

Анализ укомплектованности подразделений по защите информации в субъектах Российской Федерации на основе регрессионного моделирования

С.И. Носков, А.П. Медведев

Иркутский государственный университет путей сообщения

Аннотация: Эффективным методом исследования и прогнозирования численности сотрудников структурных подразделений является регрессионный анализ, основанный на использовании статистических данных и их обработке специальными методами. В данной работе на основе статистической информации по 81 региональному отделению Социального фонда России проведен регрессионный анализ укомплектованности отдельных подразделений по защите информации с учетом общей площади и численности населения регионов. Показано, что для ряда субъектов имеет место недоукомплектованность штатного кадрового состава, а для некоторых из них, напротив, его перенасыщенность.

Ключевые слова: защита информации, регрессионная модель, критерии адекватности, прогнозирование, анализ штатной численности, подразделения по защите информации.

Использование методов математического моделирования является весьма эффективным средством анализа и прогнозирования сложных многофакторных объектов. Подобные методы позволяют выявлять и интерпретировать различного рода неявные связи между составными факторами, что, в конечном итоге, способствует повышению качества управления подобными объектами. Так, в работе [1] описывается разработка математической модели планирования человеческих ресурсов в производственном процессе. В статье [2] производится расчет их оптимального распределения при проектировании, эксплуатации и техническом обслуживании основных производственных фондов крупной компании.

Методы регрессионного анализа являются хорошо содержательно интерпретируемыми формальными средствами, основанными на построении математических моделей по статистическим данным и, как утверждается в статье [3], уже давно занимают центральное место в области обработки экономической статистики. Подходы, основанные на построении линейных и

нелинейных регрессий на основе статистических данных, могут использоваться как для анализа внутренних процессов в структурном подразделении, например, определения его численности в разрезе функциональных обязанностей сотрудников, так и для выявления неявных зависимостей между такими факторами, как, в частности, выручка и объем выпускаемой продукции. Пример подобной модели рассматривается в статье [4]. В работе [5] показаны основные преимущества метода построения нелинейных регрессий и приведены основные подходы к определению некоторых ключевых параметров. В [6] описывается пример использования подобных методов для оценки эффективности деятельности учебного заведения. Необходимо отметить, что построение и использование регрессионных моделей часто сопровождается анализом свойств ошибок аппроксимации. Так, в работе [7] предложен метод оценивания параметров модели путем минимизации средней и максимальной относительных ошибок аппроксимации.

Как было замечено ранее в работе [8], для достоверной оценки и прогнозирования численности подразделений по защите информации необходимо выполнение условий однородного функционала подразделений, общей сферы деятельности организаций и единого периода наблюдения.

В настоящей работе в качестве информационной базы моделирования использованы данные о штатной численности подразделений по защите информации органов Социального фонда России (далее – СФР) в 81 субъекте Федерации, а также площадь и численность населения в них по состоянию на 2024 г. [9].

Как это уже было отмечено выше, одним из этапов оценки качества регрессионных моделей является анализ свойств их ошибок аппроксимации. С этой целью в соответствии с подходом, представленным в работе [10], поставим задачу максимизации числа наблюдений выборки, для которых

модуль ошибки аппроксимации в регрессионной модели не превышает некоторой допустимой величины. Дадим краткое описание этого подхода.

Рассмотрим линейное регрессионное уравнение (модель) вида:

$$y_k = \sum_{i=1}^m \alpha_i x_{ki} + \varepsilon_k, k = \overline{1, n}, \quad (1)$$

где y – зависимая, а x_i – i -я независимая переменная, α_i – i -й подлежащий оцениванию параметр, ε_k – ошибки аппроксимации, k – номер наблюдения, n – их число.

Поставим задачу максимизации числа наблюдений выборки, относительные ошибки аппроксимации для которых допустимы:

$$|S| \rightarrow \max, \quad (2)$$

Здесь $S = \{k \in \{1, 2, \dots, n\} \mid |\varepsilon_k| \leq dy_k\}$ – множество номеров наблюдений выборки, относительные ошибки аппроксимации для которых допустимы, d – допустимый уровень средней относительной ошибки аппроксимации, $|S|$ – число элементов в множестве S , его мощность.

