

ЧЕБЫШЕВСКИЙ СБОРНИК

Том 23. Выпуск 3.

УДК 519.2

DOI 10.22405/2226-8383-2022-23-3-255-261

О плотности распределения вероятностей агрегированной случайной величины для оценки функционирования сложных систем: трехмерный случай¹

Р. А. Жуков, Н. О. Козлова

Жуков Роман Александрович — кандидат физико-математических наук, Тульский филиал Финансового университета при Правительстве РФ (г. Тула).

e-mail: pluszh@mail.ru

Козлова Надежда Олеговна — кандидат технических наук, Тульский филиал Финансового университета при Правительстве РФ (г. Тула).

e-mail: 95kno@mail.ru

Аннотация

Построена плотность распределения вероятностей агрегированной случайной величины, используемая для оценки параметров агрегированной производственной функции, определяемой квадратичной сверткой производственных функций, характеризующих частные результаты функционирования элементов сложной системы. Получены соотношения в квадратурах для трехмерного случая.

Ключевые слова: плотность распределения вероятностей, производственная функция, агрегирование, модель, сложная система.

Библиография: 6 названий.

Для цитирования:

Р. А. Жуков, Н. О. Козлова. О плотности распределения вероятностей агрегированной случайной величины для оценки функционирования сложных систем: трехмерный случай // Чебышевский сборник, 2022, т. 23, вып. 3, с. 255–261.

¹Исследование выполнено за счет гранта Российского научного фонда № 22-28-20061, <https://rscf.ru/project/22-28-20061/> и Тульской области

CHEBYSHEVSKII SBORNIK

Vol. 23. No. 3.

UDC 519.2

DOI 10.22405/2226-8383-2022-23-3-255-261

On the probability distribution densities of an aggregated random variable for evaluating the functioning of complex systems: a three-dimensional case²

R. A. Zhukov, N. O. Kozlova

Zhukov Roman Aleksandrovich — candidate of physics and mathematics sciences, Financial University under the Government of the Russian Federation (Tula Branch) (Tula).

e-mail: pluszh@mail.ru **Kozlova Nadezhda Olegovna** — candidate of technical sciences, Financial University under the Government of the Russian Federation (Tula Branch) (Tula).

e-mail: 95kno@mail.ru

Abstract

The probability distribution density of an aggregated random variable is constructed, which is used to estimate the parameters of an aggregated production function determined by a quadratic convolution of production functions characterizing the particular results of the functioning of elements of a complex system. The relations in quadratures for the three-dimensional case are obtained.

Keywords: probability distribution density, production function, aggregation, model, complex system.

Bibliography: 6 titles.

For citation:

R. A. Zhukov, N. O. Kozlova, 2022, “On the probability distribution densities of an aggregated random variable for evaluating the functioning of complex systems: a three-dimensional case”, *Chebyshevskii sbornik*, vol. 23, no. 3, pp. 255–261.

1. Введение

При поиске параметров производственных функций (ПФ), применяемых при моделировании функционирования элементов сложных систем [1, 2], используют ряд методов, одним из которых является метод максимального правдоподобия (ММП, ML, MLE – maximum likelihood estimation) [3]. При использовании MLE необходимо знать плотность распределения вероятностей случайной величины и случайных величин, в случае, если рассматривается сложная система, которая оценивается по нескольким результативным признакам [4, 5], агрегированным в один интегральный (обобщенный) признак [6]. Когда агрегирование осуществляется простой суммой частных показателей, то построение плотности распределения суммы случайных величин, имеющих нормальный закон распределения, хорошо известно. Однако, в прикладных исследованиях при специальном агрегировании результативных признаков с учетом их взаимного влияния вопрос построения соответствующей им плотности распределения вероятностей агрегированной случайной величины остается открытым. Для двумерного случая получено аналитическое выражение соответствующей плотности распределения вероятностей [6]. В статье рассматривается случай трех переменных.

