y denote the total count in a future sample of size m from the same Poisson distribution so that $Y \sim$ Poisson $(m\lambda)$. The problem is to predict the value of Y based on an observed value x of X.

3.1 The exact prediction interval

Note that the conditional distribution of X, conditionally given $X+Y = s$, is binomial with the number of trials s and the success probability $n/(n+m)$, binomial $(s, n/(m+n))$. On the basis of this conditional distribution, Krishnamoorthy and Peng (2011) have proposed the following PI. Let x be an observed value of X. The smallest integer L that satisfies

$$\sum_{i=x}^{x+L} \binom{x+L}{i} \left(\frac{n}{n+m}\right)^i \left(\frac{m}{n+m}\right)^{x+L-i} > \alpha \tag{11}$$

is the $1-\alpha$ lower prediction limit for Y. The $1-\alpha$ upper prediction limit U is the largest integer for which

$$\sum_{i=0}^{x} \binom{x+U}{i} \left(\frac{n}{n+m}\right)^i \left(\frac{m}{n+m}\right)^{x+U-i} > \alpha. \tag{12}$$

For $X = 0$, the lower prediction limit is defined to be zero, and the upper prediction limit is determined by (12).

3.2 The prediction interval based on the joint sampling approach

Let $\widehat{\lambda}_{xy} = \frac{X+Y}{m+n}$. As in the binomial case, the quantity

$$\frac{m\widehat{\lambda}_{xy} - Y}{\sqrt{\widehat{\mathrm{var}}(m\widehat{\lambda}_{xy} - Y)}} = \frac{(mX - nY)}{\sqrt{mn(X+Y)}},$$

whose asymptotic distribution is the standard normal. In order to handle the zero count, we take X to be 0.5 when it is zero. The $1-2\alpha$ PI is determined by the roots (with respect to Y) of the quadratic equation $(m\widehat{\lambda} - Y)^2/[\widehat{\lambda}_{xy} m(1 + m/n)] = z_{1-\alpha}^2$. Based on these roots, the $1-2\alpha$ PI is given by

$$[\lceil L\rceil, \lfloor U\rfloor], \text{ with } [L, U] = \widehat{Y} + \frac{mz_{1-\alpha}^2}{2n} \pm z_{1-\alpha}\sqrt{m\widehat{Y}\left(\frac{1}{m}+\frac{1}{n}\right) + \frac{m^2 z_{1-\alpha}^2}{4n^2}}, \tag{13}$$

where $\lceil x\rceil$ is the ceiling function, $\lfloor x\rfloor$ is the floor function and

$$\widehat{Y} = \begin{cases} X\frac{m}{n} & \text{if } X = 1, 2, \ldots, \\ .5\frac{m}{n} & \text{if } X = 0. \end{cases}$$

3.3 Equal-tailed and HPM fiducial PIs

Let x be an observed value of $X \sim$ Poisson $(n\lambda)$. Fiducial distributions for a Poisson parameter can be obtained along the lines described in Sect. 2.1 for the binomial parameter. To obtain fiducial distributions for the Poisson case, we use the relations

$$P(X \geq x|\lambda) = P\left(\frac{\chi_{2x}^2}{2n} < \lambda\right) \text{ and } P(X \leq x) = P\left(\frac{\chi_{2x+2}^2}{2n} > \lambda\right),$$

where χ_a^2 is the chi-square random variable with degrees of freedom (df) a. On the basis of the above relations, we have a pair of fiducial distributions for λ, $\chi_{2x}^2/(2n)$ and $\chi_{2x+2}^2/(2n)$. As χ_a^2 is stochastically increasing in a, the random variable $\chi_{2x+1}^2/(2n)$, which stochastically lies between two fiducial distributions, is an approximate fiducial quantity for λ (Cox 1953). This approximate fiducial distribution is also the posterior distribution of λ with the improper prior $1/\sqrt{\lambda},\ 0 < \lambda < \infty$.