Как показано в работе [10] задача (2) эквивалентна следующей задаче линейно-булева программирования с ограничениями:

$$\sum_{i=1}^m \alpha_i x_{ki} + u_k - v_k = y_k, k = \overline{1, n}, \quad (3)$$

$$u_k \geq 0, v_k \geq 0, k = \overline{1, n}, \quad (4)$$

$$u_k + v_k + M\sigma_k \leq M + dy_k. \quad (5)$$

$$\sigma_k \in \{0, 1\}, k = \overline{1, n}. \quad (6)$$

и целевой функцией

$$\sum_{k=1}^n \sigma_k - \delta \sum_{k=1}^n (u_k + v_k) \rightarrow \max, \quad (7)$$

где δ – наперед заданное малое положительное число, сравнимое с нулем, а M – заданная большая положительная константа.

Применим описанный способ идентификации параметров линейной регрессии для моделирования численности подразделений по защите информации региональных органов СФР.

Введем следующие обозначения:

y – штатная численность подразделения, чел.;

x_1 – площадь территории, млн. км²;

x_2 – численность населения, млн. чел.

Разумеется, полный состав влияющих на зависимую переменную факторов шире данного, но очевидно, что именно выделенные независимые переменные должны быть определяющими при формировании штатной численности подразделений по защите информации.

С помощью изложенного выше подхода будем строить линейную двухфакторную модель без свободного члена вида:

$$y_k = \alpha_1 x_{k1} + \alpha_2 x_{k2} + \varepsilon_k, k = \overline{1,81}. \quad (8)$$

При этом положим $d = 0,2$, поскольку, как отмечено, например, в [11], можно считать относительно приемлемыми ошибки аппроксимации до 20%.

В результате решения задачи (3) – (7) получим:

$$y = 5,12x_1 + 8,38x_2 + \varepsilon_k, k = \overline{1,81}. \quad (9)$$

$$E = 608, \tilde{E} = 65\%,$$

где

$$E = \sum_{k=1}^n |\varepsilon_k| \text{ – сумма модулей относительных ошибок аппроксимации,}$$

$$\tilde{E} = \frac{E}{\sum_{k=1}^n y_k} 100\% \text{ – относительная процентная ошибка.}$$

Низкое качество модели (9), в соответствии со значениями используемых критериев адекватности, указывает на то, что в настоящее время формирование штатной численности подразделений по защите

информации в регионах России часто осуществляется произвольным образом, без должного научного обоснования и учета объективных обстоятельств.

Стоит отметить, что попытка введения дополнительных квадратичных и кубических переменных $x_1^2, x_2^2, x_1^3, x_2^3$ в модель (8) не приводит к ее заметному улучшению.

В табл. 1 приведен состав множества S для модели (9).

Таблица № 1

Состав множества S для модели (9)

d	Номера регионов в множестве S
0,2	1,5,7,9,17,19,21,22,24,30,32,34,38,39,40,41,43,47, 48,51,52,54,55,56,59,64,67,70,74,75,77,78,79

В результате анализа этого множества выявлены регионы с наибольшей отрицательной и положительной ошибкой, их номера входят в состав множества $\{1,2,\dots,n\} \setminus S$. При этом положительное значение ошибки аппроксимации свидетельствует об избытке кадров в подразделении соответствующего региона, а отрицательные, напротив, о его недоукомплектованности.

Так, ниже приведены номера пяти регионов с наиболее характерными показателями избытка (табл. 2) и недоукомплектованности (табл. 3) кадрового состава подразделений по защите информации органов СФР.

Информация, содержащаяся в табл. 1 – 3, может послужить серьезным основанием для принятия соответствующими органами федерального уровня управленческих кадровых решений.

Таблица № 2

Региональные подразделения с избыточным
кадровым составом

Номер региона	$y_{набл}$, чел.	$y_{расчет}$, чел.	$y_{набл} - y_{расчет}$, чел.
2	24	10	14
80	14	4	10
44	15	7	8
18	13	5	8
62	5	2	3

Таблица № 3

Региональные подразделения с недоукомплектованностью
кадрового состава

Номер региона	$y_{набл}$, чел.	$y_{расчет}$, чел.	$y_{набл} - y_{расчет}$, чел.
46	3	12	-9
73	7	24	-17
35	6	27	-21
45	5	34	-29
10	36	183	-147

Выводы

В работе предложен подход к анализу кадрового состава подразделений по защите информации в регионах России, основанный на применении метода регрессионного анализа, сводящийся к максимизации числа допустимых относительных ошибок аппроксимации линейной регрессионной модели. В результате проведенного исследования показано, что в некоторых региональных подразделениях по защите информации имеет

место избыток кадрового состава, а некоторые их них, напротив, существенно недоукомплектованны. Полученные результаты могут иметь практическую значимость для руководства СФР при планировании численности указанных подразделений.

