Closed Form Prediction Intervals Applied for Disease Counts

Hsiuying Wang

Institute of Statistics

National Chiao Tung University

Hsinchu, Taiwan

wang@stat.nctu.edu.tw

Abstract

The prediction interval is an important tool in medical applications for predicting the number of times a disease will occur in a population. The performance of the existing prediction intervals, however, are unsatisfactory when the true proportion is near a boundary. Since the true proportion can be very small in real applications, in this paper, we propose improved prediction intervals with better coverage probability than the existing methods. Their predictive distributions are compared in terms of the Kullback-Leibler distance and the intervals are compared using a hearing screening medical example.

Key words: binomial distribution, coverage probability, prediction interval, predictive distribution

1 Introduction

The prediction interval (PI) is a very useful tool to predict future observations. We consider predicting the disease count in a population for medical applications. Since the number of diseased patients in a population follows a binomial distribution, in this paper, we investigate prediction intervals for the binomial distribution.

The construction of prediction intervals for continuous distributions has been extensively studied in the literature (Basu, Ghosh and Mukerjee 2003; Hall and Rieck 2001; Hamada, Johnson and Moore 2004; Lawless and Fredette 2005; Olive 2007; Cai, Tian, Solomon and Wei 2008; Patel 1989). However, compared with the continuous distributions, there are fewer investigations for discrete distributions. The most widely used closed form prediction interval for a binomial random variable was proposed by Nelson (1982). Another prediction interval with a closed form was proposed by Bain and Patel (1993). In addition, prediction intervals with associated numerical calculation to achieve a desired coverage probability were introduced in Patel and Samaranayake (1991) and Wang (2008). Although the last two approaches can provide accurate coverage probabilities for the prediction intervals, they heavily rely on numerical calculations and can not provide closed forms. Since a prediction interval with a closed form can be easily

employed in applications, in this paper, we explore approximate prediction intervals with a closed form.

The coverage probabilities for the Nelson interval and the Bain and Patel interval do not perform well when the true binomial proportion is near the boundaries because their coverage probabilities are much lower than the nominal level as the binomial proportion goes to 0 or 1. In addition, the average coverage probabilities of these two intervals, averaged over the parameter space, are also unsatisfactory. When the sample size is not large, the average coverage probabilities of these two intervals are much lower than the nominal level based on a simulation study.

In this paper, two improved prediction intervals are proposed by inverting the score test and by adjusting an existing interval. The coverage probabilities of these two proposed prediction intervals are significantly higher than those of the existing intervals when the true proportion is close to the boundaries.

In addition, the two new intervals are evaluated by comparing their corresponding predictive distributions in terms of the Kullback-Leibler distance. The calculation results show that the distance between the score predictive distribution and the binomial distribution is smaller than that between the adjusted predictive distribution and the binomial distribution.

2   Existing prediction interval

We present several existing prediction intervals in this section. The first of these is the prediction interval for a binomial random variable constructed by Nelson (1982), which is reviewed in Hahn and Meeker (1991). Suppose that the past data consist of X successes out of n trials from a $B(n, p)$ distribution with a success probability p, $0 < p < 1$. Let Y be the future number of successes out of m trials from a $B(m, p)$ distribution. A large-sample approximate level γ two-sided prediction interval $(L(X), U(X))$ for the future number Y of occurrences based on the observed value of the number X of the past occurrences for the binomial distribution constructed by Nelson (1982) is

$$\hat{Y} \pm z_{(1+\gamma)/2}(m\hat{p}(1 - \hat{p})(m + n)/n)^{1/2} \tag{1}$$

where $\hat{p} = X/n$ and $\hat{Y} = m\hat{p}$ when $X, n - X, Y$ and $m - Y$ all are large. Here $z_{(1+\gamma)/2}$ denotes the upper $(1 + \gamma)/2$ quantile of the standard normal distribution. Note that the true coverage probability of the interval $(L(X), U(X))$ at $p = p_0$ is defined as the probability $P_{p_0}(L(X) < Y < U(X))$.

The second level γ prediction interval was proposed by Patel and Samaranayake (1991). This uses the form $(0, X + d)$ as an upper prediction interval or $(X - d, m)$ as a lower prediction interval for Y, where d is a positive integer. To guarantee that the coverage

4

probability of the upper prediction interval $(0, X + d)$ is greater than or equal to γ, the exact coverage probability of the interval is derived and it is necessary to find a d such that its coverage probability is greater than or equal to γ for all p. It turns out that the derivation of d is to find the smallest integer d satisfying

$$Inf_{0 \leq p \leq 1} \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x} (\sum_{y=0}^{x+d} \binom{m}{y} p^y (1-p)^{m-y}) \geq \gamma.$$

The value of d can be exactly derived only for the case of $m = n$ and an approximated value of d can be obtained numerically for the case of $m \neq n$. A similar argument is applied for the lower prediction bound.

The third approximate level γ prediction interval was proposed by Bain and Patel (1993). This approach considers a conditional distribution for some functions of X and Y to eliminate the unknown parameter, and then uses the conditional distribution to derive the predictive limits.

The interval has the form

$$(T_L - X, T_U - X), \tag{2}$$

where

$$T_L = \frac{(2X_1 v + sw) - \sqrt{s^2 w^2 + 4X_1 w(n - X_1)}}{2(v^2 + w)},$$

$$T_U = \frac{(2X_2 v + sw) + \sqrt{s^2 w^2 + 4X_2 w(n - X_2)}}{2(v^2 + w)},$$

$s = n + m, v = n/s, w = z^2_{(1+\gamma)/2} v(1 - v)/(s - 1), X_1 = X - 1/2$ and $X_2 = X + 1/2$.

