Comparison of the Bootstrap and Delta Method Variances of the Variance Estimator of the Bernoulli Distribution

Ying-Ying Zhang[1*], Teng-Zhong Rong[1] and Man-Man Li[1]

[1]*Department of Statistics and Actuarial Science, College of Mathematics and Statistics, Chongqing University, Chongqing, China.*

Original Research Article

Abstract

It is interesting to calculate the variance of the variance estimator of the Bernoulli distribution. Therefore, we compare the Bootstrap and Delta Method variances of the variance estimator of the Bernoulli distribution in this paper. Firstly, we provide the correct Bootstrap, Delta Method, and true variances of the variance estimator of the Bernoulli distribution for three parameter values in Table 2.1. Secondly, we obtain the estimates of the variance of the variance estimator of the Bernoulli distribution by the Delta Method (analytically), the true method (analytically), and the Bootstrap Method (algorithmically). Thirdly, we compare the Bootstrap and Delta Methods

*Corresponding author: E-mail: robertzhangyying@qq.com, robertzhang@cqu.edu.cn;

in terms of the variance estimates, the errors, and the absolute errors in three figures for 101 parameter values in [0, 1], with the purpose to explain the differences between the Bootstrap and Delta Methods. Finally, we give three examples of the Bernoulli trials to illustrate the three methods.

Keywords: Bernoulli distribution; bootstrap; delta method; variance estimate.

2010 Mathematics Subject Classification: 62F10; 62F12; 62F40.

1 Introduction

The Bootstrap Method, a resampling technique used to obtain estimates of summary statistics, is widely applied, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. And the Delta Method is a result concerning the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator [13, 14].

Moreover, Casella and Berger [13] is a worldwidely used textbook for the courses of Statistical Inference or Advanced Mathematical Statistics for first-year graduate students majoring in statistics or in a field where a statistics concentration is required. In Example 10.1.21 of Casella and Berger [13], they compared the Bootstrap and Delta Method variances of $\hat{p}\,(1-\hat{p})$, which is the variance estimator of the Bernoulli distribution. However, the variances of the Bootstrap method and the true method are wrongly calculated.

In this paper, we compare the Bootstrap and Delta Method variances of the variance estimator of the Bernoulli distribution. Firstly, we provide the correct Bootstrap, Delta Method, and true variances of the variance estimator of the Bernoulli distribution for three parameter values in Table 2.1. Secondly, we obtain the estimates of the variance of the variance estimator of the Bernoulli distribution by the Delta Method (analytically), the true method (analytically), and the Bootstrap Method (algorithmically). Thirdly, we compare the Bootstrap and Delta Methods in terms of the variance estimates, the errors, and the absolute errors in three figures for 101 parameter values in [0, 1], with the purpose to explain the differences between the Bootstrap and Delta Methods. Finally, we give three examples of the Bernoulli trials to illustrate the three methods.

The rest of the paper is organized as follows. In Section 2, we provide the right variances in Table 2.1. We also provide the estimates of the variance of $\hat{p}\,(1-\hat{p})$ by the Delta Method (analytically), the true method (analytically), and the Bootstrap Method (algorithmically). Moreover, we compare the Bootstrap and Delta Methods in terms of the variance estimates, the errors, and the absolute errors in three figures for 101 parameter values in $[0, 1]$. Section 3 provides three examples of the Bernoulli trials. Section 4 concludes.

2 Main Results

The correct Bootstrap and Delta Method variances of $\hat{p}\,(1-\hat{p})$ are given in Table 2.1, where sample size $n = 24$, and bootstrap sample size $B = 1000$. In particular, the variance of the Delta Method corresponds to $p = 2/3$ should be rounded to 0.00103, since the variance is calculated as 0.001028807. From Table 2.1 we see that, the estimate of the variance of $\hat{p}\,(1-\hat{p})$ by the Bootstrap Method is better than the first-order Delta Method at $p \neq 1/2$, but is worse than the second-order Delta Method at $p = 1/2$.

