

Представлены результаты исследования генератора Джеффи при разном количестве базовых генераторов на основе регистров сдвига с линейным обратной связью и разной степени их полиномов, которое проведено с использованием статистических тестов NIST. Важное значение среди генераторов псевдослучайных последовательностей занимает генератор Джеффи, однако его качественные характеристики являются малоисследованными. Полученные результаты позволяют оптимизировать параметры генератора при заданных параметрах исходной импульсной последовательности.

Приведены принципы оптимизации параметров структурных элементов генератора Джеффи. Качество такой оптимизации подтверждена пакетом статистических тестов NIST STS.

Ключевые слова: генератор псевдослучайных чисел, статистические характеристики, генератор Джеффи.

Otenko V.I., Harasymchuk O.I., Zhuravel I.M., Kostiv Yu.M., Pastukh A.Yu. The Optimization of Parameters Related to Geffe Generator Structural Elements

The results of the study of Geffe generator at different base number of generators based on linear feedback shift register and varying degrees of polynomials, conducted using statistical tests NIST, are presented. Geffe generator plays an important role within pseudorandom sequence generators, although its quality characteristics are scarcely explored. The obtained results allow optimizing generator parameters for a given output pulse sequence parameters. Some principles for parameter optimization of structural elements of Geffe generator are provided. The quality of such optimization package was confirmed by statistical NIST STS tests.

Key words: pseudorandom generator, statistic characteristics, Geffe generator.

УДК 629.1 *Доц. В.А. Газетдінов, канд. техн. наук – Черкаський ДТУ*
**НЕЙРОМЕРЕЖЕВЕ ПРОГНОЗУВАННЯ РИНКУ НЕРУХОМОСТІ
У КРИЗОВИХ УМОВАХ**

Здійснено прогнозування ринку нерухомості у кризових умовах за допомогою нейромережі. Для нейромережевого прогнозування ринку нерухомості у кризових умовах створено зведену таблицю з даними про об'єкти нерухомості. Наведено схематичний процес навчання нейромережі. На підставі даних про об'єкти нерухомості проведено навчання нейромережі. Застосовано методи математичної статистики, засновані на сукупності певних правил для точних цифрових даних узагальненого характеру. Загальний прогноз за нейромережевого прогнозування ринку нерухомості у кризових умовах знайдено шляхом множення значення детермінації на кожне індивідуальне значення, отримане на основі результату нейромережевого моделювання.

Ключові слова: нейромережа, навчання, ринок нерухомості, прогнозування, нейрон, комірка.

Постановка проблеми. В умовах сьогодення ринок нерухомості перебуває не у кращому стані. Це є впливом економічної кризи, що склалася у межах держави, нестійким попитом, що є наслідком цього, та нестабільною позицією. На сьогодні питання прогнозування ринку нерухомості для багатьох є актуальним питанням, що потребує досліджень та інновацій. Аналітична робота у сфері нерухомості складна та багатогранна, що вимагає дослідження методологій у підходах та правилах інтерпретації даних і алгоритму вироблення рекомендацій.

На сьогодні, в умовах зростання соціальної інфраструктури, ділової активності, потреб забезпеченості житлом, недостатність аналітичних оцінок у

рамках ринку нерухомості є прямим чинником впливу на територіальний розвиток, що, водночас, приводить до більшості криз місцевих громад.

Упродовж останніх кількох років, на основі програми розвитку державного регулювання ринку нерухомості діє методика формування інформаційної бази. В її основу закладено отримання інформації про ринок нерухомості, житловий фонд, ринок будівельних матеріалів та ін. Проте ця методика є доволі складною та потребує вдосконалення. Тому проблема прогнозування ринку нерухомості є актуальною в умовах сьогодення.

Ступінь дослідження в науковій літературі. Фундаментальну основу в розвитку теорії нейрокомп'ютинга і його застосування у фінансовій сфері склали вчені країн Заходу і США. Це насамперед: Д.-Е. Бестенс, Ван ден Берг [1], D.E. Rummelhart, G.E. Hinton, R.G. Williams [2], R.B. Berrens, M. McKee [3], J.B. Ramsey [4], D.F. Specht [5] та ін. До вітчизняних вчених варто віднести: А. Єжова [6], Б. Одинцова, А. Романова, С. Шумського, В.А. Бившого, А.І. Богомолова, В.І. Костюніна [7] та ін., котрі розробляють і впроваджують нейромережеві технології у галузі економіки.

