|  |  |
| --- | --- |
| Группа ПИИКТ 1.2 | К работе допущен |
| Студент Мухамедьяров Артур Альбертович, Ларионов Владислав Васильевич, Шубин Егор Вячеславович | Работа выполнена |
| Преподаватель Рудель Алена Евгеньевна | Отчет принят |

**Рабочий протокол и отчет по  
лабораторной работе №1**

Исследование распределения случайной величины

1. Цель работы.

Исследование распределения случайной величины на примере количества слов в текстах треков Eminem.

2. Задачи, решаемые при выполнении работы.

1. Провести многократные подсчёты количества слов в различных текстах треков Eminem
2. Построить гистограмму распределения результатов измерения.
3. Вычислить среднее значение и дисперсию полученной выборки.
4. Сравнить гистограмму с графиком функции Гаусса с такими же, как и у экспериментального распределения средним значением и среднеквадратичным отклонением.

3. Объект исследования.

Случайная величина – результат подсчёта количества слов.

4. Метод экспериментального исследования.

Многократное измерение количества слов в различных треках Eminem и проверка закономерностей распределения значений этой случайной величины.

5. Рабочие формулы и исходные данные.

* Выборочное среднее как среднее арифметическое всех результатов измерений:
* Выборочное среднеквадратичное отклонение:
* Значение плотности вероятности
* Максимальное значение плотности распределения
* Функция Гаусса
* Среднеквадратичное отклонение среднего значения
* Доверительный интервал
* Доверительный интервал для измеряемого промежутка времени
* Формулы для вычисления приближенной вероятности попадания каждого измерения t в интервал [t1; t2]
* Соотношение вероятности попадания результата измерения в интервал [t1; t2]
* Значения вероятности попадания результата каждого измерения 𝑡 в интервал [𝑡1; 𝑡2] в стандартных (наиболее употребительных на практике) интервалах при условии реализации нормального распределения случайной величины

(11)

* Относительное отклонение экспериментальной вероятности попадания результата измерения в интервал [𝑡1; 𝑡2] от теоретической при условии реализации нормального распределения случайной величины

6. Измерительные приборы.

*Таблица 1. Измерительные приборы.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *№ п/п* | *Наименование* | *Тип прибора* | *Используемый диапазон* | *Погрешность прибора* |
| *1* | *Компьютер с программой–счётчиком слов* | *Программный (Python-скрипт)* | *0–2000 слов* | *±1 слово* |

7. Схема установки (*перечень схем, которые составляют Приложение 1*).

1. Источник данных: тексты песен Eminem (страничка артиста на сайте Genius <https://genius.com/artists/Eminem>).
2. Измерительный прибор: программа-счётчик слов.
3. Далее: Excel/Python для статистики и построения графиков.

8. Результаты прямых измерений и их обработки (*таблицы, примеры расчетов*).

Результаты прямых измерений приведены в Таблице 2 Приложения 1. *(Таблица 2. Результаты прямых измерений и их обработки.)*

9. Расчет результатов косвенных измерений (*таблицы, примеры расчетов*).

* Среднее арифметическое значение результатов измерений (1):
* Выборочное среднеквадратичное отклонение (2):
* Максимальное значение плотности распределения (4):

*Таблица 3. Стандартные доверительные интервалы.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Интервал* | *Начало интервала, слова* | *Конец интервала, слова* |  |  | *P, %* |
|  | 280 | 575,6 | 36 | 0,72 | 0,6827 |
|  | 132,2 | 723,4 | 48 | 0,96 | 0,9545 |
|  | -15,6 | 871,2 | 49 | 0,98 | 0,9973 |

* Расчет попадания результата измерения в интервал (10):
* Расчет попадания результата измерения в интервал (10):
* Расчет попадания результата измерения в интервал (10):
* Среди полученных данных tmin = 155 c и tmах = 953 с разобьем промежуток [155; 953] на равных частей: 114

*Таблица 4. Данные для построения гистограммы.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *№* | *Границы интервала* | *ΔN* | *ρ* | *t* | *ρ(t)* |
| 0 | 155-268 | 6 | 0,001053 | 211,5 | 0,000925 |
| 1 | 269-382 | 12 | 0,002105 | 325,5 | 0,002124 |
| 2 | 383-496 | 20 | 0,003509 | 439,5 | 0,002691 |
| 3 | 497-610 | 6 | 0,001053 | 553,5 | 0,001880 |
| 4 | 611-724 | 4 | 0,000702 | 667,5 | 0,000725 |
| 5 | 725-838 | 1 | 0,000175 | 781,5 | 0,000154 |
| 6 | 839-953 | 1 | 0,000175 | 896,0 | 0,000018 |

* Пример расчета значения плотности вероятности для интервала №1 (3):
* Пример расчета значения функции Гаусса для интервала №1 (5):

10. Расчет погрешностей измерений (*для прямых и косвенных измерений*).

