⏱️

Лабораторная работа №1.01

Исследование распределения случайной величины

выполнили: студенты группы М3205 Бородина Ирина, Патолицына Анастасия

преподаватель: Хуснутдинова Наира Рустемовна

Цели работы:

1. Исследование распределения случайной величины на примере многократных измерений определённого интервала времени.

Задачи:

1. Провести многократные измерения определенного интервала времени (5 секунд).
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 дисперсией.

Экспериментальная установка:

В работе используются *часы с секундной стрелкой* - прибор, в котором происходит периодический процесс с частотой порядка нескольких десятых долей герца, и *цифровой секундомер*, с ценой деления не более 0,01 *с*. Первый прибор задает интервал времени, который многократно измеряется цифровым секундомером.

Теория:

- ***Распределением Гаусса*** - распределение вероятностей, которое в одномерном случае задаётся функцией плотности вероятности, совпадающей с функцией Гаусса:

$$\rho(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\langle t\rangle)^2}{2\sigma^2}\right)$$

Вид нормального распределения определяется значениями двух параметров: ***математическим ожиданием*** $\langle t\rangle$ и ***среднеквадратичным (стандартным) отклонением*** σ.
Нормальное распределение симметрично относительно вертикали с абсциссой $\langle t\rangle$. Параметр σ характеризует «ширину» распределения.

- Закономерности в распределении значений случайной величины можно ***надежно выявить при соблюдении двух условий***:
- ***Гистограмма*** - способ представления табличных данных в графическом виде - в виде столбчатой диаграммы. Количественные соотношения некоторого показателя представлены в виде прямоугольников, площади которых пропорциональны.

В нашем случае - показывающую, как часто при измерениях появляются значения, попадающие в тот или иной из равных интервалов $\triangle t$, лежащих между наименьшим t_{min} и наибольшим t_{max} из измеренных значений величины t. (см. рис. 1)

Ось абсцисс
(x) - измеряемая величина t
Ось ординат
(y) - значения $\frac{\triangle N}{N\triangle t}$.

N - полное количество измерений, $\triangle N$ - количество результатов, попавших в интервал $[t;\ t+\triangle t]$.

$\frac{\triangle N}{N}$ - доля результатов, попавших в указанный интервал, и характеризует вероятность попадания в него результата отдельного измерения. Отношение этой величины к ширине интервала $\triangle t$ характеризует некоторую «плотность вероятности».

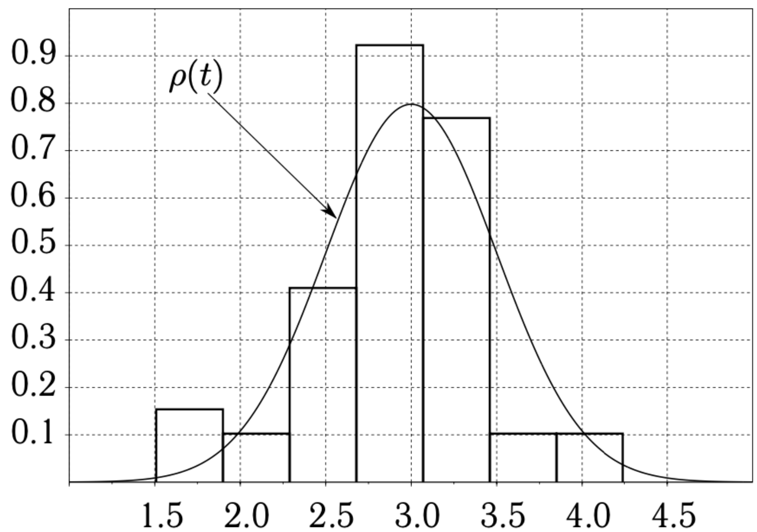

Рис. 1. Гистограмма и функция Гаусса.

При очень большом числе измерений $(N \to \infty)$ весь диапазон значений t в принципе можно разбить на «бесконечно малые» интервалы dt и подсчитать число результатов dN в каждом из них. Тогда вместо ступенчатой гистограммы получится плавная кривая соответствующая новой функции t:

$$\rho(t) \ = \ \lim_{\substack{N\to\infty \\ \triangle t \to 0}} \ \frac{\triangle N}{N \triangle t} \ = \frac{1}{N}\frac{dN}{dt}$$

- Параметры $\langle t \rangle$ и σ могут быть определены только на основе результатов бесконечно большого числа измерений (генеральной совокупности).
  Однако, в соответствии с теорией вероятности из выборочной совокупности, содержащей N значений случайной величины $(t_1, \ t_2, t_3, \ ... \ t_N)$, можно найти приближенные значений этих параметров:

выборочное среднее как среднеарифметическое всех результатов измерений:

$$\langle t \rangle_N \ = \ \frac{1}{N}\,(t_1 + t_2 + ... + t_N) \ = \ \frac{1}{N}\sum_{i=1}^{n} t_i$$

и ***выборочное среднеквадратичное отклонение:***

$$\sigma_N \ = \ \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}\left(t_i - \langle t \rangle_N\right)^2}$$

В пределе величины $\langle t \rangle_N$ и σ_N стремятся к $\langle t \rangle$ и σ, соответственно.