З методологічної точки зору, моделювання динаміки макроекономічних показників у площині статистичного моделювання і прогнозування вивчено вітчизняними авторами, зокрема Г.М. Стерніком [8], Ж.А. Морозовою [9]. Механізм побудови класифікацій методів прогнозування та методик прийняття рішення про вибір оптимального методу прогнозування досліджено у роботах таких науковців, як Н. Wittkemper, M. Steiner [10] та ін.

Мета роботи. Здійснити прогнозування ринку нерухомості у кризових умовах за допомогою нейромережі. Навести схематичний процес навчання нейромережі. На підставі даних про об'єкти нерухомості провести навчання нейромережі. Застосувати методи математичної статистики, засновані на сукупності певних правил для точних цифрових даних узагальненого характеру.

Виклад основного матеріалу. На сьогодні, у рамках сформованої ситуації на Сході України, і наслідків, що спричинили крах української економіки, прогнозування ринку нерухомості є досить актуальною проблемою. На основі проведеного дослідження методологічної бази прогнозування ринку нерухомості [8, 9] на перше місце виходить багатофакторне прогнозування – нейронне моделювання. Цей метод моделювання заснований на багатофакторній моделі нейронних мереж. В основу цього методу покладено структурний апарат, що має у своєму складі шість окремих параметрів, які певною мірою взаємопов'язані. Перший – це таблиця, вона містить певний склад індикаторів стану ринку по всіх його сегментах, який визначається на основі даних про об'єкти нерухомості (площа, термін введення в експлуатацію, якість, місце розташування).

Комірка дискретної моделі – це дані по одному конкретному сегменту у локальному ринку нерухомості. Ще одним, не менш важливим параметром, є просторова вісь, що визначає для кожної конкретної комірки її індивідуальне місце розташування або ж його адресу.

Параметр комірки описує безпосередньо набір даних про ціну як максимальну, так і мінімальну, про загальну кількість об'єктів, що потрапили у зону вибірки, про коливання ціни даного сегмента, про статистичні параметри попиту та пропозиції. Так само є таке поняття як набір комірок, це насамперед визначений шар із загальної картини, тобто якийсь набір комірок, схожих за вибірковими даними. В основі описаної моделі знаходиться штучні нейронні мережі, це свого роду програмно-апаратні реалізації математичних моделей, в основі побудови яких лежить принцип мереж нервових клітин живого організму.

Нейронна мережа – це безліч нейронів, зв'язки між якими задаються ваговими коефіцієнтами. Водночас, нейрон – це структурно-функціональна одиниця, яка реалізує операцію нелінійного перетворення суми вхідних сигналів на вагові коефіцієнти [1]. Важливим моментом є можливість штучних нейронних мереж піддаватися навчанню. Основою навчання є налаштування її вагових зв'язків для ефективного виконання конкретного завдання, а також топології.

Навчання нейромережі в рамках прогнозування ринку нерухомості у кризових умовах ґрунтується на послідовній генерації за певною процедурою вхідних векторів з одночасним підстроюванням ваг, щоб внаслідок роботи нейромережі для певної множини входів надавала бажану безліч виходів. Головною метою навчання мережі є максимальна мінімізація помилки у процесі формування виходів із заданих входів у нейронну мережу. Загальний алгоритм навчання нейронних мереж для успішного прогнозування ринку нерухомості складається з низки періодів, що взаємодіють поетапно у процесі навчання (рис. 1).

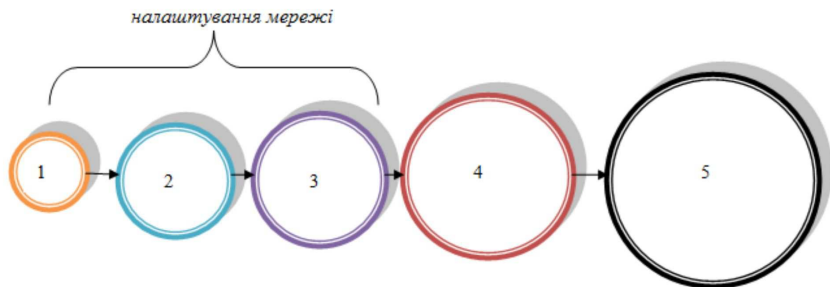


Рис. 1. Навчання нейронних мереж для успішного прогнозування ринку нерухомості. Поетапна структура (розробка автора на основі джерел [1, 5])

На першому етапі відбувається навчання першої безлічі мереж. На другому відбувається тестова вибірка і власне прогноз на третій етап, який полягає в утворенні регресійної моделі у зв'язки прогнозів відібраних мереж з фактичними даними. Четвертий етап формує фактичні дані для порівняння з третім. У разі виявлення певних розбіжностей відбувається донавчання мереж, з першого ж по третій етапи відбувається початкове або базове навчання. Так само четвертим етапом формується прогноз на п'ятий етап. І так далі, процес триває до моменту повного навчання мережі.

