Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования «Московский государственный технический университет имени Н. Э. Баумана
(национальный исследовательский университет)» (МГТУ им. Н. Э. Баумана)
Факультет «Робототехника и комплексная автоматизация»
Кафедра «Системы автоматизированного проектирования»

Домашнее задание
по дисциплине

«Теория вероятности и математическая статистика»

Вариант 6

Выполнил: студент группы РК6-32Б Журавлев Н. В.

Проверил: Берчун Ю. В.

Москва
2020

Типовой расчёт №4

Задание 1.

1.1) Постройте свой генератор с параметрами a = R1, c = G1, X0 = B1, m = 100 (здесь и далее числовые значения берутся из таблиц исходных данных к первому домашнему заданию). Составьте таблицу элементов последовательности до первого повторения, определите период генератора.

1.2) Постройте свой генератор с рационально выбранными параметрами a и c (согласно таблицам ниже), X0 = B1, m = 100. Составьте таблицу элементов последовательности до первого повторения, убедитесь в достижении максимального периода генератора.

Группа	Параметр a
РК6-32Б	41

Вариант	Параметр c
6.	57

1.3) Возьмите первые n = 50 значений из ранее полученной таблицы. Разбейте отрезок [0;99] на r = 10 равных частей [0;9], [10;19], …, [90;99]. Определите число элементов усечённой последовательности ni, попавших в соответствующий диапазон и постройте гистограмму.

1.4) Рассчитаем значение коэффициента χ^2 по n = 50 точкам:

$$\chi_n^2 = \frac{\sum_{i=1}^{r}(n_i - n \cdot p_i)^2}{n \cdot p_i}$$

где pi – вероятность попадания случайной величины в соответствующий диапазон (численно соответствует площади под графиком плотности распределения для рассматриваемого диапазона).

Для равномерного распределения $p_i = const = \frac{1}{r} = 0{,}1$, и поэтому в рассматриваемой задаче $n \cdot p_i = 5$.

1.5) Требуется определить такое значение уровня значимости, с которым можно принять гипотезу о том, что статистическая выборка соответствует равномерному распределению. Полученный уровень значимости можно будет рассматривать как характеристику качества работы генератора случайных чисел, с помощью которого была получена статистическая выборка.

Таблицы критических значений распределения χ_n^2 в часто ограничены представлением уровней значимости, близкими к 0 или к 1. Поэтому в рамках решаемой задачи рекомендуется пользоваться расширенным вариантом этой таблицы, в котором представлены и промежуточные значения (приводится ниже).

1.6) Требуется рассчитать выборочные характеристики (выборочное среднее, смещённую и исправленную оценки выборочной дисперсии) для n = 5, 10, 25 и 50 и сравнить их с соответствующими характеристиками теоретического равномерного распределения (математическим ожиданием и дисперсией). Результаты свести в таблицу, с указанием величины отклонений от теоретических значений.

2.1) Требуется провести 100 экспериментов, меняя значение rnd. Результаты моделирования оформляются в виде таблицы, в которой предусматриваются следующие столбцы:

- коэффициент загрузки первого кассира;
- коэффициент загрузки второго кассира;
- средняя длина первой очереди;
- средняя длина второй очереди.

Рассчитайте выборочные средние и исправленные выборочные оценки дисперсии для каждой собранной характеристики при n = 10, 25, 50, 100.

На основе полученных выборок для n = 100 построить гистограммы. Ширину интервалов выбирать не более половины исправленной оценки среднеквадратичного отклонения соответствующей величины. При попадании в крайние интервалы менее 5 значений объединять их с соседними.

2.2) Для каждой пары собранных характеристик рассчитайте выборочные ковариации и коэффициенты корреляции (для значений n = 10, 25, 50, 100).

2.3) Для тех же значений n = 10, 25, 60 требуется рассчитать доверительные интервалы для математических ожиданий каждой из собранных характеристик с уровнями значимости α = 0,1 и 0,01 (для двусторонней симметричной области).

