УДК 33.338 DOI: 10.17073/2072-1633-2019-2-186-190

Оптимальность оценки вероятности случайного события

© 2019 г. А.П. Смирнов

Национальный исследовательский технологический университет «МИСиС», 119049, Москва, Ленинский просп., д. 4

Одной из важнейших проблем проектирования производственных систем является обеспечение их надежности. При этом имеется в виду не только техническая надежность технологического оборудования, но и влияние внешних и внутренних случайных факторов, приводящих к сбоям производственного процесса.

В большинстве работ, связанных с исследованием закономерностей случайных процессов в производственных и других технических системах, применяется известная аксиоматика. Соответствующий ей аналитический аппарат позволяет решать задачи оценки вероятности выполнения производственных заданий (например, суточного графика выплавки стали или месячного плана производства).

Известная формула оценки вероятности некоторого случайного события, определяющая указанную вероятность как отношение числа удачных опытов к числу всех опытов, обычно принимается в качестве аксиомы. Нет сомнения, что данная аксиома справедлива, поскольку она всегда подтверждается опытным путем. Однако интересно получить это подтверждение аналитическим путем.

В данной работе приведен аналитический вывод указанной формулы.

В соответствии с частотной аксиоматикой теории вероятностей оценка вероятности некоторого события p^* определяется отношением числа реализаций этого события y в серии из n независимых испытаний к числу этих испытаний.

Величина y имеет биноминальное распределение, которое при достаточно большом n по теореме Лапласа стремится к нормальному с теми же параметрами. Практически нормальным законом можно пользоваться уже при $n > 15$.

Запишем модель оценки p^* как относительную частоту, то есть отношение y к n.

Поскольку p^* есть линейная функция y, то закон распределения p^* также асимптотически нормальный.

Числовые характеристики распределения p^* вычисляются по известным формулам для линейных функций случайных величин. Возникают вопросы: оптимальна ли оценка p^*? Чтобы ответить на этот вопрос, сформулирована оптимизационная задача, для решения которой использован метод множителей Лагранжа.

Решение этой задачи показало, что используемые аксиоматические значения множителей $1/n$ являются оптимальными в смысле минимума дисперсии оценки вероятности случайного события.

Ключевые слова: вероятность случайного события, область интегрирования, функция работоспособности системы, коэффициент готовности системы, биноминальное распределение, числовые характеристики закона распределения, оптимизационная задача

Введение

Среди существующих производственных систем можно выделить простые системы и сложные системы. Простота или сложность здесь связаны с проблемами управления. Простые системы, как правило, характеризуются наличием детерминированных процессов функционирования. В противоположность простым системам сложные системы подвержены действию случайных факторов. Именно такие системы характерны для металлургического производства. Построение алгоритмов управления производственными системами в металлургии требует исследования закономерностей случайных процессов, возникающих при функционировании системы.

Исследование надежности управления производственными процессами в сложных производственных системах требует, как правило, применения аксиом теории вероятностей [1–9].

Практическое применение указанных аксиом, безусловно, подтверждает их справедливость. Однако весьма интересно доказать их справедливость аналитически.

Смирнов А.П. – канд. техн. наук, доцент, alex.p.smirnov@gmail.com– аспирант, rnbafoev@mail.ru

Аналитический вывод формулы оценки вероятности случайного события

В соответствии с частотной аксиоматикой теории вероятностей оценка вероятности некоторого события p^* определяется отношением числа реализаций этого события в серии из n независимых испытаний к числу этих испытаний:

$$\rho^* = \frac{1}{n}\sum_{i=1}^{n} S_i = \frac{y}{n}. \qquad (1)$$

Здесь через S_i обозначены значения функции работоспособности некоторой системы, принимающие значения: $S_i = 1$ (система работала в i-м эксперименте), $S_i = 0$ (система отказала в i-м эксперименте).

Данное определение применяется во всех известных источниках, использующих понятие вероятности случайного события для постановки и решения теоретических и практических проблем, связанных с необходимостью оценки указанной вероятности и принятия решений по управлению некоторой стохастической системой [10–25].