- Из формулы функции Гаусса видно, что плотность нормального распределения имеет максимум при $t = \langle t \rangle$ и симметрична относительно этого значения.
- Формула сравнения максимальной «высоты» гистограммы и $\rho_{max}(t)$:

$$\rho_{max} = \frac{1}{\sigma\sqrt{2\pi}}$$

- Соотношение для вероятности попадания результата измерения в интервал $[t_1,\ t_2]$:

$$P\left(t_1 < t < t_2\right) = \int\limits_{t_1}^{t_2} \rho(t)dt \approx \frac{N_{12}}{N}$$

  В случае наиболее употребительных на практике интервалов (так называемых стандартных) эта вероятность при условии реализации нормального распределения случайной величины имеет следующие значения:

$$\begin{gathered} t \in [\langle t \rangle - \sigma,\ \langle t \rangle + \sigma], \quad P_\sigma \approxeq 0,683 \\ t \in [\langle t \rangle - 2\sigma,\ \langle t \rangle + 2\sigma], \quad P_{2\sigma} \approxeq 0,954 \\ t \in [\langle t \rangle - 3\sigma,\ \langle t \rangle + 3\sigma], \quad P_{3\sigma} \approxeq 0,997 \end{gathered}$$

  Из экспериментальной выборки объема N можно найти приближенные значения вероятностей, как отношения $\frac{\triangle N_\sigma}{N}$, $\frac{\triangle N_{2\sigma}}{N}$, $\frac{\triangle N_{3\sigma}}{N}$, где $\triangle N_\sigma$, $\triangle N_{2\sigma}$, $\triangle N_{3\sigma}$ - количество результатов измерений, попавших в интервалы:

$$\begin{gathered} [\langle t \rangle_N - \sigma_N,\ \langle t \rangle_N + \sigma_N] \\ [\langle t \rangle_N - 2\sigma_N,\ \langle t \rangle_N + 2\sigma_N] \\ [\langle t \rangle_N - 3\sigma_N,\ \langle t \rangle_N + 3\sigma_N] \end{gathered}$$

- ***Среднеквадратичное отклонение среднего значения:***

$$\sigma_{\langle t \rangle} = \sqrt{\frac{1}{N\left(N-1\right)}\sum_{i=1}^{n}\left(t_i - \langle t \rangle_N\right)^2}$$

- ***Доверительным*** называется ***интервал*** значений , в который попадает истинное значение измеряемой величины с заданной вероятностью.

  α - доверительная вероятность;

  $\triangle t = t_{\alpha,N}\sigma_{\langle t \rangle}$;

  $t_{\alpha,N}$ - табличное значение коэффициента Стьюдента;

$\sigma_{\langle t \rangle}$ - среднеквадратичное отклонение среднего значения;

$$\alpha = P(t \in [\langle t \rangle - \triangle t, \langle t \rangle + \triangle t])$$

Ход работы:

Мы выбрали устанавливаемый по часам промежуток времени в 5 секунд, после чего многократно устанавливая этот промежуток времени, провели 50 измерений. Результаты каждого измерения занесли в Табл. 1.

Далее мы выбрали из Табл. 1 наименьший t_{min} и наибольшим t_{max} из результатов измерений: $t_{min} = 4,20$, $t_{max} = 5,66$

После нахождения t_{min} и t_{max} мы разбили промежуток $[t_{min};\ t_{max}] = [4,20;\ 5,66] = [4,10;\ 5,85]$ на $m = 7$ равных интервалов $\triangle t = \frac{5,85-4,10}{7} = 0,25$, соблюдая условия, что:

1. m должно быть целым, близким к $\sqrt{N} = \sqrt{5}0 \approx 7,07$, N - полное число измерений,
2. Измеренные значения t_{min} и t_{max} должны попадать внутрь «крайних» интервалов;
3. Граничные значения, разделяющие соседние интервалы, должны быть по возможности «круглыми» числами - это облегчит построение гистограммы, соответственно мы решили взять интервал $[4,10;\ 5,85]$

Границы выбранных интервалов занесли в Табл. 2

Мы посчитали число результатов измерений $\triangle N_i$, попавших в каждый из интервалов $\triangle t = 0,25$ и внесли данные в Табл. 2;
Следующим этапом мы вычислили опытное значение плотности вероятности $\frac{\triangle N}{N \triangle t}$ и внесли данные вычислений в Табл. 2;

После мы построили гистограмму Рис. 1.

