ON ACCURACY OF TESTS FOR BERNOULLI TRIALS

A.L. Kostevich, A.S. Shmuratko
National Research Center for Applied Problems of Mathematics and Informatics
Belarussian State University
Minsk, Belarus
e-mail: kostevich@bsu.by, shmuratko@bsu.by

Abstract

Consider a family of tests based on the number of successes in a binary sample. We investigate the error in the calculation of the type I error probability and P-value of the tests arising from the usage of the asymptotic distribution of the statistics instead of the exact one.

1 Introduction

Testing for i.i.d. Bernoulli trials with the given success probability is one of the problems of cryptography. Usually a test is based on an asymptotic distribution of a statistic: a threshold or P-value are calculated using a known asymptotic distribution instead of an unknown exact one. Since the sample size is fixed, this results in an error in the calculation of type I and II error probabilities and a P-value [1].

In the paper we consider a family of tests based on the number of successes in a binary sample. Such tests as the monobit test [2], the autocorrelation test, the overlapping and non-overlapping template matching tests [3] belong to this family. For these tests, we investigate the error in the calculation of the type I error probability and P-value arising from the usage of the asymptotic standard normal distribution.

Let a binary sequence $X_1, \ \ldots, X_n$ be observed. Consider the following hypothesis:

$$\begin{aligned} \mathcal{H}_0: \quad & \{X_t\} \text{ independent identically distributed Bernoulli random variables,} \\ & \mathbf{P}\{X_t = 1\} = p, \quad \mathbf{P}\{X_t = 0\} = q = 1 - p, \qquad t \in \{1, 2, \ldots, n\}. \end{aligned}$$

The standardized number of successes in the Bernoulli trials is

$$S_n = \frac{1}{\sqrt{npq}} \left(\sum_{t=1}^{n} X_t - np \right).$$

According to the central limit theorem (CLT), the statistic S_n has the asymptotically normal probability distribution $\mathcal{N}(0,1)$: $F_n(x) = \mathbf{P}\{S_n < x\} \to \Phi(x)$.

Consider a two-sided test based on S_n for testing $\mathcal{H}_0$ against some alternative $\mathcal{H}_1$:

$$\text{decide} \begin{cases} \mathcal{H}_0, & \text{if } -\Delta \leqslant S_n < \Delta, \\ \mathcal{H}_1, & \text{otherwise,} \end{cases} \qquad \Delta = \Delta(\alpha) = \Phi^{-1}\left(1 - \frac{\alpha}{2}\right), \tag{1}$$

where α is the asymptotic type I error probability, $\Delta(\alpha)$ is the asymptotic threshold of the test. The exact type I error probability of the test (1) is

$$\alpha_n = \mathbf{P}\left\{S_n < -\Delta\right\} + \mathbf{P}\left\{S_n \geqslant \Delta\right\} = F_n(-\Delta) + 1 - F_n(\Delta).$$

Clearly, $\alpha_n \to \alpha$. We investigate the absolute and relative error

$$\varepsilon_n = \alpha_n - \alpha, \qquad \delta_n = \frac{\varepsilon_n}{\alpha} = \frac{\alpha_n - \alpha}{\alpha}.$$

Define the exact and asymptotic P-value of the test (1)

$$P_n = F_n(-|S_n|) + 1 - F_n(|S_n|), \quad P = 2 - 2\Phi\left(|S_n|\right).$$

Then $P_n - P \to 0$ uniformly in $|S_n|$. We investigate the absolute and relative error of the P-value calculation for $|S_n| \in [a, b]$, where $0 \leqslant a \leqslant b \leqslant \sqrt{n}$:

$$\varepsilon_n^{sup} = \sup_{|S_n|\in[a,b]} |\varepsilon_n'| = \sup_{|S_n|\in[a,b]} |P_n - P|, \qquad \delta_n^{sup} = \sup_{|S_n|\in[a,b]} |\delta_n'| = \sup_{|S_n|\in[a,b]} \left|\frac{P_n}{P} - 1\right|.$$

The task of what follows is 1) to provide theoretical estimates (obtained from the CLT and theory of large deviations [4, 5, 6]) and 2) to carry out a practical study (for $p = 0.5$) of the errors mentioned. We examine the behaviour of the errors both over n and over α (or Δ).

