Лабораторная работа №1.01
"Исследование распределения случайной величины"

Группа М3203
Кравченкова Елизавета
Лакеев Георгий

Преподаватель
Хуснутдинова Наира Рустемовна

Физические основы компьютерных и сетевых технологий
Университет ИТМО
Санк-Петербург, Россия

23 сентября 2023 г.

Оглавление

Цели и Задачи

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

Задачи:

1. Провести многократные измерения интервала в 7 секунд.
2. По результатам полученных данных посторить гистограмму, отражающую распеределение этих данных.
3. Вычислить среднее значение и среднеквадратичное отклонение.
4. Построить график полученных данных и сравнить его с функцией Гаусса.
5. Проверить точность полученных данных со стандартными значениями вероятности и занести результаты в таблицу.

Установка

2.1 Объект исследования

Распределение случайной величины.

2.2 Метод экспериментального исследования

Многократное измерение промежутка времени в 7 секунд с помощью цифрового секундомера.

2.3 Измерительные приборы

Наименование	**Тип прибора**	**Используемый диапазон, с**	**Погрешность прибора, с**
Секундомер	Цифровой	[0,8]	0.001

Используются 2 одинаковых секундомера. Первый задает интервал времени, который многократно измеряется вторым секундомером.

Теория и рабочие формулы

Случайной называют величину, которая в результате испытания примет одно и только одно числовое значение, зависящее от случайных факторов и заранее непредсказуемое.

Из-за совместного действия многочисленных неконтролируемых причин результат отдельного измерения подвергается искажениям и становится величиной случайной. Часто принимается, что результаты многократных измерений описываются функцией Гаусса (2) - так называемым законом нормального распределения.

Функция Гаусса:

$$\rho(t) = \lim_{\substack{x \to \infty \\ \Delta t \to 0}} \frac{\Delta N}{N \Delta t} = \frac{1}{N} \frac{dN}{dt}, \tag{1}$$

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

Она зависит от таких параметров как $\langle t \rangle$ и σ (математическое ожидание и среднеквадратичное (стандартное) отклонение).

Строго говоря, 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 с помощью (3) и (4).
В пределе величины $\langle t \rangle_N$ и σ_N стремятся к $\langle t \rangle$ и σ, соответственно.

Среднее арифметическое всех результатов измерений:

$$\langle t \rangle_N = \frac{1}{N} \sum_{i=1}^{N} t_i \tag{3}$$

Выборочное среднеквадратичное отклонение:

$$\sigma_N = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (t_i - \langle t \rangle_N)^2} \tag{4}$$

Гистограмма - ступенчатая диаграмма, показывающая, как часто при измерениях появляются значения, попадающие в тот или иной из равных интервалов Δt, лежащих между t_{min} и наибольшим t_{max} из измеренных значений величины t. Гистограмму строят в следующих координатах: ось абсцисс — измеряемая величина t, ось ординат — значения $\frac{\Delta N}{N \Delta t}$.

Здесь N - полное количество измерений, ΔN - количество результатов, попавших в интервал $[t; t+\Delta t]$. Частное $\frac{\Delta N}{N}$ есть доля результатов, попавших в указанный интервал, и характеризует вероятность попадания в него результата отдельного измерения. Отношение этой величины к ширине интервала Δt характеризует некоторую «плотность вероятности».

Опытное значение плотности вероятности:

$$\rho(t) = \frac{\Delta N}{N \Delta t} \tag{5}$$

Из формулы (2) сразу видно, что плотность нормального распределения имеет максимум при $t = \langle t \rangle$ и симметрична относительно этого значения: следует ожидать, что примерно так же будет выглядеть гистограмма. Можно сравнить максимальную «высоту» гистограммы и ρ_{max}:

Максимальное значение плотности распределения:

$$\rho_{max} = \frac{1}{\sigma\sqrt{2\pi}} \tag{6}$$

Вероятность попадания результата каждого измерения t в интервал $[t_1; t_2]$ в стандартных(наиболее употребительных на практике) интервалах при условии реализации нормального распределения случайной величины имеет следующие значения:

$$[\langle t \rangle_N - \sigma_N; \langle t \rangle_N + \sigma_N], P_\sigma \cong 0,683 \tag{7.1}$$

$$[\langle t \rangle_N - 2\sigma_N; \langle t \rangle_N + 2\sigma_N], P_{2\sigma} \cong 0,954 \tag{7.2}$$

$$[\langle t \rangle_N - 3\sigma_N; \langle t \rangle_N + 3\sigma_N], P_{3\sigma} \cong 0,997 \tag{7.3}$$

Формулы для вычисления приближённой вероятности попадания каждого измерения t в интервал $[t_1; t_2]$:

$$[\langle t \rangle_N - \sigma_N; \langle t \rangle_N + \sigma_N], P_{\sigma_N} = \frac{\Delta N_\sigma}{N} \tag{8.1}$$

$$[\langle t \rangle_N - 2\sigma_N; \langle t \rangle_N + 2\sigma_N], P_{2\sigma_N} = \frac{\Delta N_{2\sigma}}{N} \tag{8.2}$$

$$[\langle t \rangle_N - 3\sigma_N; \langle t \rangle_N + 3\sigma_N], P_{3\sigma_N} = \frac{\Delta N_{3\sigma}}{N} \tag{8.3}$$

Среднеквадратичное отклонение среднего значения:

$$\sigma_{\langle t \rangle} = \sqrt{\frac{1}{N \cdot (N-1)} \sum_{i=1}^{N} (t_i - \langle t \rangle_N)^2} \tag{9}$$

Доверительный интервал для измеряемого в работе промежутка времени:

$$\Delta t = t_{\alpha,N} \cdot \sigma_{\langle t \rangle} \tag{10}$$

$$\alpha = P(t \in [\langle t \rangle - \Delta t, \langle t \rangle + \Delta t]) \tag{11}$$

Расчет результатов косвенных измерений (Ход работы)

4.1 Построим гистограмму, выполнив для этого следующие операции:

4.1.1

В ходе лабораторной работы требовалось замерить ровно 7 секунд с помощью двух секундомеров. Данные, полученные в ходе работы записаны в первом столбце Таблицы 1.
В результате анализа были найдены: $t_{min} = 6.801$ и $t_{max} = 7.951$.

4.1.2

Разобьем отрезок $[t_{min}, t_{max}]$ на m равных отрезков Δt, соблюдая условие $m \approx \sqrt{N}$.
$m \approx \sqrt{N} = \sqrt{51} \approx 7$.
Далее найдем размер одного интервала:

$$\Delta t = (t_{max} - t_{min})/m = (7.951 - 6.801)/7 \approx 0.17$$

Все получившиеся интервалы занесем в первый столбец Таблицы 2.

