ЛЕКЦИЯ 23
СТАТИСТИЧЕСКАЯ ПРОВЕРКА ГИПОТЕЗ

1. По критерию Фишера проверяется адекватность математической модели. Критерий Фишера показывает, во сколько раз уравнение регрессии предсказывает результаты опыта лучше, чем среднее по опытам.

Если число опытов в каждой точке равно одному, а общее число точек равно (n), то

$$F = \frac{S_a^2}{S_y^2} \leq F_{кр(\varphi_1;\varphi_2;p)}$$

$$\varphi_2 = n - 1$$

где φ_2 - число степеней свободы дисперсии эксперимента,

$$\varphi_1 = n - 2$$

φ_1 - число степеней свободы дисперсии адекватности (остаточной дисперсии)

$F_{кр(\varphi_1;\varphi_2;p)}$ - табличное (критическое) значение критерия Фишера при уровне значимости «*p*».

Если число опытов в каждой точке разное и равно k_i, число точек равно (n), а общее число экспериментов

$$N = \sum_i^n k_i$$

всех опытах , то расчетное значение критерия Фишера и условие адекватности определяется из формулы

$$F = \frac{S_a^2/\varphi_1}{S_y^2/\varphi_2} \leq F_{кр(\varphi_1;\varphi_2;p)}$$

Где

$$S_y^2 = \sum_{i=1}^{n}\sum_{j=1}^{k}\frac{(y_{ij} - \bar{y}_i)^2}{k_j - 1}$$

- дисперси

$$S_a^2 = \sum_{i=1}^{n} k_j \frac{(\bar{y}_i - \check{y}_i)^2}{n - 2}$$

$$\varphi_2 = n - N$$

$$\varphi_1 = n - 2$$

Пример: в табл. 1 приведены значения исследуемого процесса. В каждой точке эксперимента (x_i) проводилось неравное количество опытов.

Таблица 1

Исходные данные для расчета исследуемого процесса

Номер	x_i	y_{i1}	y_{i2}	y_{i3}	y_{i4}	y_{i5}	k_i	$\bar{y}_i$	$(y_i - \bar{y}_i)^2{}_i$	$\check{y}_i$	$\sum(\bar{y}_i - \check{y}_j)^2 \cdot k$
1	-1	12	-	-	-	-	1	12	0	12	0
2	0	10,3	10,0	9,8	9,9	10,5	5	10,1	0,34	10	0,05
3	1	7,8	7,9	8,0	7,9	-	4	7,9	0,02	8	0,04
4	3	3,9	4,1	4,1	4,3	-	4	4,1	0,08	4	0,04
5	4	2,1	1,7				2	1,9	0,08	2	0,02
$\sum$							16	36	0,52	36	0,15

После обработки таблицы получено уравнение

$$\check{y} = 10 - 2x$$

$$S_a^2 = 0{,}15, \quad S_y^2 = \left(0 + \frac{0{,}34}{5-1} + \frac{0{,}02}{4-1} + \frac{0{,}08}{4-1} + \frac{0{,}08}{2-1}\right) = 0{,}198$$

$$F = \frac{S_a^2/\varphi_1}{S_y^2/\varphi_2} = \left(\frac{0{,}15/3}{0{,}198/11}\right) = 2{,}78 \le F_{\text{кр}(3;11;5\%)} = 3{,}59$$

при

$$\varphi_1 = n - 2 = 5 - 2 = 3 \qquad \varphi_2 = \sum k - n = 16 - 5 = 11$$

Так как F=2,78<Fкр=3,59, то пол[illegible]екватно описывает реальный процесс

2.Критерий Кохрена

По критерию Кохрена проверяется однородность дисперсий, т.е. возможность воспроизводства эксперимента в другой промежуток времени. Расчетное значение критерия Кохрена равно отношению максимальной дисперсии в одной точке к сумме дисперсий всего эксперимента. Если расчетное значение критерия при заданном уровне значимости должно быть меньше критического ($G_{табл}$), то дисперсии однородны. В противном случае опыты поставлены некорректно, т.е. имеется неслучайный фактор, влияющий на результат.

$$G = \frac{S_{imax}^2}{s_y^2} \le G_{кр(\varphi_1;\varphi_2;p)}$$

где $\varphi_1 = k - 1$ - число степеней свободы;

S_{imax}^2 число опытов в каждой точке

- max из дисперсий в одной из точек эксперимента

$\varphi_2 = n$ - число точек эксперимента.