For a given $X = x$, the fiducial PI for $Y \sim$ Poisson $(m\lambda)$ can be obtained from the predicting fiducial distribution determined by

$$Y^*|W \sim \text{Poisson}\,(mW) \quad \text{and } W \sim \frac{\chi_{2x+1}^2}{2n}. \tag{14}$$

Equal-tailed fiducial PI

The lower and upper 100α percentiles of Y^* form an equal-tailed PI for a future number of event from a Poisson $(m\lambda)$ distribution. The percentiles can be obtained using Monte Carlo simulation in a straightforward manner. The equal-tailed PI can also be evaluated in an exact manner using probability mass function (pmf) of Y^* (see "Appendix B")

$$P(Y^* = y) = \frac{\Gamma(x+y+.5)}{\Gamma(x+.5)\Gamma(y+1)}\left(\frac{m}{m+n}\right)^y\left(\frac{n}{m+n}\right)^{x+.5}, \quad y = 0, 1, 2, \ldots \tag{15}$$

Let L^* be the smallest integer for which $\sum_{y=0}^{L^*} P(Y^* = y) > \alpha$ and let U^* be the largest integer for which $\sum_{y=U^*}^{\infty} P(Y^* = y) > \alpha$. Then $[L^*, U^*]$ is the $100(1-2\alpha)\%$ equal-tailed fiducial PI for a future observation from a Poisson $(m\lambda)$ distribution.

HPM prediction interval

The HPM-PI can be obtained along the lines for the binomial case described earlier. Let $y_{\text{m}} = E(Y^*) = m(x + .5)/n$, the mean of Y^* in (14). The HPM–PI is formed by the set

$$\mathbf{S} = \{y_{\text{m}}, y_{\text{m}} - 1, y_{\text{m}} + 1, y_{\text{m}} - 2, y_{\text{m}} + 2, \ldots\}$$

so that $P(Y^* \in \mathbf{S}) \geq 1 - \alpha$. Such PI can be constructed by sequentially searching for values starting from y_{m} until the probability content just exceeds $1-\alpha$. The following algorithm provides computational details of calculating the HPM-PI for a Poisson distribution.

Algorithm 3

For a given (x, n, m, α),

1. Set the mean $y_{\text{m}} = \lceil m(x + .5)/n \rceil$.
2. If $x = 0$, set $L = 0$ and find the smallest integer U such that $\sum_{k=0}^{U} P(Y^* = k) \geq 1 - \alpha$, where $P(Y^* = y)$ is given in (15). Return $[L, U]$.
3. Calculate

$$S = \sum_{k=y_{\text{m}}-j}^{y_{\text{m}}+j} P(Y^* = k), \quad \text{for} \quad j = 1, 2, \ldots$$

   Stop calculation if $S \geq 1 - \alpha$, and set $[L, U] = [m - j, m + j]$. If $S - P(Y^* = L) > 1 - \alpha$ then return the PI $[L + 1, U]$; else if $S - P(Y^* = U) > 1 - \alpha$, then return the PI $[L, U - 1]$.
4. In Step 3, if $y_{\text{m}} - j = 0$ and $S < 1 - \alpha$, add the probabilities $P(Y^* = y_{\text{m}} + j + i)$, $i = 1, 2, \ldots$ to S until it becomes greater than or equal to $1 - \alpha$. In this case $L = 0$ and U is the value of $y_{\text{m}} + j + i$ at which $S \geq 1 - \alpha$.

The interval $[L, U]$ thus obtained is the $100(1-\alpha)\%$ HPM prediction interval for a future observation from a Poisson $(m\lambda)$ distribution. The R code based on Algorithm 3 is available at the journal website and the first author's homepage www.ucs.louisiana.edu/~kxk4695.

3.4 Coverage probabilities of poisson prediction intervals

All PIs, except the exact PI, are approximate, and so we shall evaluate their accuracy in terms of coverage probability and precision. For a given (n, m, λ, α), the exact coverage probability of a Poisson PI $[L(x, n, m, \alpha), U(x, n, m, \alpha)]$ can be evaluated using the expression

$$\sum_{x=0}^{\infty} \sum_{y=0}^{\infty} f_X(x|n\lambda) f_Y(y|m\lambda) I_{[L(x,n,m,\alpha), U(x,n,m,\alpha)]}(y), \tag{16}$$

where $f_X(x|\eta)$ is the pmf of the Poisson (η) distribution.