Для нейромережевого прогнозування ринку нерухомості у кризових умовах необхідно створити зведену таблицю з даними про об'єкти нерухомості (рис. 2).

Місто	Загальна кількість	Квартира																	
		Розмір																	
		1 к				2 к				3 к				4 к				до 50	
Панель	Цегла	Моноліт	Панель	Цегла	Моноліт	Панель	Цегла	Моноліт	Панель	Цегла	Моноліт	Панель	Цегла	Моноліт	Панель	Цегла	Моноліт		
Київ	16678	991	1396	1098	876	991	131	1396	798	276	991	1396	254	1396	1122	1396	774		
Деснянський	3456	23	345	46	56	23	45	345	567	78	23	345	23	345	456	345	46		
Оболонь	1741	234	56	4	456	234	4	56	78	9	234	56	78	56	6	56	68		
Подол	1894	45	47	456	4	45	7	47	48	90	45	47	48	47	456	47	68		
Шулявка	3113	455	234	56	34	455	48	234	8	78	455	234	8	234	56	234	56		
Дорогожичі	1260	45	12	456	46	45	7	12	7	6	45	12	7	12	68	12	456		
Боршаговка	4613	55	657	35	234	55	8	657	78	3	55	657	78	657	35	657	35		
Дарниця	901	134	45	45	46	134	12	45	12	12	134	45	12	45	45	45	45		
Запоріжжя	17613	1510	1738	303	1396	538	753	991	1396	1396	798	1396	1274	580	1181	991	274		
Космічний	3807	351	579	23	345	346	46	23	345	345	567	345	23	345	12	23	43		
Ленінський	1933	678	89	45	56	4	54	234	56	56	78	56	34	56	1121	234	78		
Комунарський	1947	35	456	67	47	5	4	45	47	47	47	47	47	47	456	45	48		
Осіпенківський	3138	67	34	8	234	57	534	455	234	234	8	234	456	8	56	455	8		
Бабурка	1833	34	532	67	12	45	67	45	12	7	12	12	12	12	156	45	7		
Кічкас	4132	343	45	6	657	5	45	55	657	657	78	657	657	67	35	55	78		
Южний	823	2	3	87	45	76	3	134	45	45	12	45	45	45	45	134	12		

Рис. 2. Дані про об'єкти нерухомості

Схематичний процес навчання нейромережі зображено на рис. 3.

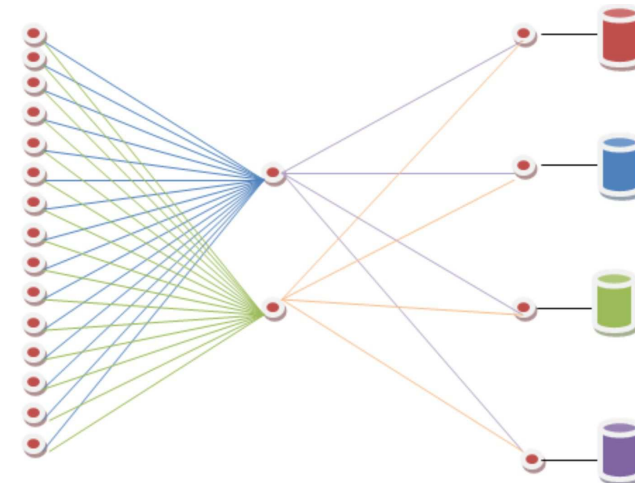


Рис. 3. Схематичний процес навчання нейромережі (розробка автора на основі джерел [1, 6])

На підставі даних про об'єкти нерухомості відбувається навчання нейромережі, при заданні параметрів навчання на вхід мережі подаються всі значення параметрів комірок, що входять у перший шар. А на виході отримаємо такий цільовий параметр, значення якого буде відомо тільки надалі. Таким чином, маємо модель навчання, в якій число входів буде дорівнювати кількості комірок, які увійшли в один шар, помноженому на кількість параметрів, а число виходів дорівнює кількості комірок в одному шарі. Отже, під час оброблення даних про об'єкти нерухомості отримаємо чотири нейронні мережі, навчання яких пройде на однаковому наборі початкових даних з рис. 2, проте спосіб введення даних буде значно відрізнятися згідно з параметрами. Під час проведення навчання на виходах мереж деякі дані будуть збігатися, це результат збігу вторинних параметрів.