4.1.3

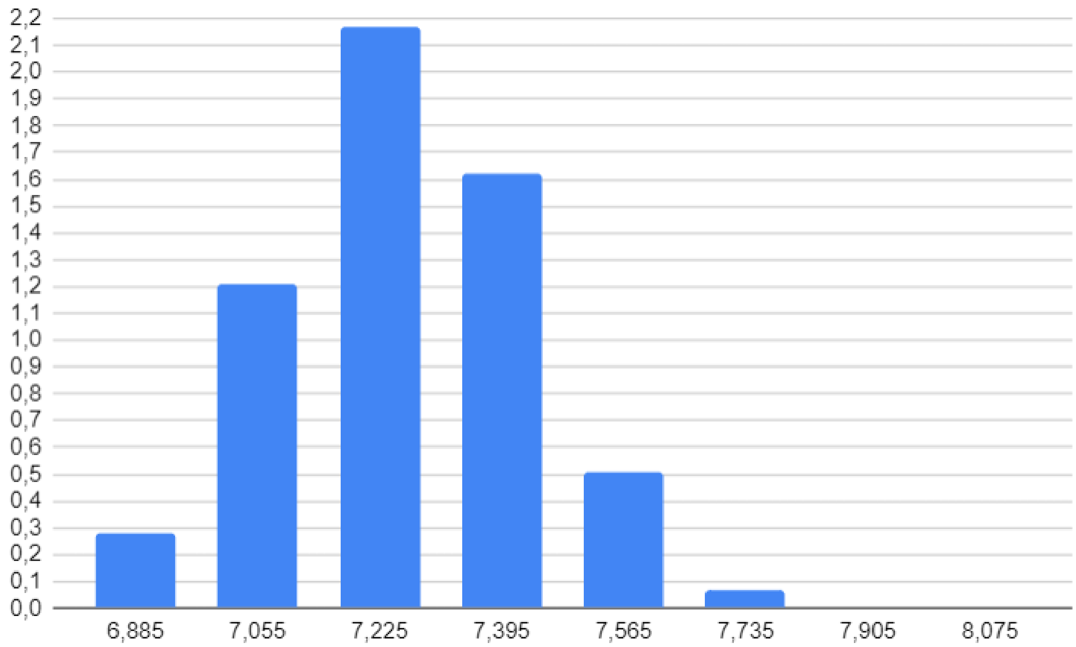
Для каждого интервала посчитаем количество значений из измерений ΔN_i,Таблицы 1, попадающих в этот отрезок. Соответствующие данные внесем во второй столбец Таблицы 2.

4.1.4

По указанной формуле (5) расчитаем опытные значения плотности вероятности (значения по оси ординат для гистограммы). Получившиеся значения внесем в третий столбец Таблицы 2.

4.1.5

По полученным данным построим гистрограмму.



4.2 Вычислим выборочное значение среднего и выборочное среднеквадратичное отклонение

По данным Таблицы 1 с помощью формул (3) (4) вычислилим выборочное значение среднего и выборочное среднеквадратичное отклонение .

$$\langle t \rangle_N = \frac{1}{N}\sum_{i=1}^{N} t_i = 7,253$$

$$\sigma_N = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(t_i - \langle t \rangle_N)^2} = 0,1817489887$$

4.3 Проверим $\langle t \rangle_N$.

Стоит проверить правильность подсчета $\langle t \rangle_N$ с помощью вычисления

$$\sum_{i=1}^{N}(t_i - \langle t \rangle_N) = 0$$

Это можно объяснить тем, что

$$\sum_{i=1}^{N} t_i - \langle t \rangle_N = \sum_{i=1}^{N} t_i - \sum_{i=1}^{N} \langle t \rangle_N = \sum_{i=1}^{N} t_i - N \cdot \langle t \rangle_N = \sum_{i=1}^{N} t_i - N \cdot (\frac{1}{N} \sum_{i=1}^{N} t_i) = 0$$

Запишем результаты в «подвал» Таблицы 1.

4.4 Вычислим максимальное значение плотности распределения ρ_{max}

Вычисление максимального значения плотности по формуле (6):