Максимальная дисперсия в одной из точек (в данном эксперименте она во второй точке) находится из формулы:

$$S_{max} = \sum_{1}^{k_j} \frac{(y_{ij} - \bar{y}_i)^2}{k - 1}$$

Пример: Проверить однородность дисперсии эксперимента, результаты которого приведены в таблице 1.

$$G = \frac{0{,}34}{5 - 1}/0{,}198 = 0{,}43 > G_{табл}$$

$$\varphi_1 = k = 5 \qquad \varphi_2 = N = 5 \qquad G_{табл(5;5;5\%)} = 0{,}5$$

Вывод: дисперсии однородны.

Таким образом, полученная модель $\check{y} = 10 - 2x$ может быть использована для прогнозирования ещё не известных результатов процесса, который она описывает.

На рис. 1 приведены экспериментальные точки и уравнение регрессии $\breve{y} = 10 - 2x$

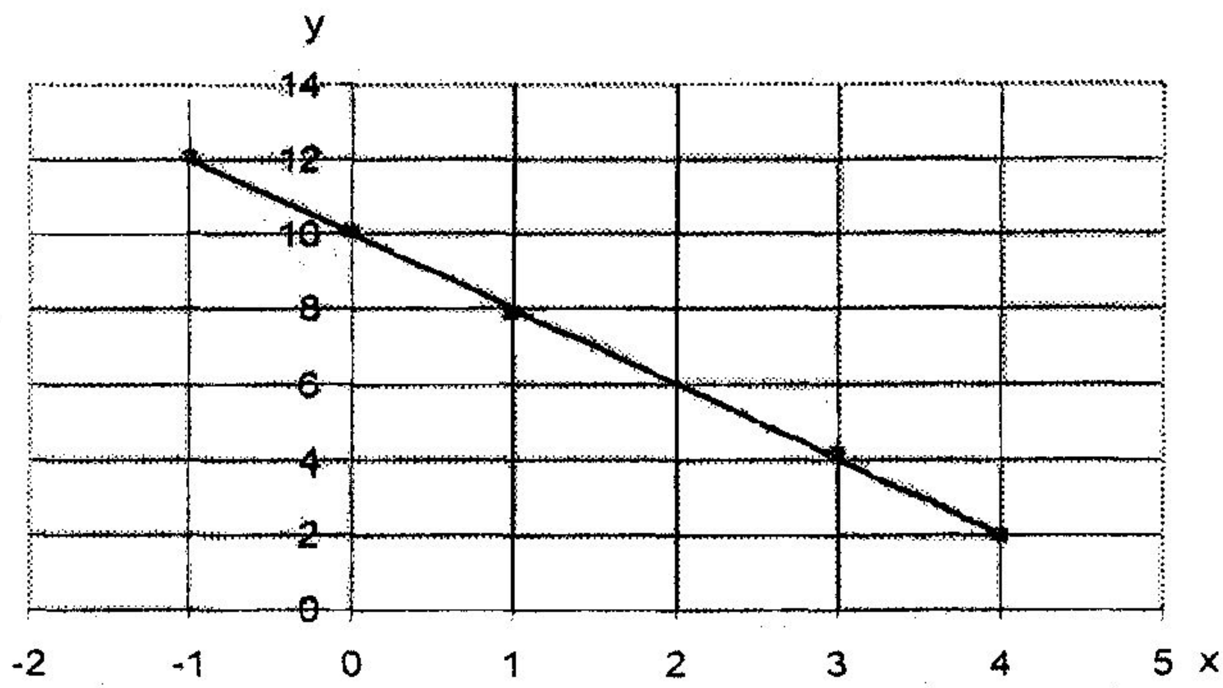

Рис. 1. Зависимость $\breve{y} = 10 - 2x$

Таблица 2

Значение критерия Кохрена при доверительной вероятности 95 %

N	f					
	1	2	3	16	36	∞
2	2	0,99	0,94	0,83	0,73	0,66
3	3	0,97	0,80	0,65	0,55	0,48
4	4	0,91	0,68	0,54	0,44	0,37
5	5	0,84	0,60	0,46	0,36	0,31
6	6	0,78	0,53	0,40	0,31	0,26
7	7	0,73	0,48	0,35	0,28	0,23
8	8	0,68	0,44	0,32	0,27	0,20
9	9	0,64	0,40	0,29	0,22	0,18
10	10	0,60	0,37	0,27	0,20	0,16
15	15	0,47	0,28	0,19	0,14	0,11
20	20	0,39	0,22	0,15	0,11	0,09
30	30	0,29	0,16	0,11	0,08	0,06
60	60	0,17	0,09	0,06	0,04	0,03
∞	0	0	0	0	0	0

Примечание, f - число степеней свободы сравниваемой дисперсии; N – число сравниваемых дисперсий.