²The study was carried out at the expense of a grant from the Russian science Foundation № 22-28-20061, <https://rscf.ru/project/22-28-20061/> and Tula region

2. MLE и плотность распределения вероятностей агрегированной случайной величины

MLE подразумевает максимизацию функции правдоподобия вида:

$$\ln L(y^* | C_{i,j}^*, x_{i,j}^*, \sigma_{y^*}) = \sum_{k=1}^K \sum_{t=1}^T f(y^* | C_{i,j}^*, x_{i,j}^*, \sigma_{y^*}) \rightarrow \max. \quad (1)$$

Максимум (1) ищется по набору $C_{i,j}^*$ (параметры стандартизованной i -той ПФ) и σ_{y^*} (стандартное отклонение y_i^*), y_i^* – стандартизованное значение фактического результата функционирования i -того элемента сложной системы y_i , которое определяется по формуле:

$$y_i = f_i(C_{i,j}, x_{i,j}) + \epsilon_i, \quad (2)$$

где i – индекс случайной величины ($i = 1..m \in N$), m – число результативных признаков, $C_{i,j}$ – параметры функций $f_i(\cdot) = \hat{y}_i$ (ожидаемое значение, норматив), $x_{i,j}$ – факторные признаки, ϵ_i – стохастические случайные составляющие (предполагается $N(0; \sigma_{\epsilon_i}^2)$, дисперсия $\sigma_{\epsilon_i}^2$ неизвестна [6]).

Построим плотность распределения случайной величины ϵ , являющейся комбинацией случайных величин ϵ_i из (2), используемых при моделировании функционирования сложных систем с построением агрегированных производственных функций (АПФ), характеризующих подсистемы или систему в целом посредством квадратичной свертки ПФ результатов функционирования элементов системы. Для стандартизованной (центрированной и нормированной) агрегированной случайной величины ϵ^* плотность вероятности будет определяться как:

$$f_p(\epsilon) = \frac{1}{(2\pi)^{m/2} \sqrt{\Delta}} \frac{d}{dy^*} \int \cdots \int_D \exp\left(-\frac{1}{2\Delta} \sum_{i=1}^m \sum_{j=1}^m A_{i,j} \epsilon_i^* \epsilon_j^*\right) dD, \quad (3)$$

где Δ , $A_{i,j}$ – соответственно определитель и алгебраические дополнения корреляционной матрицы $\|r_{i,j}\|$, элементами которой являются парные коэффициенты корреляции;

$$\epsilon_i^* = \frac{(y_i^* - \hat{y}_i^*)^2}{2\sigma_{y_i^*}^2}; \quad (4)$$

D зависит от комбинации ϵ_i^* и соответственно y^* , являющейся комбинацией y_i^* ; $\sigma_{y_i^*}^2$ – дисперсия y_i^* . При этом агрегированные фактические и ожидаемые значения стандартизованного результативного признака подсистемы или системы в целом определяются по формулам:

$$y_k^*(t) = \sqrt{\sum_{i_1=1}^I \sum_{i_2=1}^I r_{i_1, i_2} y_{i_1, k}^*(t) y_{i_2, k}^*(t)}, \quad (5)$$

$$\hat{y}_k^*(t) = \sqrt{\sum_{i_1=1}^I \sum_{i_2=1}^I \hat{r}_{i_1, i_2} \hat{y}_{i_1, k}^*(t) \hat{y}_{i_2, k}^*(t)}, \quad (6)$$

где r_{i_1, i_2} , $\hat{r}_{i_1, i_2}$ – соответствующие значения парного коэффициента корреляции Пирсона между i_1 -тыми $y_{i_1}^*$, $\hat{y}_{i_1}^*$ и i_2 -тыми $y_{i_2}^*$, $\hat{y}_{i_2}^*$ переменными ($i_1, i_2 = 1..I$) для k -той подсистемы в период t . Выражение (6), записанное для переменной $\hat{y}^*$, принимающей значения $\hat{y}_k^*(t)$, представляет собой стандартизованную АПФ.

В трехмерном случае требуемую плотность вероятности $f_p(y^*)$ можно найти, используя соотношение:

$$\begin{aligned} f_p(y^*) &= \frac{d}{dy^*}(F(y^*)) = Pr\left\{\sum_{i=1}^3 \sum_{j=1}^3 r_{i,j} y_i^* y_j^* \leq (y^*)^2\right\} = \\ &= \frac{d}{dy^*} \left(\iiint_{\sum_{i=1}^3 \sum_{j=1}^3 r_{i,j} y_i^* y_j^* \leq (y^*)^2} f_p(y_1^*, y_2^*, y_3^*) dy_1^* dy_2^* dy_3^* \right). \end{aligned} \quad (7)$$

Здесь $F(y^*)$ – функция распределения случайной величины y^* . Для приведения квадратичной формы $\sum_{i=1}^3 \sum_{j=1}^3 r_{i,j} y_i^* y_j^*$ к каноническому виду можно найти собственные числа и собственные векторы матрицы $\|r_{i,j}\|$ или воспользоваться методом Лагранжа.

