Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Лабораторная работа №1 по дисциплине

«**Математическая статистика**»

Вариант 5

Выполнили:

Боринский Игорь Дмитриевич, группа P3214

Болорболд Аригуун, группа P3211

Преподаватель:

Милованович Екатерина Воиславовна

Санкт-Петербург, 2024 г.

Название: Оценивание, проверка статистических гипотез

Цель работы: На основании анализа выборки:

1. Построить интервальный ряд исследуемой случайной величины.
2. Построить полигон частот, выборочную функцию распределения и гистограмму.
3. Найти точечные оценки мат. ожидания и дисперсии.
4. Найти доверительные интервалы для мат. ожидания и дисперсии с доверительной вероятностью 0,95.
5. Проверить статистическую гипотезу о виде нормального распределения.

Исходные данные:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| -0.602 | 2.301 | -0.432 | 0.273 | -0.802 | -0.322 | 0.459 | -0.023 | 0.361 | 0.557 |
| -0.993 | -0.27 | -0.194 | 2.646 | -0.456 | -0.703 | 0.66 | 0.134 | -2.058 | -0.18 |
| 1.188 | 0.502 | 0.985 | -0.053 | 0.193 | -0.744 | 1.124 | 2.408 | -2.332 | -0.035 |
| 2.388 | -0.119 | 0.468 | 0.472 | 0.889 | 0.371 | 0.979 | 0.901 | -0.37 | 1.934 |
| 2.265 | -0.001 | -1.364 | -2.08 | -1.591 | 1.437 | -1.316 | 0.076 | 1.285 | 1.305 |
| -0.355 | -2.735 | 1.194 | -1.038 | 0.586 | -0.213 | 1.143 | 0.454 | 0.097 | -0.016 |
| -0.327 | -0.535 | 0.743 | 0.628 | 1.525 | 0.492 | 0.979 | -1.417 | -0.226 | 0.449 |
| 0.083 | 2.209 | -0.121 | 0.867 | 2.143 | -0.323 | 0.492 | -0.919 | -0.317 | -0.522 |
| 0.433 | -0.605 | -0.031 | 2.071 | -0.746 | 0.822 | 1.257 | -1.448 | 0.634 | -1.055 |
| -1.435 | -1.003 | -0.594 | -1.531 | 1.414 | 0.594 | -1.481 | 0.039 | -0.047 | 1.152 |

Расчетные формулы с пояснениями, графики

Найдём максимальное и минимальное значение выборки:

Определим размах выборки:

Вычислим шаг разбиения по формуле Стерджеса:

В итоге разбили ряд на 8 интервалов с длиной 0,7

Далее определим частоты и относительные частоты для каждого интервала.

Полученные данные занесём в таблицу:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Номера интервалов** | **Границы интервалов** | xi​ | *ni*​ | *P\*i* |
| 1 | (-2.8; -2.1) | -2.45 | 2 | 0.02 |
| 2 | (-2.1; -1.4) | -1.75 | 8 | 0.08 |
| 3 | (-1.4; -0.7) | -1.05 | 11 | 0.11 |
| 4 | (-0.7; 0.0) | -0.35 | 27 | 0.27 |
| 5 | (0.0; 0.7) | 0.35 | 24 | 0.24 |
| 6 | (0.7; 1.4) | 1.05 | 16 | 0.16 |
| 7 | (1.4; 2.1) | 1.75 | 5 | 0.05 |
| 8 | (2.1; 2.8) | 2.45 | 7 | 0.07 |

Построим полигон частот, а также гистограмму, в которой высота столбца на интервале равна



Изображение выглядит как снимок экрана, прямоугольный, Прямоугольник, линия

Автоматически созданное описание

Далее построим эмпирическую функцию распределения. Она имеет вид:

Отсюда график эмпирической функции распределения имеет вид:



Найдем точечные оценки математического ожидания и дисперсии. В качестве таких оценок возьмем среднее выборочное значение и выборочную дисперсию .

Результаты занесём в таблицу:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № интервала | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |  |
|  | -2,45 | -1,75 | -1,05 | -0,35 | 0,35 | 1,05 | 1,75 | 2,45 | - |
|  | 0,02 | 0,08 | 0,11 | 0,27 | 0,24 | 0,16 | 0,05 | 0,07 | - |
|  | -0,049 | -0,14 | -0,116 | -0,095 | 0,084 | 0,168 | 0,088 | 0,172 | 0,112 |
|  | -0,12 | 0,245 | 0,121 | 0,033 | 0,029 | 0,176 | 0,153 | 0,42 | 1,299 |

Получаем, что

Построим доверительный интервал:

Доверительный интервал – *γ* = 0,95

Ввиду большого объема выборки доверительный интервал имеет вид:

Параметр t определяется из равенства:

*,*

где .

Для определения t при использовании функции Лапласа будем иметь следующее уравнение .

*.*

Для рассматриваемого примера будем иметь при , откуда t = 1.95, поэтому в нашем примере имеем:

Таким образом, доверительный интервал для математического ожидания имеет вид:

Доверительный интервал для дисперсии имеет следующий вид:

Здесь значения и находятся по таблицам -распределения с степенью свободы, причем в таблице мы ищем и .

По таблице -распределения находим , , число степеней свободы

и определяем критические точки и .

Искомый 95% — доверительный интервал имеет вид:

Доверительный интервал для дисперсии в итоге имеет следующий вид:

Проверим гипотезу о том, что генеральная совокупность, из которой произведена выборка, имеет нормальный закон распределения:

Применим критерий согласия χ2 (Пирсона).

где — число элементов выборки в каждом из k интервалов; pi – теоретическая вероятность попадания случайной величины в i -й интервал, которая определяется по формуле.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Интервалы | (; -2,1) | (-2,1; -1,4) | (-1,4; -0,7) | (-0,7; 0) | (0; 0,7) | (0,7; 1,4) | (1,4; 2,1) | (2,1; ) | Примечания |
|  | -1,941 | -1,229 | -0,518 | 0 | 0,614 | 1,229 | 1,843 |  |  |
|  | 0,0261 | 0,1095 | 0,3022 | 0,5 | 0,7304 | 0,8905 | 0,9673 | 1 |  |
|  | 0,0261 | 0,0834 | 0,1927 | 0,1978 | 0,2304 | 0,1601 | 0,0768 | 0,0327 |  |
|  | 2 | 8 | 11 | 27 | 24 | 16 | 5 | 7 |  |
|  | 4 | 64 | 121 | 729 | 576 | 256 | 25 | 49 |  |
|  | 2,61 | 8,34 | 19,27 | 19,78 | 23,04 | 16,01 | 7,68 | 3,27 |  |
|  | 1,53 | 7,67 | 6,28 | 36,86 | 25 | 15,99 | 3,26 | 14,98 |  |

при уровне значимости 0,05 , полученное значение меньше заданного, следовательно, гипотеза о нормальном распределении генеральной совокупности не отвергается.

Итоги:

В ходе анализа выборки были построены визуализации в виде интервального ряда, полигона частот, гистограммы и функции распределения, которые наглядно продемонстрировали распределение данных. Также были вычислены точечные оценки для математического ожидания и дисперсии, что позволило количественно охарактеризовать данные. Доверительные интервалы для этих параметров были определены с доверительной вероятностью 0,95, что дало представление о диапазонах их возможных значений. Проверка гипотезы о нормальном распределении подтвердила применимость статистических методов для дальнейшего анализа. В целом, проведённый статистический анализ показал важность глубокого понимания данных для обоснования дальнейших исследований и принятия решений.