Важливо зазначити, що цей метод прогнозування ринку нерухомості у кризових умовах передбачає прогнозування, за якого нейронна мережа є проміжним етапом аналізу. У представленій моделі дані на виході нейронних мереж – це сукупність прогнозів по кожній комірці.

Тобто для точних цифрових даних узагальненого характеру є потреба застосування методів математичної статистики, заснованих на сукупності певних правил, а саме регресійного аналізу

$$f(X^*) = E(y | X - X^*), \quad (1)$$

де: Ey – математичне очікування випадкової величини y , а $E(y | X^*)$ – умовне математичне очікування випадкової величини y , обчислене за умови, що значення пояснюючих змінних X зафіксовані на рівні X^* .

Загальний результат нейромережевого прогнозування ринку нерухомості у кризових умовах представлено у табл. 1.

Табл. 1. Нейромережеве прогнозування ринку нерухомості у кризових умовах

1 мережа (за площею)	2 мережа (за матеріалами)	3 мережа (за місцем розташування)	4 мережа (за конструкцією)
34,32	31,30	45,76	31,44
43,56	41,16	39,26	48,16
31,65	34,65	31,45	34,15
42,4	46,41	32,41	42,41
42,8	33,8	40,81	38,8
39,87	40,17	39,87	34,16
40,43	42,43	40,43	29,16
35,77	39,17	35,77	31,45
38,64	32,14	38,64	32,41
38,77	34,77	38,77	40,81

Виходячи з отриманих даних, із застосуванням методів математичної статистики, отримуємо результат прогнозування у числовому еквіваленті (табл. 2). Загальний прогноз з нейромережевого прогнозування ринку нерухомості у кризових умовах здійснюємо шляхом множення значення детермінації на кожне індивідуальне значення, отримане на основі результату нейромережевого моделювання з табл. 1. Отримуємо результат (табл. 3).

Висновки. На сьогодні представлена модель для нейромережевого прогнозування ринку нерухомості у кризових умовах є основою низки напрямів дослідження ринку нерухомості. На підставі цієї моделі можливий аналіз кон'юнктури попиту та пропозиції, оборотів і ємності ринку нерухомості, ліквідності об'єктів і т. ін.

Табл. 2. Результат прогнозування із застосуванням методів математичної статистики

	1 мережа (за площею)	2 мережа (за матеріалами)	3 мережа (за місцем розташування)	4 мережа (за конструкцією)
Кореляція	0,876	0,877	0,881	0,891
Похибка	0,123	0,112	0,221	0,151
Детермінація	0,342	0,541	0,142	0,211
Вплив змін	0,032	0,052	0,022	0,026

Табл. 3. Загальний прогноз з нейромережевого прогнозування ринку нерухомості у кризових умовах

1 мережа (за площею)	2 мережа (за матеріалами)	3 мережа (за місцем розташування)	4 мережа (за конструкцією)
11,73744	16,9333	6,49792	6,63384
14,89752	22,26756	5,57492	10,16176
10,8243	18,74565	4,4659	7,20565
14,5008	25,10781	4,60222	8,94851
14,6376	18,2858	5,79502	8,1868
13,63554	21,73197	5,66154	7,20776
13,82706	22,95463	5,74106	6,15276
12,23334	21,19097	5,07934	6,63595
13,21488	17,38774	5,48688	6,83851
13,25934	18,81057	5,50534	8,61091