В первом случае достаточно составить характеристическое уравнение и получить его корни непосредственным вычислением с помощью, например, формулы Кордано. В соответствии с известной теоремой о вещественности собственных значений вещественной симметричной матрицы можно утверждать, что собственные числа матрицы $R = \|r_{i,j}\|$ также будут вещественными. В качестве дополнения к теореме достаточно показать, что в формуле Кордано выражение для определителя кубического уравнения $Q \leq 0$, то есть:

$$Q = r_{12}^2 r_{13}^2 r_{23}^2 - \frac{1}{27} (r_{12}^2 + r_{13}^2 + r_{23}^2)^3 < 0. \quad (8)$$

Для этого найдем экстремумы соответствующей функции и покажем, что максимальное значение не превосходит 0. Определим стационарные точки и воспользуемся критерием Сильвестра.

Найдем первые частные производные и приравняем их к 0:

$$\begin{cases} \frac{\partial Q}{\partial r_{12}} = r_{12} (9r_{13}^2 r_{23}^2 - (r_{12}^2 + r_{13}^2 + r_{23}^2)^2) \\ \frac{\partial Q}{\partial r_{13}} = r_{13} (9r_{12}^2 r_{23}^2 - (r_{12}^2 + r_{13}^2 + r_{23}^2)^2) \\ \frac{\partial Q}{\partial r_{23}} = r_{23} (9r_{12}^2 r_{13}^2 - (r_{12}^2 + r_{13}^2 + r_{23}^2)^2). \end{cases} \quad (9)$$

Из (9) следует, что соотношения выполняются, если

$$(\pm r_{12}) = (\pm r_{13}) = (\pm r_{23}) = r. \quad (10)$$

Следовательно стационарные точки r_{0i} имеют координаты $(\pm r_{12}; \pm r_{13}; \pm r_{23})$.

Найдем вторые производные и составим матрицу Гессе с учетом (10):

$$H(Q) = \begin{pmatrix} -12r^4 & \pm 6r^4 & \pm 6r^4 \\ \pm 6r^4 & -12r^4 & \pm 6r^4 \\ \pm 6r^4 & \pm 6r^4 & -12r^4 \end{pmatrix}. \quad (11)$$

В соответствии с критерием Сильвестра форма является отрицательно определенной, если ее главные миноры M_i имеют чередующиеся знаки, начиная с отрицательного, и в этом случае найденная точка является максимумом.

В данном случае $M_1 < 0$, $M_2 > 0$, $M_3 = 0 \forall r \in [-1; 1]$. Следовательно, квадратичная форма $H(r) = \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial^2 Q}{\partial r_i \partial r_j} r_i r_j$ ($r_1 = r_{12}, r_2 = r_{13}, r_3 = r_{23}$) неположительна. Можно доказать, что в точке, определенной соотношением (10), Q будет иметь максимум. Достаточно рассмотреть ее значение в окрестности точки r_{0i} . Составим разность

$$\Delta Q = Q(r_{12}, r_{13}, r_{23}) - Q(r_{12} + h_1, r_{13} + h_2, r_{23} + h_3), \quad (12)$$

где $h_1, h_2, h_3 > 0$, $h_1, h_2, h_3 < h$ и $h_1 \ll r_{12}, h_2 \ll r_{13}, h_3 \ll r_{23}$. С учетом (10) первым ненулевым коэффициентом в (12) будет являться коэффициент при r^4 , то есть:

$$\Delta Q = \frac{18}{27}((h_1 - h_2)^2 + (h_1 - h_3)^2 + (h_2 - h_3)^2)r^4 + O(h^3)r^3, \quad (13)$$

где $O(h^3)$ – бесконечно малая третьего порядка малости.

Ограничиваясь этим членом, можно заключить, что $\Delta Q > 0, \forall h_1, h_2, h_3 < h$. Следовательно, точка r_{0i} является максимумом. На рис. 1 изображены значения Q при различных значениях r_{12}, r_{13}, r_{23} . Формулы Кордано для вычисления собственных чисел неудобны в силу

