1 Построить статистический ряд на основании собранной статистики

Совокупность всех наблюдений исследуемого процесса (в нашем примере продолжительности выполнения операции

называется

Простая статистическая совокупность

32	15	21	20	34	16	18	24	31	16	25	14	43	30	26
18	20	28	23	25	15	18	16	21	25	30	24	24	33	18
24	23	28	31	34	23	16	19	25	18	23	17	26	24	18
34	41	30	26	19	25	28	16	24	20	18	34	25	31	20
15	21	30	25	20	21	29	35	26	33	29	16	25	20	24
21	18	26	35	27	35	40	24	28	36	15	24	20	18	35
19	31	40	15	23	32	28	18	25	20	17	27	23	20	28
24	33	26	31	24	22	34	25	19	21	24	23	35	28	23

При большом количестве наблюдений пользоваться простой статистической совокупностью затруднительно

Поэтому она оформляется в

Статистический ряд

Диапазон наблюдений, разбитый на интервалы, оформленный в таблицу с указанием границ интервалов и частоты попадания наблюдений в каждый интервал

п/п	ΔI	Число наблюдений, m_i	Частость попадания (Доля интервалов от общего числа), $P_i = гр3 / n$
1	2	3	5
1	23-30	29	0,29
2	30-37	23	0,23
3	37-44	17	0,17
4	44-51	12	0,12
5	51-58	12	0,12
6	58-65	5	0,07
7	65-72	2	
Итого		100	1

Порядок построения статистического ряда

1 Выявить минимальное (I_{min}) и максимальное (I_{max}) значение операций.

44	39	34	28	43	35	68	30	28	34	23	70	31	25	43
45	48	28	38	41	29	45	28	25	31	39	44	48	23	62
35	25	29	41	49	63	53	58	50	40	62	51	33	28	30
25	43	54	33	58	28	31	49	56	34	25	44	51	28	55
64	28	39	35	43	45	27	34	25	47	37	58	64	63	29
36	30	29	49	56	40	30	34	26	34	28	45	49	28	58
51	65	32	29	38	43	40	55	34	33	31	64	27	36	44
55	28	36	37	44	29	36	58	61	29	45	55	40	69	34

2 Определить величину разряда по формуле

$$\Delta I = \frac{I_{max} - I_{min}}{1 + 3{,}2 \lg n} \quad \Rightarrow \quad \Delta I = \frac{70 - 23}{1 + 3{,}2 \lg 100} = 7$$

3 Определить границы каждого интервала (к Imin последовательно прибавлять ΔI).
Конец предыдущего интервала является началом последующего

п/п	ΔI
1	2
1	23-30
2	30-37
3	37-44
4	44-51
5	51-58
6	58-65
7	65-72
Итого	

4 Подсчитать число интервалов n_i, попадающих в каждый разряд. В каждом разряде должно быть не менее пяти наблюдений

п/п	ΔI	Число наблюдений, m_i
1	2	3
1	23-30	29
2	30-37	23
3	37-44	17
4	44-51	12
5	51-58	12
6	58-65	5
7	65-72	2
Итого		100

5 Определяется частность попадания анализируемой величины в каждом разряде

$$P_i^* = \frac{m_i}{n}.$$

п/п	ΔI	Число наблюдений, m_i	Частость попадания $P_i = гр3 / n$
1	2	3	5
1	23-30	29	0,29
2	30-37	23	0,23
3	37-44	17	0,17
4	44-51	12	0,12
5	51-58	12	0,12
6	58-65	5	0,07
7	65-72	2	
Итого		100	1

2 Рассчитать вероятностные числовые характеристики

1

1 Математическое ожидание случайной величины $M(x)$ – истинное среднее значение случайной величины, определяемое как сумма произведений всех возможных значений случайной величины на вероятности этих значений.

2

2 Дисперсия случайной величины $D(x)$ – числовая характеристика, определяющая степень рассеивания значений случайной величины относительно ее математического ожидания. Степень рассеивания определяется как разность между значениями случайной величины и ее математическим ожиданием, взятая в квадрате.

