**Трифакторний варіансний аналіз**

Нехай дано вибірку  незалежних спостережень над деякою одновимірною нормально розподіленою мінливою величиною, елементи якої можна згрупувати за рівнями трьох виділених факторів *A, B* та *C*: на  груп за ознакою *А*, на  груп за ознакою *В* і на груп за ознакою *С*. Отримаємо  класифікаційних груп. Припустимо, що в кожній групі є тільки одне спостереження. Позначимо через  – спостереження. в *і*-тій групі за ознакою *А*, в 𝑗-тій групі за ознакою *B* та в -тій групі за ознакою *C*. Всі  спостережень можна розмістити в  таблицях вигляду двофакторного варіансного аналізу . У кожній з – таблиць третій індекс  сталий, . Перший індекс – індекс довготи, другий – широти, третій – глибини . Введемо середні відповідних класифікаційних підгруп:

;  ;  ;  ;  ;  ;  (∗)

де наприклад:

, , .

Спостереження  та середні (∗) пов̕’язані алгебраїчною тотожністю

(∗∗)

Тотожність (∗∗) випливає з тотожності

  
і того, що сума відхилень вибіркових даних від їх середнього арифметичного у кожній відповідній групі дорівнює нулю. Співвідношення (∗∗) вказує на те, що повна девіація (ліва частина рівності (\*\*)) розкладається на сім девіацій (справа рівності (\*\*)): перші три характеризують відповідно мінливість між групами ознак *А*, *В,* *С*, наступні три оцінюють взаємодію добутків *АВ*, *АС*, *ВС* (інтеракції) відповідно, остання виражає залишкову мінливість, яку можна вважати взаємодією другого порядку, *АВС*. Тотальну мінливість назвемо мінливістю нульового порядку,; мінливість між групами – мінливістю першого порядку ; мінливість взаємодії – мінливістю другого порядку ; залишкова мінливість – мінливістю третього порядку . У тотожності (∗∗) сума зліва має  ступенів вільності. Для семи сум справа маємо відповідно таке число ступенів вільності:

  

;

;

;

.

Число ступенів вільності девіацій, що виступають у тотожності (∗∗) утворюють наступну тотожність:   
Зліва даної тотожності маємо число ступенів вільності нульового порядку; справа – три сукупності ступенів вільності першого, другого і третього порядків відповідно.

Якщо поділити тотожність (∗∗) на , отримаємо, що повна варіанса є лінійною комбінацією варіанс між групами, інтеракцій (взаємодії) та залишкової варіанси:



Тут наприклад,

,  


Аналогічно записують інші варіанси:  ,  ,  , . Перелічені варіанси можна використати при доведенні гіпотези однорідності за допомогою критерію Фішера, для чого порівняємо варіанси між групами або варіанси взаємодій із залишковою. Отже, при доведенні відповідних гіпотез розглядаємо відповідні статистики. Зокрема, для аналізу можливого впливу факторів *A, B,C* на генеральну сукупність використовуємо статистики:

(\*\*\*)

відповідно для аналізу сумісного впливу факторів *AB, AC, BC* використовуємо такі статистики

, , . (\*\*\*\*)

Якщо гіпотеза про однорідність даних в нормальній популяції – вірна, то статистики (∗∗∗) та (\*\*\*\*\*) мають розподіл Фішера відповідно до чисел ступеня вільності:

, , 

для виразу (\*\*\*) та

, ,

.

для виразу (\*\*\*\*) відповідно.

Обчислення при статистичному доведенні однорідності зручно розмістити в таблиці трифакторного варіансного аналізу при одному спостереженні в кожній групі:

|  |  |  |  |
| --- | --- | --- | --- |
| Мінливість | Девіація | *d.f.* | Варіанса |
| між групами *А* |  |  |  |
| між групами *В* |  |  |  |
| між групами *С* |  |  |  |
| взаємодія *АВ* |  |  |  |
| взаємодія *АС* |  |  |  |
| взаємодія *ВС* |  |  |  |
| залишкове |  |  |  |
| Повна |  |  |  |

**Приклад**. Перевірити дію 2 добрив на 5 сортів якогось збіжжя. Проводимо експеримент у  дослідних центрах. Дістанемо 40 даних про врожаї, які класифікуємо в таблиці  (чотири центри, п’ять сортів збіжжя і два типи добрив). Нехай врожаї, виміряні від якоїсь зручної вихідної вартості та виражені в означених одиницях, будуть такі, як у таблиці 1, де  і - два типи добрив.