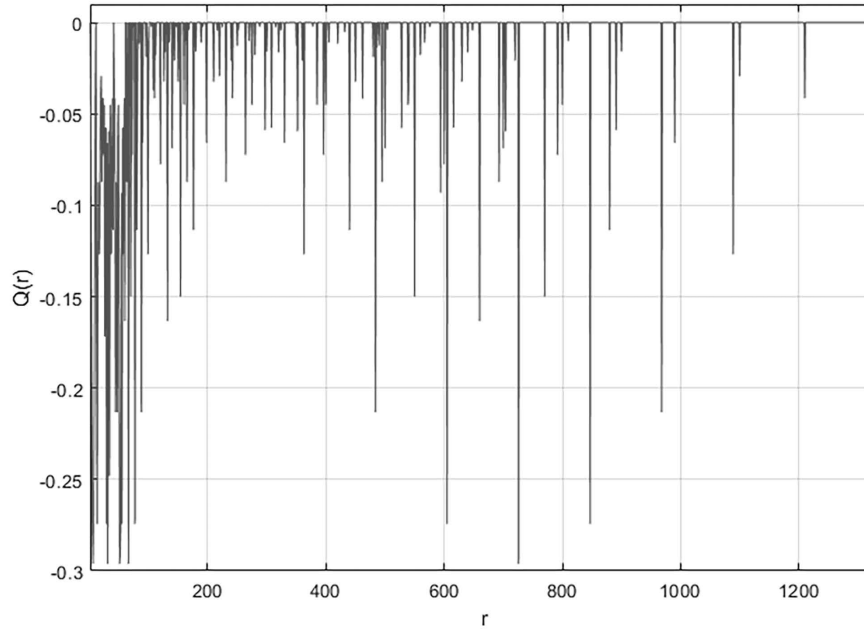


Рис. 1: Значение определителя Q для некоторых значений r_{12}, r_{13}, r_{23} (r_{12}, r_{13}, r_{23} последовательно изменяются в интервале $[-1;1]$ с шагом 0,2. Число точек 1331)

их громоздкости и дальнейшего использования при преобразовании плотности распределения $f_p(y_1^*, y_2^*, y_3^*)$ в (7).

Целесообразно воспользоваться методом Лагранжа для приведения области интегрирования к каноническому виду. Сделаем следующую замену переменных:

$$\begin{aligned} y_1^* &= u_1^* - \frac{r_{12}}{\sqrt{K_{33}\Delta}}u_2^* - \left(\frac{r_{12}K_{23}}{\sqrt{K_{33}\Delta}} + r_{13}\sqrt{K_{33}}\right)u_3^*, \\ y_2^* &= \frac{r_{12}}{\sqrt{K_{33}\Delta}}(u_2^* + K_{23}\sqrt{\Delta}u_3^*), \\ y_3^* &= \sqrt{K_{33}}u_3^*. \end{aligned} \quad (14)$$

Здесь Δ определяется по формуле:

$$\Delta = 1 + 2r_{12}r_{13}r_{23} - r_{12}^2 - r_{13}^2 - r_{23}^2. \quad (15)$$

Тогда область интегрирования в (7) преобразуется к виду:

$$(u_1^*)^2 + (u_2^*)^2 + (u_3^*)^2 \leq (y^*)^2, \quad (16)$$

а якобиан в подынтегральном выражении будет равен:

$$J = \begin{vmatrix} 1 & -\frac{r_{12}}{\sqrt{1-r_{12}^2}} & \frac{r_{12}r_{23}-r_{13}}{\sqrt{1-r_{12}^2}\Delta} \\ 0 & \frac{1}{\sqrt{1-r_{12}^2}} & \frac{r_{23}-r_{12}r_{13}}{\sqrt{1-r_{12}^2}\Delta} \\ 0 & 0 & \frac{\sqrt{1-r_{12}^2}}{\sqrt{\Delta}} \end{vmatrix}. \quad (17)$$

Переходя к сферической системе координат и взяв производную по радиусу r , получим выражение для плотности вероятности $f_p(y^*)$:

$$\begin{aligned} f_p(y^*) = & \frac{1}{(2\pi)^{3/2}\sqrt{\Delta}} \int_0^{2\pi} \int_0^\pi (y^*)^2 \sin\theta \exp[-\frac{1}{2}(y^*)^2 \cdot \\ & \cdot (c_{11}\cos^2\phi\sin^2\theta + c_{22}\sin^2\phi\sin^2\theta) + c_{22}\cos^2\theta + \\ & + c_{12}\sin 2\phi\sin^2\theta + c_{13}\cos\phi\sin 2\theta + c_{23}\sin\phi\sin 2\theta] d\theta d\phi, \end{aligned} \quad (18)$$