3

3 Среднеквадратическое отклонение σ_x характеризует степень рассеивания случайной величины относительно среднего значения и представляет собой квадратный корень из дисперсии.

4

4 Коэффициент вариации ν_x характеризует степень рассеивания значений непрерывной случайной величины относительно математического ожидания и представляет собой отношение среднеквадратического отклонения к математическому ожиданию случайной величины.

Порядок расчета вероятностных числовых характеристик наблюдаемой операции

1

Будем считать значение случайной величины в каждом разряде постоянным и равным среднему значению разряда $\bar{x}_i$:

$$\bar{x}_i = \frac{x_i + x_{i+1}}{2}.$$

п/п	ΔI	Число наблюдений, m_i	Частость попадания $P_i = гр3/n$	Среднее значение, $I_i^{ср}$
1	2	3	5	4
1	23-30	29	0,29	27
2	30-37	23	0,23	34
3	37-44	17	0,17	41
4	44-51	12	0,12	48
5	51-58	12	0,12	55
6	58-65	5	0,07	65
7	65-72	2		
Итого		100	1	

2

Рассчитать для каждого интервала произведение частости попадания на среднее значение интервала. Найти сумму полученных значений, которая будет соответствовать математическому ожиданию

$$M(I_i) = \sum I_i^{ср} * P_i^* = 39{,}07$$

п/п	ΔI	Число наблюдений, m_i	Частость попадания $P_i = гр3/n$	Среднее значение, $I_i^{ср}$	Математическое ожидание, $M(I) = гр4 \cdot гр5$
1	2	3	5	4	6
1	23-30	29	0,29	27	7,69
2	30-37	23	0,23	34	7,71
3	37-44	17	0,17	41	6,89
4	44-51	12	0,12	48	5,7
5	51-58	12	0,12	55	6,54
6	58-65	5	0,07	65	4,55
7	65-72	2			
Итого		100	1		39,07

3

Рассчитать для каждого интервала $(I\text{ср} - M(I))^2 * P_i$. Найти сумму полученных значений, которая будет соответствовать дисперсии случайной величины

$$D(I) = \sum (I\text{ср} - M(I))^2 * P_i = 137{,}47$$

п/п	ΔI	Число наблюдений, m_i	Частость попадания $P_i = гр3/n$	Среднее значение, $I_i^{ср}$	Математическое ожидание, $M(I) = гр4 \cdot гр5$	Дисперсия, $Д(I) = (гр4 - M(I_i))^2 \cdot гр5$
1	2	3	5	4	6	7
1	23-30	29	0,29	27	7,69	45,78
2	30-37	23	0,23	34	7,71	7,12
3	37-44	17	0,17	41	6,89	0,35
4	44-51	12	0,12	48	5,7	8,54
5	51-58	12	0,12	55	6,54	28,59
6	58-65	5	0,07	65	4,55	47,08
7	65-72	2				
Итого		100	1		39,07	137,47

4

Рассчитать квадратическое отклонение

$$\sigma = \sqrt{Д} = \sqrt{137{,}47} = 11{,}72$$

5

Определяется коэффициент вариации

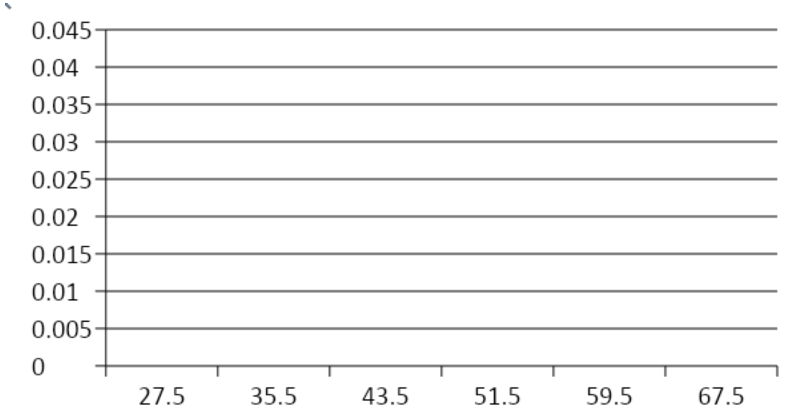
$$\upsilon = \frac{\sqrt{Д}}{\sum M(I)} = \frac{\sqrt{137{,}47}}{39{,}07} = 0{,}3$$