The coverage probabilities of the exact PI, equal-tailed PI, JS-PI and the HPM-PI are plotted in Fig. 4 for some values of (n, m) and λ. We first observe from these plots that the exact PIs are quite conservative with coverage probabilities much larger than the nominal level .95. Among all four PIs, the exact PI is too conservative followed by the equal-tailed PI, and the HPM-PI performs better than other three PIs. The coverage probabilities of the HPM-PI and JS-PI are close to the nominal level for moderate to large λ or both $m\lambda$ and $n\lambda$ are large, and they are expected to perform similar for large values of $m\lambda$ and $n\lambda$. We also notice that the coverage probabilities of the JS-PI are appreciably smaller than the nominal level .95 when $(n, m) = (10, 1)$. In general, for small values of λ, the coverage probabilities of HPM-PI are closer to the nominal level than those of the JS-PI, and so the former is expected to be less conservative than the latter.

To judge the precision of the PIs for the Poisson case, we calculated the expected widths of all four PIs noted in the preceding paragraph, and plotted them in Fig. 5. Examination these plots indicates that the exact PI is much wider than all other three PIs even for large Λ. The HPM-PI is expected to be shorter than the JS-PI for small to moderate values of λ. In general, the HPM-PIs are shorter than or same as JS-PIs. In fact, this comparison of expected widths is in agreement with the coverage comparison in the preceding paragraph where we noted that the JS-PIs are conservative for small values of λ, as a result, they are wider than the HPM-PIs.

We also evaluated coverage probabilities and expected widths of PIs with confidence coefficients 0.90 and 0.99. The results of comparison are similar to those of 95% PIs, and so they are not reported here. On an overall basis, the HPM-PI is preferable to other PIs for small to moderate values of λ. For other cases, the JS-PI could be preferred for its simplicity.

4 Examples

Example 1 This example is adapted from Example 3.13 of Krishnamoorthy (2015). A manufacturer of an expensive piece of medical equipment has sold 40 units in the past year, and found that two required repair/services over a period of 1 year. The manufacturer, who has currently supplied 60 orders for the equipment, is concerned on the number of service calls that he may receive from hospitals that will use the equipment in the forthcoming year. If it is assumed that the probability that an equipment requires a repair/service is the same for all equipments, then the problem is to find a prediction interval for $Y \sim$ binomial $(60, p)$ based on the observed value 2 of $X \sim$ binomial $(40, p)$, where p is the probability of receiving a repair/service call for an equipment.

The 95% PIs based on various methods are as follows. The exact PI is **[0, 12]**, the equal-tailed PI is **[0, 11]**, and the one based on Wang et al. (2012) fiducial distribution (3) is also **[0, 11]**. The 95% HPM-PI is **[0, 9]** and the PI based on the joint sampling approach is **[0, 12]**. Notice that among all four PIs, the HPM-PI is the shortest and the exact PI is the widest.

Example 2 To illustrate the binomial prediction methods for large samples, we shall use the hearing screen outcome data collected over a 4-years period (January 1, 1993–