Литература

1. Kostic A., Maric B., Kustura M, Timotic V. Mathematical model for human resources planning in the production process // 32nd DAAAM international symposium on intelligent manufacturing and automation, 2021, pp. 4-9.

2. Shen L., Liu K., Chai J., Ma W., Guo X., Li Y., Zhao P., Liu B. Research on the mathematical model for optimal allocation of human resources in the operation and maintenance units of a heavy haul railway // Mathematics, 2022, No. 10, pp. 1-18.

3. Sykes A. An Introduction to Regression Analysis, 1993, 34 p.

4. Носков С.И., Афонин М.В., Бычков Ю.А., Медведев А.П., Торопов В.Д. Нелинейная регрессионная модель функционирования горно-металлургической компании // Инженерный вестник Дона, 2024, №4. URL: ivdon.ru/ru/magazine/archive/n4y2024/9119

5. Ruckstuhl A. Introduction to Nonlinear Regression, 2010, 30 p.

6. Eric Ng, Hong C., Zheng L., Ramesh K. M. H. M. An Effectiveness of Human Resource Management Practices on Employee Retention in Institute of Higher learning: -A Regression Analysis // International Journal of Business Research and Management (IJBRM), 2012, Vol. 3, pp. 60-79.

7. Носков С.И. Минимизация средней и максимальной относительных ошибок аппроксимации регрессионной модели // Известия Тульского государственного университета. Технические науки, 2023, № 1, С. 340-343.



8. Носков С.И., Медведев А.П. Регрессионное моделирование штатной численности подразделений по защите информации // Инженерный вестник Дона, 2024, № 6. URL:ivdon.ru/ru/magazine/archive/n6y2024/9283.

9. Официальный сайт Федеральной службы государственной статистики. URL:rosstat.gov.ru (дата обращения: 11.11.2024).

10. Носков С.И., Шахуров А.Н. Максимизация числа допустимых ошибок аппроксимации при построении линейной регрессионной модели // Вестник Югорского государственного университета, 2024, Т.20, № 3, С. 57-62.

11. White I.R. Simsum: Analyses of simulation studies including Monte Carlo error // The Stata Journal, 2010, Vol. 10, No. 3, pp. 369-385.

References

1. Kostic A., Maric B., Kustura M, Timotic V., 32nd DAAAM international symposium on intelligent manufacturing and automation, 2021, pp. 4-9.

2. Shen L., Liu K., Chai J., Ma W., Guo X., Li Y., Zhao P., Liu B. Mathematics, 2022, No. 10, pp. 1-18.

3. Sykes A. An Introduction to Regression Analysis, 1993, 34 p.

4. Noskov S.I., Afonin M.V., Bychkov Yu.A., Medvedev A.P., Toropov V.D. Inzhenernyj vestnik Dona, 2024, №4. URL: ivdon.ru/ru/magazine/archive/n4y2024/9119.

5. Ruckstuhl A. Introduction to Nonlinear Regression, 2010, 30 p.

6. Eric Ng, Hong C., Zheng L., Ramesh K. M. H. M. International Journal of Business Research and Management (IJBRM), 2012, Vol. 3, pp. 60-79.

7. Noskov S.I. Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki, 2023, № 1, P. 340-343.

8. Noskov S.I, Medvedev A.P., Inzhenernyj vestnik Dona, 2024, № 6. URL:ivdon.ru/ru/magazine/archive/n6y2024/9283.



9. The official website of the Federal State Statistics Service of Russia.
URL:rosstat.gov.ru (accessed: 11.11.2024).

10. Noskov S.I., Shakhurov A.N., Vestnik Yugorskogo gosudarstvennogo universiteta, 2024, Vol.20, № 3, pp. 57-62.

11. White I.R., The Stata Journal, 2010, Vol. 10, No. 3, pp. 369-385.

Дата поступления: 9.11.2024

Дата публикации: 15.12.2024