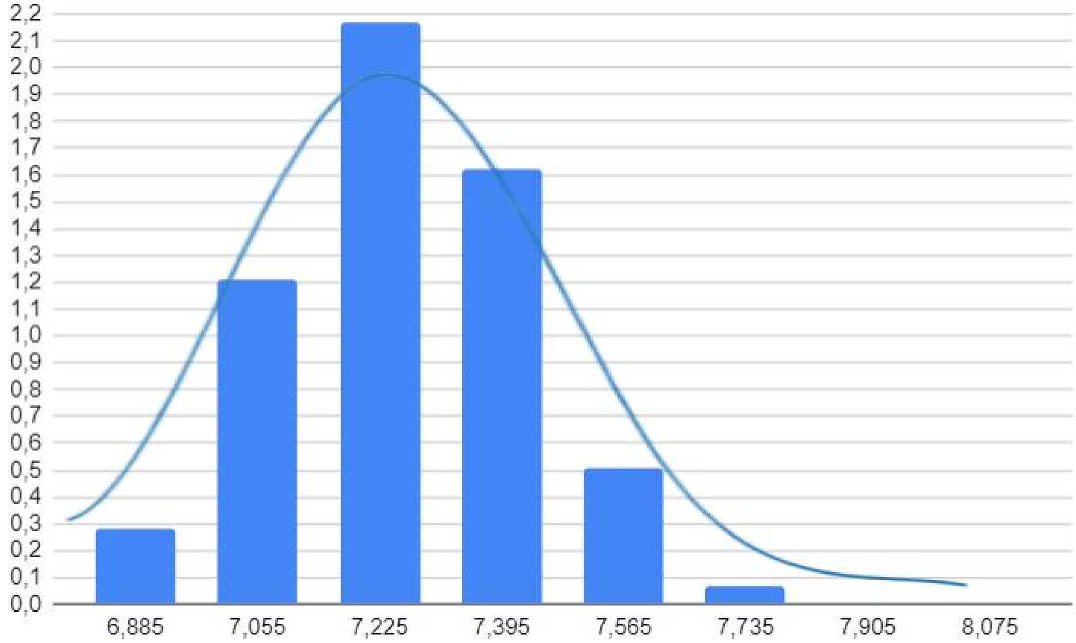
$$\rho_{max} = 2,195017883$$

Запишем результат в «подвал» Таблицы 1.

4.5 Найдем середины интервалов t, а также значения плотности распределения $\rho(t)$. Построение функции плотности распределения

Найдем значения t соответствующие серединам выбранных ранее итервалов, занесем в четвертый столбец Таблицы 2.
Для этих значений, используя параметры $\langle t \rangle_N$ и σ_N в качестве $\langle t \rangle$ и σ, вычислилим по формуле значения плотности распределения $\rho(t)$ по формуле (2).

Занесем значения $\rho(t)$ в пятый столбец Таблицы 2.
Нанесем все расчетные точки на график, на котором изображена гистограмма, и проведем через них плавную кривую.



4.6 Проверим, насколько точно выполняется соотношение между стандарными значениями вероятности $P_{i\sigma}$ и долями $\frac{\Delta N_\sigma}{N}$,$\frac{\Delta N_{2\sigma}}{N}$, $\frac{\Delta N_{3\sigma}}{N}$.

Для найденных ранее значений $\langle t\rangle_N$ и σ_N вычислим границы интервалов (7.1)-(7.3):

$$[\langle t\rangle_N - \sigma_N; \langle t\rangle_N + \sigma_N] : [7,253 - 0,1817; 7,253 + 0,1817], [7,071; 7,435]$$

$$[\langle t\rangle_N - 2\sigma_N; \langle t\rangle_N + 2\sigma_N] : [7,253 - 0,3634; 7,253 + 0,3634], [6,890; 7,617]$$

$$[\langle t\rangle_N - 3\sigma_N; \langle t\rangle_N + 3\sigma_N] : [7,253 - 0,5451; 7,253 + 0,5451], [6,708; 7,798]$$

Подсчитаем $\Delta N_{i\sigma}$ - число попавших в i-ый интервал значений.
Занесем эти данные в Таблицу 3.
Подчитаем $\frac{\Delta N_\sigma}{N}$,$\frac{\Delta N_{2\sigma}}{N}$, $\frac{\Delta N_{3\sigma}}{N}$. Занесем эти данные в Таблицу 3.

Можно заметить, что эти значения достаточно близки к стандарными значениями вероятности $P_{i\sigma}$. См последние два столбца Таблицы 3.

4.7 Расчитаем среднеквадратичное отклонение от среднего значения

Расчитаем среднеквадратичное отклонение от среднего значения по формуле (9):

$$\sigma_{\langle t \rangle} = \sqrt{\frac{1}{N \cdot (N-1)} \sum_{i=1}^{N} (t_i - \langle t \rangle_N)^2} = 0,02544994892$$

4.8 Найдем значение коэффициента Стьюдента $t_{\alpha,N}$. Запишем доверительный интервал для измеряемого в работе промежутка времени.

