Проверка гипотезы по критерию Колмогорова-Смирнова

Задание.

В течение месяца выборочно осуществлялась проверка торговых точек города по продаже овощей. Результаты двух проверок по недовесам покупателям одного вида овощей приведены в таблице:

№	1	2	3	4	5	6	7	8	9
Интервалы,г	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
n_{1i}	3	10	15	20	12	5	25	15	5
n_{2i}	5	12	8	25	10	8	20	7	5

Можно ли считать при уровне значимости 0,05, что недовесы овощей являются устойчивым и закономерным процессом при продаже овощей в данном городе (т.е. описываются одной и той же функцией распределения)?

Решение.

Используем критерий Колмогорова-Смирнова (проверяем гипотезу $H_0 : F_1(x) = F_2(x)$ - о том, что данные описываются одной и той же функцией распределения). Вычислим накопленные частоты для обоих выборок и значения эмпирических функций (относительные накопленные частоты). Расчеты будем вести в таблице.

интер-вал	n_1	n_2	$n_1^{\text{íàê}}$	$n_2^{\text{íàê}}$	$F_1^*(x) = \frac{n_1^{\text{íàê}}}{n_1}$	$F_2^*(x) = \frac{n_2^{\text{íàê}}}{n_2}$	$\lvert F_1^*(x) - F_2^*(x) \rvert$
1	3	5	3	5	0,027	0,050	0,023
2	10	12	13	17	0,118	0,170	0,052
3	15	8	28	25	0,255	0,250	0,005
4	20	25	48	50	0,436	0,500	0,064
5	12	10	60	60	0,545	0,600	0,055
6	5	8	65	68	0,591	0,680	**0,089**
7	25	20	90	88	0,818	0,880	0,062
8	15	7	105	95	0,955	0,950	0,005
9	5	5	110	100	1,000	1,000	0,000

Найдем наибольшее отклонение, затем вычисляем значение критерия:

$$\lambda = \max_{x_i} \lvert F_1^*(x) - F_2^*(x) \rvert \cdot \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = 0,089 \cdot \sqrt{\frac{100 \cdot 110}{100 + 110}} \approx 0,644 .$$

Так как $\lambda < \lambda_{0,05} = 1,36$, то гипотеза принимается, можно считать, что недовесы овощей описываются одной и той же функцией распределения.

.