2 Type I error probability approximation

Define $C_0 = 0.7655$, $\mu = \frac{q-p}{\sqrt{pq}}$,

$$C_1 = C_1(\Delta, p) = 2\frac{p^2 + q^2}{\sqrt{pq}} \min\left\{C_0, \frac{C_0 + 8(1+e)}{1 + \Delta^3}\right\}.$$

Theorem 1. *[4, 5, 6]. Under hypothesis $\mathcal{H}_0$ the following estimates hold:*

$$|\varepsilon_n| \leqslant \frac{C_1(\Delta, p)}{\sqrt{n}}, \qquad |\delta_n| \leqslant \frac{C_1(\Delta, p)}{\alpha\sqrt{n}} \quad \textit{for all } n \textit{ and } \Delta > 0;$$

$$|\varepsilon_n| \leqslant \left(\frac{|\mu|}{3}\Delta^2 + C_\varepsilon\right)\frac{e^{-\Delta^2/2}}{\sqrt{n}}, \quad |\delta_n| \leqslant \left(\frac{|\mu|}{6}\Delta^3 + C_\delta(\Delta+1)\right)\frac{1}{\sqrt{n}} \quad \textit{for } \Delta = \mathcal{O}\left(n^{1/6}\right);$$

where $C_\varepsilon > 0$ and $C_\delta > 0$ depend only on p and on the constant in $\mathcal{O}(\cdot)$.

Using Theorem 1 one can easily calculate the sample size required to attain the specified accuracy of the type I error probability approximation. E.g., the percentage error of the approximation need not exceed 1%. Then introduce

$$n_{min}(\Delta, p) = \min\{k \in \mathbb{N} : |\delta_n| \cdot 100\% \leqslant 1\% \text{ when } n \geqslant k\} -$$

the minimal sample size for which we can guarantee that the percentage error does not exceed 1%. Theorem 1 provides the upper bound for n_{min}:

$$n_{min} \leqslant n_1 = \left\lceil \left(\frac{C_1 \cdot 100}{\alpha} \right)^2 \right\rceil,$$

where $\lceil x \rceil$ is the smallest integer greater than or equal to x.

Now put $p = 0.5$. Choose $\alpha = 0.05,\ 0.01,\ 0.001,\ 0.0001,\ 0.00001$ and $n \in I = \{100, 121, \ldots, 2000\}$. Define the constants

$$\hat{C}_1(\Delta) = \sup_{n \in I} |\varepsilon_n| \sqrt{n}, \qquad \hat{C}_\varepsilon(\Delta) = \sup_{n \in I} |\varepsilon_n| \sqrt{n} e^{\Delta^2/2} = \hat{C}_1 e^{\Delta^2/2},$$

$$\hat{n}_1 = \left\lceil \left(\frac{\hat{C}_1}{\alpha} 100 \right)^2 \right\rceil, \qquad \hat{C}_\delta(\Delta) = \sup_{n \in I} |\delta_n| \frac{\sqrt{n}}{\Delta + 1} = \frac{\hat{C}_1}{\alpha(\Delta + 1)},$$

which serve as practical analogues to C_1, C_ε, C_δ and n_1.

Table 1 contains the values of the constants mentioned.

Table 1: Theoretical and "practical" constant values, $p = 0.5$

α	C_1	$\hat{C}_1$	C_1/α	$\hat{C}_1/\alpha$	$\hat{C}_\varepsilon$	$\hat{C}_\delta$	n_1	$\hat{n}_1$
0.05	1.531	0.12	30.62	2.43	0.83	0.82	9375844	59172
0.01	1.531	0.029	153.1	2.89	0.8	0.81	234396100	83479
0.001	1.531	0.0037	1531	3.68	0.82	0.86	23439610000	135059
0.0001	1.019	0.00047	10189.1	4.66	0.9	0.95	1038185508334	217547
0.00001	0.7	0.000056	69992.9	5.56	0.96	1.03	48990009851824	309064

Since even $\hat{n}_1$ is rather big, we judge that the quality of the type I error probability approximation studied is unsatisfactory.