R1	G1	B1
6	8	11

Решение

1.1)

Генератор строится по формуле:

$$X_i = (a \cdot X_{i-1} + c) \; mod \; m$$

Используя excel построим таблицу и найти период генератора:

i	Xi		
0	11		
1	74		
2	52		
3	20		
4	28		
5	76		
6	64		
7	92		
8	60		
9	68		
10	16		
11	4		**Период генератора**
12	32	=>	27
13	0		
14	8		
15	56		
16	44		
17	72		
18	40		
19	48		
20	96		
21	84		
22	12		
23	80		
24	88		
25	36		
26	24		
27	52		

1.2)

Выбрав рационально параметры, определим период нового генератора:

i	xi
0	11
1	8
2	85
3	42
4	79
5	96
6	93
7	70
8	27
9	64
10	81
11	78
12	55
13	12
14	49
15	66
16	63
17	40
18	97
19	34
20	51
21	48
22	25
23	82
24	19
25	36
26	33
27	10
28	67
29	4
30	21
31	18
32	95
33	52
34	89
35	6
36	3
37	80
38	37
39	74
40	91
41	88
42	65
43	22
44	59
45	76
46	73
47	50
48	7
49	44
50	61

=>

Период генератора
100

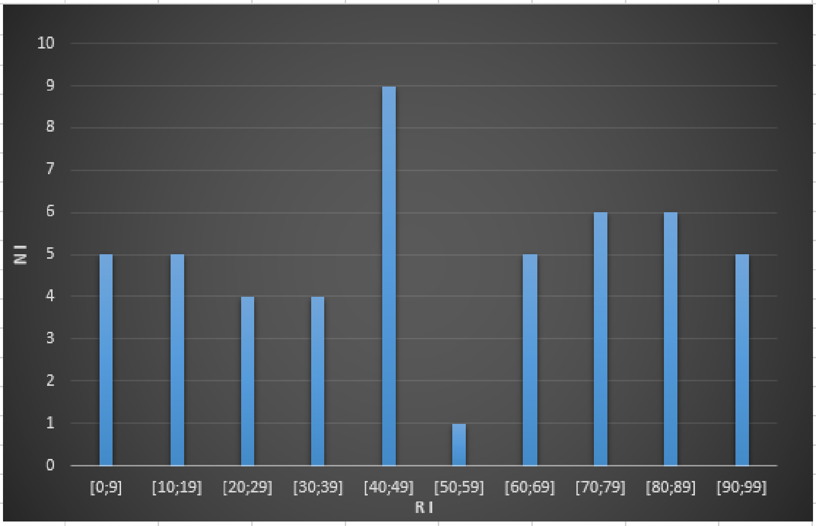
50	61
51	58
52	35
53	92
54	29
55	46
56	43
57	20
58	77
59	14
60	31
61	28
62	5
63	62
64	99
65	16
66	13
67	90
68	47
69	84
70	1
71	98
72	75
73	32
74	69
75	86
76	83
77	60
78	17
79	54
80	71
81	68
82	45
83	2
84	39
85	56
86	53
87	30
88	87
89	24
90	41
91	38
92	15
93	72
94	9
95	26
96	23
97	0
98	57
99	94
100	11

Как мы видим увелся период генератора и стал равный 100.

1.3)

Возьмем первые n = 50 значений из ранее полученной таблицы. Разобьем отрезок [0;99] на r = 10 равных частей [0;9], [10;19], …, [90;99]. Определим число элементов усечённой последовательности n_i, попавших в соответствующий диапазон и построим гистограмму:

i	r i	ni	Σ
1	[0;9]	5	50
2	[10;19]	5	
3	[20;29]	4	
4	[30;39]	4	
5	[40;49]	9	
6	[50;59]	1	
7	[60;69]	5	
8	[70;79]	6	
9	[80;89]	6	
10	[90;99]	5	



1.4)

Если бы мы имели дело с идеальным генератором, то на каждый отрезок попало бы ровно по 5 значений. В нашем же случае высота столбиков разная, значит нужно оценить, насколько критичны эти отклонения.

Воспользуемся методом проверки статистических гипотез на основе критерия Пирсона:

Численно оценим совокупную величину отклонения элементов выборки от теоретически ожидаемых результатов, для этого рассчитаем значение коэффициента X_n по n = 50 точкам:

$$X_n^2 = \frac{\sum_{i=1}^{r}(n_i - n \cdot p_i)^2}{n \cdot p_i}$$

Для равномерного распределения $p_i = const = \frac{1}{r} = 0{,}1$, и поэтому в рассматриваемой задаче $n \cdot p_i = 5$.

