

Пример 1.2. (часть 1). Распределение "взгляда" лабораторной крысы*

Чтобы определить цветовые предпочтения у лабораторной крысы, перед ней в клетке были повешены 4 цветных круга (красный, синий, зеленый и желтый). К голове крысы были прикреплены сенсоры фиксации глаз. Исследователем за полчаса эксперимента были получены результаты, представленные в табл. А:

Таблица А – Распределение "взгляда" лабораторной крысы
между 4-мя цветными кругами

Цвет	Красный $n_1^{\text{экс}}$	Синий $n_2^{\text{экс}}$	Зеленый $n_3^{\text{экс}}$	Желтый $n_4^{\text{экс}}$	Всего "взглядов"
Количество "взглядов"	14	5	8	5	32

Полученные эмпирические частоты необходимо сопоставить с теоретическими. Если крыса не отдает предпочтения ни одному цвету, то данное распределение показателя направленности "взгляда" не будет отличаться от равномерного распределения: на всех цветовых кругах задержка взгляда будет примерно одинаковой частоты.

Нулевая гипотеза: распределение "взглядов" лабораторной крысы между цветами не отличается от равномерного распределения.

* Исходные данные примера взяты из учебно-методического пособия Д.Е. Гавриков. Введение в биологическую статистику / Иркутск: Изд-во Иркутск. гос. пед. ун-та, 2002. –72 с.

Теоретическая частота "взгляда" при равномерном распределении определяется просто. Если бы все "взгляды" крысы распределялись равномерно между 4-мя кругами, то, очевидно, каждый из них получил бы по $1/4$ всех ее "взглядов".

Теоретическая частота при сопоставлении эмпирического распределения с равномерным определяется формулой: $n_{\text{теор}} = n/M$, где $n = \sum n_i^{\text{эк}}$ – количество наблюдений; M – количество классов признака ($M = 4$ – рассматриваются четыре цветовых круга).

В данном случае признак – "взгляд" крысы, направленный на какой-либо из цветов; количество классов признака равно 4 – это направления "взгляда" по количеству цветовых кругов; количество наблюдений – 32. Таким образом, в рассматриваемом случае: $n_{\text{теор}} = 32/4 = 8$.

Далее с этой теоретической частотой сравниваются все эмпирические частоты. Заметим, что по табл. А отмечается явное предпочтение красного цвета.

Расчеты производятся в среде MS Excel по схеме [рис. А](#).

1. Заносятся необходимые поясняющие (текстовые) данные.
2. Заносятся исходные данные – эмпирические частоты (B5:B8) и уровни значимости (C11:C12).
3. В ячейку B9 заносится формула подсчета общего числа наблюдений =СУММ(B5:B8) {=SUM(B5:B8)}.
4. В ячейку C5 заносится формула, определяющая теоретическое значение частоты $\text{=B\$9/4}$. Механизмом автозаполнения по C5 вводятся значения C6:C8.

	A	B	C	D	E	F	G	H	I	J
1	Нулевая гипотеза: распределение "взглядов" лабораторной крысы									
2	между цветами не отличается от равномерного распределения									
3										
4		Π _i эксп	Π _i теор	χ^2		=B\$9/4				
5	красный	14	8	4,500		= (B5-C5)^2/C5				
6	синий	5	8	1,125			=COUNT(B5:B8)-1			
7	зеленый	8	8	0,000		df= 3	=СЧЁТ(B5:B8)-1			
8	желтый	5	8	1,125						
9	итого	32	32	6,750		=СУММ(D5:D8)	$\chi^2_{\text{эмп}}$			
10						=SUM(D5:D8)				
11	$\chi^2_{\text{кр}}$	при $\alpha = 0,05$	0,05	7,815		=ХИ2ОБР(C11;G\$7)				
12		при $\alpha = 0,01$	0,01	11,345		=СНПИНV(C11;G\$7)				
13										
14										
15		поскольку $\chi^2_{\text{эмп}} < \chi^2_{\text{кр}}$ то								
16										
17	Ответ: нулевая гипотеза принимается. Распределение "взгляда" лабораторной									
18	крысы между цветными кругами не отличается от равномерного распределения									

Рис. А. Скриншот расчетного листа MS Excel к примеру 1.2 (для части 1)