In addition to these existing prediction intervals, Wang (2008) proposed procedures to calculate the minimum coverage probability and average coverage probability for a prediction interval. Based on those procedures, the factor $z_{(1+\gamma)/2}$ can be adjusted to obtain the prediction interval with either a desirable minimum coverage probability or a desirable average coverage probability.

As mentioned in the introduction section, in this paper we mainly focus on the intervals with closed forms. The performance of the two existing prediction intervals with closed forms (1) and (2) in terms of their coverage probabilities are discussed as follows.

Figures 1 and 2 give the coverage probabilities and expected lengths of the Nelson and the Bain and Patel prediction intervals for different sample size n when m is fixed at 50. It is seen that the coverage probabilities of these existing intervals are far from the nominal level when p is near the boundaries. Since the true binomial proportion in real applications may be close to the boundaries, the behavior near a boundary is important. When p is not close to the boundaries, the coverage probability of the Nelson interval is lower than the nominal level 0.95. In contrast, the coverage probability of the Bain and Patel interval is higher than the nominal level 0.95 when p is not near a boundary, but it is lower than 0.95 for p near boundaries when the sample size is not large enough. Overall, in

addition to the poor performance for p near the boundaries, the existing methods cannot achieve the desirable coverage probability or are too conservative.

Analyzing the Nelson's interval, the form is derived from the fact that

$$\frac{Y - m\hat{p}}{\sqrt{\hat{p}(1 - \hat{p})m(m + n)/n}} \tag{3}$$

is approximately $N(0, 1)$ distributed. This is similar to the construction of the Wald confidence interval for a binomial proportion p, which is

$$\hat{p} \pm z_{(1+\gamma)/2}\sqrt{\hat{p}(1 - \hat{p})/n}. \tag{4}$$

It is well known that the coverage probability of the Wald interval is much lower than the nominal level for a binomial distribution when the true proportion is close to a boundary (Wang 2007). This unsatisfactory property also occurs at the prediction interval construction if we simply employ the Wald approach. To obtain prediction intervals with better performance when the true proportion is near a boundary, we can use similar approaches, such as the score approach or the Agresti-Coull approach (Agresti and Coull 1998) for improving the coverage probabilities of confidence intervals (Brown, Cai and DasGupta 2001), to solve the problem. Agresti and Caffo (2000) and Pires and Amado (2008) also provide some discussions and comparisons of the confidence intervals for the binomial proportion. In the next section, two improved confidence intervals in the literature for the

binomial distribution are introduced, and improved prediction intervals based on similar

approaches are proposed.

3 Improved prediction intervals

In this section, we introduce two alternative confidence intervals for a binomial proportion

and use similar approaches to construct improved prediction intervals for a binomial

random variable.

The two alternative confidence intervals discussed in Agresti and Coull (1998), Brown,

Cai and DasGupta (2002), Wilson (1927) and Wang (2007) are as follows.

1. The Wilson interval. Let $\tilde{X} = X + z^2_{(1+\gamma)/2}/2$ and $\tilde{n} = n + z^2_{(1+\gamma)/2}$. Let $\tilde{p} = \tilde{X}/\tilde{n}$,

   $\tilde{q} = 1 - \tilde{p}$, $\hat{p} = X/n$ and $\hat{q} = 1 - \hat{p}$. The level γ Wilson interval has the form

   $$CI_W(X) = \tilde{p} \pm \frac{z_{(1+\gamma)/2}n^{1/2}}{\tilde{n}} \left(\hat{p}\hat{q} + \frac{z^2_{(1+\gamma)/2}}{4n} \right)^{1/2}.$$

2. The Agresti-Coull interval. The level γ Agresti-Coull interval is

   $$CI_{AC}(X) = \tilde{p} \pm z_{(1+\gamma)/2} \left(\tilde{p}\tilde{q} \right)^{1/2} \tilde{n}^{-1/2},$$

where the notations are as in the case 1 for the Wilson interval.

The Wilson and Agresti-Coull intervals successfully increase the coverage probability

for p near boundaries, compared with the Wald confidence interval. The Wilson interval

8

is derived by replacing $\hat{p}$ by p in the Wald interval, and then solving p from the equation

$p = \hat{p} \pm z_{(1+\gamma)/2}\sqrt{p(1-p)/n}$, which is the inversion of the score test. The Agresti-Coull

interval uses the approach of adding two successes and two failures to adjust the Wald

interval.

Remark 1. There are two other confidence intervals, likelihood ratio and Bayesian

credible intervals, discussed in Brown et al. (2002). Since the likelihood ratio interval

does not have a closed form and the minimum coverage probability of the credible interval

is zero (Wang 2007), we do not consider these two intervals here.

To construct the first proposed prediction interval, we employ an approach similar

to the construction of the Wilson interval. We replace $\hat{p}$ by $(X + Y)/(m + n)$ in the

denominator of (3) and use the fact that the random variable

$$\frac{Y - m\hat{p}}{\sqrt{\frac{(X+Y)}{(n+m)}\left(1 - \frac{(X+Y)}{(n+m)}\right)\frac{m(m+n)}{n}}} \tag{5}$$

is approximately $N(0,1)$ distributed. To avoid the poor coverage probability when the

parameter is near the boundaries, we invert

$$\{y : y = m\hat{p} \pm z_{(1+\gamma)/2}\sqrt{W(x,y)}\}, \tag{6}$$

to derive the prediction limits instead of inverting

$$\{y : y = m\hat{p} \pm z_{(1+\gamma)/2}\sqrt{\frac{(x+y)}{(n+m)}\left(1 - \frac{(x+y)}{(n+m)}\right)\frac{m(m+n)}{n}}\}, \tag{7}$$

9

where

$$W(x,y) = \frac{(x + z^2_{(1+\gamma)/2}/2 + y)}{(n + z^2_{(1+\gamma)/2} + m)}(1 - \frac{(x + z^2_{(1+\gamma)/2}/2 + y)}{(n + z^2_{(1+\gamma)/2} + m)})\frac{m(m+n)}{n}.$$

Note that the form of $W(x,y)$ adds $z^2_{(1+\gamma)/2}/2$ to x and $z^2_{(1+\gamma)/2}$ to n in the square root

term in (7). This modification prevents the interval (6) from shrinking to the empty set

when $x = y = 0$.