- инструментальная погрешность

Найдем табличное значение коэффициента Стьюдента где доверительная вероятность α = 0,95, а число измерений N = 50:

Рассчитаем среднеквадратичное отклонение среднего значения, для этого воспользуемся уже найденной величиной и формулой, которую можно вывести из формул (2) и (6):

слов

Далее воспользовавшись формулой (8), рассчитаем доверительный интервал измеряемой величины:

20,903 = 42,01 слов

Рассчитаем абсолютную погрешность с учетом доверительного интервала и инструментальной погрешности

Рассчитаем относительную погрешность:

В данном случае случайная погрешность много больше инструментальной в связи с чем значение абсолютной погрешности практически не отличается от доверительного интервала.

Рассчитаем относительное отклонение экспериментальной вероятности попадания результата измерения в стандартные интервалы от теоретической при условии реализации нормального распределения измеряемой величины по формуле (12) (обозначим их как , и для интервалов , и соответственно):

Исходя из полученных расчетов можем заметить, что полученное распределение достаточно точно соответствует нормальному.

11. Графики (*перечень графиков, которые составляют Приложение 2*).

Гистограмма распределения измеряемой величины и график нормального распределения приведены в Приложении 2 на рисунке 1.

12. Окончательные результаты.

13. Выводы и анализ результатов работы.

В ходе лабораторной работы была проведена серия измерений количества слов в текстах треков Eminem. На основе 50 полученных значений рассчитаны выборочное среднее и выборочное среднеквадратичное отклонение, построена гистограмма распределения и проведено сравнение с теоретической кривой нормального распределения. Анализ показал, что распределение экспериментальных данных в целом близко к нормальному, особенно во втором и третьем стандартных интервалах, хотя в первом интервале наблюдается некоторое превышение экспериментальной вероятности над теоретической, что может быть связано с ограниченностью выборки и особенностями текстов песен. Абсолютная погрешность измерений составила 42 слова, а относительная -- 10%, что свидетельствует о значительном влиянии случайной погрешности. Полученные результаты подтверждают корректность выбранного метода исследования и позволяют сделать вывод о характере распределения изучаемой случайной величины.14. Дополнительные задания.

15. Выполнение дополнительных заданий.

16. Замечания преподавателя (*исправления, вызванные замечаниями преподавателя, также помещают в этот пункт*).

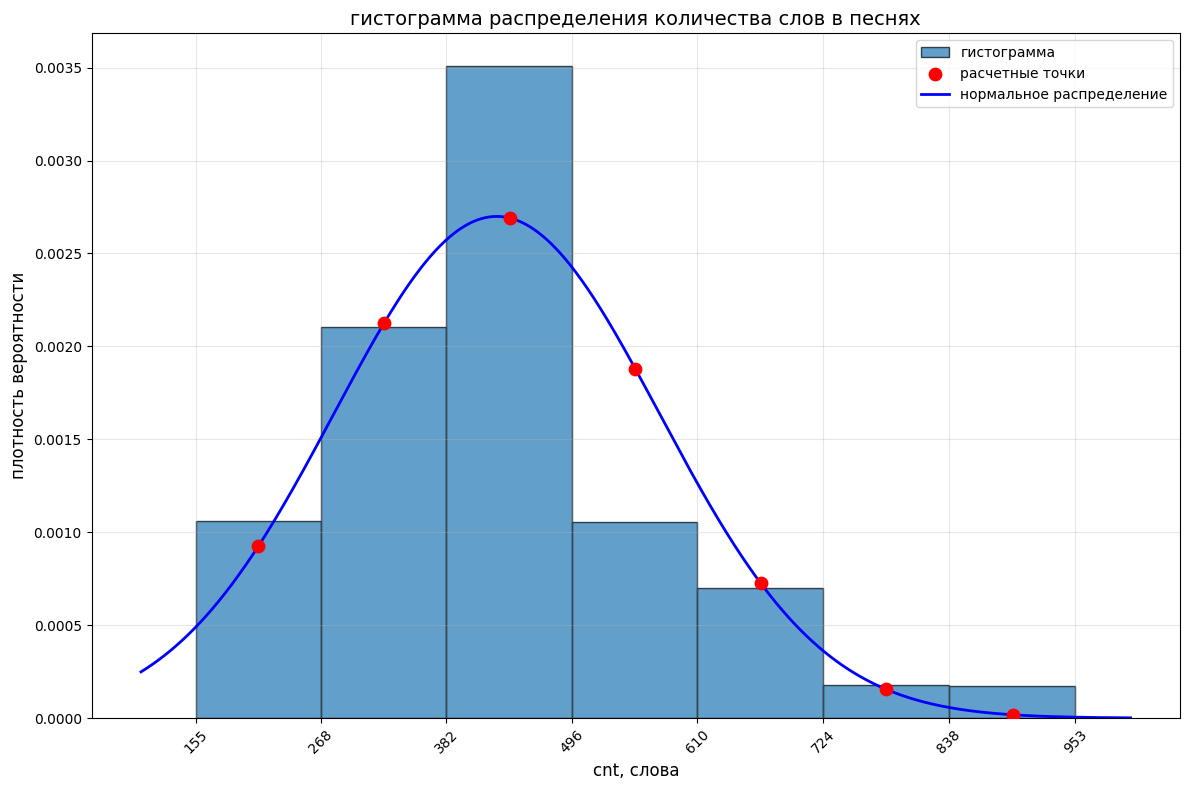
|  |  |
| --- | --- |
| ***Примечание:*** | 1. *Пункты 1-6,8-13 Протокола-отчета* ***обязательны*** *для заполнения.* |
|  | 1. *Необходимые исправления выполняют непосредственно в протоколе-отчете.* |
|  | 1. *При ручном построении графиков рекомендуется использовать миллиметровую бумагу.* |
|  | 1. *Приложения 1 и 2 вкладывают в бланк протокола-отчета.* |