Furthermore, the original values from Table 10.1.1 in Casella and Berger [13] are provided in Table 2.2 so that potential readers do not need to be referred to the book.

Table 2.1. Bootstrap and Delta Method variances of $\hat{p}(1-\hat{p})$. The second-order Delta Method is used when $p = 1/2$. The Delta Method variance is calculated numerically assuming that $\hat{p} = p$

	$p = 1/4$	$p = 1/2$	$p = 2/3$
Bootstrap	0.00190	0.00025	0.00108
Delta Method	0.00195	0.00022	0.00103
True	0.00191	0.00021	0.00111

Table 2.2. Table 10.1.1 in Casella and Berger [13]) Bootstrap and Delta Method variances of $\hat{p}(1-\hat{p})$. The second-order Delta Method is used when $p = 1/2$. The Delta Method variance is calculated numerically assuming that $\hat{p} = p$

	$p = 1/4$	$p = 1/2$	$p = 2/3$
Bootstrap	0.00508	0.00555	0.00561
Delta Method	0.00195	0.00022	0.00102
True	0.00484	0.00531	0.00519

The potential reasons why the estimates in Table 10.1.1 of Casella and Berger [13] are not right are summarized as follows. Firstly and the most importantly, the exact expression for $\mathrm{Var}_p(\hat{p}(1-\hat{p}))$ may be wrongly calculated by Casella and Berger. Secondly, the Bootstrap procedure is wrongly programmed by Casella and Berger.

The estimate of the variance of $\hat{p}(1-\hat{p})$ by the first-order Delta Method is (see Casella and Berger [13] Example 10.1.15)

$$\widehat{\mathrm{Var}}_p^{\mathrm{Delta1}}(\hat{p}(1-\hat{p})) = \frac{\hat{p}(1-\hat{p})(1-2\hat{p})^2}{n} = f_1(\hat{p}).$$

Since $\widehat{\mathrm{Var}}_p^{\mathrm{Delta1}}(\hat{p}(1-\hat{p}))\Big|_{\hat{p}=1/2} = 0$, a clear underestimate of the variance of $\hat{p}(1-\hat{p})$. Therefore, when $\hat{p} = 1/2$, we need to use a second-order Delta Method. When $\hat{p} = 1/2$, the estimate of the variance of $\hat{p}(1-\hat{p})$ by the second-order Delta Method is

$$\widehat{\mathrm{Var}}_p^{\mathrm{Delta2}}(\hat{p}(1-\hat{p})) = \frac{2\hat{p}^2(1-\hat{p})^2}{n^2} = \frac{1}{8n^2}. \tag{2.1}$$

The derivation of (2.1) can be found in the appendix. Therefore, the estimate of the variance of $\hat{p}(1-\hat{p})$ by the Delta Method is formed by combining the first-order and the second-order Delta Method, and is given by

$$\widehat{\mathrm{Var}}_p^{\mathrm{Delta}}(\hat{p}(1-\hat{p})) = \begin{cases} \hat{p}(1-\hat{p})(1-2\hat{p})^2/n, & \text{if } \frac{1}{2} \neq \hat{p} \in [0,1], \\ 2\hat{p}^2(1-\hat{p})^2/n^2 = 1/\left(8n^2\right), & \text{if } \hat{p} = \frac{1}{2}. \end{cases}$$

The true variance of $\hat{p}(1-\hat{p})$ is

$$\mathrm{Var}_p(\hat{p}(1-\hat{p})) = \frac{1}{n^4}\left[\begin{array}{c} 2n(n-1)(3-2n)p^4 + 4n(n-1)(2n-3)p^3 \\ +n(n-1)(7-5n)p^2 + n(n-1)^2 p \end{array}\right] = f(p). \tag{2.2}$$

The derivation of (2.2) can be found in the appendix. We note that the exact expression for $\mathrm{Var}_p(\hat{p}(1-\hat{p}))$ is required to be calculated in Exercise 10.10 in Casella and Berger [13]. However, in their solution manual, there is no solution for this exercise.

