

УДК 336.774.3(243.2)

ЗАСТОСУВАННЯ ДИСПЕРСІЙНОГО АНАЛІЗУ В КРЕДИТУВАННІ

Дар'я Семенова , Наталія Смирнова
Національний авіаційний університет, Київ

Науковий керівник – Роксолана Ткаченко, асистент

Ключові слова: однофакторний дисперсійний аналіз, критерій Фішера-Снедекора, загальна і факторна сума квадратів відхилень.

Початково дисперсійний аналіз був призначений для статистичного аналізу сільськогосподарських дослідів. Зараз його використовують також у економічних, технічних та соціальних експериментах. Основна ідея цього аналізу полягає в тому, що загальну дисперсію ознаки, яка досліджується, розбивають на окремі складові, які зумовлені впливом конкретних факторів і порівнюється факторна дисперсія (породжується впливом фактора) і залишкова (зумовлена випадковими причинами).

В нашій роботі досліджується чи впливає квартал на обсяг видачі кредитів на основі застосування дисперсійного аналізу.

Таблиця 1. Сумарні дані щодо наданих кредитів комерційними банками України з 2019 по 2023 роки.

Фактори j	F_1 (1 квартал)	F_2 (2 квартал)	F_3 (3 квартал)	F_4 (4 квартал)	
2019	3305618	3285400	3237296	3189167	
2020	3104378	3243142	3162099	3004416	
2021	2897582	2941652	3023683	3154052	
2022	3232392	3215837	3286336	3250885	
2023	3078824	2977150	2968323	3027790	
\bar{x}_j гр. \bar{x}	3123758,8	3132636,2	3135547,4	3125262	3129301,1

Допишемо до таблиці додатковий рядок із середнім значенням для кожного кварталу, що визначає рівень фактору (F_1, F_2, F_3, F_4). Знайдемо загальне середнє: $\bar{x} = \frac{3123758,8 + 3132636,2 + 3135547,4 + 3125262}{4} = 3129301,1$. Віднімаємо від кожного спостережуваного значення x_{ij} , загальне середнє \bar{x} і перейдемо до зменшених величин $y_{ij} = x_{ij} - 3129301,1$, наприклад: $y_{11} = x_{11} - 3129301,1 = 3305618 - 3129301,1 = 176316,9$

Таблиця 2. Розрахункова таблиця

	y_{i1}	y_{i2}	y_{i3}	y_{i4}
2019	176316,9	156098,9	107994,9	59865,9
2020	-24923,1	113840,9	32797,9	-124885,1
2021	-231719,1	-187649,1	-105618,1	24750,9
2022	103090,9	86535,9	157034,9	121583,9
2023	-50477,1	-152151,1	-160978,1	-101511,1
$S_j = \sum y_{ij}^2$	98578222731	220177221045	74467692216	44879969398
$T_j = \sum y_j$	-27711,5	16675,5	31231,5	-20195,5
T_j^2	767927232,3	278072300,2	975406592,2	407858220,3

Отже, $\sum T_j = -0,0000000019$, $\sum T_j^2 = 2429264345$, $\sum S_j = 321103105389,8$.

Знайдемо загальну, факторну і залишкову суми квадратів відхилень, враховуючи, що число рівнів фактора $p = 4$, число випробувань на кожному рівні $q = 5$.

$$S_{\text{заг}} = \sum_{j=1}^p S_j - \frac{[\sum_{j=1}^p T_j]^2}{p \cdot q} = 321103105389,8; \quad S_{\text{факт}} = \frac{\sum_{j=1}^p T_j^2}{q} - \frac{[\sum_{j=1}^p T_j]^2}{p \cdot q} = 485852869;$$

$$S_{\text{зал}} = S_{\text{заг}} - S_{\text{факт}} = 320617252520,8.$$

Знайдемо факторну дисперсію, поділивши $S_{\text{факт}}$ на число ступенів свободи $[p - 1 = 3]$:

$$S_{\text{факт}}^2 = \frac{S_{\text{факт}}}{p-1} = 161950956,33; \text{ залишкову дисперсію, поділивши } S_{\text{зал}} \text{ на число ступенів свободи}$$

$$[p(q - 1) = 4 \cdot (5 - 1) = 16]: S_{\text{зал}}^2 = \frac{S_{\text{зал}}}{p \cdot (q-1)} = 20038578282,55 \text{ Порівняємо факторну і залишкову}$$

дисперсії (критерій Фішера-Снедекора): $K_{\text{емп}} = \frac{S_{\text{факт}}^2}{S_{\text{зал}}^2} = 123,73$. Враховуючи $k_1=16$, $k_2=3$ і рівень значущості $\alpha=0,01$, за таблицею розподілу Фішера знайдемо критичну точку: $k_{\text{кр}}(0,01; 16; 3)=27,1$.

Висновок

Оскільки $K_{\text{емп}} > k_{\text{кр}}$ нульову гіпотезу про рівність групових середніх не приймаємо. Іншими словами, групові середні суттєво різняться, тобто номер кварталу суттєво впливає на обсяг видачі кредитів.

Список використаних джерел:

1. Показники діяльності банків. URL: <https://index.minfin.com.ua/ua/banks/stat/2023-12/> (дата звернення 17.03.2024)
2. Ластівка І. О. Теорія ймовірностей та математична статистика: практикум. Київ: «НАУ-друк», 2009. 236 с.