The two solutions of y in (6) are the proposed lower prediction limit $L_s(X)$ and the

upper prediction limit $U_s(X)$, which are

$$\frac{A}{C} \pm \frac{B}{C}, \tag{8}$$

where

$$A = mn[2xz^2_{(1+\gamma)/2}(n + z^2_{(1+\gamma)/2} + m) + (2x + z^2_{(1+\gamma)/2})(m+n)^2]$$

$$B = (mn(m+n)z^2_{(1+\gamma)/2}(m+n+z^2_{(1+\gamma)/2})^2(2(n-x)[n^2(2x + z^2_{(1+\gamma)/2}) + 4mnx + 2m^2x]$$

$$+nz^2_{(1+\gamma)/2}[n(2x + z^2_{(1+\gamma)/2}) + 3mn + m^2]))^{1/2}$$

and

$$C = 2n[(n + z^2_{(1+\gamma)/2})(m^2 + n(n + z^2_{(1+\gamma)/2})) + mn(2n + 3z^2_{(1+\gamma)/2})].$$

Since this approach is similar to constructing the score confidence interval, we call

this interval the score prediction interval.

In addition to the above approach, to avoid the poor performance of p near the boundaries, we can adjust the usual prediction interval (1) by replacing $\hat{p}$ with $\tilde{p}$, which leads to the second proposed interval $(L_a(X), U_a(X))$:

$$\hat{Y} \pm z_{(1+\gamma)/2}(m\tilde{p}(1-\tilde{p})(m+n)/n)^{1/2}. \tag{9}$$

Note that here we do not consider replacing $\hat{p}$ in $\hat{Y}$ by $\tilde{p}$ because the expectation $E_p(Y - m\tilde{p})$ is not zero. If we replace $\hat{p}$ in $\hat{Y}$ by $\tilde{p}$, the Kullback -Leibler distance discussed in Section 4 diverges as the sample size increases. This interval basically uses a method similar to the Agresti and Coull confidence interval, where $\tilde{p}$ is used as an estimator of p instead of $\hat{p}$ to overcome the problem of the poor behavior of the Wald interval. We call the second proposed interval the adjusted prediction interval.

The performance of the score and adjusted prediction intervals are presented in Figures 3 and 4. The coverage probabilities of the proposed intervals are decreasing in p when the proportion is near 0 and are increasing in p when the proportion is near 1. The coverage probabilities are close to the nominal level for p in an interval with a center at $p = 0.5$. The proposed intervals have the advantage of higher coverage probability when p is near the boundaries in which case the performance of the coverage probabilities of the existing intervals are unsatisfactory. In addition, the score interval has shorter expected length than the other intervals.

Remark 2. The coverage probabilities presented in Figures 1-4 are the exact coverage probabilities calculated by the definition. Since the performance of the coverage probabilities are significantly different for different intervals when p goes to the boundaries, to clarify the presentation, we use different scales for the y-axis in these figures.

Remark 3. Since the value of Y is from 0 to m, suitable modifications for the intervals (8) and (9) are $[max(0, L_s(X)), min(U_s(X), m)]$ and $[max(0, L_a(X)), min(U_a(X), m)]$, respectively. However, since the existing intervals do not use a modified form, for a fair comparison, we still use the original form of the proposed interval for investigation in this study.

4 Predictive distribution

The new prediction intervals can be evaluated by the criterion of the predictive distribution estimation. The true distribution of Y is the binomial distribution. Since the two proposed intervals are constructed using the normal approximation, the degree of approximation can be measured by comparing these normal approximations with the true binomial distribution.

There is a large literature on the predictive distribution estimation, for example, Aitchison (1995), Murray (1997), Ng (1980), Lejeune and Faulkenberry (1982), Harris

(1989) and Lawless and Fredette (2005). One method of constructing a predictive distribution from a predictive limit is treating α prediction limits as the α quantiles in the predictive distribution function.

Note that the true probability mass function of the future observation Y is

$$f_p(y) = \binom{m}{y} p^y (1-p)^{m-y}. \tag{10}$$

Based on (6) and (9), let $f^s(y|x)$ and $f^a(y|x)$ denote the predictive densities derived by the score and adjusted predictive limits using the plug-in estimators, which indicates that $f^s(y|x)$ and $f^a(y|x)$ are the density functions of the normal distributions $N(m\hat{p}, W(x, y))$ and $N(m\hat{p}, \tilde{p}(1-\tilde{p})m(m+n)/n)$.

An approach to evaluate a predictive distribution is to measure the goodness of the predictive distribution in terms of the Kullback-Leibler distance between $\tilde{f}(y|x)$ and $f_p(y)$,

$$E_X(\sum_{y=0}^{m} f_p(y) log\{\frac{f_p(y)}{\tilde{f}(y|x)}\}), \tag{11}$$

where $\tilde{f}(y|x)$ is a predictive density estimator. See, for example, Lawless and Fredette (2005).