Найдем табличное значение коэффициента Стьюдента $t_{\alpha,N}$ для доверительной вероятности $\alpha = 0,95$ и N - числа измерений.
Из таблицы `http://old.exponenta.ru/educat/referat/XIkonkurs/student5/tabt-st.pdf` узнали, что

$$t_{0,95;51} = 2,0075838$$

Запишем доверительный интервал для измеряемого в работе промежутка времени по формулам (10)-(11).

$$\Delta t = 0,05109290516262$$

$$\alpha = P(t \in [7,2019; 7,3040])$$

Результаты

5.1 Таблица 1

№	t_i, с	$t_i - t_N$, с	$(t_i - t_N)^2$, с^2
1	7.351	0.351	0.123201
2	7.400	0.400	0.160000
3	7.352	0.352	0.123904
4	7.301	0.301	0.090601
5	7.300	0.300	0.090000
6	7.401	0.401	0.160801
7	7.401	0.401	0.160801
8	7.101	0.101	0.010201
9	7.050	0.050	0.002500
10	6.801	-0.199	0.039601
11	7.400	0.400	0.160000
12	7.250	0.250	0.062500
13	7.360	0.360	0.129600
14	7.050	0.050	0.002500
15	7.251	0.251	0.063001
16	7.400	0.400	0.160000
17	7.351	0.351	0.123201
18	7.301	0.301	0.090601
19	7.300	0.300	0.090000
20	7.160	0.160	0.025600
21	7.478	0.478	0.228484
22	7.300	0.300	0.090000
23	7.301	0.301	0.090601
24	7.050	0.050	0.002500
25	7.351	0.351	0.123201
26	7.151	0.151	0.022801
27	7.301	0.301	0.090601
28	7.351	0.351	0.123201
29	7.351	0.351	0.123201
30	7.150	0.150	0.022500
31	7.300	0.300	0.090000
32	7.407	0.407	0.165649
33	7.251	0.251	0.063001
34	7.301	0.301	0.090601
35	6.951	-0.049	0.002401
36	7.051	0.051	0.002601
37	7.100	0.100	0.010000
38	6.951	-0.049	0.002401
39	7.362	0.362	0.131044
40	7.351	0.351	0.123201
41	7.401	0.401	0.160801
42	7.101	0.101	0.010201
43	7.200	0.200	0.040000
44	7.951	0.951	0.904401
45	7.356	0.356	0.126736
46	7.101	0.101	0.010201
47	7.051	0.051	0.002601
48	7.310	0.310	0.096100
49	7.300	0.300	0.090000
50	6.950	-0.050	0.002500
51	7.100	0.100	0.010000
	$\langle t \rangle_N = 7,253$ с	$\sum\limits_{i=1}^{N} (t_i - \langle t \rangle_N) = 0$ с	$\sigma_N = 0,1817$ с $\rho_{max} = 2,1950\ c^{-1}$