Figures 1, 3 illustrate the behaviour of δ_n over n and over α.

From Fig. 1 we gather that the speed of covergence of δ_n in n, $n^{-1/2}$, seems to be true. We also see that, for fixed α, the value of δ_n decreases oscillating around zero. So taking larger samples does not guarantee obtaining smaller relative errors. The same is true about the absolute error since in this case (fixed α) its behaviour is essentially the same.

For fixed n – see Fig. 3 – as α decreases, the oscillation of δ_n around zero gradually changes into an entire worsening of its behaviour. For the larger values of n, however, this effect "softens" and moves to the smaller values of α.

3 P–value approximation

Since $\varepsilon'_n(|S_n|) = \varepsilon_n(|S_n|)$, $\delta'_n(|S_n|) = \delta_n(|S_n|)$, many of the results here can be derived from the results of the previous section. However, a special choise of $\Delta = \Delta_n = |S_n|$ brings some new results.

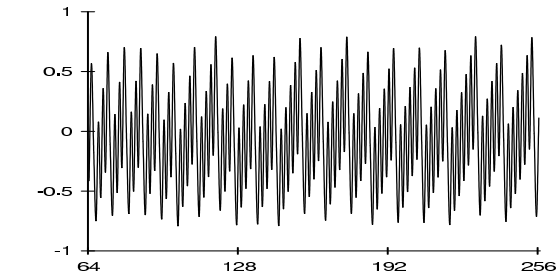


Figure 1: Plot of $\delta_n \frac{\sqrt{n}}{\Delta+1}$ over $n \in [64, 256]$ when $\alpha = 0.01$

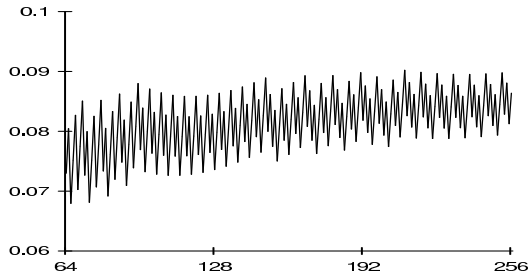
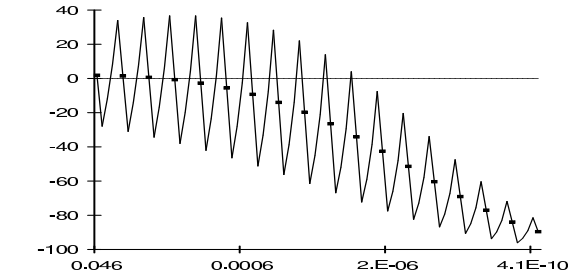
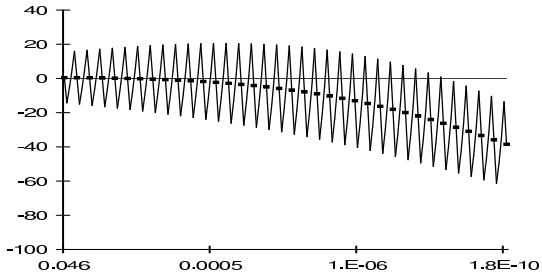


Figure 2: Plot of $\delta_n^{sup} n(1+b)^4 \Phi(-b)$ over $n \in [64, 256]$ when $[a, b] = [1.96, 4.42]$ (or $P \in [0.00001, 0.05]$)



$n = 64$



$n = 256$

Figure 3: Plots of $\delta_n \cdot 100\%$ over $\alpha \in [10^{-10}, 0.05]$ (or $\Delta \in [1.96, 6.47]$); Dotted lines — plots of $\delta'_n \cdot 100\%$ over P (or $|S_n|$) in the same intervals.

The behaviour of ε'_n and δ'_n over $|S_n|$ (or over P) is studied in the same way as above. Besides, for ε'_n and δ'_n one can obtain the estimates similar to those in statement 3 of Theorem 2.