Пусть в каждом опыте производится индикация изучаемого события, которое либо происходит, либо не происходит. Например, это индикация функции S работоспособности системы, принимающей значения 1 или 0. Закон распределения дискретной величины S таков:

$$Bep(S = 1) = \rho, \quad Bep(S = 0) = 1 - \rho.$$

Здесь p – неизвестное число, подлежащее оценке. Оно выражает коэффициент готовности системы. Выразим математическое ожидание и дисперсию случайной величины S через вероятность p:

$$M[S] = 0 \cdot (1-\rho) + 1 \cdot \rho = \rho,$$

$$D[S] = (0-\rho)^2 (1-\rho) + (1-\rho)^2 \rho = \rho(1-\rho). \qquad (2)$$

Вычислим теперь математическое ожидание и дисперсию абсолютной частоты y:

$$y = \sum_{i=1}^{n} S_i, \qquad (3)$$

где S_i – реализация функции S в i-м опыте.

По известным из теории вероятностей формулам [9–13] для числовых характеристик суммы независимых случайных величин имеем:

$$M[y] = \sum_{i=1}^{n} M[S] = n\rho,$$

$$D[y] = \sum_{i=1}^{n} D[S] = n\rho(1-\rho). \qquad (4)$$

Величина y имеет биноминальное распределение, которое при достаточно большом n по теореме Лапласа стремится к нормальному с теми же параметрами. Практически нормальным законом можно пользоваться уже при $n > 15$.

Запишем теперь модель оценки p^* как относительную частоту:

$$\rho^* = \frac{1}{n}\sum_{i=1}^{n} S_i = \frac{y}{n}. \qquad (5)$$

Закон распределения p^* как линейной функции y также асимптотически нормальный.

Числовые характеристики распределения p^* также вычисляются по известным формулам [9–16] для линейных функций случайных величин и в соответствии с формулами (4):

$$M\left[\rho^*\right] = \frac{1}{n} M[y] = \rho,$$

$$D\left[\rho^*\right] = \frac{1}{n^2} D[y] = \frac{\rho(1-\rho)}{n}. \qquad (6)$$

В соответствии с (6) оценка p^* (5) является несмещенной (имеет математическое ожидание p) и состоятельной (дисперсия при увеличении n стремится к нулю).

Возникают вопросы: оптимальна ли оценка p^* (5) и как оценить ее дисперсию, чтобы затем вычислить $\sigma^*[p^*]$? Для ответа на первый вопрос запишем модель (5) в более общем виде:

$$\rho^* = \sum_{i=1}^{n} \alpha_i S_i \qquad (7)$$

Потребуем таких значений коэффициентов α_i, чтобы оценка p^* была несмещенной (то есть $M[p] = p$) и имела минимальную дисперсию $D[p^*]$. Итак, будем искать оптимальную оценку вероятности p^* в смысле минимума ее дисперсии. Оптимизационную задачу сформулируем следующим образом: найти минимум дисперсии

$$D\left[\rho^*\right] = D[S] \cdot \sum_{i=1}^{n} \alpha_i^2 \qquad (8)$$

при ограничении:

$$M\left[\rho^*\right] = M[S] \cdot \sum_{i=1}^{n} \alpha_i = \rho \qquad (9)$$

Учитывая, что $M[S] = \rho$, а $D[S] = \rho(1-\rho)$, перепишем (8) и (9) в виде:

$$D\left[\rho^*\right] = \rho(1-\rho)\sum_{i=1}^{n} \alpha_i^2 \to \min_{\alpha_i} \qquad \sum_{i=1}^{n} \alpha_i = 1, \qquad (10)$$

Для поиска минимума $D[p^*]$ по α_i используем метод множителей Лагранжа и запишем следующую целевую функцию:

$$\varphi = \rho(1-\rho)\sum_{i=1}^{n} \alpha_i^2 - \lambda\left(\sum_{i=1}^{n} \alpha_i - 1\right), \qquad (11)$$

где λ – множитель Лагранжа.

Приравняем к нулю производные по неизвестным коэффициентам α_i:

$$\frac{\partial\varphi}{\partial\alpha_i} = \rho(1-\rho) \cdot 2\alpha_i - \lambda = 0, \; i = 1, n \qquad (12)$$

$$\frac{\partial\varphi}{\partial\lambda} = -\left(\sum_{i=1}^{n} \alpha_i - 1\right) = 0.$$

Просуммировав по i первые n уравнений, получим:

$$2\rho(1-\rho)\sum_{i=1}^{n}\alpha_i - n\lambda = 0. \quad (13)$$

Подставив сюда последнее из уравнений (12), имеем:

$$2\rho(1-\rho) - n\lambda = 0. \quad (14)$$

Отсюда

$$\lambda = \frac{2\rho(1-\rho)}{n}. \quad (15)$$

Подставив (15) в i-е уравнение (12), запишем:

$$2\alpha_i\rho(1-\rho) = \frac{2\rho(1-\rho)}{n}. \quad (16)$$

Отсюда

$$\alpha_i = \frac{1}{n} \quad (17)$$

Полученные значения α_i показали, что оценка значения вероятности случайного события (1, 5) оптимальна в смысле минимума дисперсии.

