УДК 004.312.43.052

С. Л. Волков

ОЦІНКА БІНОМІАЛЬНОГО ПАРАМЕТРА R ПО МЕТОДУ КЛОПЕРА — ПІРСОНА

Приводяться лема, її доказ та уточнені визначення наукових понять, які стосуються оцінки біноміального параметра R по методу Клопера — Пірсона

Ключові слова: *випробування, біноміальний параметр, статистика, довірчий інтервал*

1. Вступ

У системі випробувань Бернуллі параметр $R = P(A_i)$ функції розподілу $B_i(n,R,x)$ завжди вважається відомим. Він має сенс імовірності успішного функціонування технічної системи в одному i-му випробуванні, якщо подія A_i визначена відповідним чином, а саме — якщо A_i складається в успішному функціонуванні системи в i-му випробуванні. Однак на практиці часто імовірність R невідома та підлягає оцінюванню за результатами випробувань. Як показала практика, найбільш точні результати дають випробування, які проводяться по методу, який отримав назву «Метод Клопера — Пірсона». Однак різні вчені та дослідники при практичних дослідженнях використовують різну термінологію, що в достатньому ступені утруднює узагальнення результатів.

2. Постановка проблеми

Відповідно до сказаного, пропонуються узагальнені визначення деяких наукових та технічних термінів, що стосуються оцінки біноміального параметра R при проведенні випробувань технічних об'єктів, наприклад, на надійність та живучість, з використанням методу, запропонованого Клопером і Пірсоном.

3. Основна частина

3.1. Аналіз літературних джерел по темі дослідження. Стосовно до теорії випробувань на надійність та живучість технічних засобів одноразового та короткочасного використання, а також аналогічні проблеми стосовно до теорії функціонування складних систем, освітлені в багатьох доступних літературних джерелах. Серед вчених проблемами випробувань технічних систем, включаючи метод Клопера — Пірсона, займалися як вітчизняні, так і зарубіжні теоретики та практики. Серед них — Р. В. Судаков, Г. А. Птіцин, Е. Ю. Барзилович, В. А. Каштанов, В. І. Борщ, Б. В. Гнеденко, Ю. К. Беляев, Ф. А. Мірталібов, О. О. Скопа, Н. Ф. Казакова та ін. Деякі результати приведені в [1—5].

3.2. Результати досліджень. У теорії надійності для параметра R біноміального розподілу використовують термін «імовірність R безвідмовної роботи в одному випробуванні». Термін «*імовірність успішного функціонування*» може розглядатися як синонім, але в дійсності він включає згаданий термін з теорії надійності як окремий випадок, оскільки у виразі $R = P(A_i)$ події A_i може надаватися довільний смисл, аби малася можливість у n біноміальних випробуваннях реєструвати число r подій $\bar{A}_i$ (або число $n-r$ подій A_i).

Як результат випробування n-серії Бернуллі пропонується приймати число r подій $\bar{A}_i$. Якщо вважати, що $\bar{A}_i$ — подія, яка складається у виникненні відмовлення в i-м випробуванні, то r — число відмовлень у n біноміальних випробуваннях. Приведемо пропоновані визначення.

Визначення 1. Будь-яка невідома константа θ, що підлягає оцінюванню по результатах $\omega \in \Omega$ випробувань (Ω — сукупність всіх результатів ω), називається *параметром*.

Визначення 2. Всяка функція $g(\omega)$, яка залежить від результатів $\omega \in \Omega$ випробувань, називається *статистикою*.

Визначення 3. Сукупність з n біноміальних випробувань Бернуллі називається *n-серією Бернуллі* або *серією обсягу n*.

Визначення 4. Випадкову величину $\hat{R} = 1 - \frac{r}{n}$ назвемо *точковою оцінкою для параметра R*, який являє собою невідому константу.

При кожному конкретному результаті $r = k$, що отриманий після проведення n випробувань (наприклад, при одержанні $r = 2$ в $n = 10$ випробуваннях), статистика $\hat{R}$ приймає також конкретне, невипадкове значення $\hat{R} = 1 - k/n$. Однак до проведення випробувань значення $\hat{R}$ непередбачене і можна лише стверджувати, що $0 \le \hat{R} \le 1$.

В силу випадковості $\hat{R}$ ці нерівності виконуються кожна з деякою імовірністю, а саме

$$P(\hat{R} < R) = P(r > n(1-R)) = 1 - P(r \le nq) =$$

$$= 1 - \sum_{k \le nq} \binom{n}{k} R^{n-k} (1-R)^k,$$

$$P(\hat{R} \ge R) = \sum_{k \le nq} \binom{n}{k} R^{n-k} (1-R)^k,$$

де $q = 1 - R$, причому $P(\hat{R} < R) \le P(\hat{R} \ge R) \le \frac{1}{2}$.

Для практики більший інтерес представляють оцінки для R, які можна позначати як $\underline{R}$ і $\overline{R}$. Вони мають властивості гарантованості в тім смислі, що, як правило, $\underline{R}$ не перевищує R (тобто $\underline{R} \le R$), а $\overline{R}$, навпаки, як правило перевищує R (тобто $\overline{R} \ge R$). При цьому фраза «як правило» означає, «з наперед заданою і досить великою імовірністю γ», наприклад, при $\gamma = 0{,}90$.

Таким чином нас цікавлять статистики $\underline{R}$ і $\overline{R}$ такі, для яких при заданій γ і невідомому R виконуються нерівності:

$$P\left(\underline{R} \le R\right) \ge \gamma, \quad P\left(\overline{R} \ge R\right) \ge \gamma. \tag{1}$$

Визначення 5. Статистику $\underline{R}$ з (1) назвемо γ — *нижньою границею для імовірності* R, а число γ буде називатися *довірчою імовірністю*.