The estimate of the variance of $\hat{\theta} = \hat{p}(1-\hat{p})$ by the Bootstrap Method is calculated as follows [13, 4, 14].

Step 1. Given $p \in [0,1]$, generate an $n \times B$ matrix

$$\mathbf{X}^* = \left(x_{ki}^*\right)_{n\times B}, \;\; x_{ki}^* \sim \text{Bernoulli}\,(p)\,.$$

Step 2. Calculate

$$\hat{p}_i^* = \bar{x}_i^* = \frac{1}{n}\sum_{k=1}^{n} x_{ki}^*, \;\; i = 1, \ldots, B.$$

Step 3. Calculate

$$\hat{\theta}_i^* = \hat{p}_i^*\left(1 - \hat{p}_i^*\right), \;\; i = 1, \ldots, B.$$

Step 4. Calculate the bootstrap approximator

$$\text{Var}_B^*\left(\hat{\theta}\right) = \frac{1}{B-1}\sum_{i=1}^{B}\left(\hat{\theta}_i^* - \overline{\hat{\theta}^*}\right)^2,$$

where

$$\overline{\hat{\theta}^*} = \frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_i^*.$$

In Figs. 2.1-2.3, there are 101 p values $[0, 0.01, 0.02, \ldots, 0.99, 1]$; sample size is $n = 24$; and bootstrap sample size increases to $B = 10000$.

Three p values of the three methods are compared in Table 2.1. However, what are the variance comparisons at other p values in $[0,1]$? Let us see Fig. 2.1. In Fig. 2.1, the red full line is the variance estimated by the Bootstrap Method, the blue dashed line is the variance estimated by the Delta Method, and the black dotted line is the true variance. We see that the variance estimated by the Bootstrap Method approximates the true variance very well. Moreover, the second-order Delta Method at $p = \frac{1}{2}$ has a good variance estimate, while the first-order Delta Method at $p = \frac{1}{2}$ has a very bad variance estimate. Furthermore, the true variance curve has two peaks and one valley, and the true variances are equal to 0 at the two end points. We also see that the three curves exhibit symmetries about $p = \frac{1}{2}$. It is easy to check that the first-order Delta Method curve is symmetric about $p = \frac{1}{2}$ by checking

$$f_1\left(\frac{1}{2} + \varepsilon\right) = f_1\left(\frac{1}{2} - \varepsilon\right).$$

We can also check that the true variances curve is symmetric about $p = \frac{1}{2}$ by checking

$$f\left(\frac{1}{2} + \varepsilon\right) = f\left(\frac{1}{2} - \varepsilon\right).$$

We can exploit the Mathematica software to do this job. And the Mathematica codes can be found in the supplemental file "TestSymmetric.nb". We see that the Bootstrap Method curve is numerically symmetric about $p = \frac{1}{2}$.

The error comparison of the two methods is shown in Fig. 2.2, and the error curve is calculated by the difference of the estimated curve and the true curve. Thus the error will be positive if the estimated variance is higher than the true variance, and negative otherwise. From Fig. 2.2 we see that the Bootstrap Method has small errors on the whole interval $[0,1]$, while the Delta Method has a sine shaped error (in fact the curve is a polynomial of p of order 4) with an exception at $p = \frac{1}{2}$, as expected. The error curves of the two methods are symmetric about $p = \frac{1}{2}$.