Заключение

Проблема вычисления вероятности случайного события является типичной при исследовании проблем и поиске закономерностей различных случайных явлений, возникающих при проектировании производственных процессов в металлургии и поиске эффективных алгоритмов управления такими процессами. Нет сомнения, что аксиоматика теории вероятностей представляет верную формулировку для экспериментальной оценки указанной вероятности, подтвержденную неоднократно опытным путем. Однако аналитический вывод формулы оценки вероятности случайного события отсутствовал.

В данной работе классическое аксиоматическое определение формулы для вычисления вероятности случайного события получило свое аналитическое подтверждение в результате представленного доказательства.

Библиографический список

1. *Ананьев Б.И.* Оптимизация оценивания статистики неопределенной системы // Автоматика и телемеханика. 2018. № 1. С. 18–32.

2. *Ефросинин Д.В., Фархадов М.П., Степанова Н.В.* Исследование управляемой системы массового обслуживания с ненадежными неоднородными приборами // Автоматика и телемеханика. 2018. № 2. С. 80–105.

3. *Наумов В.А., Самуйлов К.Е.* Анализ сетей ресурсных систем массового обслуживания // Автоматика и телемеханика. 2018. № 5. С. 59–68.

4. *Кан Ю.С., Соболь В.Р.* Асимптотический доверительный интервал для условной вероятности принятия решений // Автоматика и телемеханика. 2017. № 10. С. 130–138.

5. *Гамкрелидзе Н.Г.* Об одном свойстве симметризованных распределений // Теория вероятн. и ее примен. 2017. Т. 62. № 1. С. 68–71. DOI: 10.4213/tvp5100

6. *Баврин И.И.* Теория вероятностей и математическая статистика. М.: Высшая школа, 2005. 160 с.

7. *Вентцель Е.С., Овчаров Л.А.* Задачи и упражнения по теории вероятностей. М.: Издательский центр «Академия», 2003. 448 с.

8. *Виленкин Н.Я., Виленкин А.Н., Виленкин П.А.* Комбинаторика. М.: ФИМА, МЦНМО, 2006. 400 с.

9. *Вуколов Э.А., Ефимов А.В., Земсков В.Н., Поспелов А.С.* Сборник задач по математике для втузов. Ч. 4. М.: Физматлит, 2004. 432 с.

10. *Гмурман В.Е.* Руководство к решению задач по теории вероятностей и математической статистике. М.: Высшая школа, 2004. 404 с.

11. *Гмурман В.Е.* Теория вероятностей и математическая статистика: учебное пособие для вузов. М.: Высшая школа, 2009. 478 с.

12. *Гнеденко Б.В.* Курс теории вероятностей. М.: Изд-во ЛКИ, 2007. 447 с.

13. *Кибзун А.И.* Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами . М.: Физматлит, 2002. 224 с.

14. *Кибзун А.И., Горяинова Е.Р., Наумов А.В.* Теория вероятностей и математическая статистика. Базовый курс с примерами и задачами. М.: Физматлит, 2007. 231 с.

15. *Кремер Н.Ш.* Теория вероятностей и математическая статистика. М.: ЮНИТИ-ДАНА, 2004. 573 с.

16. *Максимов Ю.Д.* Вероятностные разделы математики. СПб: Иван Федоров, 2001. 592 с.

17. *Горяинов В.Б., Павлов И.В., Цветкова Г.М., Тескин О.И.* Математическая статистика. М.: Изд-во МГТУ им. Н.Э. Баумана, 2001. 424 с.

18. *Письменный Д.Т.* Конспект лекций по теории вероятностей, математической статистике и случайным процессам. М.: Айрис-пресс, 2008. 288 с.

19. *Пугачев В.С.* Теория вероятностей и математическая статистика. М.: Физматлит, 2002. 496 с.

20. *Сидняев Н.И.* Теория планирования эксперимента и анализ статистических данных. М.: Юрайт, 2012. 399 с.

21. *Печинкин А.В., Тескин О.И., Цветкова Г.М.* и др.Теория вероятностей. М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. 456 с.