Remark 4. Note that the variances of the two normal approximations are not close to that of the binomial distribution $B(n, p)$ when n is not large enough. It is mainly because the mean $m\hat{p}$ is a random variable, but not a constant mp. Since the mean of

$m\hat{p}$ is mp, we still can use the Kullback-Leibler distance between a predictive distribution and the binomial distribution to evaluate the performance of the predictive distribution.

The Kullback-Leibler distances of $f^s(y|x)$ and $f^a(y|x)$ to (10) are

$$E_X(\sum_{y=0}^{m} f_p(y)log\{\frac{f_p(y)}{f^s(y|x)}\}) \tag{12}$$

and

$$E_X(\sum_{y=0}^{m} f_p(y)log\{\frac{f_p(y)}{f^a(y|x)}\}). \tag{13}$$

Comparisons of the Kullback-Leibler distances for different sample sizes are shown in Figure 5. It can be seen that the predictive distribution derived from the score intervals can approximate the true binomial distribution more accurately than that derived from the adjusted interval.

Theorem 1 shows that the variance of the distribution with respect to the density function $f^s(y|x)$ is closer to the true variance than that of the distribution with respect to $f^a(y|x)$. This can provide an intuitive explanation for the results in Figure 5.

Theorem 1 *The variance of the true distribution for Y, $mp(1-p)$, is closer to the expectation of the variance estimator $W(X,Y)$ than to the expectation of $\tilde{p}(1-\tilde{p})m(m+n)/n$. That is,*

$$|E(W(X,Y)) - mp(1-p)| < |E(\tilde{p}(1-\tilde{p})m(m+n)/n) - mp(1-p)|. \tag{14}$$

14

The proof of Theorem 1 can be obtained by straightforward calculations.

Note that here we do not list the Kullback-Leibler distance of the predictive distribution derived from the Nelson interval because its Kullback-Leibler distance is divergent. Since the predictive density function derived from it is

$$\frac{1}{\sqrt{2\pi\hat{p}(1-\hat{p})m(m+n)/n}}e^{\frac{-(Y-m\hat{p})^2}{2\hat{p}(1-\hat{p})m(m+n)/n}}, \tag{15}$$

when $x = 0$, the denominator of (15) is equal to zero. Thus, it leads to an infinite Kullback-Leibler distance. From the Kullback-Leibler distance criterion, the proposed intervals with finite Kullback-Leibler distances are better than the Nelson interval. In addition, since the derivation of Bain and Patel interval is not directly derived by the normal approximation, we cannot directly obtain its predictive distribution.

5 Applications

In this paper, we take the example of a hearing screening program for all births with transient evoked otoacoustic emissions in all 8 maternity hospitals in the state of Rhode Island over a 4-year period during 1993-1996 as an application of the binomial prediction interval. The goal of this hearing screening program is to ensure that all infants and toddlers with hearing loss are identified as early as possible and provided with timely and

appropriate audiological, educational, and medical intervention. This example contains hearing screening data collected prospectively for 47991 normal nursery liveborns born in Rhode Island between January 1, 1993 and December 31, 1996 (Vohr, et al. 1998). The prediction interval can be used to predict the number of children with hearing loss for future years. Since the time period considered here is not large, we can assume that the number of children with hearing loss follows the same binomial distribution in each year.

Table 1 lists the numbers of all births and infants with permanent hearing loss, respectively for each year during 1993-1996.

Table 1. Screening demographics between 1993 and 1996

Year	1993	1994	1995	1996	Total
Normal nursery liveborns	9885	13176	12694	12236	47991
Identified with permanent hearing loss	11	12	20	18	61

To compare the performance of the prediction intervals, we use the observations of the two years 1993 and 1994 for the normal nursery liveborns to predict the number of infants with hearing loss for the future two years 1995 and 1996. The total number of the normal nursery liveborns for 1993 and 1994 is 23061, and the total number of the infants with hearing loss for these two years is 23. Assume that the number of the infants with hearing loss follows a binomial distribution. The level 0.9 Nelson interval,

Bain and Patel interval, score interval and adjusted interval, based on the first two year observations, for the number of the infants with hearing loss for the future two years 1995 and 1996 are $(13.07, 36.66)$, $(13.52, 39.36)$, $(14.27, 38.36)$ and $(12.73, 37.00)$, respectively, where $z_{(1+\gamma)/2} = 1.64$ in these prediction intervals. However, according to the data, the true total number of the infants with hearing loss of the future two years 1995 and 1996 was 38, which does not belong to the Nelson interval or the adjusted interval, but it does fall into the Bain and Patel interval and the score interval.

To predict the number of the infants with hearing loss for the year 1995 based on the data from 1993 and 1994, we obtain that the 0.9 level Nelson prediction interval, Bain and Patel interval, score prediction interval and adjusted prediction interval are $(5.4, 19.92), (5.4, 21.55), (5.96, 20.83)$ and $(5.19, 20.13)$, respectively. The Bain and Patel, score and adjusted intervals cover the true number 20, but the Nelson interval does not cover the true number 20.

It reveals that the performance of the score predictive interval is better than the Nelson interval in this application which assumes that the model that the binomial distribution in each year is the same is true. A comparison of the score and adjusted prediction intervals reveals that the theoretical comparison of Kullback-Leibler distances for the two predictive distributions is consistent with the comparison from this application example.