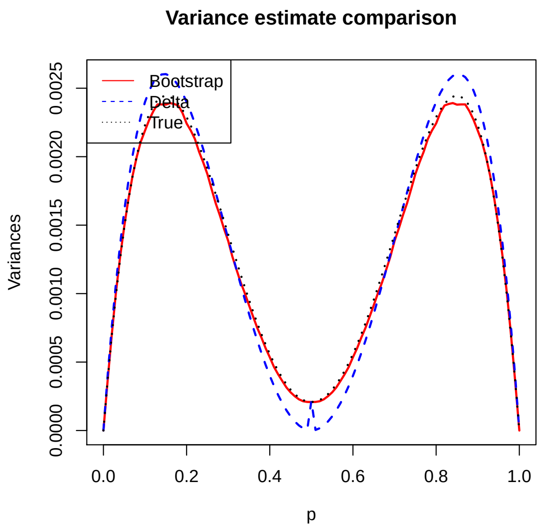

Fig. 2.1. Variance estimate comparisons of the three methods

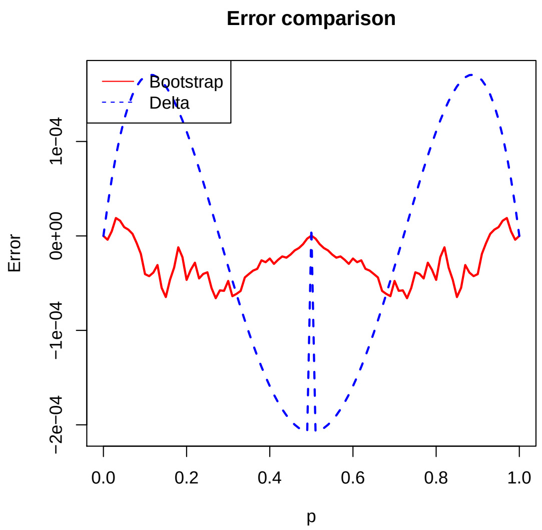

Fig. 2.2. Error comparisons of the Bootstrap and Delta Method

Fig. 2.3 shows the absolute error comparison of the two methods. By absolute error, we mean the absolute value of the error. We see that in most cases, the Bootstrap Method outperforms the Delta Method. The magnitude of the absolute errors of the Bootstrap Method is smaller than that of the Delta Method. The absolute error curves of the two methods are symmetric about $p = \frac{1}{2}$. The variances of the two methods agree with the true variances at the two endpoints 0 and 1. The variances of the Delta Method agree with the true variances at two intermediate values, and thus it behaves better than the Bootstrap Method near the two values. While the variance of the Bootstrap Method seems to agree with the true variance at $p = \frac{1}{2}$, and thus it is better than the

Delta Method at the point. This phenomenon seems to contradict with that phenomenon which is seen in Table 2.1. However, there is no contradiction because in Table 2.1, $B = 1000$, while in Fig. 2.3, $B = 10000$.

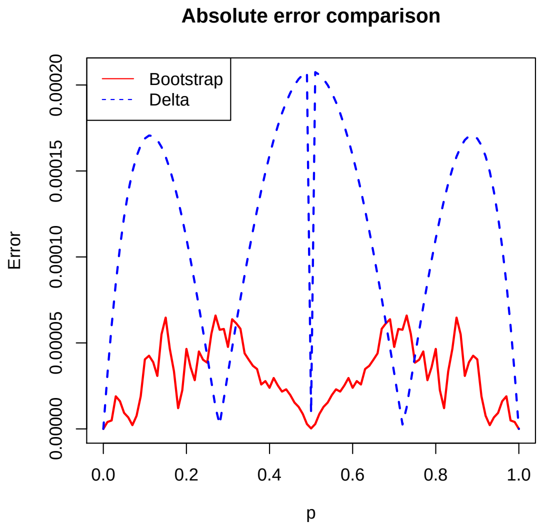

Fig. 2.3. Absolute error comparisons of the Bootstrap and Delta Method

3 Examples

In this section, we provide three examples of the Bernoulli trials. Then we calculate the estimates of the variance of the variance estimator of the Bernoulli distribution by means of the Bootstrap Method, the Delta Method, and the true method. In the three examples, the bootstrap sample size is $B = 10000$. Note that the sample size n in the three examples are different.