Література

1. Бэстенс Д.-Э. Нейронные сети и финансовые рынки: принятие решений у торговых операциях / Д.-Э. Бэстенс, ван ден В.-М. Берг, Д. Вуд. – М.: Изд-во ТВП, 1997. – 324 с.
2. Rummelhart D.E. Learning representation by back – propagation error / D.E. Rummelhart, G.E. Hinton, R.G. Williams // Nature. – 1986. – Vol. 323, № 6088. – Pp. 533-536.
3. Berrens R.B. What price nondisclosure? The effects of nondisclosure of real estate sales prices / R.B. Berrens, M. McKee // Social Science Quarterly. – June 2004. – Vol. 85. – No. 2. – Pp. 509-520.
4. Ramsey J.B. Tests for specification errors in classical linear least-squares regression analysis / J.B. Ramsey // Journal of the Royal Statistical Society. Series B (Methodological). – 1969. – Vol. 31. – No. 2. – Pp. 350-371.
5. Specht D.F. A general regression neural network / D.F. Specht // IEEE Transactions on Neural Networks. – Vol. 2. – Issue 6. – Pp. 568-576.
6. Ежов А.А. Нейрокомпьютинг и его применения в экономике и бизнесе / А.А. Ежов, С.А. Шумский. – М.: Изд-во МИФИ, 1998. – 224 с.
7. Бышев В.А. Оптимальное комбинирование прогнозов различных моделей массовой оценки стоимостных показателей объектов недвижимости / В.А. Бышев, А.И. Богомолов, В.И. Костюнин // Актуальные проблемы математического моделирования у финансово-экономической области: сб. науч. стат. / под ред. д-р техн. наук, проф. В.А. Бышева. – М.: Изд-во "Финакадемия". – 2008. – Вып. 7. – С. 23-37.
8. Стерник Г.М. Методические рекомендации по анализу рынка недвижимости / Г.М. Стерник. – М.: Изд-во РГР, 1999. – 62 с.
9. Морозова Ж.А. Оценка недвижимости / Ж.А. Морозова. – М.: Изд-во "Финансы и статистика", 2007. – 496 с.
10. Wittkemper H. Portfolio optimization with a neural network implementation of the coherent market hypothesis / H. Wittkemper, M. Steiner // European Journal of Operational Research. – 1997. – Vol. 100. – Pp. 27-40.

Тазетдинов В.А. Нейросетевое прогнозирование рынка недвижимости в кризисных условиях

Осуществлено прогнозирование рынка недвижимости в кризисных условиях с помощью нейросети. Для нейросетевого прогнозирования рынка недвижимости в кризисных условиях создана сводная таблица с данными об объектах недвижимости. Приведен схематический процесс обучения нейросети. На основании данных об объектах недвижимости проведено обучение нейросети. Применены методы математической статистики, основанные на совокупности определенных правил для точных цифровых данных обобщенного характера. Общий прогноз по нейросетевому прогнозированию рынка недвижимости в кризисных условиях найден путем умножения значения, детерминации на каждое индивидуальное значение полученное на основе результата нейросетевого моделирования.

Ключевые слова: нейросеть, обучение, рынок недвижимости, прогнозирование, нейрон, ячейка.

Tazetdinov V.A. Neural Network Forecast of Real Estate Market in Crisis Conditions

Forecasting for real estate market in the crisis conditions using neural networks is carried. A pivot table with data on real property is compiled in order to produce the neural network forecasting of the real estate market in the crisis conditions. The schematic process of neural network learning is provided. Training neural network is conducted on the basis of real estate data. Statistical methods based on a set of specific rules for accurate digital data for general information only are applied. The overall outlook for the neural network forecasting of real estate market in crisis conditions is obtained by multiplying the value of determination for each individual value obtained based on the result of neural network modelling.

Key words: neural network, training, real estate market, forecasting, neuron cell.

УДК 519.765 *Асист. І.Ю. Хомицька; проф. В.М. Теслюк, д-р техн. наук – НУ "Львівська політехніка"*

СТАТИСТИЧНИЙ АНАЛІЗ АНГЛІЙСЬКИХ ПОЕТИЧНИХ ТЕКСТІВ

Проаналізовано статистичні структури текстів поезії Дж. Г. Байрона і Т. Мура. Доведено, що тексти містять спільні елементи, зумовлені приналежністю до одного історичного періоду і літературного напрямку. Зроблено спробу визначити статистичним методом дію фактора манери авторського викладу в зіставлених текстах на фонологічному рівні. Ступінь дії зазначеного фактора встановлено за кількістю груп приголосних фонем, за якими визначено істотні відмінності між зіставленими текстами. На основі отриманих результатів побудовано модель, яка репрезентує дію авторського фактора в межах підстилю поезії художнього стилю англійської мови.

Ключові слова: середня частота груп приголосних фонем, нормальний розподіл, критерій Ст'юдента, статистична структура.

Вступ. Застосування методів математичної статистики в гуманітарних галузях науки, зокрема в лінгвістиці, не втрачає своєї актуальності. Позаяк багато лінгвістичних проблем залишаються невирішеними, структурна та математична лінгвістики пропонують нові підходи до вивчення лінгвістичних об'єктів і явищ та успішно дають відповідь на запитання, які неможливо вирішити на рівні філологічної компетенції та інтуїції. На основі отриманих числових даних будують лінгвістичні моделі, які наочно розкривають суть досліджуваних явищ.

