

Пример 1.6. В Тегульдетском районе Томской области в 1999 и 2002 годах В. П. Перевозкин выполнил сбор материала и получил распределение сочетаний хромосомных инверсий комаров *Anopheles messeae* (данные в таблице).



Распределение кариотипов *Anopheles* в различные годы

год выборки	кариотипы				
	XL11 2R11	XL22 2R11	XL11 2R01	XL00 2R00	XL11 2R00
1999	38	26	16	62	69
2002	37	18	14	78	93

Вопрос: изменилось или нет распределение по сочетанию инверсий *Anopheles messeae* за три года? Иными словами – выборки 1999 и 2002 года статистически одинаковы или различны?

Используется следующий порядок вычислений.

На первом этапе определяются "теоретические" частоты по соотношению

$$T_{ij} = \frac{\sum_{q=1}^2 Z_{qj} \times \sum_{p=1}^M Z_{ip}}{\sum_{q=1}^2 \sum_{p=1}^M Z_{qp}},$$

или, если угодно, $T_{ij} = \frac{(\text{сумма значений } i\text{-той строки}) \times (\text{сумма значений } j\text{-того столбца})}{(\text{сумма всех значений})}$.

На втором этапе определяются критическое значение $\chi_{кр}^2$ и расчетное $\chi_{эмп}^2$ по формуле

$$\chi_{эмп}^2 = \sum_i \sum_j \frac{(Z_{ij} - T_{ij})^2}{T_{ij}}.$$

Сравниваются значения критерия и делается заключение.

На [рис. А](#) представлен скриншот решения данной задачи, содержащий следующие этапы вычислений.

1. В ячейки A5:F10 заносятся пояснительные надписи, в ячейку I5 заносится величина уровня значимости и в диапазон B6:F7 собственно данные по кариотипам.
2. В ячейку B11 вводится формула =СУММ (B\$6:B\$7) / СУММ (\$B\$6:\$F\$7) * СУММ(\$B6:\$F6) {=SUM (B\$6:B\$7)/ SUM (\$B\$6:\$F\$7)* SUM (\$B6:\$F6)}, определяющая **ожидаемое (опорное) значение** по первому кариотипу за 1999 год; для остальных – тиражируется автозаполнением на ячейки C11:F11, B12:F12.
3. В ячейку B16 вводится формула =(B6-B11)^2/B11 – значение χ^2 по первому кариотипу за 1999 год; для остальных – тиражируется автозаполнением на ячейки C16:F16, B17:F17.
4. В ячейку I17 вводится формула =СУММ(B16:F17) {=SUM(B16:F17)} подсчета суммы значений χ^2 для всех кариотипов, т.е. **определяется $\chi_{эмп}^2$** .
5. В ячейке B15 формулой =ХИ2ОБР(I5;СЧЁТ(В6:F6)-1) {=CHIINV(I5;СЧЁТ(В6:F6)-1)} определяется критическое значение $\chi_{кр}^2$.

	A	B	C	D	E	F	G	H	I	J	
1	Нулевая гипотеза: Распределение <i>Anopheles</i> по кариотипам										
2	в Тегульдете с 1999 по 2002 год не изменилось										
3											
4	фактические данные										
5	кариотип	XL11 2R11	XL22 2R11	XL11 2R01	XL00 2R00	XL11 2R00		$\alpha =$	0,05		
6	1999 год	38	26	16	62	69					
7	2002 год	37	18	14	78	93		$\chi^2_{кр} =$	9,49	←	
8							=ХИ2ОБР(И5;СЧЁТ(В6:F6)-1)				
9	ожидаемые "теоретические" значения										
10	кариотип	XL11 2R11	XL22 2R11	XL11 2R01	XL00 2R00	XL11 2R00					
11	1999 год	35,1	20,6	14,0	65,5	75,8					
12	2002 год	39,9	23,4	16,0	74,5	86,2					
13	=СУММ(В\$6:В\$7)/СУММ(\$В\$6:\$F\$7)*СУММ(\$B6:\$F6)										
14											
15	=(B6-B11)^2/B11										
16	χ^2	0,24	1,42	0,27	0,19	0,61		=СУММ(В16:F17)			
17		0,21	1,25	0,24	0,16	0,54		$\chi^2_{эмп} =$	5,14	←	
18											
19	Эмпирическое значение $\chi^2_{эмп} = 5,14 < 9,49 = \chi^2_{кр}$ критического значения, следовательно										
20	гипотеза о неизменности распределения <i>Anopheles</i> подтверждается										

Рис. А. Скриншот расчетного листа MS Excel к примеру 1.6.

6. Сравниваются расчетное и критическое значений χ^2 . Поскольку в данном случае $\chi^2_{эмп} < \chi^2_{кр}$ ($5,14 < 9,49$), то можно сделать вывод: гипотеза о неизменности распределения *Anopheles* подтверждается.