Оценивание недостающих данных

Подготовила: Романова Ю.В.
Группа: 154-341
Руководитель: Петухов С.Л.

Содержание:

1. Рандомизированное полноблочное планирование
2. Статистический анализ
3. Оценивание недостающих данных

1. Рандомизированное полноблочное планирование

	Образец			
	1	**2**	**3**	**4**
Тип острия	4	2	3	1
	3	1	2	2
	1	4	1	4
	2	3	4	3

Таблица 1. Рандомизированный полноблочный план эксперимента по проверке твердости

2. Статистический анализ

Тип острия	Образец (блок)			
	1	2	3	4
1	9,3	9,4	9,6	10,0
2	9,4	9,3	9,8	9,9
3	9,2	9,4	9,5	9,7
4	9,7	9,6	10,0	10,2

Таблица 2 – исходные данные

Тип острия	Образец (блок)				$y_{i.}$
	1	2	3	4	
1	−2	−1	1	5	3
2	−1	−2	3	4	4
3	−3	−1	0	2	−2
4	2	1	5	7	15
$y_{.j}$	−4	−3	9	18	$20 = y_{..}$

Таблица 3 – преобразованные исходные данные с дополнительными вычислениями

Найдем суммы квадратов:

$$SS_{общ} = \sum_{i=1}^{4}\sum_{j=1}^{4} y_{ij}^2 - \frac{1}{N} y_{..}^2 = 154,00 - \frac{1}{16} 20^2 = 129,00;$$

$$SS_{обр} = \frac{1}{b}\sum_{i=1}^{4} y_{i.}^2 - \frac{1}{N} y_{..}^2 = \frac{1}{4}[3^2 + 4^2 + (-2)^2 + 15^2] - \frac{20^2}{16} = 38,50;$$

$$SS_{бл} = \frac{1}{a}\sum_{j=1}^{4} y_{.j}^2 - \frac{1}{N} y_{..}^2 = \frac{1}{4}[(-4)^2 + (-3)^2 + 9^2 + 18^2] - \frac{20^2}{16} = 82,50;$$

$$SS_{ош} = SS_{общ} - SS_{обр} - SS_{бл} = 129,00 - 38,50 - 82,50 = 8,00.$$

Источник изменчивости	Сумма квадратов	Степени свободы	Средний квадрат	F_0
Обработки (тип острия)	38,50	3	12,83	14,42
Блоки (образцы)	82,50	3	27,50	
Ошибка	8,00	9	0,89	
Сумма	129,00	15		

Таблица 4 – результаты ДА

3. Оценивание недостающих данных

Тип острия	Образец (блок)			
	1	2	3	4
1	—2	—1	1	5
2	—1	—2	x	4
3	—3	—1	0	2
4	2	1	5	7

Таблица 6 – Исходные данные с неизвестным наблюдением

В общем случае пусть $y_{..}'$— общая сумма наблюдений при одном недостающем данном, $y_{i.}'$— сумма для обработки с одним недостающим данным и $y_{.j}'$— сумма для блока с одним недостающим данным.

Предположим, что мы хотим выбрать оценку недостающего данного так, чтобы x давало наименьший вклад в сумму квадратов ошибки. Поскольку

$$SS_{\text{ош}} = \sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2,$$

то это эквивалентно выбору значения *x*, минимизирующего

$$SS_{\text{ош}} = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - \frac{1}{b} \sum_{i=1}^{a} \left(\sum_{j=1}^{b} y_{ij} \right)^2 - \frac{1}{a} \sum_{j=1}^{b} \left(\sum_{i=1}^{a} y_{ij} \right)^2 + \frac{1}{ab} \left(\sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij} \right)^2$$

или

$$SS_{\text{ош}} = x^2 - \frac{1}{b} (y'_{i.} + x)^2 - \frac{1}{a} (y'_{.j} + x)^2 + \frac{1}{ab} y'_{..} + x)^2 + R,$$

где R включает в себя все слагаемые, не зависящие от *x*.
Из условия *d*SSош/*d*x = 0 получаем выражение

$$x = \frac{a y'_{i.} + b y'_{.j} - y'_{..}}{(a-1)(b-1)},$$

которое является оценкой недостающего данного.
Для данных предыдущей таблицы находим, что $y_{2.}'=1$, $y_{.3}'=6$ и $y_{..}'=17$. Следовательно, в соответствии с соотношением

$$x = y_{23} = \frac{4 \cdot 1 + 4 \cdot 6 - 17}{3 \cdot 3} = 1{,}22.$$

Теперь можно провести обычный ДА, используя y_{23} = 1,22 и уменьшив число степеней свободы для ошибки на единицу. Результаты такого анализа приведены в таблице.

Источник изменчивости	Сумма квадратов	Степени свободы	Средний квадрат	F_0
Тип острия	39,98	3	13,33	17,12 *
Образцы (блоки)	79,53	3	26,51	
Ошибка	6,22	8	0,78	
Сумма	125,73	14		

* Значимо при 5 процентах.