Example 3.1. A coin is flipped 100 times, and there are 52 heads up. Therefore, the proportion of heads up is

$$\hat{p} = \frac{52}{100} = 0.52.$$

The estimates of the variance of the variance estimator of the Bernoulli distribution by the three methods are respectively given by

$$\begin{aligned} \mathrm{Var}_B^* \left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 1.6183\text{e-}05, \\ \widehat{\mathrm{Var}}_p^{\mathrm{Delta}} \left(\hat{p}\left(1-\hat{p}\right)\right) &= 3.9936\text{e-}06, \\ \mathrm{Var}_p \left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 1.6250\text{e-}05. \end{aligned}$$

Example 3.2. There is a batch of vegetable seeds, and each seed sprouts with probability p. We sample 500 seeds at random. After seed soaking treatment with seed coating agent, there are 445 seeds sprouted. Therefore, the sprouting rate is

$$\hat{p} = \frac{445}{500} = 0.89.$$

The estimates of the variance of the variance estimator of the Bernoulli distribution by the three methods are respectively given by

$$\begin{aligned}
\mathrm{Var}_B^*\left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 0.0001172489,\\
\widehat{\mathrm{Var}}_p^{\mathrm{Delta}}\left(\hat{p}\left(1-\hat{p}\right)\right) &= 0.0001191247,\\
\mathrm{Var}_p\left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 0.0001187252.
\end{aligned}$$

Example 3.3. 400 newborns in the local area are observed in one hospital. There is only one newborn with chromosomal abnormality. Therefore, the rate of chromosomal abnormality in the local area is

$$\hat{p} = \frac{1}{400} = 0.0025.$$

The estimates of the variance of the variance estimator of the Bernoulli distribution by the three methods are respectively given by

$$\begin{aligned}
\mathrm{Var}_B^*\left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 6.078563\text{e-}06,\\
\widehat{\mathrm{Var}}_p^{\mathrm{Delta}}\left(\hat{p}\left(1-\hat{p}\right)\right) &= 6.172187\text{e-}06,\\
\mathrm{Var}_p\left(\hat{p}\left(1-\hat{p}\right)\right)\big|_{p=\hat{p}} &= 6.141442\text{e-}06.
\end{aligned}$$

4 Conclusions

We compare the Bootstrap and Delta Method variances of $\hat{p}\left(1-\hat{p}\right)$, which is the variance estimator of the Bernoulli distribution. First, we provide the right Bootstrap, Delta Method, and true variances of $\hat{p}\left(1-\hat{p}\right)$ in Table 2.1. The parameter values of Table 2.1 and Table 10.1.1 in Casella and Berger [13] are the same. Then, we provide the estimates of the variance of $\hat{p}\left(1-\hat{p}\right)$ by the Delta Method, the true method, and the Bootstrap Method. The derivations of the estimates of the variance of $\hat{p}\left(1-\hat{p}\right)$ by the second-order Delta Method and the true method are given in the appendix. Moreover, we compare the Bootstrap and Delta Methods in terms of the variance estimates, the errors, and the absolute errors in three figures for 101 parameter values in $[0, 1]$. It is worth noting that, the three variance estimate curves exhibit symmetries about $p = \frac{1}{2}$, and in most cases, the Bootstrap Method outperforms the Delta Method. Finally, three examples of the Bernoulli trials are given to illustrate the three methods.

Acknowledgements

The authors gratefully acknowledge the constructive comments offered by the referees. Their comments improve the quality of the paper significantly. The research was supported by the Fundamental Research Funds for the Central Universities (CQDXWL-2012-004; 106112016CDJXY100002), China Scholarship Council (201606055028), National Natural Science Foundation of China (11671060), and MOE project of Humanities and Social Sciences on the west and the border area (14XJC910001).

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 1985; **39**:783–791.

[2] Meyer JS, Ingersoll CG, McDonald LL, Boyce MS. Estimating uncertainty in population growth rates: Jackknife vs. bootstrap techniques. *Ecology* 1986; **67**:1156–1166.

[3] Hillis DM, Bull JJ. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 1993; **42**:182–192.