22. Информатика / под ред. проф. Н.В. Макаровой. СПб: Питер, 2010. 768 с.

23. *Втюрин В.А.* Компьютерные технологии в области автоматизации и управления. СПб: СПбГЛТУ, 2011. 103 с.

24. *Норенков И.П.* Основы автоматизированного проектирования. М.: МГТУ им. Н.Э. Баумана, 2002. 333 с.

25. *Гнеденко Б.В.* Курс теории вероятностей: учебник для студентов математических специальностей университетов. М.: Изд-во ЛКИ, 2007. 447 с.

Ekonomika v promyshlennosti = Russian Journal of Industrial Economics
2019, vol. 12, no. 2, pp. 186–190
ISSN 2072-1633 (print)
ISSN 2413-662X (online)

The optimal estimates of the probability of a random event

A.P. Smirnov – Associated Professor, alex.p.smirnov@gmail.com.

National University of Science and Technology «MISiS», 4 Leninskiy Prospekt, Moscow 119049, Russia

Abstract. One of the most important problems of designing production systems is to ensure their reliability. At the same time, it is meant not only the technical reliability of technological equipment, but also the influence of external and internal random factors that lead to failures of the production process.

In most works related to the study of regularities of random processes in production and other technical systems, the known axiomatics is used. The corresponding analytical tool allows to solve the problem of assessing the probability of production tasks (for example, daily schedule of steel smelting or monthly production plan).

The known formula for estimating the probability of some random event, which determines the specified probability as the ratio of the number of successful experiments to the number of all experiments, is usually accepted as an axiom.

There is no doubt that this axiom is fair, as it is always confirmed by experience. However, it is interesting to obtain this confirmation analytically.

This paper presents an analytical conclusion of this formula.

According to the frequency axiomatics of probability theory, the probability of some event p^* is determined by the ratio of the number of realizations of this event y in a series of n independent tests to the number of these tests.

The value of y has a binomial distribution, which at sufficiently large n by Laplace's theorem tends to normal with the same parameters. Almost normal law can be used already at $n > 15$.

We write the evaluation model p^* as the relative frequency, that is, the ratio y to n.

The distribution law of p^* as a linear function of y is also asymptotically normal.

The numerical characteristics of the distribution of p^* are calculated by known formulas for linear functions of random variables.

Questions arise: is the estimate of p^* optimal and how to estimate its variance to then compute $\sigma^*[p^*]$? To answer these questions, an optimization problem is formulated, for which the Lagrange multiplier method is used.

The solution of this problem has shown that the axiomatic values of the $1/n$ multipliers used are optimal in the sense of the minimum variance of the random event probability estimate.

Keywords: probability of a random event, integration area, system operability function, system availability coefficient, binomial distribution, numerical characteristics of the distribution law, optimization problem

References

1. Anan'ev B.I. Optimizing estimation of a statistically undefined system. *Avtomatika i telemekhanika = Automation and remote*. 2018. No. 1. Pp. 18–32. (In Russ.)
2. Efrosinin D.V., Farkhadov M.P., Stepanova N.V. Study of a Controllable Queueing System with Unreliable Heterogeneous Servers. *Avtomatika i telemekhanika = Automation and remote*. 2018. No. 2. Pp. 80–105. (In Russ.)
3. Naumov V.A., Samuilov K.E. Analysis of Networks of the Resource Queuing Systems. *Avtomatika i telemekhanika = Automation and remote*. 2018. No. 5. Pp. 59–68. (In Russ.)
4. Kang Y.S., Sobol V.R. Asymptotic confidence interval for conditional probability at decision making. *Avtomatika i telemekhanika = Automation and remote*. 2017. No. 10. Pp. 130–138. (In Russ.)
5. Gamkrelidze N.G. On One Property of Symmetrized Distributions. *Theory Probab. Appl*. 2018. Vol. 62. No. 1. Pp. 55–57. DOI: 10.1137/S0040585X97T988484
6. Bavrin I.I. *Teoriya veroyatnostei i matematicheskaya statistika* [Probability theory and mathematical statistics]. Moscow: Vysshaya shkola, 2005. 160 p. (In Russ.)
7. Venttsel' E.S., Ovcharov L.A. *Zadachi i uprazhneniya po teorii veroyatnostei* [Problems and exercises in probability theory]. Moscow: Izdatel'skii tsentr «Akademiya», 2003. 448 p. (In Russ.)
8. Vilenkin N.Ya., Vilenkin A.N., Vilenkin P.A. *Kombinatorika* [Combinatorics]. Moscow: FIMA, MTsNMO, 2006. 400 p. (In Russ.)
9. Vukolov E.A., Efimov A.V., Zemskov V.N., Pospelov A.S. *Sbornik zadach po matematike dlya vtuzov, ch. 4* [Collection of problems in mathematics for technical colleges]. Moscow: Fizmatlit, 2004. 432 p. (In Russ.)
10. Gmurman V.E. *Rukovodstvo k resheniyu zadach po teorii veroyatnostei i matematicheskoi statistike* [Guide to solving problems in probability theory and mathematical statistics]. Moscow: Vysshaya shkola, 2004. 404 p. (In Russ.)
11. Gmurman V.E. *Teoriya veroyatnostei i matematicheskaya statistika* [Probability theory and mathematical statistics] Moscow: Vysshaya shkola, 2009. 478 p. (In Russ.)