По данным Табл. 1 с помощью формул $\langle t \rangle_N = \frac{1}{N}(t_1 + t_2 + ... + t_N) = \frac{1}{N}\sum_{i=1}^{n} t_i$ и $\sigma_N = \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}(t_i - \langle t \rangle_N)^2}$ мы вычислили выборочное значение среднего $\langle t \rangle_N = 5,00c$ и выборочное среднеквадратичное отклонение $\sigma_N = 0,23c$.

Полученные значения мы внесли в самый низ Табл. 1.

Для контроля того, правильно ли мы вычислили $\langle t \rangle_N$, мы вычислили $\sum_{i=1}^{n}(t_i - \langle t \rangle_N) = 0,00$с.

Это значение мы также внесли в низ Табл. 1.

Следующим этапом мы вычислили максимальное значение плотности распределения ρ_{max}, соответствующее $t = \langle t \rangle$. Для этого мы использовали формулу $\rho_{max} = \frac{1}{\sigma\sqrt{2\pi}}$.

Так как в пределе величины $\langle t \rangle_N$ и σ_N стремятся к $\langle t \rangle$ и σ, соответственно, то формула приобретает вид $\rho_{max} = \frac{1}{\sigma_N\sqrt{2\pi}} = 1,73c^{-1}$

Это значение мы вносим в низ Табл. 1.

Далее мы нашли значения t, соответствующие серединам выбранных ранее интервалов из Табл. 2 использовав формулу нахождения среднего арифметического значения, полученные значения внесли в соответствующий столбец Табл. 2.

После нам потребовалось вычислить для полученных
t значения плотности распределения $\rho(t)$ по формуле $\rho(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\langle t \rangle)^2}{2\sigma^2}\right)$, где мы приняли параметры $\langle t \rangle_N$ и σ_N в качестве $\langle t \rangle$ и σ.

Внесем полученные значения в Табл. 2.

На гистограмму Рис. 1 мы нанесли все расчетные точки и провели через них плавную кривую.

Следующим этапом мы проверили, насколько точно выполняется в наших опытах соотношение между вероятностями:

$$
\begin{aligned}
t \in [\langle t \rangle - \sigma,\ \langle t \rangle + \sigma], \quad P_\sigma \approxeq 0,683 \\
t \in [\langle t \rangle - 2\sigma,\ \langle t \rangle + 2\sigma], \quad P_{2\sigma} \approxeq 0,954 \\
t \in [\langle t \rangle - 3\sigma,\ \langle t \rangle + 3\sigma], \quad P_{3\sigma} \approxeq 0,997
\end{aligned}
$$

и долями $\frac{\triangle N_\sigma}{N}$, $\frac{\triangle N_{2\sigma}}{N}$, $\frac{\triangle N_{3\sigma}}{N}$.

Для этого мы занесли в Табл. 3 вычисленные границы интервалов $\langle t \rangle_N \pm \sigma_N$, $\langle t \rangle_N \pm 2\sigma_N$, $\langle t \rangle_N \pm 3\sigma_N$.

Далее нам потребовалось по данным Табл. 1 подсчитать и занести в Табл. 3 количество $\triangle N$ измерений, попадающих в каждый из вычисленных ранее для Табл. 3 интервалов, и отношение $\frac{\triangle N}{N}$ этого количества к общему числу измерений.

Значения нормального распределения случайной величины $P_\sigma \approxeq 0,683$, $P_{2\sigma} \approxeq 0,954$, $P_{3\sigma} \approxeq 0,997$ мы также внесли в Табл. 3, где наглядно видно насколько точно выполняется в наших опытах соотношение между вероятностями.

Потом мы рассчитали ***среднеквадратичное отклонение среднего значения*** по формуле:

$$\sigma_{\langle t \rangle} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^{n} (t_i - \langle t \rangle_N)^2}$$

Получили:

$$\sigma_{\langle t \rangle} = \sqrt{\frac{1}{50(50-1)} 2,5756} = \sqrt{0,00105127} = 0,0324233\ c$$

Далее мы нашли табличное значение коэффициента Стьюдента $t_{\alpha,N}$ для доверительной вероятности $\alpha = 0,95$. Рис. 2.