6　Conclusion

This paper proposes two improved prediction intervals, the score prediction interval and the adjusted prediction interval, with closed forms for predicting disease count. Both of them can increase the coverage probability when p is close to the boundaries compared with the existing prediction intervals. A simulation study shows the score interval has the shortest expected length of these intervals. The two new intervals are also evaluated in terms of the Kullback-Leibler distance through the predictive distributions. The comparison shows the predictive distribution corresponding to the score interval can approximate the binomial distribution better than that corresponding to the adjusted prediction interval.

In addition, to obtain more accurate results, we can employ the procedure of Wang (2008) to derive an appropriate value of $z_{(1+\gamma)/2}$ such that the prediction intervals can achieve either a desired minimum coverage probability or a desired average coverage probability.

References

[1] Aitchison, J. (1975), "Goodness of prediction fit," *Biometrika*, 62, 547-554.

[2] Agresti, A. and Caffo, B. (2000), "Simple and effective confidence intervals for proportions and difference of proportions result by adding two successes and two failures," *The American Statistican*, 54, 280-288.

[3] Agresti, A. and Coull, B. (1998), "Approximate is better than 'exact' for interval estimation of binomial proportions," *The American Statistican*, 52, 119-126.

[4] Bain, L. J. and Patel, J. K. (1993), "Prediction intervals based on partial observations for some discrete distributions," *IEEE Transactions on Reliability*, 42, 459-463.

[5] Basu, R., Ghosh, J. K., Mukerjee, R. (2003), " Empirical Bayes prediction intervals in a normal regression model: higher order asymptotics," *Statistics and Probability Letters*, 63, 197-203.

[6] Brown, L. D., Cai, T., DasGupta, A. (2000), " Confidence intervals for a binomial and asymptotic expansions," *Annals of Statistics*, 30, 160-201.

[7] Brown, L. D., Cai, T., DasGupta, A. (2001), " Interval Estimation for a Binomial Proportion," *Statistical Science*, 16, 101-133.

[8] Cai, T., Tian, L., Solomon, S. D. and Wei, L. J. (2008), " Predicting future responses based on possibly mis-specified working models", *Biometrika*, 95, 75-92.

[9] Hahn, G. J., Meeker, W. Q. (1991), "Statistical Intervals: A Guide for Practitioners", *Wiley Series*.

[10] Hall, P., Rieck, A. (2001), "Improving coverage accuracy of nonparametric prediction intervals", *Journal of Royal Statistical Society, Series B*, 63, 717-725.

[11] Hamada, M., Johnson, V., Moore, L. M. Wendelberger, J. (2004), " Bayesian prediction intervals and their relationship to tolerance intervals", *Technometrics*, 46, 452-459.

[12] Harris, I. R. (1989)," Predictive fit for natural exponential families", *Biometrika*, 76, 675-684.

[13] Lawless, J. F., Fredette, M. (2005), "Frequentist prediction intervals and predictive distributions", *Biometrika*, 92, 529-542.

[14] Lejeune, M., Faulkenberry, D. G. (1982)," A simple predictive density function", *Journal of the American Statistical Association*, 77, 654-657.

[15] Murray, G. D. (1977), " A note on the estimation of probability density functions", *Biometrika*, 64, 150-152.

[16] Nelson, W. "Applied life data analysis", *Wiley, N. Y.* 1982.

[17] Ng, V. M. (1980). "On the estimation of parametric density functions", *Biometrika*, 67, 505-506.

[18] Olive, D. J. (2007), " Prediction intervals for regression models", *Computational Statistics and Data Analysis*, 51, 3115-3122.

[19] Patel, J. K. (1989), " Prediction intervals—a review". *Communications in Statistics: Theory and Methods*, 18 , 2393-2465.

[20] Patel, J., Samaranayake, V.A. (1991), "Prediction Intervals for Some Discrete Distributions", *Journal of Quality Technology*, 23, 270-278.

[21] Pires, A. M. and Amado, C. (2008), " Interval estimators for a binomial proportion: comparison of twenty method", *REVSTAT-Statistical Journal*, 6, 165-197.

[22] Vohr, B.R., Carty, L. M., Moore, P. E., Letourneau, K. (1998), " The Rhode Island Hearing Assessment Program: Experience with statewide hearing screening (1993-1996)", *The Journal of Pediatrics*, 133(3), 353-357.

[23] Wang, H. (2007), " Exact confidence coefficients of confidence intervals for a binomial proportion", *Statistica Sinica*, 17, 361-368.

[24] Wang, H. (2008), " Coverage probability of prediction intervals for discrete random variables," *Computational Statistics and Data Analysis*, 53, 17-26.