5. В ячейку D5 заносится формула расчета значение $\chi^2 = (B5-C5)^2/C5$. Механизмом автозаполнения по D5 вводятся значения D6:D8.
6. От ячейки B9 автозаполнением определяются C9 и D9.
7. В ячейке G7 формулой $=СЧЁТ(B5:B8)-1$ $\{=COUNT(B5:B8)-1\}$ определяем число степеней свободы df .
8. В ячейке D11 определяется значение $\chi_{кр}^2$ для первого (C11) уровня значимости $\alpha=0,05$ формулой $=ХИ2ОБР(C11;G\$7)$ $\{=CHIINV (C11;G\$7)\}$. Выполняется автозаполнение для ячейки D12.
9. Сравниваем $\chi_{эмп}^2$ со значениями $\chi_{кр}^2$. Поскольку $\chi_{эмп}^2$ меньше критических значений, расхождения между распределениями статистически недостоверны на обоих уровнях значимости.

Если $\chi_{эмп}^2$ равно критическому значению или превышает его, расхождения между распределениями статистически достоверны (см. [рис.В](#)).

По "оси значимости" ясно, что чем больше отклонения эмпирических частот от теоретических, тем больше будет величина χ^2 . Поэтому зона значимости располагается справа, а зона незначимости – слева. На основании этих данных получаем $\chi_{эмп}^2 < \chi_{кр}^2$.



Рис. В. Анализ значимости различий

Вывод: нулевая гипотеза принимается. Распределение "взгляда" лабораторной крысы между цветными кругами не отличается от равномерного распределения.

Пример 1.2. (часть 2). Допустим, что исследование на этом не остановили и получили следующее распределение предпочтений: красный – 15 "взглядов", синий – 6 "взглядов", зеленый – 9 "взглядов", желтый – 6 "взглядов".

Очевидно, что красный цвет остался предпочтительным. У исследователя есть следующий путь, чтобы доказать это статистически: суммировать количества "взглядов" в первом опыте по каждому цвету: количество "взглядов" на красный цвет в опыте 1 + количество "взглядов" на красный цвет в опыте 2, и сопоставить полученное распределение с равномерным. Поскольку количество наблюдений возросло, есть вероятность, что различия окажутся достоверными.

Создается новая таблица эмпирических частот (табл. В), в которой суммированы все "взгляды" лабораторной крысы.

Таблица В – Распределение взгляда лабораторной крысы
между 4-мя цветными кругами

Цвет	Красный	Синий	Зеленый	Желтый	Всего "взглядов"
Количество "взглядов"	29	11	17	11	68

Формулируется нулевая гипотеза: распределение проявлений цветовых предпочтений крысы не отличается от равномерного распределения. Все расчеты производятся аналогично предыдущему случаю, схема на [рис. С](#).

Таким образом, имеется:

$$n_{\text{теор}} = 68/4 = 17$$

$$df = 3$$

$$\chi^2_{\text{кр}} = \begin{cases} 7,815 (P \leq 0,05) \\ 11,345 (P \leq 0,01) \end{cases} \quad \begin{matrix} \chi^2_{\text{эмп}} = 12,706 \\ \chi^2_{\text{эмп}} > \chi^2_{\text{кр}} \end{matrix}$$

Вывод: нулевая гипотеза отклоняется. Распределение "взгляда" лабораторной крысы между цветными кругами отличается от равномерного распределения ($\alpha < 0,01$).

На этом примере можно убедиться, что увеличение числа наблюдений повышает достоверность результата, если, конечно, в новых наблюдениях воспроизводится прежняя тенденция различий.

	A	B	C	D	E	F	G	H	I
1	Нулевая гипотеза: распределение "взглядов" лабораторной крысы								
2	между цветами не отличается от равномерного распределения								
3									
4		Pi эксп	Pi теор	χ^2		=B\$9/4			
5	красный	29	17	8,471		= (B5-C5)^2/C5			
6	синий	11	17	2,118					
7	зеленый	17	17	0,000		df= 3			
8	желтый	11	17	2,118			=СЧЁТ(B5:B8)-1		
9	итого	68	68	12,706		=СУММ(D5:D8)	$\chi^2_{\text{эмп}}$		
10									
11	$\chi^2_{\text{кр}}$	при $\alpha = 0,05$	0,05	7,815		=ХИ2ОБР(C11;G\$7)			
12		при $\alpha = 0,01$	0,01	11,345					
13									
14	поскольку $\chi^2_{\text{эмп}} > \chi^2_{\text{кр}}$ то Ответ: нулевая гипотеза отклоняется.								
15	Распределение "взгляда" лабораторной крысы между цветными								
16	кругами отличается от равномерного распределения								

Рис. С. Скриншот расчетного листа MS Excel к примеру 1.2 (для части 2)