

Математическая статистика

Домашняя работа № 4

Проверка статистических гипотез

Попов Юрий, СКБ-172

ОГЛАВЛЕНИЕ

4.1 Геометрическое распределение	5
4.2 Экспоненциальное распределение	7

Предисловие

Немного теории

Определение 1 *Статистическая гипотеза* - это некоторое предположение о виде или параметрах распределений

Определение 2 *Статистический критерий* - это правило, по которому каждой реализации выборки ставится в соответствие решение: принимаем гипотезу H_0 или отвергаем ее (то есть принимаем гипотезу H_1)

Определение 3 *Уровень значимости статистического теста* - допустимая для данной задачи вероятность ошибки первого рода, то есть вероятность отклонить нулевую гипотезу, когда на самом деле она верна

Определение 4 В случае, когда H_0 и H_1 - простые гипотезы, $P(x \in \mathcal{X} | H_0) = \lambda$ - ошибка 1 рода

$\beta = P(H_0 | H_1)$ - ошибка второго рода

Определение 5 Если H_0 состоит из одного определения, то говорят, что H_0 - *простая гипотеза*, иначе H_0 *сложная гипотеза*

Определение 6 Если H_1 состоит из одного определения, то говорят, что H_1 - *простая гипотеза*, иначе H_1 *сложная гипотеза*

Перейдем к практике

4.1 Геометрическое распределение

В качестве критерия я взял критерий Пирсона. Этот метод позволяет оценить статистическую значимость различий двух или нескольких относительных показателей. Это наиболее часто употребляемый критерий для проверки гипотезы о принадлежности наблюдаемой выборки некоторому теоретическому закону распределения.

Этот критерий универсальный

Для каждой выборки составим расчетную таблицу:

Расчетная таблица для 5 реализаций выборки, объема 1000:

	0	1	2	3	4	5	6	7
freq	100.0	87.00	81.00	89.00	58.0	47.00	56.0	42.0
$x_i \cdot \frac{n_i}{n}$	0.0	87.00	162.00	267.00	232.0	235.00	336.0	294.0
$(\hat{p}_i(x) - x_i)^2 \cdot \frac{n_i}{n}$	1651225.0	1414293.75	1296182.25	1401772.25	899014.5	716855.75	840350.0	620014.5

3 rows x 51 columns

Выборочное среднее равно: 128.5
Выборочная исправленная дисперсия равна: 14376.928
Выборочное исправленное среднее квадратическое отклонение равно 119.904

	0	1	2	3	4	5	6
freq	96.0	89.00	84.0	93.00	69.00	79.00	57.00
$x_i \cdot \frac{n_i}{n}$	0.0	89.00	168.0	279.00	276.00	395.00	342.00
$(\hat{p}_i(x) - x_i)^2 \cdot \frac{n_i}{n}$	1585176.0	1446806.25	1344189.0	1464773.25	1069517.25	1204927.75	855356.25

3 rows x 49 columns

Выборочное среднее равно: 128.5
Выборочная исправленная дисперсия равна: 14530.416
Выборочное исправленное среднее квадратическое отклонение равно 120.542

	0	1	2	3	4	5	6
freq	96.0	97.00	80.0	75.00	66.0	57.00	50.0
$x_i \cdot \frac{n_i}{n}$	0.0	97.00	160.0	225.00	264.0	285.00	300.0
$(\hat{p}_i(x) - x_i)^2 \cdot \frac{n_i}{n}$	1585176.0	1576856.25	1280180.0	1181268.75	1023016.5	869378.25	750312.5

3 rows x 49 columns

Выборочное среднее равно: 128.5
Выборочная исправленная дисперсия равна: 14406.972
Выборочное исправленное среднее квадратическое отклонение равно 120.029

46	48	49	51	52	58	62	Сумма
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1000.0
46.00	48.00	49.00	51.00	52.00	58.00	NaN	8877.0
6806.25	6480.25	6320.25	6006.25	5852.25	4970.25	4422.25	14377928.0

47	49	51	58	68	Сумма
1.00	1.00	1.00	1.00	1.00	1000.0
47.00	49.00	51.00	58.00	NaN	8204.0
642.25	6320.25	6006.25	4970.25	3660.25	14531416.0

62	63	70	71	Сумма
1.00	1.00	1.00	1.00	1000.0
62.00	63.00	70.00	NaN	8768.0
1422.25	4290.25	3422.25	3306.25	14407972.0

Выдвинем гипотезу H_0 : распределение генеральной совокупности X подчинено геометрическому закону с параметром p . Проверим эту гипотезу по критерию Пирсона при уровне значимости $\lambda = 0.05$

	0	1	2	3	4	5	6
freq	91.00	88.0	72.0	74.0	68.0	63.00	46.0
$x_i \cdot n_i$	0.00	88.0	144.0	222.0	272.0	315.00	276.0
$(\hat{x} - x_i)^2 \cdot n_i$	1502614.75	1430550.0	1152162.0	1165518.5	1054017.0	960891.75	690287.5

3 rows × 51 columns

Выборочное среднее равно: 128.5
 Выборочная исправленная дисперсия равна: 14305.844
 Выборочное исправленное среднее квадратическое отклонение равно 119.607

	0	1	2	3	4	5	6
freq	99.00	85.00	80.0	69.00	76.0	54.0	51.00
$x_i \cdot n_i$	0.00	85.00	160.0	207.00	304.0	270.0	306.00
$(\hat{x} - x_i)^2 \cdot n_i$	1634712.75	1381781.25	1280180.0	1086767.25	1178019.0	823621.5	765318.75

3 rows × 47 columns

Выборочное среднее равно: 128.5
 Выборочная исправленная дисперсия равна: 14358.978
 Выборочное исправленное среднее квадратическое отклонение равно 119.829

59	69	72	76	Сумма
1.00	1.00	1.00	1.00	1000.0
59.00	69.00	72.00	NaN	9196.0
830.25	3540.25	3192.25	2756.25	14306844.0

45	50	51	Сумма
3.00	2.0	1.00	1000.0
135.00	100.0	NaN	8961.0
10916.75	12324.5	6006.25	14359978.0

4.2 Экспоненциальное распределение

Литература

- [1]
- [2] [ссылка1](#)
- [3] [ссылка2](#)
- [4] [// ссылка3](#)
- [5] [// ссылка4](#)