[25] Wilson, E. B. (1927), " Probable inference, the low of succession, and statistical inference", *Journal of the American Statistical Association*, 22, 209-212.
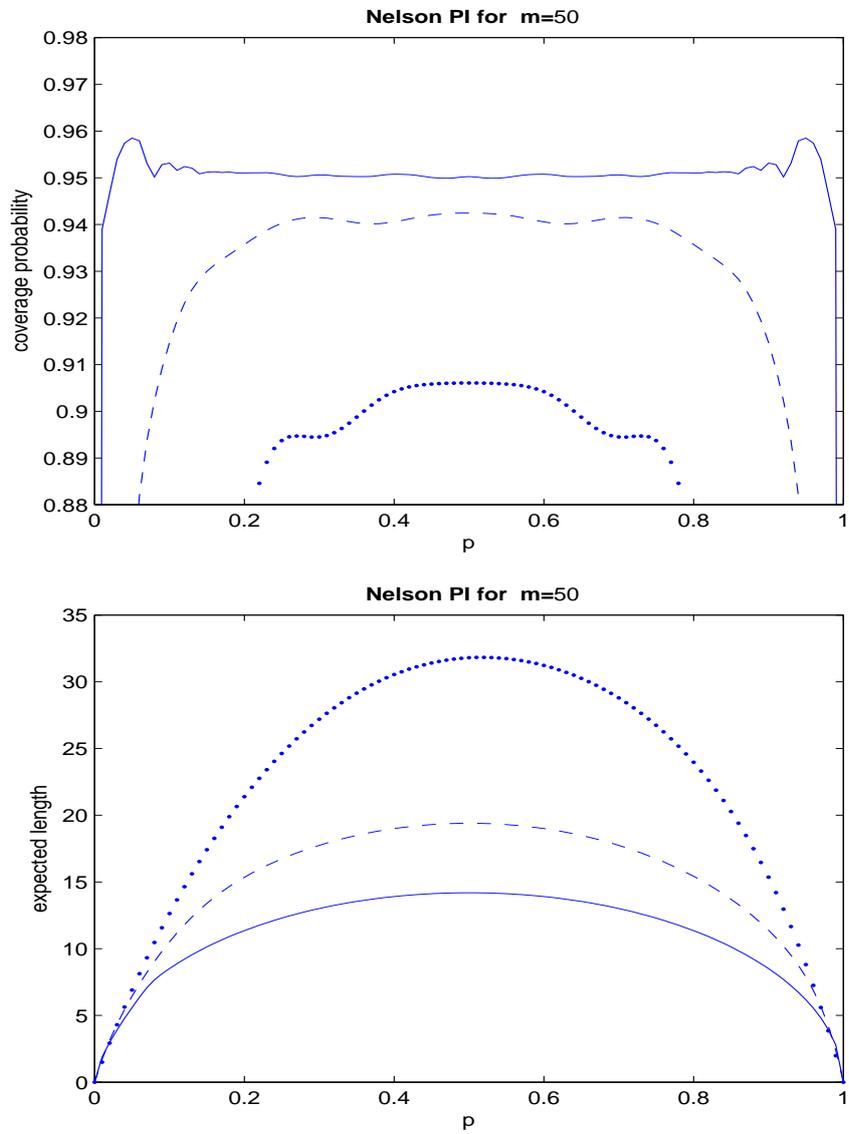
Figure 1: Coverage probabilities and expected lengths of the 95% level Nelson prediction intervals for the Binomial distributions with $n = 10$(dotted line), $n = 50$(dashed line) and $n = 1000$(solid line).
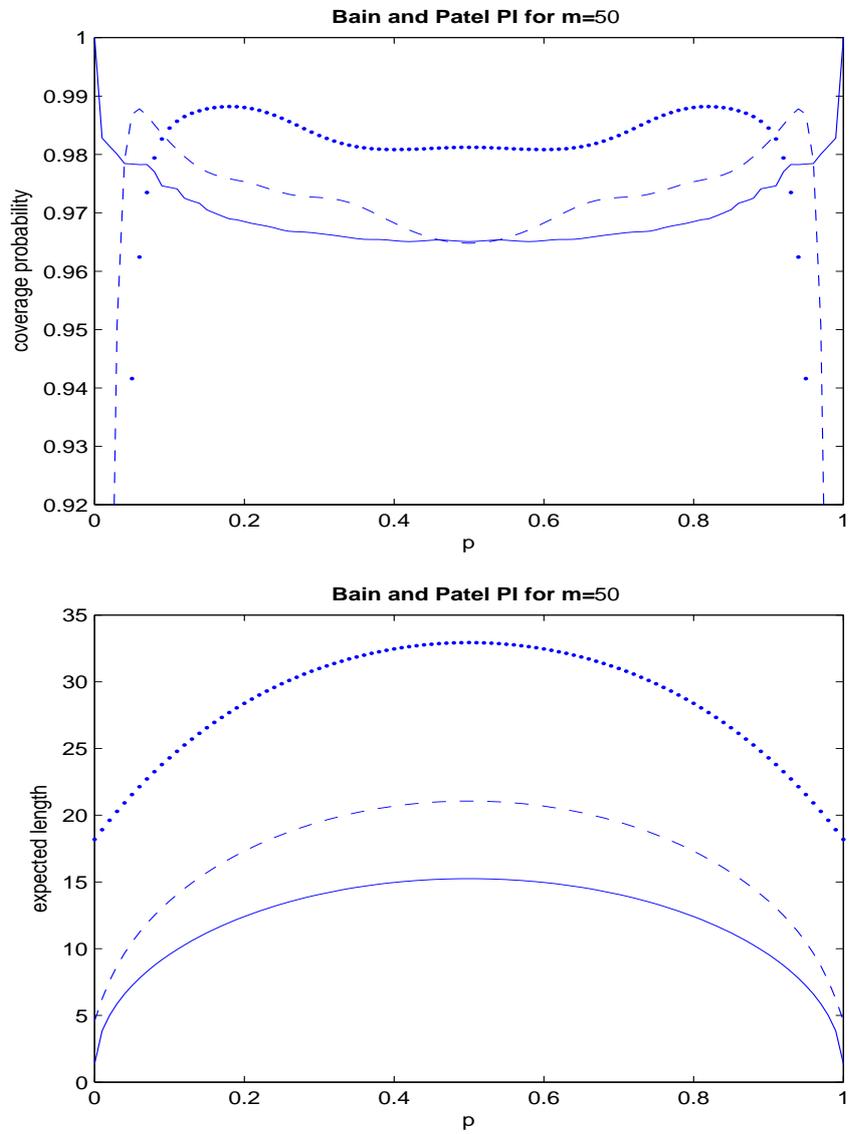
