**Математическая статистика**

**Обработка одномерной выборки**

Строим вариационный ряд:

0.70 0.79 2.09 2.04 0.71 0.49 0.94 0.35 1.85 0.18 0.50 1.75 0.81 0.56 1.77 3.15 0.69 4.36 0.20 4.07 0.61 2.10 1.50 3.31 0.66 1.98 3.53 0.92 0.48 1.03 0.54 0.18 0.70 0.23 0.88 6.19 1.41 0.05 4.53 0.88 2.20 0.16 1.74 0.30 0.78 0.03 0.28 1.99 2.45 1.02 0.77 0.42 5.91 0.08 0.04 0.34 0.39 0.28 3.04 0.99 2.09 0.64 0.91 0.47 0.28 5.72 0.39 1.32 0.14 0.91 0.67 1.28 0.59 0.28 2.19 1.37 0.31 0.24 0.83 2.24 1.39 0.05 0.06 0.40 1.24 0.08 1.66 0.82 2.91 1.88 0.61 0.03 0.26 0.03 0.65 0.27 4.86 0.95 5.35 1.19

Строим график эмпирической функции F\*(x).

Количество интервалов М, необходимое для построения гистограмм, определим по объему выборки: .

Для равноинтервальной гистограммы рассчитаем величины Aj, Bj, hj, Р\*j и fj по формулам: , , , ,  и заполним все колонки таблицы.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| j | Aj | Bj | hj | νj | P\*j | fj |
| 1 | 0.03 | 0.646 | 0.616 | 38 | 0.38 | 0.6169 |
| 2 | 0.646 | 1.262 | 0.616 | 25 | 0.25 | 0.4058 |
| 3 | 1.262 | 1.878 | 0.616 | 11 | 0.11 | 0.1786 |
| 4 | 1.878 | 2.494 | 0.616 | 11 | 0.11 | 0.1786 |
| 5 | 2.494 | 3.11 | 0.616 | 2 | 0.02 | 0.0325 |
| 6 | 3.11 | 3.726 | 0.616 | 3 | 0.03 | 0.0487 |
| 7 | 3.726 | 4.342 | 0.616 | 1 | 0.01 | 0.0162 |
| 8 | 4.342 | 4.958 | 0.616 | 3 | 0.03 | 0.0487 |
| 9 | 4.958 | 5.574 | 0.616 | 1 | 0.01 | 0.0162 |
| 10 | 5.574 | 6.19 | 0.616 | 3 | 0.03 | 0.0487 |

Равноинтервальная гистограмма имеет вид

Для равновероятностной гистограммы рассчитаем величины Aj, Bj, hj, Р\*j и

fj по формулам: , , , ,  и заполним все колонки таблицы.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| j | Aj | Bj | hj | νj | P\*j | fj |
| 1 | 0.03 | 0.15 | 0.12 | 10 | 0.1 | 0.8333 |
| 2 | 0.15 | 0.28 | 0.13 | 10 | 0.1 | 0.7692 |
| 3 | 0.28 | 0.445 | 0.165 | 10 | 0.1 | 0.6061 |
| 4 | 0.445 | 0.645 | 0.2 | 10 | 0.1 | 0.5 |
| 5 | 0.645 | 0.8 | 0.155 | 10 | 0.1 | 0.6452 |
| 6 | 0.8 | 0.97 | 0.17 | 10 | 0.1 | 0.5882 |
| 7 | 0.97 | 1.455 | 0.485 | 10 | 0.1 | 0.2062 |
| 8 | 1.455 | 2.065 | 0.61 | 10 | 0.1 | 0.1639 |
| 9 | 2.065 | 3.23 | 1.165 | 10 | 0.1 | 0.0858 |
| 10 | 3.23 | 6.19 | 2.96 | 10 | 0.1 | 0.0338 |

Равновероятностная гистограмма имеет вид

Вычислим точечную оценку математического ожиданияпо формуле:

.

Вычислим точечную оценку дисперсиипо формуле:

.

Построим доверительный интервал для математического ожидания с надёжностью . Для этого в таблице функции Лапласа найдём значение, равное , и определим значение аргумента, ему соответствующее: . Получим доверительный интервал формуле : .

Построим доверительный интервал для дисперсии с надёжностью . По формуле: .

По виду графика эмпирической функции распределения *F*\*(*x*) и гистограмм выдвигаем двухальтернативную гипотезу о законе распределения случайной величины:

H0 – величина *X* распределена по нормальному закону:

, ;

H1 - величина *X* не распределена по нормальному закону:

, .