Получаем результат: $X_n = 7{,}2$

1.5)

Определим такое значение уровня значимости, с которым можно принять гипотезу о том, что статистическая выборка соответствует равномерному распределению.

Посмотрев по таблице получаем: $\alpha \approx 0{,}7$

1.6)

Рассчитаем выборочные характеристики (выборочное среднее, смещённую и исправленную оценки выборочной дисперсии) для n = 5, 10, 25 и 50:

Выборочное среднее определяется как среднее арифметическое элементов выборки:

$$\underline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Выборочная дисперсия:

$$\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Исправленная оценка выборочной дисперсии:

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \underline{x}\right)^2$$

Математическое ожидание равномерного распределения:

$$M = \frac{a+b}{2}$$

Дисперсия равномерного распределения:

$$D = \frac{(b-a)^2}{12}$$

Сравним их с характеристиками теоретического равномерного распределения и запишем результаты в таблицу:

n	x̄	σ^2	S^2	M	D	|x̄-M|	|σ^2-D|	|S^2-D|
5	45	1058	1322,5	49,5	816,75	4,5	241,25	505,75
10	57,5	994,25	1104,722	49,5	816,75	8	177,5	287,9722
25	55	748	779,1667	49,5	816,75	5,5	68,75	37,58333
50	51,5	844,25	861,4796	49,5	816,75	2	27,5	44,72959

2.1)

Проведём 100 экспериментов, рассматривая имитационную модель системы массового обслуживания на GPSS, меняя значение rnd. Результаты моделирования запишем в виде таблицы, со столбцами:

- Значение коэффициента rnd
- Коэффициент загрузки первого кассира;
- Коэффициент загрузки второго кассира;
- Средняя длина первой очереди;
- Средняя длина второй очереди.

rnd	UTIL. (1)	UTIL. (2)	AVE.CONT. (1)	AVE.CONT. (2)
1	0,781	0,577	0,201	0,058
2	0,81	0,555	0,219	0,103
3	0,793	0,532	0,166	0,049
4	0,798	0,535	0,173	0,041
5	0,794	0,525	0,193	0,051
6	0,825	0,574	0,242	0,083
7	0,758	0,526	0,167	0,058
8	0,789	0,565	0,23	0,082
9	0,78	0,531	0,199	0,068
10	0,749	0,533	0,203	0,087
11	0,809	0,62	0,331	0,197
12	0,8	0,577	0,244	0,155
13	0,784	0,521	0,193	0,094
14	0,775	0,555	0,16	0,046
15	0,787	0,568	0,197	0,058
16	0,794	0,581	0,219	0,065
17	0,795	0,564	0,213	0,079
18	0,802	0,577	0,214	0,075
19	0,765	0,556	0,197	0,064
20	0,749	0,555	0,165	0,055
21	0,743	0,487	0,124	0,031
22	0,79	0,595	0,281	0,106
23	0,809	0,571	0,237	0,098
24	0,786	0,59	0,274	0,138
25	0,795	0,601	0,192	0,075
26	0,793	0,577	0,268	0,116
27	0,783	0,57	0,172	0,05
28	0,781	0,56	0,2	0,074
29	0,73	0,523	0,133	0,033
30	0,782	0,58	0,251	0,099
31	0,806	0,574	0,211	0,065
32	0,801	0,567	0,283	0,093
33	0,758	0,479	0,145	0,026
34	0,802	0,642	0,323	0,137
35	0,799	0,562	0,201	0,057
36	0,795	0,592	0,282	0,113
37	0,779	0,576	0,164	0,04
38	0,772	0,546	0,147	0,034
39	0,793	0,572	0,207	0,081
40	0,77	0,522	0,174	0,03
41	0,771	0,559	0,173	0,048
42	0,769	0,557	0,14	0,02
43	0,784	0,561	0,234	0,076
44	0,766	0,523	0,185	0,067
45	0,802	0,574	0,328	0,214
46	0,826	0,614	0,22	0,083
47	0,743	0,525	0,129	0,032
48	0,778	0,569	0,246	0,113
49	0,794	0,594	0,233	0,085
50	0,793	0,573	0,216	0,077
50	0,793	0,573	0,216	0,077
51	0,765	0,524	0,158	0,037
52	0,812	0,597	0,227	0,072
53	0,808	0,548	0,172	0,03
54	0,748	0,532	0,147	0,042
55	0,757	0,568	0,24	0,09
56	0,783	0,569	0,222	0,066
57	0,784	0,569	0,221	0,056
58	0,748	0,559	0,176	0,064
59	0,765	0,548	0,183	0,056
60	0,799	0,588	0,223	0,077
61	0,779	0,546	0,152	0,034
62	0,801	0,591	0,2	0,056
63	0,764	0,581	0,182	0,057
64	0,768	0,593	0,198	0,071
65	0,76	0,555	0,161	0,031
66	0,803	0,567	0,195	0,082
67	0,767	0,546	0,183	0,072
68	0,802	0,57	0,218	0,103
69	0,82	0,601	0,263	0,095
70	0,778	0,518	0,145	0,038
71	0,76	0,582	0,151	0,038
72	0,795	0,587	0,252	0,111
73	0,765	0,575	0,192	0,057
74	0,792	0,562	0,207	0,069
75	0,771	0,582	0,192	0,066
76	0,808	0,618	0,228	0,099
77	0,79	0,598	0,214	0,07
78	0,782	0,546	0,218	0,097
79	0,799	0,587	0,235	0,108
80	0,802	0,573	0,221	0,069
81	0,767	0,569	0,175	0,061
82	0,776	0,566	0,185	0,054
83	0,803	0,58	0,181	0,046
84	0,791	0,589	0,244	0,094
85	0,77	0,574	0,172	0,057
86	0,773	0,579	0,153	0,057
87	0,77	0,551	0,176	0,047
88	0,773	0,541	0,156	0,053
89	0,792	0,603	0,23	0,077
90	0,802	0,553	0,228	0,089
91	0,796	0,557	0,211	0,058
92	0,787	0,578	0,227	0,076
93	0,749	0,503	0,124	0,036
94	0,782	0,588	0,16	0,048
95	0,793	0,59	0,221	0,062
96	0,792	0,571	0,218	0,068
97	0,766	0,542	0,238	0,085
98	0,799	0,618	0,242	0,106
99	0,799	0,606	0,253	0,088
100	0,79	0,557	0,195	0,048