Получилось $t_{\alpha,N} = 2,009575234489209$

Следом за этим, мы записали доверительный интервал для измеряемого в работе промежутка времени (5 секунд).

Для этого нам необходимо было сначала вычислить $\triangle t = t_{\alpha,N}\sigma_{\langle t \rangle} = 2,009575234489209 \times 0,0324233 = 0,0651570607 \approx 0,066$

Тогда $\alpha = P(t \in [\langle t \rangle - \triangle t, \langle t \rangle + \triangle t])$ имеет вид $0,95 = P(t \in [4,934,\ 5,066])$

Результаты:

Таблица 1: Результаты прямых измерений

№	$t_i,\ c$	$t_i - \langle t \rangle_N,\ c$	$(t_i - \langle t \rangle_N)^2,\ c^2$
1	5,08	0,08	0,0064
2	4,99	-0,01	0,0001
3	4,20	-0,80	0,6400
4	5,02	0,02	0,0004
5	4,78	-0,22	0,0484
6	5,40	0,40	0,1600
7	4,90	-0,10	0,0100
8	5,14	0,14	0,0196
9	5,01	0,01	0,0001
10	4,46	-0,54	0,2916
11	5,66	0,66	0,4356
12	5,00	0,00	0,0000
13	4,81	-0,19	0,0361
14	5,05	0,05	0,0025
15	4,93	-0,07	0,0049
16	4,76	-0,24	0,0576
17	4,93	-0,07	0,0049
18	5,07	0,07	0,0049
19	5,20	0,20	0,0400
20	5,00	0,00	0,0000
21	5,11	0,11	0,0121
22	4,81	-0,19	0,0361
23	5,03	0,03	0,0009
24	4,91	-0,09	0,0081
25	5,16	0,16	0,0256
26	4,98	-0,02	0,0004
27	5,08	0,08	0,0064
28	4,76	-0,24	0,0576
29	5,28	0,28	0,0784
30	5,05	0,05	0,0025
31	4,87	-0,13	0,0169

№	$t_i,\ c$	$t_i - \langle t \rangle_N,\ c$	$(t_i - \langle t \rangle_N)^2,\ c^2$
32	5,09	0,09	0,0081
33	4,83	-0,17	0,0289
34	5,28	0,28	0,0784
35	4,94	-0,06	0,0036
36	4,99	-0,01	0,0001
37	4,73	-0,27	0,0729
38	5,15	0,15	0,0225
39	4,98	-0,02	0,0004
40	5,13	0,13	0,0169
41	5,01	0,01	0,0001
42	4,85	-0,15	0,0225
43	5,36	0,36	0,1296
44	5,26	0,26	0,0676
45	5,11	0,11	0,0121
46	5,03	0,03	0,0009
47	5,04	0,04	0,0016
48	5,17	0,17	0,0289
49	4,98	-0,20	0,0400
50	4,82	-0,18	0,0324
	$\langle t \rangle_N = 5,00\ c$	$\sum_{i=1}^{n} (t_i - \langle t \rangle_N) = 0,00\text{с}$	$\sigma_N\ =\ 0,23\ c$ $\rho_{max}\ =\ 1,73\ \text{с}^{-1}$

Таблица 2: Данные для построения гистограммы

$\text{Границы интервалов},\ c$	$\triangle N$	$\frac{\triangle N}{N \triangle t},\ c^{-1}$	$t,\ c$	$\rho,\ c^{-1}$
от: 4,10 до: 4,35	1	0,08	4,225	0,00594
от: 4,35 до: 4,60	1	0,08	4,475	0,12823
от: 4,60 до: 4,85	9	0,72	4,725	0,84898

Границы интервалов, c	$\triangle N$	$\frac{\triangle N}{N \triangle t}$, c^{-1}	t, c	ρ, c^{-1}
от: 4,85 до: 5,10	25	2	4,975	1,72478
от: 5,10 до: 5,35	11	0,88	5,225	1,07520
от: 5,35 до: 5,60	2	0,16	5,475	0,20566
от: 5,60 до: 5,85	1	0,08	5,725	0,01158

Таблица 3: Стандартные доверительные интервалы

	Интервал, c от	Интервал, c до	$\triangle N$	$\frac{\triangle N}{N}$	P
$\langle t \rangle_N \pm \sigma_N$	4,77	5,23	44	0,88	0,685
$\langle t \rangle_N \pm 2\sigma_N$	4,54	5,46	47	0,94	0,954
$\langle t \rangle_N \pm 3\sigma_N$	4,31	5,69	49	0,98	0,997