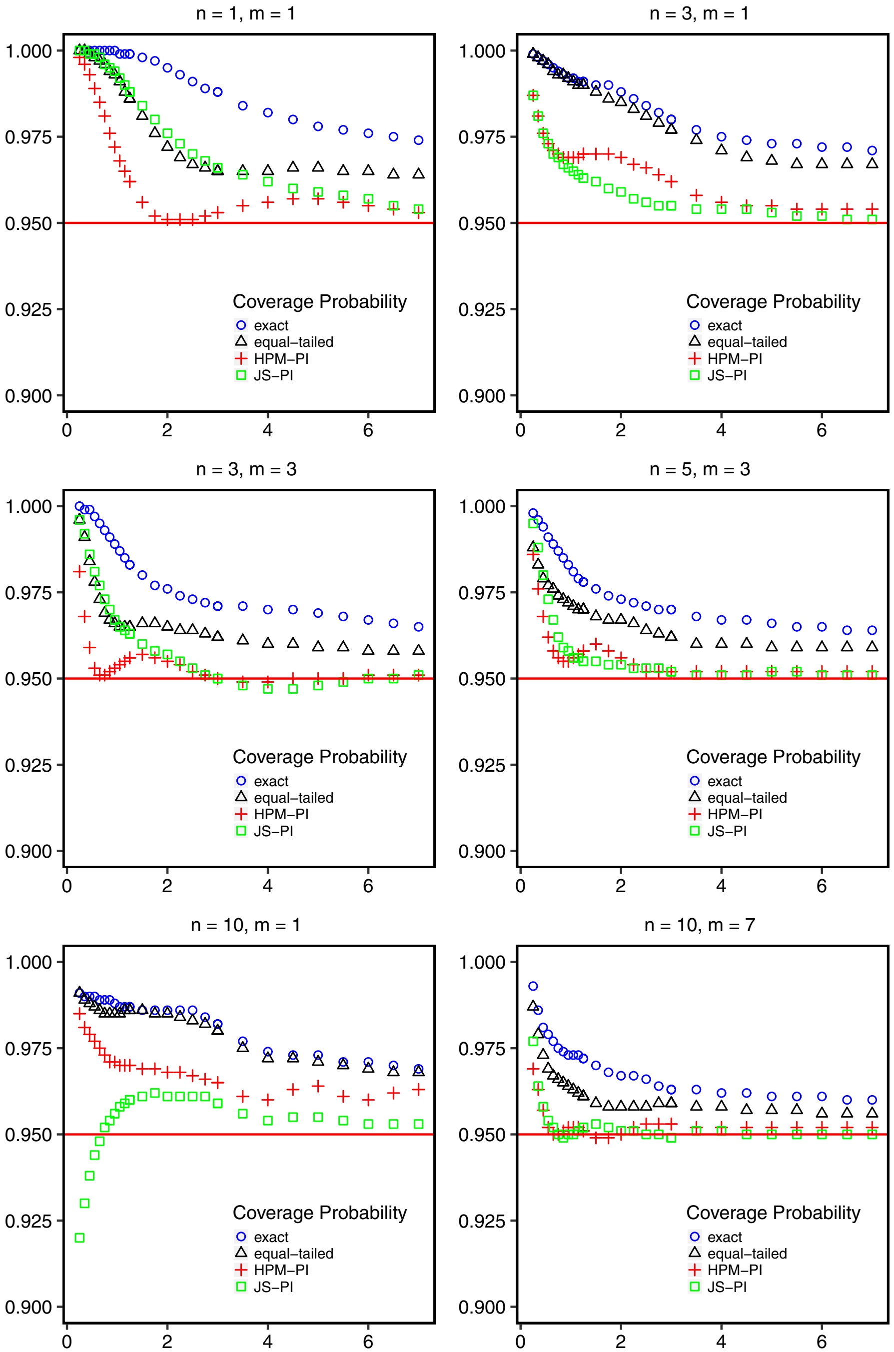

Fig. 4 Coverage probabilities of 95% PIs for a Poisson distribution as a function of λ