Рассчитаем выборочные средние и исправленные выборочные оценки дисперсии для каждой собранной характеристики при n = 10, 25, 50, 100:

n= 10

UTIL. (1)

i	$\bar{x}$	$(xi-\bar{x})$^2	S^2
1	0,7877	4,489E-05	0,00051
2		0,00049729	
3		2,809E-05	
4		0,00010609	
5		3,969E-05	
6		0,00139129	
7		0,00088209	
8		1,69E-06	
9		5,929E-05	
10		0,00149769	

UTIL. (2)

i	$\bar{x}$	$(xi-\bar{x})$^2	S^2
1	0,5453	0,00100489	0,00041
2		9,409E-05	
3		0,00017689	
4		0,00010609	
5		0,00041209	
6		0,00082369	
7		0,00037249	
8		0,00038809	
9		0,00020449	
10		0,00015129	

AVE.CONT. (1)

i	$\bar{x}$	$(xi-\bar{x})$^2	S^2
1	0,1993	2,89E-06	0,00067
2		0,00038809	
3		0,00110889	
4		0,00069169	
5		3,969E-05	
6		0,00182329	
7		0,00104329	
8		0,00094249	
9		9E-08	
10		1,369E-05	

AVE.CONT. (2)

i	$\bar{x}$	$(xi-\bar{x})$^2	S^2
1	0,068	1E-04	0,0004
2		0,001225	

n= 25

UTIL. (1)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,7864	2,916E-05	0,000421
2		0,00055696	
3		4,356E-05	
4		0,00013456	
5		5,776E-05	
6		0,00148996	
7		0,00080656	
8		6,76E-06	
9		4,096E-05	
10		0,00139876	
11		0,00051076	
12		0,00018496	
13		5,76E-06	
14		0,00012996	
15		3,6E-07	
16		5,776E-05	
17		7,396E-05	
18		0,00024336	
19		0,00045796	
20		0,00139876	
21		0,00188356	
22		1,296E-05	
23		0,00051076	
24		1,6E-07	
25		7,396E-05	