For $p = 0.5$ the plot of $\delta'_n \cdot 100\%$ over P is given in Fig. 3 — see the dotted lines in the middle. Unlike δ_n, the value of δ'_n remains steadily close to zero up to a certain point. The bigger n, the father this point.

Now we move to the study of ε_n^{sup} and δ_n^{sup}. From Theorem 1 and [5] one obtains

Theorem 2. *Under hypothesis $\mathcal{H}_0$ the following estimates hold:*

1) $|\varepsilon_n^{sup}| \leqslant \dfrac{C_1(a,p)}{\sqrt{n}}$, $\quad |\delta_n^{sup}| \leqslant \dfrac{C_1(b,p)}{2\Phi(-b)\sqrt{n}}$ *for all n;*

2) $|\varepsilon_n^{sup}| \leqslant C_\varepsilon (a^2+1)^{\chi(p)} \dfrac{e^{-a^2/2}}{\sqrt{n}}$, $\quad |\delta_n^{sup}| \leqslant \left(\dfrac{|\mu|}{6} b^3 + C_\delta (b+1) \right) \dfrac{1}{\sqrt{n}}$ *for $b = \mathcal{O}\left(n^{1/6}\right)$;*

3) $|\varepsilon_n^{sup}| \leqslant \dfrac{C}{n(1+a)^4}$, $\quad |\delta_n^{sup}| \leqslant \dfrac{C}{n(1+b)^4 \Phi(-b)}$ *for $p = 0.5$ and all n;*

where C_1 is defined in Theorem 1, $C_\varepsilon > 0$ and $C_\delta > 0$ depend only on p and on the constant in $\mathcal{O}(\cdot)$, $C > 0$ is an absolute constant; $\chi(p) = 0$ for $p = 0.5$, $\chi(p) = 1$ for $p \neq 0.5$.

Put $p = 0.5$. Choose $n \in I = \{100, 101, \ldots, 2000\}$, $|S_n| \in [a, b] = [1.96, 4.42]$ (or $P \in [0.00001, 0.05]$). Similarly, we have calculated the values of the following constants:

$$C_1(b, 0.5) = 0.7, \qquad n_1 = \left\lceil \left(\frac{C_1}{2\Phi(-b)} 100 \right)^2 \right\rceil = 48990009851182,$$

$$\hat{C}_1 = \sup_{n \in I} |\delta_n^{sup}| \sqrt{n} 2\Phi(-b) = 0.00002, \quad \hat{n}_{11} = \left\lceil \left(\frac{\hat{C}_1}{2\Phi(-b)} 100 \right)^2 \right\rceil = 40903,$$

$$\hat{C} = \sup_{n \in I} |\delta_n^{sup}| \, n(1+b)^4 \Phi(-b) = 0.09, \quad \hat{n}_{12} = \left\lceil \frac{\hat{C}}{(1+b)^4 \Phi(-b)} 100 \right\rceil = 2127.$$

In fact, it seems that in our case $|\delta_n^{sup}| \cdot 100\% \leqslant 1\%$ when $n \geqslant 2116$.

We conclude that the quality of the P-value approximation studied is a lot better than that of the type I error probability approximation.

For the plot of $\delta_n^{sup} n(1+b)^4 \Phi(-b)$ over n see Fig. 2.

References

[1] Brown L.D., Cai T.T., DasGupta A. (2001) Interval Estimation for a Binomial Proportion. *Statistical Science*, **16**, No 2, pp. 101-133.

[2] FIPS Publication 140–2. (2001) Security Requirements for Cryptographic Modules, National Institute of Standards and Technology.

[3] NIST Special Publication 800-22. (2000) A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, National Institute of Standards and Technology.

[4] Michel R. (1981) On the constant in the non-uniform version of the Berry-Esseen theorem. *Z. Wahrscheinlichkeitstheorie verw. Geb.* Bd. **55**, No 1, pp. 109-117.

[5] Petrov V. V. (1972) *Sums of independent random variables*. Science, Moscow (in Russian). P. 211, 280, 281.

[6] Petrov V. V. (1987) *Limit theorems for the sums of independent random variables*. Science, Moscow (in Russian). P. 157, 158, 178, 283, 284.