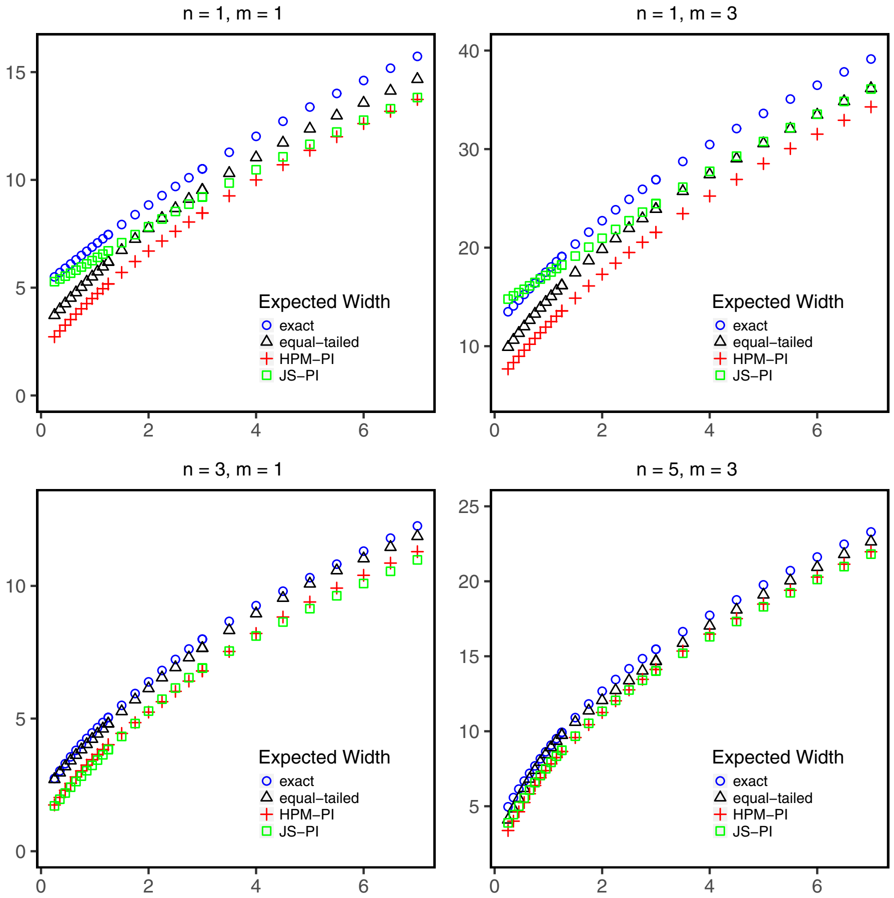

Fig. 5 Expected widths of 95% PIs for a Poisson distribution as a function of λ

December 31, 1996) in Rhode Island. The objective of the hearing screen program is to ensure that all infants with hearing loss are identified as early as possible in order to provide appropriate audiological, educational and medical intervention. The data were used in Wang (2010) for illustration purpose and originally reported in Vohr et al. (1998). A part of the data for years 1995 and 1996 are given in the following table.

	Year		Total
	1995	1996	
Number of nursery liveborn infants	12, 694	12, 236	24, 930
Infants with permanent hearing loss	20	18	38

Notice that the proportions for the years 1995 and 1996 are practically the same, and thus satisfying our model assumption. So, in order to judge the validity of the PIs, we shall use the data for the year 1995 to predict the number of infants with permanent hearing loss for the year 1996. In this case, n =12,694, x =20 and m =12,236. The 90% PIs by various methods are as follows. The exact PI is **[10, 31]**, the HPM-PI is **[9, 29]**, the equal-tailed PI is **[10, 31]**, the PI based on (3) is also **[10, 31]** and the JS-PI is **[11, 31]**. Thus, all PIs include 18, the actual number of infants with hearing loss in 1996. We also see that, even for large samples, the methods produced slightly different PIs. Furthermore, we notice that the HPM-PI and JS-PI are different with same width and are narrower than other PIs.

We shall now pool the data for years 1995 and 1996 to predict the number of infants with permanent hearing loss out of 10,000 liveborn infants in a future year. That is, we like to predict the number of infants with hearing loss out of $m = 10,000$ liveborn infants based on n =12,694+12,236=24,930 and x =20+18=38. The 90% PIs by various methods are as follows. The exact PI is **[8, 24]**, the HPM-PI is **[7, 22]**, the equal-tailed PI is **[8, 24]**, the PI by Wang et al. (2012) based on (3) is also **[8, 24]** and the JS-PI is **[8, 23]**. Among all PIs, the JS-PI and the HPM-PI are shorter than others. The 95% PIs are as follows. The exact PI is **[7, 26]**, the HPM-PI is **[7, 24]**, the equal-tailed PI is **[7, 25]**, the PI based on (3) is **[7, 25]** and the JS-PI is **[7, 25]**. For this case, the HPM-PI is the shortest.