Figure 2: Coverage probabilities and expected lengths of the 95% level Bain and Patel prediction intervals for the Binomial distributions with $n = 10$(dotted line), $n = 50$(dashed line) and $n = 1000$(solid line).
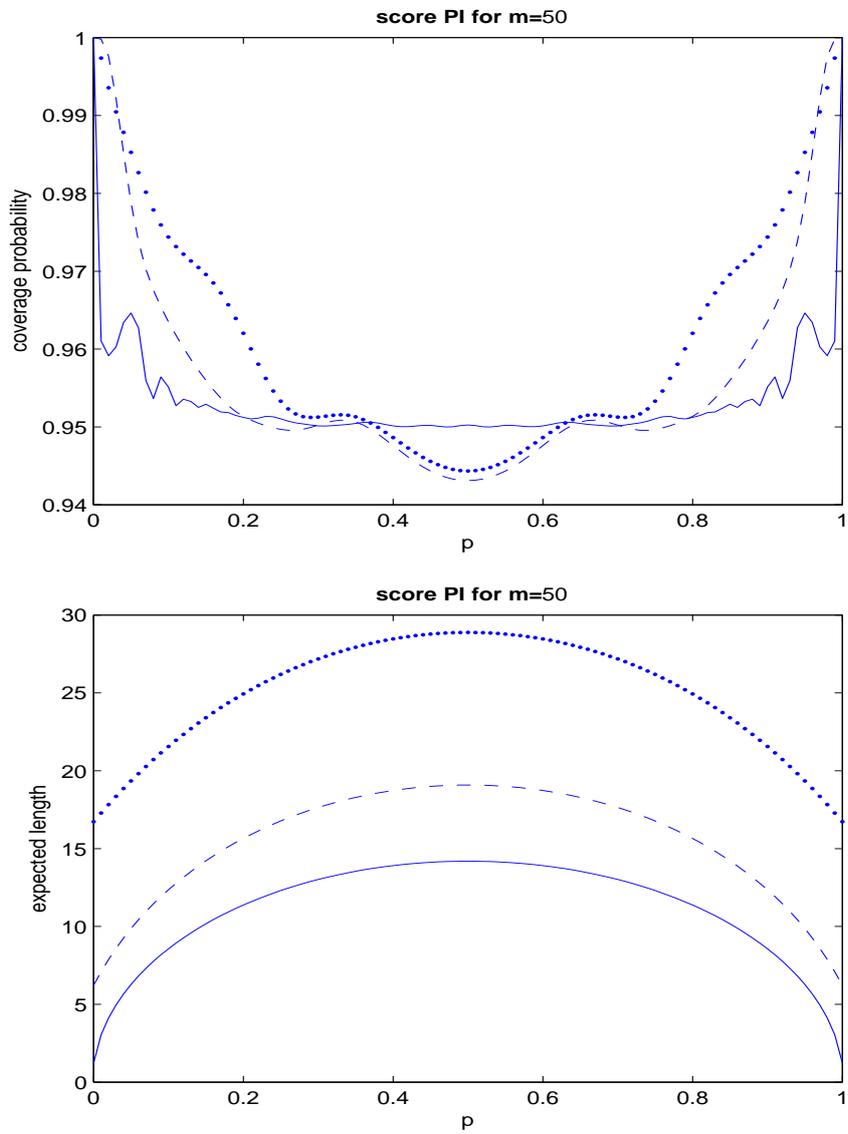
Figure 3: Coverage probabilities and expected lengths of the 95% level score prediction intervals for the Binomial distributions with $n = 10$(dotted line), $n = 50$(dashed line) and $n = 1000$(solid line).
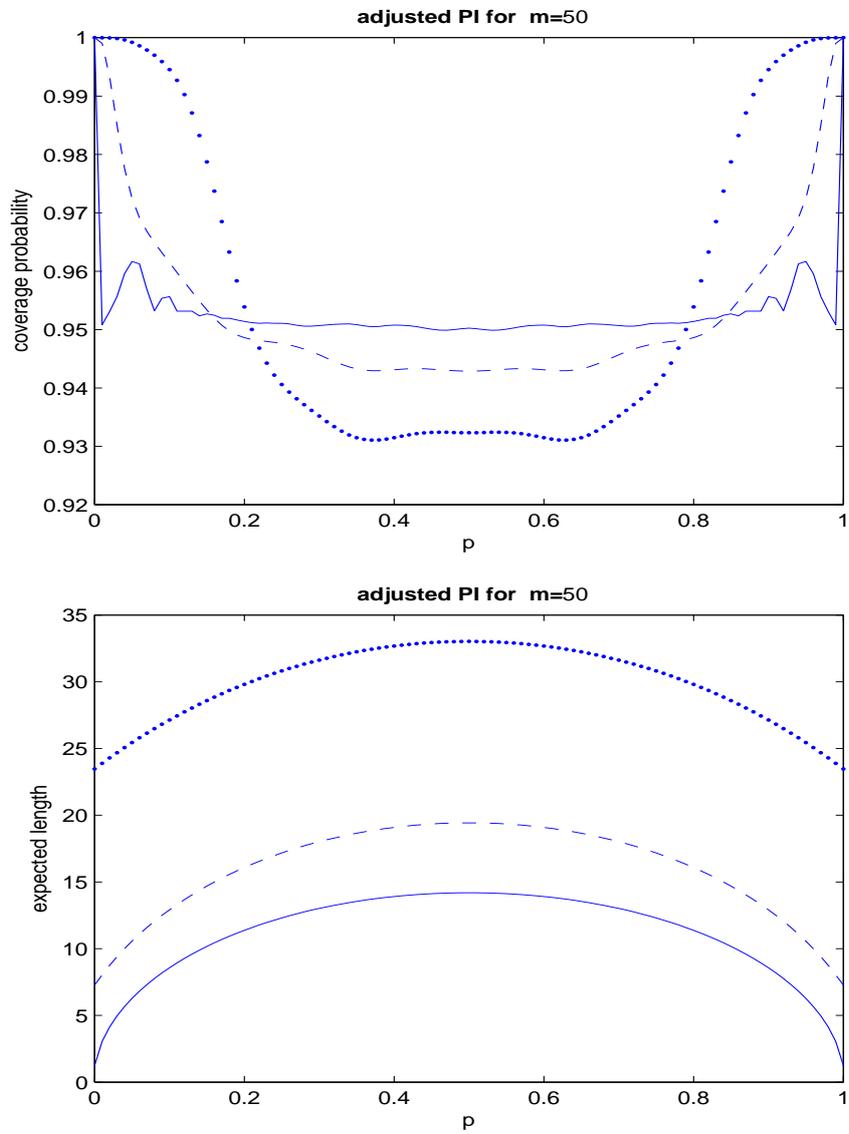