Це дослідження ґрунтується на наукових здобутках у напрямі використання математичної статистики для визначення ступеня дії фактору авторської манери викладу та моделювання статистичних структур поетичних текстів. Так, Г. Альтман, К.Б. Бектаєв, Г.Я. Мартиненко, Р.Г. Піотровський, С.В. Чебанов провели низку досліджень, в яких здійснено підрахунок і вимірювання стилістичних явищ з метою упорядкування текстів і їх частин [1, 2, 4]. У наших попередніх статтях досліджено статистичні параметри текстів з драматургії Б. Шоу [11, 12]. Розробку лінгвістичних моделей представлено у наукових розвідках І.І. Ревзіна та В.І. Перебийніс. Так, І.І. Ревзін на основі класифікації звуків певної мови в певний історичний період розробив парадигматичну звукову модель [7, 8]. Матеріалом дослідження є тексти поем Дж. Г. Байрона "Корсар",

"Шильйонський в'язень", "Гяур", "Мазепа", "Абідоська наречена", "Лара", і поем Т. Мура "Лалла Рук" та "Ірландські мелодії".

Мета роботи – вивчити взаємодію фактора підстилю поезії епохи романтизму та фактора манери авторського викладу за встановленням співвідношення кількості груп приголосних фонем, за якими, методами математичної статистики, визначити істотні відмінності за середніми частотами груп приголосних фонем.

У плані нашого дослідження середні значення частот досліджуваних груп фонем, як критерій диференціації, повинні дати таку інформацію: якою мірою ці величини, одержані з вибірки обмеженого обсягу, характеризують тексти поезії різних авторів та підстиль поезії романтизму загалом для побудови моделі визначення статистичних параметрів взаємодії зіставлених текстів з поезії Дж. Г. Байрона і Т. Мура.

Щоб відповісти на поставлене запитання, вводиться рівень значущості – імовірність появи певного значення різниці середніх значень частот порівнюваних систем, якою можна нехтувати. Якщо внаслідок досліджень одержуємо величину, імовірність появи якої рівна або менша від рівня значущості, то порівнювані системи відрізняються істотно. У протилежному випадку різниця між цими системами неістотна (у плані вибраного критерію) і зумовлена звичайним статистичним розкидом. З точки зору теорії імовірності, розподіл значень досліджуваної величини дає практично всю інформацію про цю величину.

Можна вважати апіорі, що ця функція – нормальний розподіл Лапласа-Гауса. Доведемо це твердження. Висуваємо нульову гіпотезу H_0 – одержані емпіричні дані відповідають нормальному розподілу частоти кожної з досліджуваних груп фонем. Для перевірки цієї гіпотези необхідна деяка статистична характеристика, яка оцінює розбіжність емпіричного і теоретичного розподілів. Як критерій значущості використовуємо χ^2 – розподіл Пірсона. Статистика χ^2 має корисну для нашого дослідження властивість: її закон розподілу не залежить від закону розподілу досліджуваної випадкової величини. Розподіл Пірсона задає при заданому числі степенів вільності ν імовірність $\rho(\chi^2)$ одержати значення, яке є більше від деякого χ_0^2 . Для χ^2 -розподілу Пірсона побудовано таблицю, яка дає змогу при заданих числу степенів вільності ν і рівні значущості g визначити χ_0^2 .

При здійсненні конкретних розрахунків замість χ_0^2 звичайно пишуть $\chi_g^2; \nu$. Якщо внаслідок проведеного дослідження одержуємо $\chi^2 < \chi_g^2; \nu$, гіпотеза приймається. У протилежному випадку вона відкидається. Іншими словами, це означає таке: якщо внаслідок розрахунків одержуємо значення величини χ^2 , імовірність появи якого $\rho(\chi^2)$ при числі степенів вільності ν є менша за рівень значущості g , емпіричний розподіл частоти досліджуваної групи фонем не є нормальним. Якщо ж $\rho(\chi^2) \geq g$, то маємо нормальний розподіл. Чим більше $\rho(\chi^2)$, тим краще емпіричний розподіл наближається до нормального.

Використовуємо 5 %-й рівень значущості, тобто $g = 0,05$. Результати нашого дослідження представлені у поданій нижче таблиці для груп носових фонем. З таблиці видно, що для цієї досліджуваної групи фонем задовольняється 5 %-й рівень значущості, тобто $\chi^2 < \chi_{0,05,4}^2$.