*Приложение 1.*

*Таблица 2. Результаты прямых измерений и их обработки.*

|  |  |  |  |
| --- | --- | --- | --- |
| № | ti, слова | ti - , слова | (ti - )2, слова2 |
| 1 | 412 | -15.8 | 249.64 |
| 2 | 378 | -49.8 | 2480.04 |
| 3 | 512 | 84.2 | 7089.64 |
| 4 | 456 | 28.2 | 795.24 |
| 5 | 678 | 250.2 | 62600.04 |
| 6 | 289 | -138.8 | 19265.44 |
| 7 | 345 | -82.8 | 6855.84 |
| 8 | 423 | -4.8 | 23.04 |
| 9 | 389 | -38.8 | 1505.44 |
| 10 | 467 | 39.2 | 1536.64 |
| 11 | 312 | -115.8 | 13409.64 |
| 12 | 398 | -29.8 | 888.04 |
| 13 | 512 | 84.2 | 7089.64 |
| 14 | 155 | -272.8 | 74419.84 |
| 15 | 267 | -160.8 | 25856.64 |
| 16 | 334 | -93.8 | 8798.44 |
| 17 | 189 | -238.8 | 57025.44 |
| 18 | 423 | -4.8 | 23.04 |
| 19 | 345 | -82.8 | 6855.84 |
| 20 | 740 | 312.2 | 97468.84 |
| 21 | 953 | 525.2 | 275835.04 |
| 22 | 456 | 28.2 | 795.24 |
| 23 | 278 | -149.8 | 22440.04 |
| 24 | 389 | -38.8 | 1505.44 |
| 25 | 334 | -93.8 | 8798.44 |
| 26 | 456 | 28.2 | 795.24 |
| 27 | 267 | -160.8 | 25856.64 |
| 28 | 378 | -49.8 | 2480.04 |
| 29 | 512 | 84.2 | 7089.64 |
| 30 | 678 | 250.2 | 62600.04 |
| 31 | 456 | 28.2 | 795.24 |
| 32 | 567 | 139.2 | 19376.64 |
| 33 | 423 | -4.8 | 23.04 |
| 34 | 234 | -193.8 | 37558.44 |
| 35 | 389 | -38.8 | 1505.44 |
| 36 | 512 | 84.2 | 7089.64 |
| 37 | 678 | 250.2 | 62600.04 |
| 38 | 456 | 28.2 | 795.24 |
| 39 | 345 | -82.8 | 6855.84 |
| 40 | 389 | -38.8 | 1505.44 |
| 41 | 278 | -149.8 | 22440.04 |
| 42 | 423 | -4.8 | 23.04 |
| 43 | 334 | -93.8 | 8798.44 |
| 44 | 267 | -160.8 | 25856.64 |
| 45 | 456 | 28.2 | 795.24 |
| 46 | 389 | -38.8 | 1505.44 |
| 47 | 512 | 84.2 | 7089.64 |
| 48 | 678 | 250.2 | 62600.04 |
| 49 | 423 | -4.8 | 23.04 |
| 50 | 456 | 28.2 | 795.24 |
| среднее tN: 427.8  выборочное среднеквадратичное отклонение: 147.8046145478302  контроль правильности нахождения среднего: 0.0  максимальное значение плотности распределения: 0.002699119250247314 | | | |

*Приложение 2.*



*Рисунок 1. Гистограмма распределения измеряемой величины и график нормального распределения*