Определим оценки неизвестных параметров *m* и σ гипотетического (нормального) закона распределения по формулам:

, .

Таким образом, получаем полностью определенную гипотетическую функцию распределения:

.

Проверим гипотезу о нормальном законе с помощью критерияχ2 .

Вычислим значение критерия χ2 на основе равноинтервального статистического рада по формуле:

.

Теоретические вероятности pi попадания в интервалы равноинтервального статистического ряда нормальной случайной величины вычислим по формуле:

.

Значения функции Лапласа определяем из таблицы значений для функции Лапласа. Результаты расчёта сводим в таблицу.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| j | Aj | Bj | F0(Aj) | F0(Bj) | pj | P\*j |  |
| 1 | -∞ | 0.646 | 0 | 0.323 | 0.3228 | 0.38 | 0.010151 |
| 2 | 0.646 | 1.262 | 0.3228 | 0.488 | 0.1653 | 0.25 | 0.043434 |
| 3 | 1.262 | 1.878 | 0.488 | 0.655 | 0.1674 | 0.11 | 0.019675 |
| 4 | 1.878 | 2.494 | 0.6554 | 0.8 | 0.1441 | 0.11 | 0.008079 |
| 5 | 2.494 | 3.11 | 0.7995 | 0.9 | 0.1002 | 0.02 | 0.064174 |
| 6 | 3.11 | 3.726 | 0.8997 | 0.957 | 0.0576 | 0.03 | 0.013194 |
| 7 | 3.726 | 4.342 | 0.9573 | 0.985 | 0.0273 | 0.01 | 0.010989 |
| 8 | 4.342 | 4.958 | 0.9846 | 0.995 | 0.0107 | 0.03 | 0.034637 |
| 9 | 4.958 | 5.574 | 0.9953 | 0.999 | 0.0035 | 0.01 | 0.012227 |
| 10 | 5.574 | +бескон | 0.9988 | 1 | 0.0012 | 0.03 | 0.701967 |
|  |  |  |  | Сумма | 1 | 0.98 | 0.918527 |

Проверяем выполнение контрольного соотношения для pj:

.

В результате получаем .

Вычислим число степеней свободы *k* = *M*−1−*s*=10−1−2=7, где s - число параметров, от которых зависит выбранный гипотезой H0 закон распределения (и ) и по заданному уровню значимости α =0,05 из таблицы распределения χ2 выбираем критическое значение .

Так как , то гипотеза H0 о нормальном законе распределения не принимается.