Example 3 To illustrate the methods for finding Poisson PIs, we shall use the surface defects data in steel plates given in Montgomery (1996). We use a part of the data, as considered in Wang and Tsung (2009), for constructing Poisson PIs. The counts of surface defects on 21 steel plates are

```
1, 0, 4, 3, 1, 2, 0, 2, 1, 1, 0, 0, 2, 1, 3, 4, 3, 1, 0, 2, 4.
```

The mean surface defects on the basis of 21 plates is 1.6667, and the total is 35. The 95% HPM-PI for the number of surface defects in a future steel plate is **[0, 4]** and the one based on the joint sampling approach is also **[0, 4]**. The 95% exact and equal-tailed PIs are the same, which is **[0, 5]**.

Suppose it is desired predict the number of surface defects in a sample of 10 steel plates. To find PIs, note that $n = 21$, $x = 35$ and $m = 10$. On the basis of this data, we computed the 95% HPM-PI as **[7, 26]**, the JS-PI as **[8, 27]** and the exact and equal-tailed PIs are **[8, 28]**. Notice that the HPM-PI and JS-PI have the same expected width, and they are narrower than the equal-tailed PI.

5 Concluding remarks

Even though the fiducial inference was introduced in 1930's, it was not a well recognized method until recently; see the review article by Hannig et al. (2016). This inferential method was resurfaced in the name of "generalize variable approach" by Tsui and Weerahandi (1989) and Weerahandi (1993). Numerous articles have been published since then by showing applications to many complex problems. The beta fiducial distribution that we have used in this article has been already used to find simple and satisfactory confidence intervals for various function of binomial parameters such as the difference, relative risk and odds ratio. We here showed that the beta fiducial distribution produced satisfactory prediction intervals for the binomial case. The chi-square fiducial distribution for the Poisson case has also shown to produce simple yet accurate confidence intervals for estimating functions of Poisson means such as the ratio of Poisson means (Cox 1953) and weighted average of Poisson means (Krishnamoorthy and Lee 2010). The same chi-square fiducial distribution has produced better PIs for a future observation from a Poisson distribution. Thus, this article along with published articles in the literature clearly indicate the fiducial approach is so versatile, and produces results that are appreciably accurate even for small samples.

Acknowledgements The authors are grateful to two reviewers and the editor for providing valuable comments and suggestions.

Appendix A

We shall now provide an exact expression for the probability mass function of Y^* defined in (8). Note that

$$\begin{aligned}
P(Y^*=y) &= E_U P(Y^*=y|W)\\
&= E\left\{\binom{m}{y} W^y (1-W)^{m-y}\right\}\\
&= \binom{m}{y}\frac{1}{B(x+.5,n-x+.5)}\int_0^1 w^{y+x-.5}(1-w)^{n+m-y-x-.5}dw\\
&= \binom{m}{y}\frac{B(y+x+.5,n+m-y-x+.5)}{B(x+.5,n-x+.5)}\\
&= \frac{\Gamma(x+y+.5)}{\Gamma(y+1)\Gamma(x+.5)}\frac{\Gamma(m+n-x-y+.5)}{\Gamma(n-x+.5)\Gamma(m-y+1)}\Big/\binom{m+n}{m}.
\end{aligned}$$

Appendix B

We shall now provide an exact expression for the probability mass function of Y^* defined in (14) for predicting a future observation from a Poisson distribution. We first note that χ^2_a/b is distributed as a gamma random variable with shape parameter $a/2$ and the scale parameter $2/b$. Letting $a = 2x + 1$ and $b = 2n$, we see that

$$\begin{aligned}
P(Y^*=y) &= E_W P(Y^*=y|W)\\
&= \frac{(b/2)^{\frac{a}{2}}}{\Gamma(a/2)}\int_0^\infty \frac{e^{-mw}(mw)^y}{y!}e^{-bw/2}w^{a/2-1}dw\\
&= \frac{m^y(b/2)^{\frac{a}{2}}}{y!\Gamma\left(\frac{a}{2}\right)}\int_0^\infty e^{-w(m+b/2)}w^{a/2+y-1}dw\\
&= \frac{m^y(b/2)^{\frac{a}{2}}\Gamma(a/2+y)}{y!\Gamma(a/2)(m+b/2)^{a/2+y}}\\
&= \frac{\Gamma(x+y+.5)}{\Gamma(x+.5)\Gamma(y+1)}\left(\frac{m}{m+n}\right)^y\left(\frac{n}{m+n}\right)^{x+.5}.
\end{aligned}$$