[4] Shao J, Tu D. *The Jackknife and the Bootstrap*. Springer: New York, 1995.

[5] Visscher PM, Thompson R, Haley CS. Confidence intervals in qtl mapping by bootstrapping. *Genetics* 1996; **143**:1013–1020.

[6] Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: A non-parametric approach to confidence interval estimation. *Health Economics* 1997; **6**:327–340.

[7] Carpenter J, Bithell J. Bootstrap confidence intervals: When, which, what? a practical guide for medical statisticians. *Statistics in Medicine* 2000; **19**:1141–1164.

[8] Alfaro ME, Zoller S, Lutzoni F. Bayes or bootstrap? a simulation study comparing the performance of bayesian markov chain monte carlo sampling and bootstrapping in assessing phylogenetic confidence. *Molecular Biology and Evolution* 2003; **20**:255–266.

[9] Stamatakis A, Hoover P, Rougemont J. A rapid bootstrap algorithm for the raxml web servers. *Systematic Biology* 2008; **57**:758–771.

[10] Wang DL, Hutson AD. A smooth bootstrap procedure towards deriving confidence intervals for the relative risk. *Communications in Statistics - Theory and Methods* 2014; **43**:1979–1990.

[11] Ju H. Moving block bootstrap for analyzing longitudinal data. *Communications in Statistics - Theory and Methods* 2015; **44**:1130–1142.

[12] Shen MY, Machado SG. Bioequivalence evaluation of sparse sampling pharmacokinetics data using bootstrap resampling method. *Journal of Biopharmaceutical Statistics* 2016; :1–8.

[13] Casella G, Berger RL. *Statistical Inference*. Duxbury: USA, 2002. 2nd edition.

[14] Shao J. *Mathematical Statistics*. Springer: New York, 2003. 2nd edition.

APPENDIX

The derivations of (2.1) and (2.2) are given in the appendix.

The derivation of (2.1). We have, by the Central Limit Theorem,

$$\frac{\bar{X} - \mathrm{E}\bar{X}}{\sqrt{\mathrm{Var}\left(\bar{X}\right)}} \xrightarrow{d} N\left(0,1\right), \text{ as } n \to \infty,$$

where $\bar{X}$ is the sample mean of $X_1, X_2, \ldots, X_n$, which are iid from Bernoulli$\,(p)$. The Maximum Likelihood Estimator (MLE) of p is $\hat{p} = \bar{X}$. And

$$\mathrm{E}\hat{p} = \mathrm{E}\bar{X} = \mathrm{E}X = p,$$

$$\mathrm{Var}\left(\hat{p}\right) = \mathrm{Var}\left(\bar{X}\right) = \frac{\mathrm{Var}\left(X\right)}{n} = \frac{p\left(1-p\right)}{n}.$$

Therefore,

$$\frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \xrightarrow{d} N\left(0,1\right), \text{ as } n \to \infty.$$

Rearranging, we obtain

$$\sqrt{n}\left(\hat{p} - p\right) \xrightarrow{d} \sqrt{p\left(1-p\right)}N\left(0,1\right) = N\left(0, p\left(1-p\right)\right), \text{ as } n \to \infty.$$

Let $g\left(p\right) = p\left(1-p\right)$. Then by the second-order Delta Method (Theorem 5.5.26 in [13]), we have

$$n\left[g\left(\hat{p}\right) - g\left(p\right)\right] \xrightarrow{d} \frac{g^{''}\left(p\right)}{2}\sigma^2\chi_1^2, \text{ as } n \to \infty,$$

where $\sigma^2 = p\left(1-p\right)$ and χ_1^2 is the chi-square random variable with 1 degree of freedom. Therefore, for large n,

$$\mathrm{Var}_p\left(g\left(\hat{p}\right)\right) \approx \mathrm{Var}_p\left(\frac{g^{''}\left(p\right)}{2n}\sigma^2\chi_1^2\right) = \mathrm{Var}_p\left(\frac{-2}{2n}p\left(1-p\right)\chi_1^2\right) = \frac{2p^2\left(1-p\right)^2}{n^2}.$$