AVE.CONT. (1)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,20936	6,98896E-05	0,00191
2		9,29296E-05	
3		0,00188009	
4		0,00132205	
5		0,00026765	
6		0,00106537	
7		0,00179437	
8		0,00042601	
9		0,00010733	
10		4,04496E-05	
11		0,01479629	
12		0,00119993	
13		0,00026765	
14		0,00243641	
15		0,00015277	
16		9,29296E-05	
17		1,32496E-05	
18		2,15296E-05	
19		0,00015277	
20		0,00196781	
21		0,00728633	
22		0,00513229	
23		0,00076397	
24		0,00417833	
25		0,00030137	

UTIL. (2)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,55884	0,000329786	0,000892

AVE.CONT. (2)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,08064	0,00051257	0,00142

n= 50

UTIL. (1)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,7846	1,3E-05	0,000419

AVE.CONT. (1)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,20998	8,06404E-05	0,002515

UTIL. (2)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,56124	0,000248	0,000964

AVE.CONT. (2)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2
1	0,07758	0,000383376	0,001616

n= 100

UTIL. (1)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2	S/2
1	0,78375	7,5625E-06	0,00037	0,009612

UTIL. (2)

i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2	S/2
1	0,56557	0,000131	0,000795	0,014101589

AVE.CONT. (1)

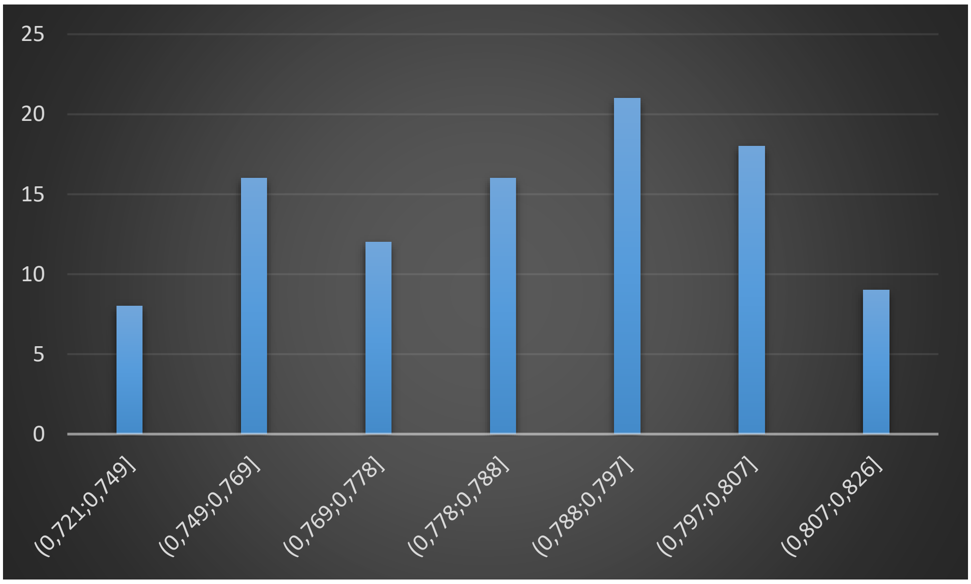
i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2	S/2
1	0,20464	1,32496E-05	0,001833	0,021408

AVE.CONT. (2)