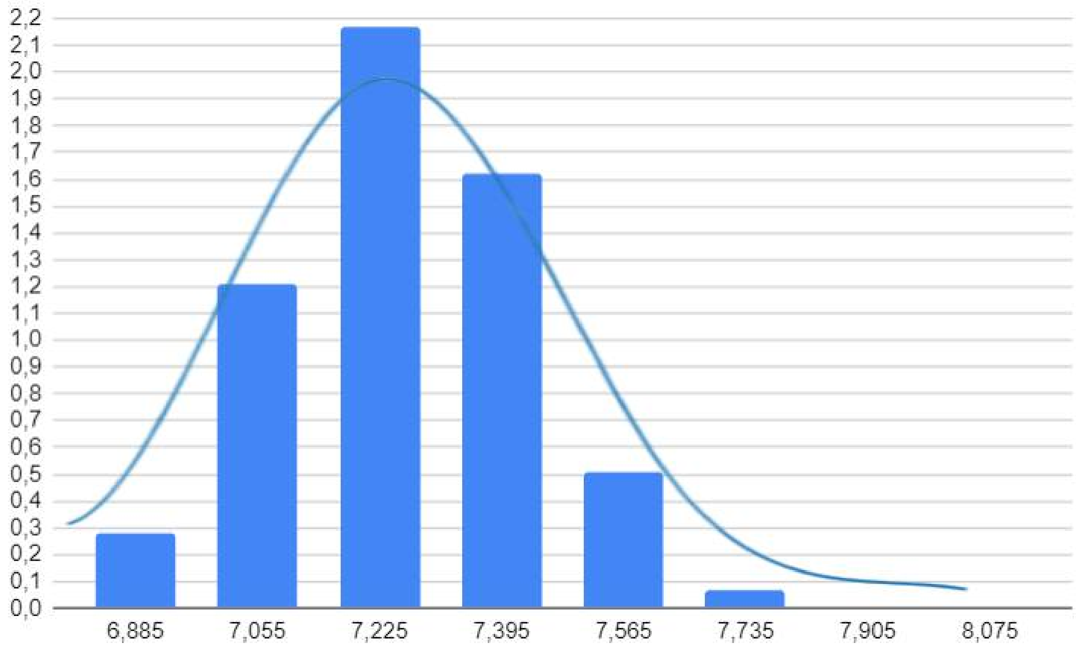
5.2 Таблица 2

Гр. интервалов, с	ΔN	$\frac{\Delta N}{N \Delta t}$, $с^{-1}$	t, с	ρ , $с^{-1}$
6,8-6,97	4	0,461361015	6,885	0,2821248441
6,97-7,14	10	1,153402537	7,055	1,211473639
7,14-7,31	19	2,191464821	7,225	2,168834283
7,31-7,48	17	1,960784314	7,395	1,618745356
7,48-7,65	0	0	7,565	0,5036982779
7,65-7,82	0	0	7,735	0,06534346515
7,82-7,99	1	0,1153402537	7,905	0,003534057581

5.3 Таблица 3

Начало, с	Конец интервала, с	ΔN	$\frac{\Delta N}{N}$	P
7,071	7,435	40	0,7843137255	0,683
6,890	7,617	49	0,9607843137	0,954
6,708	7,798	50	0,9803921569	0,997

5.4 Графики



2,2
2,1
2,0
1,9
1,8
1,7
1,6
1,5
1,4
1,3
1,2
1,1
1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0,0
6,885
7,055
7,225
7,395
7,565
7,735
7,905
8,075

Выводы

В ходе лабораторной работы была получена выборка для дискретной случайной величины, был исследован закон распределения случайной величины, для полученной выборки были вычислены доверительный интервал, среднее квадратичное отклонение, максимальное значение плотности распределения, среднеквадратичное отклонение среднего значения. Был построен график, на котором изображены гистограмма и функция плотности распределения. Полученные результаты близки к эталлоным (имеется в виду Гауссова функция и стандарные значения вероятности $P_{i\sigma}$)

Ответы на контрольные вопросы

7.1 Являются ли по вашему мнению, случайными следующие физические величины?

1. **Плотность алмаза при $20°$?** Неслучайная, так как при заданной температуре молекулы твердых веществ не изменяются, а следовательно и плотность неизменна.
2. **Напряжение сети?** Можно ответить и да, и нет. В идеальных условиях, когда цепь не изменяется, напряжение постоянно, но в повседневной жизни, когда каждый раз изменяется количество проводников в цепи, конечно случайно.
3. **Сопротивление резистора, взятого наугад из партии с одним и тем же номинальным сопротивлением?** С одной стороны, если все резисторы идеальные, то тогда неслучайное, но всегда можно допустить вероятность брака => случайная.
4. **Число молекул в $1cm^3$ при нормальных условиях.** Существует физическая (химическая) константа, описывающая это количество - число Ломшмидта. Но в тех же газах, в результате броуновского движения, это число может меняться в достаточно большом диапазоне.
5. **Примеры случайных величин:** скорость молекулы газа; температура атмосферного воздуха, дальность полета снаряда; продолжительность человеческой жизни.
6. **Примеры неслучайных величин:** скорость света, скорость звука, бросание кубика с одинаковыми гранями, заряд электрона.