References

Agresti A, Coull BA (1988) Approximate is better than “exact” for interval estimation of binomial proportion. Am Stat 52:119–125

Bain LJ, Patel JK (1993) Prediction intervals based on partial observations for some discrete distributions. IEEE Trans Reliab 42:459–463

Brown LD, Cai T, Das Gupta A (2001) Interval estimation for a binomial proportion (with discussion). Stat Sci 16:101–133

Cai T (2005) One-sided confidence intervals in discrete distributions. J Stat Plan Inference 131:63–88

Clopper CJ, Pearson ES (1934) The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26:404–413

Cox DR (1953) Some simple approximate tests for Poisson variates. Biometrika 40:354–360

Dawid AP, Stone M (1982) The functional-model basis of fiducial inference. Ann Stat 10:1054–1074

Efron B (1998) R. A. Fisher in the 21st century. Stat Sci 13:95–114

Faulkenberry GD (1973) A method of obtaining prediction intervals. J Am Stat Assoc 68:433–435

Fisher RA (1930) Inverse probability. In: Proceedings of the Cambridge philosophical society, London, vol xxvi, pp 528–535

Fisher RA (1935) The fiducial argument in statistical inference. Ann Eugen VI:91–98

Hahn GJ, Meeker WQ (1991) Statistical intervals: a guide to practitioners. Wiley, New York

Hannig J (2009) On generalized fiducial inference. Stat Sin 19:491–544

Hannig J, Iyer H, Lai RCS, Lee T (2016) Generalized fiducial inference: a review and new results. J Am Stat Assoc 111:1346–1361

Krishnamoorthy K, Lee M (2010) Inference for functions of parameters in discrete distributions based on fiducial approach: binomial and Poisson cases. J Stat Plan Inference 140:1182–1192

Krishnamoorthy K, Peng J (2011) Improved closed-form prediction intervals for binomial and Poisson distributions. J Stat Plan Inference 141:1709–1718

Krishnamoorthy K (2015) Handbook of statistical distributions with applications. Chapman & Hall/CRC, Boca Raton, FL

Krishnamoorthy K, Lee M, Zhang D (2017) Closed-form fiducial confidence intervals for some functions of independent binomial parameters with comparisons. Stat Methods Med Res 26:43–63

Montgomery DC (1996) Introduction to statistical quality control Hoboken. Wiley, New York

Nelson W (1982) Applied life data analysis. Wiley, Hoboken, NJ

Stevens WL (1950) Fiducial limits of the parameter of a discontinuous distribution. Biometrika 37:117–129

Thatcher AR (1964) Relationships between Bayesian and confidence limits for prediction. J R Stat Soc Ser. B 26:176–192

Thulin M (2014) The cost of using exact confidence intervals for a binomial proportion. Electron J Stat 8:817–840

Thulin M, Zwanzig S (2017) Exact confidence intervals and hypothesis tests for parameters of discrete distributions. Bernoulli 23:479–502

Tsui KW, Weerahandi S (1989) Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. J Am Stat Assoc 84:602–607

Vohr BR, Carty LM, Moore PE, Letourneau K (1998) The Rhode island hearing assessment program: experience with statewide hearing screening (1993–1996). J Pediatr 133:353–357

Wang H (2008) Coverage probability of Prediction intervals for discrete random variables. Comput Stat Data Anal 53:17–26

Wang H (2010) Closed form prediction intervals applied for disease counts. Am Stat 64:250–256

Wang H, Tsung F (2009) Tolerance intervals with improved coverage probabilities for binomial and Poisson variables. Technometrics 51:25–33

Wang CM, Hannig J, Iyer HK (2012) Fiducial prediction intervals. J Stat Plan Inference 142:1980–1990

Weerahandi S (1993) Generalized confidence intervals. J Am Stat Assoc 88:899–905