**ОБРАБОТКА ДВУХМЕРНОЙ ВЫБОРКИ**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| № | x | y | x2 | y2 | x\*y |
| 1 | 3.83 | 2.4 | 14.6689 | 5.76 | 9.192 |
| 2 | -0.01 | 2.61 | 0.0001 | 6.8121 | -0.0261 |
| 3 | 2.4 | 2.92 | 5.76 | 8.5264 | 7.008 |
| 4 | 1.04 | 2.04 | 1.0816 | 4.1616 | 2.1216 |
| 5 | 3.01 | 1.98 | 9.0601 | 3.9204 | 5.9598 |
| 6 | 6.39 | 2 | 40.8321 | 4 | 12.78 |
| 7 | 3.95 | 7.46 | 15.6025 | 55.6516 | 29.467 |
| 8 | 8.49 | 5.88 | 72.0801 | 34.5744 | 49.9212 |
| 9 | 1.7 | 2.62 | 2.89 | 6.8644 | 4.454 |
| 10 | -1.23 | 0.25 | 1.5129 | 0.0625 | -0.3075 |
| 11 | 4.52 | 3.47 | 20.4304 | 12.0409 | 15.6844 |
| 12 | 1.87 | 5.37 | 3.4969 | 28.8369 | 10.0419 |
| 13 | 1.15 | 7.98 | 1.3225 | 63.6804 | 9.177 |
| 14 | -0.11 | 0.8 | 0.0121 | 0.64 | -0.088 |
| 15 | 0.93 | -0.25 | 0.8649 | 0.0625 | -0.2325 |
| 16 | 4.09 | 2.49 | 16.7281 | 6.2001 | 10.1841 |
| 17 | 1.62 | 2.45 | 2.6244 | 6.0025 | 3.969 |
| 18 | -1.52 | -3.44 | 2.3104 | 11.8336 | 5.2288 |
| 19 | 4.71 | 4.8 | 22.1841 | 23.04 | 22.608 |
| 20 | 6.16 | -0.5 | 37.9456 | 0.25 | -3.08 |
| 21 | 6.81 | 9.28 | 46.3761 | 86.1184 | 63.1968 |
| 22 | 5.56 | 7.45 | 30.9136 | 55.5025 | 41.422 |
| 23 | 6.14 | 5.82 | 37.6996 | 33.8724 | 35.7348 |
| 24 | 6.44 | 6.83 | 41.4736 | 46.6489 | 43.9852 |
| 25 | 6.79 | 5.17 | 46.1041 | 26.7289 | 35.1043 |
| 26 | 5.87 | 7.77 | 34.4569 | 60.3729 | 45.6099 |
| 27 | 9.39 | 5.4 | 88.1721 | 29.16 | 50.706 |
| 28 | 3.51 | 5.36 | 12.3201 | 28.7296 | 18.8136 |
| 29 | 11.08 | 5.63 | 122.7664 | 31.6969 | 62.3804 |
| 30 | 2.52 | 7.68 | 6.3504 | 58.9824 | 19.3536 |
| 31 | -0.01 | -2.59 | 0.0001 | 6.7081 | 0.0259 |
| 32 | 0.45 | 4.9 | 0.2025 | 24.01 | 2.205 |
| 33 | -0.17 | 2.46 | 0.0289 | 6.0516 | -0.4182 |
| 34 | 5.99 | 6.77 | 35.8801 | 45.8329 | 40.5523 |
| 35 | -0.61 | -2.48 | 0.3721 | 6.1504 | 1.5128 |
| 36 | 2.3 | 4.23 | 5.29 | 17.8929 | 9.729 |
| 37 | -0.99 | 0.14 | 0.9801 | 0.0196 | -0.1386 |
| 38 | 8.52 | 4.95 | 72.5904 | 24.5025 | 42.174 |
| 39 | 2.97 | 4.2 | 8.8209 | 17.64 | 12.474 |
| 40 | 3.6 | 4.66 | 12.96 | 21.7156 | 16.776 |
| 41 | 6.14 | 11.2 | 37.6996 | 125.44 | 68.768 |
| 42 | 10.62 | 8.56 | 112.7844 | 73.2736 | 90.9072 |
| 43 | -2.01 | -0.11 | 4.0401 | 0.0121 | 0.2211 |
| 44 | 3.81 | 6.28 | 14.5161 | 39.4384 | 23.9268 |
| 45 | -0.16 | 3.32 | 0.0256 | 11.0224 | -0.5312 |
| 46 | 6.83 | 11.81 | 46.6489 | 139.4761 | 80.6623 |
| 47 | 7.27 | 10.04 | 52.8529 | 100.8016 | 72.9908 |
| 48 | 3.1 | 7.54 | 9.61 | 56.8516 | 23.374 |
| 49 | -2.25 | -0.59 | 5.0625 | 0.3481 | 1.3275 |
| 50 | 0.34 | 2.23 | 0.1156 | 4.9729 | 0.7582 |
| Средние | 3.4568 | 4.1448 | 23.17043 | 29.25787 | 21.95332 |

Значения в 3-ем, 4-ом и 5-ом столбцах вычисляются по формулам, приведённым в первой строке таблицы. В последней строке таблицы приведены средние арифметические значений каждого из столбцов.

Таким образом получены:

- оценки математических ожиданий по каждой переменной

,

;

- оценки начальных моментов второго порядка по каждой переменной:

,

;

- оценка смешанного начального момента второго порядка:

.

На основе этих данных вычислим оценки дисперсий:

,

,

и оценку корреляционного момента: .

Вычислим точечную оценку коэффициент корреляциипо формуле:

.

Вычислим интервальную оценку коэффициента корреляции с надёжностью . Для этого в таблице функции Лапласа найдём значение, равное , и определим значение аргумента, ему соответствующее: . Получим доверительный интервал по формуле , где , :

.

Проверим гипотезу об отсутствии корреляционной зависимости:

;

.

Так как объём выборки велик , то вычислим значение критерия по формуле : .

Определим значение  из таблицы функции Лапласа: .

Так как , то гипотеза  не принимается, т. е. величины X и Y коррелированны.

Оценки параметров  и  линии регрессии  вычислим по формулам:

,

.

Уравнение линии регрессии имеет вид:

.

Построим диаграмму рассеивания, изобразив значения исходной двумерной выборки, в виде точек с координатами (*хi*, *уi*) на плоскости в декартовой системе координат, и линию регрессии.