12. Gnedenko B.V. *Kurs teorii veroyatnostej: uchebnik dlya studentov matematicheskix special`nostej universitetov* [Course of probability theory: textbook for students of mathematical specialties of universities]. Moscow: Izdatel`stvo LKI, 2007. 447 p. (In Russ.)

13. Kibzun A.I. *Teoriya veroyatnostei i matematicheskaya statistika. Bazovyi kurs s primerami i zadachami* [Probability theory and mathematical statistics. Basic course with examples and tasks]. Moscow: Fizmatlit, 2002. 224 p. (In Russ.)

14. Kibzun A.I., Goryainova E.R., Naumov A.V. *Teoriya veroyatnostei i matematicheskaya statistika. Bazovyi kurs s primerami i zadachami* [Probability theory and mathematical statistics:a basic course with examples and problems]. Moscow: Fizmatlit, 2007. 231 p. (In Russ.)

15. Kremer N.Sh. *Teoriya veroyatnostei i matematicheskaya statistika* [Probability theory and mathematical statistics]. Moscow: UNITY-DANA, 2004. 573 p. (In Russ.)

16. Maksimov Yu.D. *Veroyatnostnye razdely matematiki* [Probabilistic areas of mathematics]. St. Petersburg: Ivan Fedorov, 2001. 592 p. (In Russ.)

17. Goryainov V.B., Pavlov I.V., Tsvetkova G.M., Teskin O.I. *Matematicheskaya statistika* [Mathematical statistics]. Moscow: Iedatel'stvo MGTU im. N.E. Baumana, 2001. 424 p. (In Russ.)

18. Pis'mennyi D.T. *Konspekt lektsii po teorii veroyatnostei, matematicheskoi statistike i sluchainym protsessam* [Lecture Notes on probability theory, mathematical statistics and random processes]. Moscow: Airis-press, 2008. 288 p. (In Russ.)

19. Pugachev V.S. *Teoriya veroyatnostei i matematicheskaya statistika* [The theory of probability and mathematical statistics]: Moscow: Fizmatlit, 2002. 496 p. (In Russ.)

20. Sidnyaev N.I. *Teoriya planirovaniya eksperimenta i analiz statisticheskikh dannykh* [Theory of experiment planning and analysis of statistical data]. Moscow: Yurait, 2012. 399 p. (In Russ.)

21. Pechinkin A.V., Teskin O.I., Tsvetkova G.M. and etc. *Teoriya veroyatnostei* [Probability theory]. Moscow: Iedatel'stvo MGTU im. N.E.Baumana, 2004. 456 p. (In Russ.)

22. *Informatika* [Informatics]. Ed. prof. N.V. Makarovoi. St. Petersburg: Piter, 2010. 768 p. (In Russ.)

23. Vtyurin V.A. *Komp'yuternye tekhnologii v oblasti avtomatizatsii i upravleniya* [Computer technologies in the field of automation and control]. St. Petersburg: SPbGLTU, 2011. 103 p. (In Russ.)

24. Norenkov I.P. *Osnovy avtomatizirovannogo proektirovaniya* [Fundamentals of computer-aided design]. Moscow: Iedatel'stvo MGTU im. N.E.Baumana, 2002. 333 p. (In Russ.)

25. Gnedenko B.V. *Kurs teorii veroyatnostei: uchebnik dlya studentov matematicheskikh spetsial'nostei universitetov* [Course of probability theory: textbook for students of mathematical specialties of universities]. Moscow: Izdatel'stvo LKI, 2007. 447 p.