Таким образом, можно прийти к выводу, что многие величины можно считать случайными, будет отличаться только дисперсия. Т.е у одних величин это может возникать в результате погрешности прибора, другие случайны по своей сути.

7.2 Изучая распределение ЭДС партии электрических батареек, студент использовал цифровой вольтметр. После нескольких измерений получились такие результаты (в вольтах): 1,50; 1,49; 1,50; 1,50; 1,49. Имеет ли смысл продолжать измерения? Что бы вы изменили в методике этого эксперимента?

Продолжать эксперимент стоит лишь в том случаи, если нужна большая точность вычислений. Заметим, что результаты измерений так и не стабилизировались, т.е каждый раз мы получаем одно из двух значений. Возможно, проблема заключается в использованном оборудовании, т.е погрешность прибора ≈ 0.1. Тогда для получения более точных данных, необходим другой вольтметр и большее количество измерений.

7.3 При обработке результатов измерений емкости партии конденсаторов получено: $\langle C \rangle = 1{,}1$ мкФ, $\sigma = 0{,}1$ мкФ. Если взять коробку со 100 конденсаторами из этой партии, то сколько среди них можно ожидать конденсаторов с емкостью меньше 1 мкФ? больше 1,3 мкФ?

Как мы знаем при увеличенни числа эксперименов $\sigma_N \to \sigma$ и $\langle C \rangle_N \to \langle C \rangle$. Тогда мы можем воспользоваться стандартными интервалами - эта вероятность при условии реализации нормального распределения:

$$t \in [\langle C \rangle - \sigma, \langle C \rangle + \sigma], P_\sigma \approx 0.683$$

$$t \in [\langle C \rangle - 2\sigma, \langle C \rangle + 2\sigma], P_\sigma \approx 0.954$$

Таким образом, в отрезок от $[1; 1.2]$ попадут 68.3% всех измерений. Из оставшихся 31,7% только 15.85% попадают в луч от $(-\infty; 1)$. Т.е не более 15 конденсаторов с емкостью меньше 1мкФ. Теперь рассмотрим второй отрезок $[0.9; 1.3]$. В него попадают 95.4% значений. Значит в луч $(1.3; +\infty)$ попадает половина от оставшихся, т.е 2.3%. Т.е не более 2 конденсаторов с емкостью больше 1.3мкФ.

7.4 Как изменяется коэффициент Стьюдента при возростании количества измерений?

С ростом числа отсчетов N коэффициент Стьюдента уменьшается, а значит доверительный интервал сужается, так как возрастает точность измерений.

7.5 Как зависит коэффициент Стьюдента от доверительной вероятности?

С ростом доверительной вероятности, то есть надежности значения абсолютной погрешности, коэффициент Стьюдента увеличивается. Ведь чем больше вероятность, тем большее отклонение от среднего значения мы можем позволить себе.

7.6 В чем отличие среднеквадратичного отколения среднего значения от среднеквадратичного отклонения выборки?

Дело в том, что среднеквадратичное (стандартное) отклонение σ - истинное, оно получается при бесконечно большом количестве измерений N. В подсчете среднеквадратичного отклонения выборки σ_N используются конечное число N - количество проведенных измерений. среднеквадратичное (стандартное) отклонение σ более точное, в пределе σ_N стремится к σ.

7.7 Обязательно ли в данной работе должно получиться распределение, близкое к нормальному? Почему?

В этой работе должно получится распределени близкое к нормальному, так как скорость реакции человека - это не постоянная величина, а изменяющаяся во времени. Это вероятность распределенна равномерна и максимальное количество измерений будет скапливаться вокруг значения, которые мы измеряли. Из этого следует, что распределение данной величины будет близко к нормальному.