где

$$\begin{aligned} c_{11} &= K_{11} \\ c_{22} &= \frac{1}{K_{33}\Delta} (K_{11}r_{12}^2 + K_{22} - 2K_{12}r_{12}), \\ c_{33} &= \frac{1}{K_{33}} (r_{12}K_{23} + r_{13}K_{33})(K_{11}r_{12}K_{23} + K_{11}r_{13}K_{33} - \\ & - 2K_{12}K_{23} - 2K_{13}K_{33}) + K_{33}^2 + 2K_{23}^2 + \frac{K_{22}K_{23}^2}{K_{33}}, \\ c_{12} &= \frac{1}{\sqrt{K_{33}\Delta}} (-K_{11}r_{12} + K_{12}) \\ c_{13} &= \frac{1}{\sqrt{K_{33}}} (-K_{11}r_{12}K_{23} - K_{11}r_{13}K_{33} + K_{12}K_{23} + K_{13}K_{33}), \\ c_{23} &= \frac{1}{K_{33}\sqrt{\Delta}} (K_{11}r_{12}(r_{12}K_{23} + r_{13}K_{33}) + K_{22}K_{23} - \\ & - K_{12}(2r_{12}K_{23} + r_{13}K_{33}) - r_{12}K_{13}K_{33} + K_{23}K_{33}), \end{aligned} \quad (19)$$

$K_{i,j}^{-1} = A_{ij}/\Delta$ – элементы обратной матрицы ковариаций.

3. Заключение

Полученные соотношения могут быть использованы при поиске параметров агрегированной производственной функции и уточнения параметров трех производственных функций, входящих в состав АПФ.

Предложенный алгоритм можно расширить на m -мерный случай, когда в рассматриваемую подсистему или систему в целом включены более трех элементов. Однако, уже в трехмерном случае плотность распределения вероятностей удастся построить только численно.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

1. Клейнер Г. Б. Производственные функции: теории, методы, применение. — Москва: Финансы и статистика, 1986. 239 с.

2. Жуков Р. А. Оценка эффективности функционирования социально-экономических систем на основе производственных функций: новый подход // Вестник Волгоградского государственного университета. Экономика. 2019. Том 21. № 3. С. 71-82. DOI: 10.15688/ek.jvolsu.2019.3.7.
3. Ллойд Э., Лендерман У. Справочник по прикладной статистике. Том 1. — Москва: Финансы и статистика, 1989. 510 с.
4. Айвазян С. А., Афанасьев М. Ю., Руденко В. А. Некоторые вопросы спецификации трехфакторных моделей производственного потенциала компании, учитывающих интеллектуальный капитал // Прикладная эконометрика. 2012. Том 27. № 3. С. 36–69.
5. Айвазян С. А., Афанасьев М. Ю., Руденко В. А. Исследование зависимости случайных составляющих стохастической производственной функции при оценке технической эффективности // Прикладная эконометрика. 2014. Том 34. № 2. С. 3-18.
6. Жуков Р. А. Метод оценки результатов функционирования иерархических социально-экономических систем на основе агрегированной производственной функции // Экономика и математические методы. 2021. Том 57. № 3. С. 17-31. DOI: 10.31857/S042473880016428-9.

REFERENCES

1. Kleiner G. B., 1986, “*Production functions: theories, methods, application*“, Moscow, Finance and Statistics, 239 p.
2. Zhukov R. A., 2019, “Efficiency Evaluation of Functioning of Socio-Economic Systems Based on Production Functions: New Approach“, *Journal of Volgograd State University. Economics* , vol. 21 no. 3, pp. 71-82. DOI: 10.15688/ek.jvolsu.2019.3.7.
3. Lloyd E., Lederman W., 1989, “*Handbook of Applied Statistics. Volume 1.*“, Moscow, Finance and Statistics, 510 p.
4. Ayvazyan S. A., Afanasyev M. Yu., Rudenko V. A., 2012, “Some Issues on the Specification of Three-Factor Models of the Company’s Production Potential Taking into Account Intellectual Capital“, *Applied Econometrics*, vol. 27, no. 3, pp. 36-69.
5. Ayvazyan S. A., Afanasyev M. Yu., Rudenko V. A., 2014, “Analysis of Dependence between the Random Components of a Stochastic Production Function for the Purpose of Technical Efficiency Estimation“, *Applied Econometrics*, vol. 34, no. 2, pp. 3-18.
6. Zhukov R. A., 2021, “Method for assessing the results of hierarchical socio-economic systems’ functioning based on the aggregated production function“, *Economics and Mathematical Methods*, vol.57, no. 3, pp.17-31. DOI: 10.31857/S042473880016428-9.

Получено 11.04.2022

Принято в печать 14.09.2022