Replacing p by its MLE $\hat{p}$ in the above equation, we obtain the estimate of the variance of $\hat{p}\left(1-\hat{p}\right)$ by the second-order Delta Method, namely,

$$\widehat{\mathrm{Var}}_p^{\mathrm{Delta2}}\left(\hat{p}\left(1-\hat{p}\right)\right) = \frac{2\hat{p}^2\left(1-\hat{p}\right)^2}{n^2}.$$

Therefore, (2.1) is established. □

The derivation of (2.2). We have

$$\hat{p} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}Y,$$

where $X_1, X_2, \ldots, X_n$ are iid from Bernoulli$\,(p)$ and

$$Y = \sum_{i=1}^{n} X_i \sim \mathrm{Binomial}\left(n, p\right).$$

Therefore,

$$\mathrm{Var}_p\left(\hat{p}\left(1-\hat{p}\right)\right) = \mathrm{Var}_p\left(\bar{X}\left(1-\bar{X}\right)\right) = \mathrm{Var}_p\left(\frac{Y}{n}\left(1-\frac{Y}{n}\right)\right)$$

$$= \mathrm{Var}_p\left(\frac{Y\left(n-Y\right)}{n^2}\right) = \frac{1}{n^4}\mathrm{Var}_p\left(Y\left(n-Y\right)\right).$$

Now

$$\mathrm{Var}_p\left(Y\left(n-Y\right)\right)=\mathrm{E}\left[Y^2\left(n-Y\right)^2\right]-\left\{\mathrm{E}\left[Y\left(n-Y\right)\right]\right\}^2.$$

To calculate $\mathrm{Var}_p\left(Y\left(n-Y\right)\right)$, we need to know the first four moments of Y which can be calculated by the derivative of the moment generating function evaluated at $t=0$. The first four moments of Y are given by:

$$\begin{aligned}
\mathrm{E}Y &= np,\\
\mathrm{E}Y^2 &= np+n\left(n-1\right)p^2,\\
\mathrm{E}Y^3 &= np+3n\left(n-1\right)p^2+n\left(n-1\right)\left(n-2\right)p^3,\\
\mathrm{E}Y^4 &= np+7n\left(n-1\right)p^2+6n\left(n-1\right)\left(n-2\right)p^3+n\left(n-1\right)\left(n-2\right)\left(n-3\right)p^4.
\end{aligned}$$

Now

$$\begin{aligned}
\mathrm{E}\left[Y\left(n-Y\right)\right] &= \mathrm{E}\left[nY-Y^2\right]=n\mathrm{E}Y-\mathrm{E}Y^2\\
&= n\times np-np-n\left(n-1\right)p^2=n\left(n-1\right)p\left(1-p\right),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{E}\left[Y^2\left(n-Y\right)^2\right] &= \mathrm{E}\left[Y^2\left(Y^2-2nY+n^2\right)\right]=\mathrm{E}\left[Y^4-2nY^3+n^2Y^2\right]\\
&= \mathrm{E}Y^4-2n\mathrm{E}Y^3+n^2\mathrm{E}Y^2.
\end{aligned}$$

Therefore, by calculating, we obtain

$$\begin{aligned}
\mathrm{Var}_p\left(Y\left(n-Y\right)\right) &= \mathrm{E}Y^4-2n\mathrm{E}Y^3+n^2\mathrm{E}Y^2-\left[n\left(n-1\right)p\left(1-p\right)\right]^2\\
&= 2n\left(n-1\right)\left(3-2n\right)p^4+4n\left(n-1\right)\left(2n-3\right)p^3\\
&\quad +n\left(n-1\right)\left(7-5n\right)p^2+n\left(n-1\right)^2p.
\end{aligned}$$

Dividing $\mathrm{Var}_p\left(Y\left(n-Y\right)\right)$ by n^4, we obtain (2.2). □