Визначення 6. Статистику $\overline{R}$ з (1) назвемо γ — *верхньою границею для імовірності* R.

Визначення 7. Проміжок $\left[\underline{R}, \overline{R}\right]$ будемо називати *довірчим інтервалом для параметра* R.

Відмітимо, що довжина $\overline{R} - \underline{R}$ довірчого інтервалу — випадкова величина, а сам інтервал містить («накриває») невідому константу R з імовірністю, не меншою ніж $\gamma' = 1 - 2(1-\gamma)$, якщо $\underline{R} < \overline{R}$ (тут доведення не приводиться).

Класичний результат в загальному вигляді розглянутий Клопером та Пірсоном. Конкретизацію рішення зазначеної задачі дає наступна лема.

Лема 1. У якості γ — границь $\underline{R}$ і $\overline{R}$ можна вибрати статистики

$$\underline{R} = f_2(n, r, \gamma), \quad \overline{R} = f_1(n, r, \gamma), \tag{2}$$

які є коренями рівнянь $1-\gamma = \sum_{k=0}^{r} \binom{n}{k} x^{n-k} (1-x)^k$ та $\gamma = \sum_{k=0}^{r=1} \binom{n}{k} x^{n-k} (1-x)^k$, що розв'язуються при заданих n та γ для кожного реєструємого значення r відносно $x \in [0,1]$. При цьому гарантується, що $P\left(\underline{R} \le R\right) \ge \gamma$ та $P\left(\overline{R} \ge R\right) \ge \gamma$.

Підкреслимо, що γ — границі (2) залежать від числа n випробувань.

Визначення 7. Проміжок $\left[\underline{R}, \overline{R}\right]$ випадкової довжини $\overline{R} - R$ називається *довірчим інтервалом для імовірності* R.

Строгий доказ леми 1 можна отримати за допомогою нерівності Большева.

Доказ леми 1.

Нехай $P\left(\widehat{R} \le x\right) = F(x)$.

Тоді $P\left(\widehat{R} \ge x\right) = 1 - F(x-0) = P\left(r \le c = (1-x)n\right) = I_R\left(n - [c], [c]+1\right)$. По відомій нерівності Большева, можна встановити, що

$$\gamma = P\left(F\left(\widehat{R}-0\right) \le \gamma\right) = P\left(1 - F\left(\widehat{R}-0\right) \ge 1-\gamma\right) =$$
$$= R\left(I_R(n-r, r+1) \ge 1-\gamma \equiv I_{\underline{R}}(n-r, r+1)\right) = P\left(\underline{R} \le R\right).$$

Аналогічно:

$$P\left(\overline{R} \ge\right) = P\left(I_{\overline{R}}(n-r+1, r) \ge I_R(n-r+1, r)\right) =$$
$$= P\left(\gamma \ge B_i(n, R, r-0)\right) = P\left(B_i(n, R, r-0) \le \gamma\right) \ge \gamma$$

або $P\left(\overline{R} \ge R\right) \ge \gamma$.

Лема доведена.

Література

1. Скопа О. О. Принципи вибору формальних параметрів при побудові профілей захисту інфоресурсів [Текст] / Ю. В. Щербина, С. Л. Волков, О. О. Скопа // Східно-Європейський журнал передових технологій. — 2012. — Т. 5, № 2(59). — С. 31—33.
2. Скопа О. О. Концепція контрольних випробувань резервних систем на основі біноміальної схеми [Текст] / О. О. Скопа, С. Л. Волков, А. В. Мінін // Інформаційна безпека. — Луганськ : СНУ ім. В. Даля. — 2011. — № 2(6). — С. 69—76.
3. Скопа О. О. Біноміальні моделі випробування живучості захищених інформаційних каналів [Текст] / А. В. Мінін, О. О. Скопа, М. Александер // Вісник Східноукраїнського національного університету імені Володимира Даля. — Луганськ : СНУ ім. В. Даля. — 2012. — № 8(179). — Ч. 1. — С. 42—58.
4. Казакова Н. Ф. Оцінка живучості систем моніторингу інформаційного простору [Текст] / Н. Ф. Казакова // Східно-Європейський журнал передових технологій. — 2012. — Т. 4, № 2(58). — С. 12—15.
5. Скопа О. О. Статистичне тестування симетричних криптографічних перетворень [Текст] / О. О. Скопа // Східно-Європейський журнал передових технологій. — 2011. — Т. 4, № 9(52). — С. 15—18.

ОЦЕНКА БИНОМИАЛЬНОГО ПАРАМЕТРА R ПО МЕТОДУ КЛОППЕРА — ПИРСОНА

С. Л. Волков

Приводятся лемма, ее доказательство и уточненные определения научных понятий, которые касаются оценки биномиального параметра R по методу Клопера — Пирсона.

Ключевые слова: испытания, биномиальный параметр, статистика, доверительный интервал.

***Сергей Леонидович Волков**, соискатель кафедры Информационно-измерительных технологий Одесской государственной академии технического регулирования и качества, **тел.: (050) 316-71-14, e-mail: greyw@ukr.net**.*

BINOMIAL PARAMETER R ESTIMATION IN CLOPPER — PEARSON METHOD

S. Volkov

Lemma, and also clarified the definition of scientific concepts in the evaluation parameter binomial R Clopper — Pearson method.

Keywords: test, the binomial option, statistics, confidence interval.

***Sergey Volkov**, searcher of the Department of Information and measurement technologies Odessa State Academy of Technical Regulation and Quality, **tel.: (050) 316-71-14, e-mail: greyw@ukr.net**.*