Лабораторная работа № 2

Статистическая проверка гипотезы о нормальном распределении признак X (или Y) генеральной совокупности (X, Y)

Цель работы: проверить по критерию Пирсона согласуется ли гипотеза о нормальном распределении одного из признаков генеральной совокупности при уровне значимости $\alpha = 0{,}05$. Если нет, то определить уровень значимости, при котором гипотезу о нормальном распределении можно принять. Найти интервальные оценки параметров a и σ с доверительной вероятностью $\gamma = 1 - \alpha$.

Порядок выполнения работы:

Выдвинуть гипотезу H_0 о нормальном распределении признака X (или Y).

1. Если частоты первых и последних интервалов малы (меньше пяти [1]), то объединить их с соседними интервалами. l' — новое число интервалов.
2. Вычислить по выборке значение статистики $\chi^2 = \sum_{i=1}^{l'} \frac{(n_{x_i} - n'_{x_i})^2}{n'_{x_i}}$. Случайная величина χ^2 имеет χ^2- распределение Пирсона.

n_{x_i} — наблюдаемая частота, соответствующая i – му интервалу;

n'_{x_i} — теоретическая частота, которую вычисляется по формуле $n_{x_i} = n \cdot p_{x_i}$, где p_{x_i} — теоретическая вероятность попадания случайной величины X в интервал $[\alpha_{i-1},\ \alpha_i]$

$$p_{x_i} = \Phi\left(\frac{\alpha_i - \bar{x}}{s_x}\right) - \Phi\left(\frac{\alpha_{i-1} - \bar{x}}{s_x}\right), \qquad \text{где } \Phi(u) \text{ — функция Лапласа.}$$

3. В таблице χ^2 – распределения Пирсона найти критическое значение статистики $\chi^2_{кр}(\alpha;\ k)$ по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = l' - r - 1$, где l' — число интервалов с учетом их объединения. Для нормального распределения r =2. Сравнить значение $\chi^2_{набл}$, вычисленное по выборке, с табличным критическим значением $\chi^2_{кр}$ и сделать вывод о возможности принятия гипотезы H_0 при уровне значимости $\alpha = 0{,}05$. Если окажется, что при уровне значимости $\alpha = 0{,}05$ $\chi^2_{набл} > \chi^2_{кр}$, то по таблице Пирсона найти наименьшее критическое значение статистики $\chi^2_{кр}(\alpha', k)$ и определить уровень значимости α', при котором можно считать, что данная выборка подчиняется нормальному закону распределения.

4. Найти доверительный интервал для неизвестного параметра a по формуле:

$$\bar{x} - \frac{t_\gamma s_x}{\sqrt{n}} \le a \le \bar{x} + \frac{t_\gamma s_x}{\sqrt{n}} \qquad (1),$$

где t_γ — коэффициент Стьюдента, значение которого определяется объемом выборки и надежностью $\gamma = 1 - \alpha_1$ оценки. Значение t_γ находим по таблице распределения Стьюдента [1].

Доверительный интервал для неизвестной дисперсии определяется формулой

$$s_x(1-q) \le \sigma \le s_x(1+q) \qquad (2).$$

Значение $q(\gamma, n)$ приведено в таблице (см. приложение 4 [1]).

Проверим гипотезу о нормальном распределении X по критерию Пирсона и найдем интервальные оценки a и σ.

1. Объединим пятый и шестой интервалы, так как значения частоты в шестом интервале меньше 5 (см. табл. 2).

2. Для вычисления по данным выборки значений статистики $\chi^2_{набл}$ составим расчетную таблицу 4.

Таблица 4

№	α_{i-1} — α_i	n_{x_i}	$u_i = \frac{\alpha_{i-1} - \bar{x}}{s_x}$	$\Phi(u_i)$	$p'_i = \Delta\Phi$	$n'_{x_i} = n \cdot p'_i$	Целое n'_{x_i}	$\frac{(n_{x_i} - n'_{x_i})^2}{n'_{x_i}}$
1	53 — 58	5	-2,62	-0,4591	0,0365	2,85	3	1.33
2	58 — 63	9	-1,74	-0.4591	0.1513	11,80	12	0,75
3	63 — 68	20	-0,87	-0,3078	0,3078	24,01	24	0,66
4	68 — 73	33	0	0	0,3078	24,01	24	3,37
5	73 — 83	11	0,87	0,3078	0,1878	14,63	15	1,01
$\sum$		78					78	7,12

При расчетах использовались значения, полученные в лабораторной работе №1: $\bar{x} = 68$,

$s_x = 5{,}73$, $n = 78$.

3. Наблюдаемое значение статистики $\chi^2_{набл} = 7{,}12$. По таблице критических точек распределения χ^2 (см. приложение 5 [1]) по уровню значимости $\alpha = 0{,}05$ и числу степеней свободы $k = 2$ находим $\chi^2_{кр}(0{,}05;\ 2)$=6. Так как $\chi^2_{набл} > \chi^2_{кр}$, то при уровне значимости $\alpha = 0{,}05$ нельзя признать распределение нормальным. Однако нулевая гипотеза может быть признана с уровнем значимости $\alpha' = 0{,}025$, так как $\chi^2_{кр}(0{,}025;\ 2)$=7,4. В этом случае $\chi^2_{набл} < \chi^2_{кр}$, то есть, нет оснований отвергнуть нулевую гипотезу.

4. Оценим по формуле (1) неизвестное математическое ожидание при помощи доверительного интервала с надежностью $\gamma = 1 - \alpha' = 1 - 0{,}025 = 0{,}975$. При $\gamma = 0{,}975$, $n = 80$ $t_\gamma \approx 2{,}24$ (см. приложение 3 [1]).

$$68 - \frac{2{,}24 \cdot 5{,}73}{\sqrt{78}} \le a \le 68 + \frac{2{,}24 \cdot 5{,}73}{\sqrt{78}}$$

$$66{,}55 \le a \le 69{,}45.$$

По таблице $q(\gamma, n)$ (см. приложение 4 [1]) найдем значение $q = q(0{,}975;\ 78) \approx 0{,}19$.
По формуле (2) находим интервальную оценку параметра σ:

$$5{,}73 \cdot (1 - 0{,}19) \le \sigma \le 5{,}73 \cdot (1 + 0{,}19).$$

$$4{,}64 \le \sigma \le 6{,}82.$$