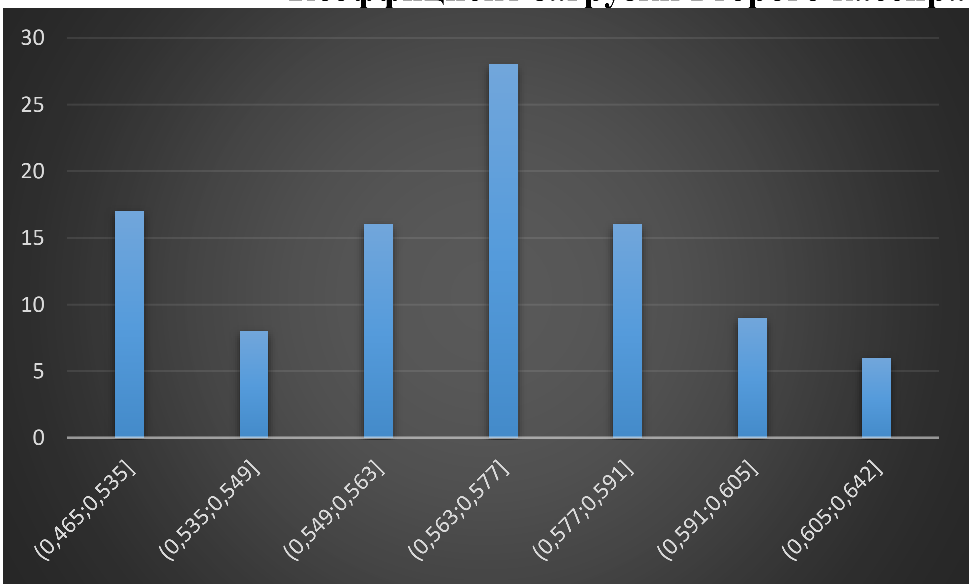
i	$\bar{x}$	(xi-$\bar{x}$)^2	S^2	S/2
1	0,07202	0,000197	0,001063	0,016305

На основе полученных выборок для n = 100 построим гистограммы. Ширину интервалов выберем не более половины исправленной оценки среднеквадратичного отклонения соответствующей величины. При попадании в крайние интервалы менее 5 значений объединим их с соседними:

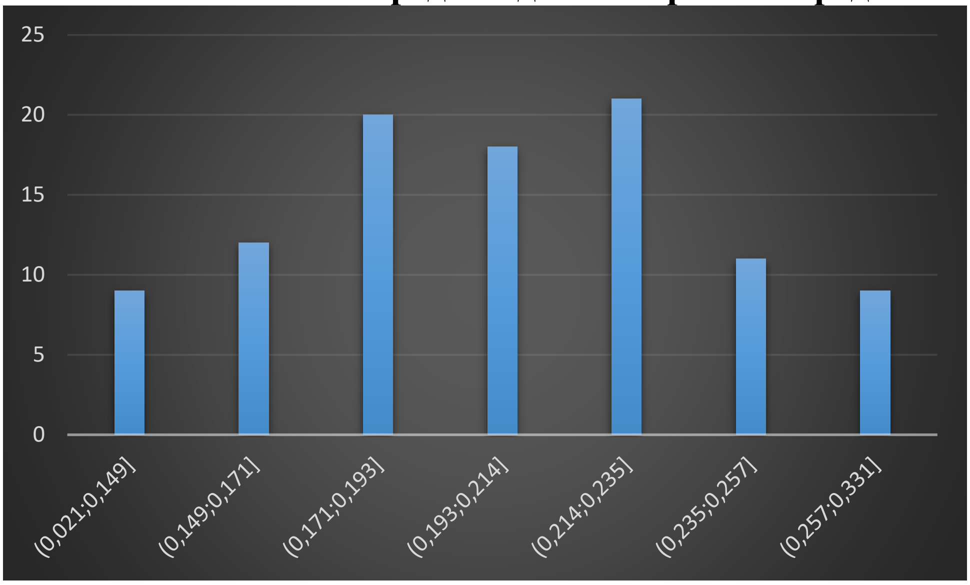
Коэффициент загрузки первого кассира



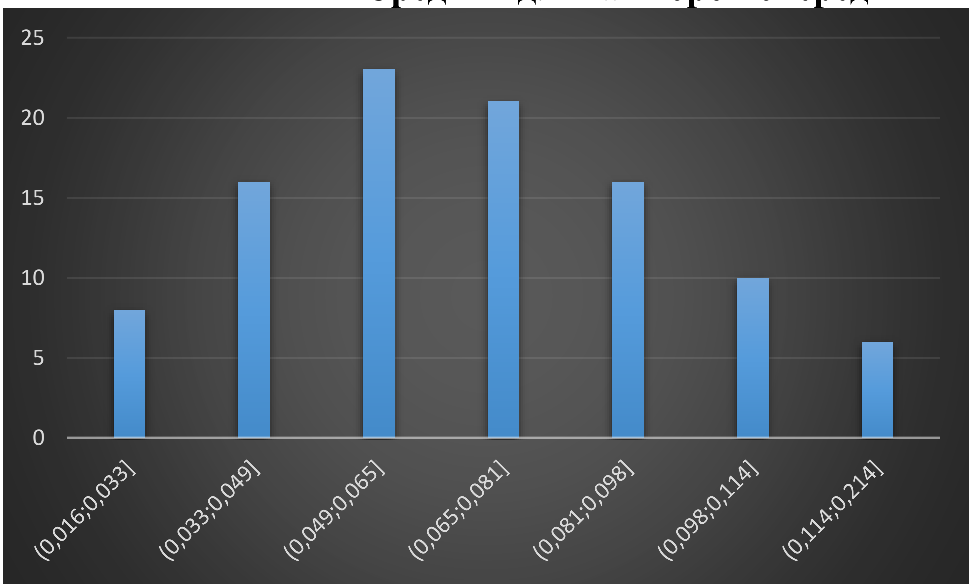
Коэффициент загрузки второго кассира



Средняя длина первой очереди



Средняя длина второй очереди



2.2)

Для каждой пары собранных характеристик рассчитаем выборочные ковариации и коэффициенты корреляции (для значений n = 10, 25, 50, 100).

Исправленная выборочная ковариация пары случайных величин:

$$cov_n(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

Исправленный выборочный коэффициент корреляции:

$$\rho(X,Y) = \frac{cov_n(X,Y)}{S_n(X)\cdot S_n(Y)}$$

В результате получаем:

n= 10

UTIL

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	-0,000212	0,000217	0,474369

AVE.COUNT.

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	-1,7E-05	0,000398	0,769627

n= 25

UTIL

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	-9,806E-05	0,000338	0,552109

AVE.COUNT.

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	0,00018927	0,001452	0,881927

n= 50

UTIL

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	-5,674E-05	0,000415	0,652617

AVE.COUNT.

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	0,000175828	0,00182	0,902434

n= 100

UTIL

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	-0,001587	0,00032	0,589694

AVE.COUNT.

i	(xi-x̅)(yi-y̅)	cov(X,Y)	ρ(X,Y)
1	5,10328E-05	0,001234	0,883983

2.3)

Для значений n = 10, 25, 60 рассчитаем доверительные интервалы для математических ожиданий каждой из собранных характеристик с уровнями значимости α = 0,1 и 0,01 (для двусторонней симметричной области):

$$\underline{x} - \frac{S_n}{\sqrt{n}} \cdot t_{n-1}(\alpha) < M < \underline{x} + \frac{S_n}{\sqrt{n}} \cdot t_{n-1}(\alpha)$$

$t_{n-1}(\alpha)$ мы ищем по таблице критических точек распределения Стьюдента.

α = 0,1

	UTIL. (1)	UTIL. (2)	AVE.CONT. (1)	AVE.CONT. (2)
n = 10	tn-1(α) 1,83	tn-1(α) 1,83	tn-1(α) 1,83	tn-1(α) 1,83
	0,774691 < M < 0,800709	0,53351 < M < 0,5571	0,184291 < M < 0,214309	0,0564486 < M < 0,079551
n = 25	tn-1(α) 1,71	tn-1(α) 1,71	tn-1(α) 1,71	tn-1(α) 1,71
	0,777602 < M < 0,791598	0,54863 < M < 0,5691	0,194415 < M < 0,224305	0,067751 < M < 0,093529
n = 60	tn-1(α) 1,67	tn-1(α) 1,67	tn-1(α) 1,67	tn-1(α) 1,67
	0,778779 < M < 0,787854	0,55467 < M < 0,5675	0,197503 < M < 0,218097	0,0662886 < M < 0,082678

α = 0,01

tn-1(α)	UTIL. (1)	UTIL. (2)	AVE.CONT. (1)	AVE.CONT. (2)
4,78	0,75372 < M < 0,82168	0,514511 < M < 0,5760893	0,160096 < M < 0,238504	0,037827 < M < 0,098173
3,74	0,769294 < M < 0,799906	0,536502 < M < 0,5811776	0,176673 < M < 0,242047	0,05245 < M < 0,10883
3,46	0,773915 < M < 0,792718	0,547806 < M < 0,5743271	0,186466 < M < 0,229134	0,057505 < M < 0,091462