Таблиця 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Дослідницькі центри | Сорти збіжжя | | | | | |
| 1 | 2 | 3 | 4 | 5 | Разом |
|  |  |  |  |  |  |
| 1  2  3  4 | -6 -4  -2 -1  3 2  3 6 | -4 -2  -5 -1  -2 3  3 2 | -10 -7  -3 -4  -4 0  6 3 | -3 -5  -4 -1  4 1  -1 5 | 1 0  -2 1  3 3  6 8 | -22 -18  -16 -6  4 9  17 24 |
| Разом | -2 3 | -8 2 | -11 -8 | -4 0 | 8 12 | -17 3 |

З умови задачі випливає, що дослідницькі центри – це фактор *A,* сорти збіжжя – фактор *B,* добрива – фактор *C,* генеральна сукупність – врожайність збіжжя.

Тому . Одно- і двоіндексні середні розмістимо в таблиці 2.

Таблиця 2

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Значення змінного індексу |  |  |  |  |  |  |  |  |  |  |  |
| 1 | -4,0 | 0,125 | -0,85 | -5,0 | -1,5 | 2,5 | 4,5 | -4,4 | -3,6 | -0,50 | 0,75 |
| 2 | -2,2 | -0,75 | 0,45 | -3,0 | -3,0 | 0,5 | 2,5 | -3,2 | -1,2 | -2,00 | 0,5 |
| 3 | 1,3 | -2,375 |  | -8,5 | -3,5 | -2,0 | 4,5 | 0,8 | 1,8 | -2,75 | -2,00 |
| 4 | 4,1 | -0,5 2,5 |  | -4,0 0,5 | -2,5 -0,5 | 2,5 3,0 | 2,0 7,0 | 3,4 | 4,8 | -1,00 2,00 | 0,00 3,00 |

Девіації, ступені вільності та варіанси для різних мінливостей розмістимо в таблиці 3.

Таблиця 3

|  |  |  |  |
| --- | --- | --- | --- |
| Мінливість | Девіація |  | Варіанса |
| Між центрами А  Між сортами В  Між добривами Т Інтеракція(взаємодія) АВ Інтеракція(взаємодія) АТ Інтеракція(взаємодія) ВТ Залишкове | 381,8  100,15  16,90  62,45  2,10  3,85  61,15 | 3  4  1  12  3  4  12 | 130,60  25,04  16,9  5,20  0,70  0,96  5,10 |
| Повна | 638,40 | 39 | - |

За допомогою варіансного аналізу перевірити гіпотезу про можливий вплив вказаних факторів на врожайність збіжжя. На основі даних таблиці 3 обчислимо емпіричні значення статистик Фішера для ітеракцій із залишковою варіансою при рівні значущості :

; ;

.

При рівні значущості  та кількостях ступенів вільності (12,12), 12,3), (12,4), відповідно, наприклад з таблиці (додаток 8) маємо:

2,69; 8,74; 5,91.

Усі три емпіричні значення статистик інтеракцій менші відповідних критичних значень цих статистик. Отже, всі ці фактори діють незалежно. Дослідимо чи впливають фактори *A, B, C* нагенеральну сукупність. Для цього вобчислимо емпіричні значення статистики Фішера для цих факторів:

; ; .

При кількості ступенів вільності (3,12), (4,12) , (1,12), відповідно, з таблиць (додаток 8) маємо:

якщо , то

3,49; 3,26; 4,75;

якщо , то

5,95; 5,41; 9,43.

Отже різниці між центрами істотні; між сортами – неістотні при рівні значущості 1%, і істотні при рівні значущості 5%, між добривами неістотні. Отже, мінливість врожаїв залежить від мінливості між центрами (і можливо між сортами), чого не можна сказати без деяких досліджень про залежність її від мінливості між добривами.