Figure 4: Coverage probabilities and expected lengths of the 95% level adjusted prediction intervals for the Binomial distributions with $n = 10$(dotted line), $n = 50$(dashed line) and $n = 1000$(solid line).
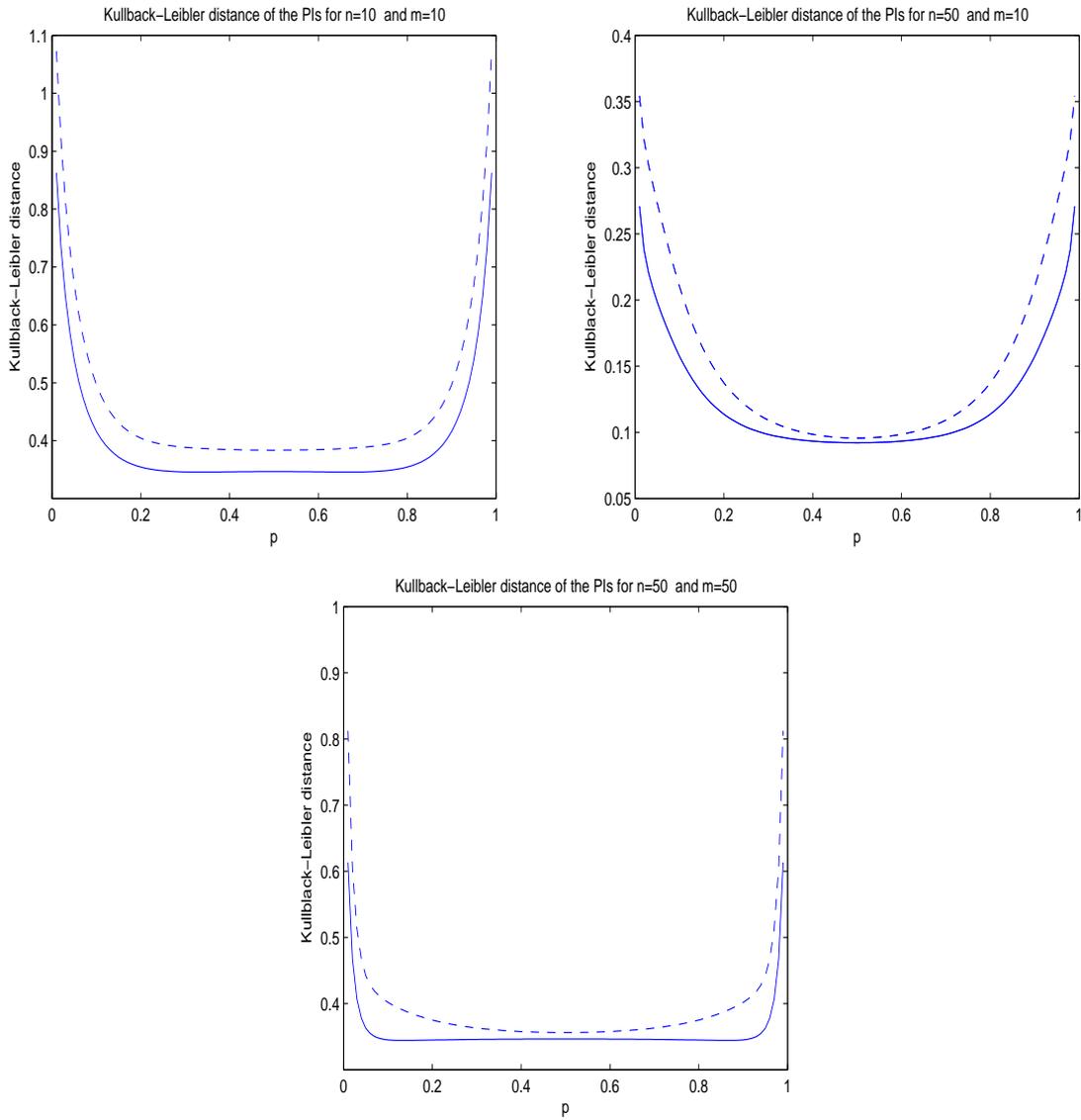
Figure 5: Kullback-Leibler distances of the score (solid line) and adjusted predictive distributions (dashed line) from the true binomial distribution when the sample sizes are (1) $n = m = 10$, (2) $n = 50, m = 10$ and (3) $n = m = 50$