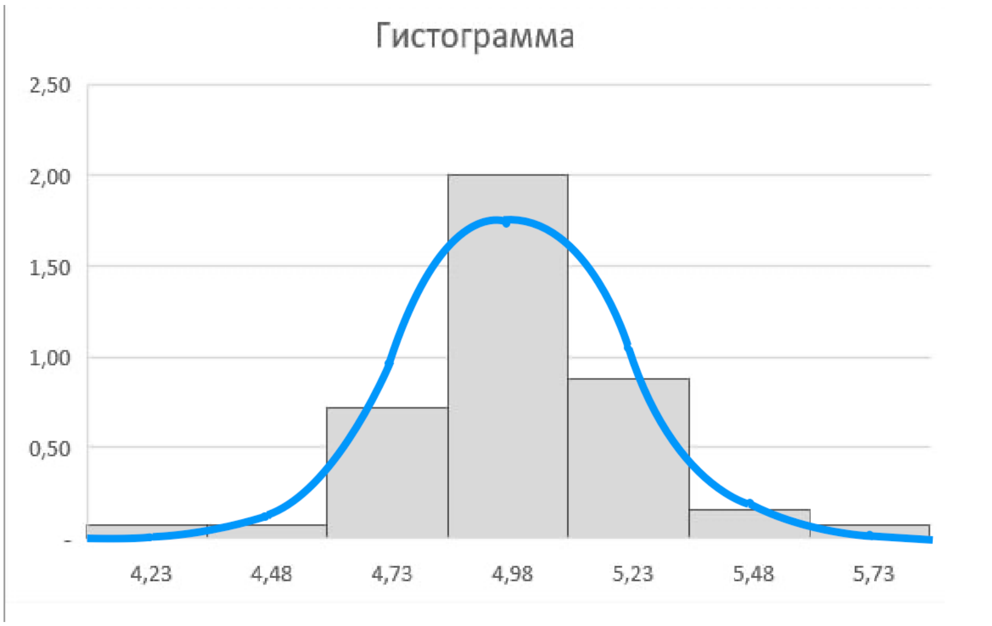

Рис. 1. Гистограмма и функция Гаусса.

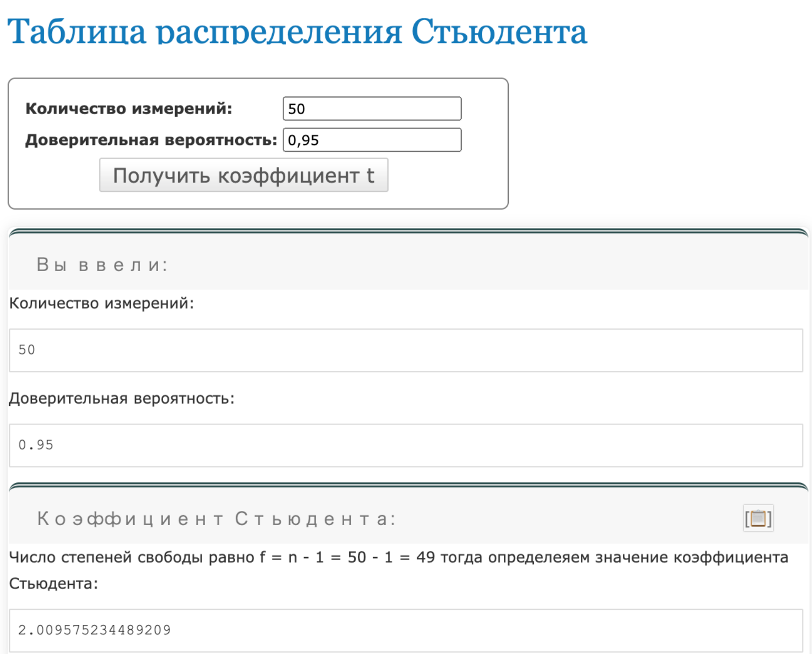

Рис. 2. Коэффициент Стьюдента для 50 измерений при доверительной вероятности 0,95

Выводы:

В данной лабораторной работе мы исследовали распределения случайной величины на примере многократных измерений определённого интервала времени, где мы замеряли этот интервал с часов с секундной стрелкой цифровым секундомером.

В ходе опытов соответствующие значения $\frac{\triangle N}{N}$ получились примерно равны соответствующим значениям плотности вероятности. Достаточно точно выполняется соотношение между вероятностями $P_\sigma \approxeq 0,683$, $P_{2\sigma} \approxeq 0,954$, $P_{3\sigma} \approxeq 0,997$ и долями $\frac{\triangle N_\sigma}{N}$, $\frac{\triangle N_{2\sigma}}{N}$, $\frac{\triangle N_{3\sigma}}{N}$.

По графику видно, что исследуемое распределение соответствует нормальному.