3 Построить гистограмму и статистическую функцию распределения

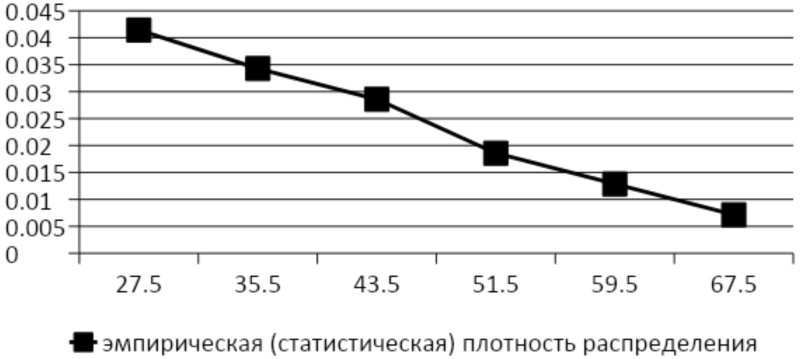
Гистограмма → **Графическое оформление статистического ряда**

Для построения гистограммы на оси абсцисс откладываются интервалы размером ΔI, на этих интервалах как на основании строятся прямоугольники, площадь которых равна P_i, с высотой $h_i = P_i / \Delta I$

п/п	ΔI	Число наблюдений, m_i	Частость попадания $P_i = гр3/n$	Среднее значение, $I_i^{ср}$
1	2	3	5	4
1	23-30	29	0,29	27
2	30-37	23	0,23	34
3	37-44	17	0,17	41
4	44-51	12	0,12	48
5	51-58	12	0,12	55
6	58-65	5	0,07	65
7	65-72	2		
Итого		100	1	



По средним значениям гистограммы строится кривая – которая называется статистической плотностью распределения



→ **Необходимость построения** →

Для наглядного представления на какую теоретическую кривую похожа статистическая плотность распределения, чтобы ее использовать для сглаживания собранных данных и для дальнейших расчетов

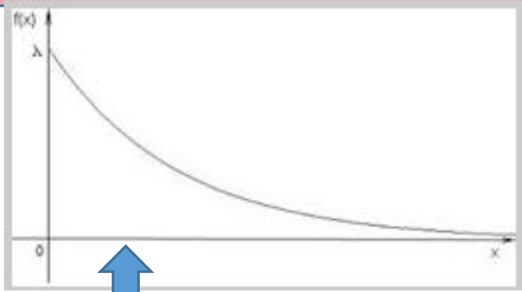
Теоретические законы распределения случайных величин

Необходимость сглаживания статистических данных с помощью теоретического закона

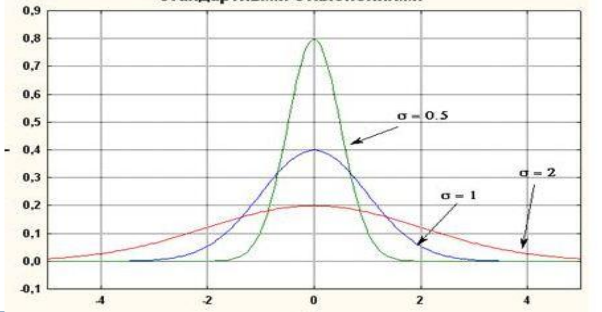
В любом статистическом распределении присутствует элемент случайности, связанный с тем, что число наблюдений ограничено. Поэтому при обработке статистического материала часто ставится задача подбора для данного статистического ряда теоретического закона распределения, выражающего лишь существенные черты распределения случайной величины, а не случайности, связанные с недостаточным объемом экспериментальных данных. Данная задача

Наиболее часто используемые теоретические законы распределения непрерывной случайной величины

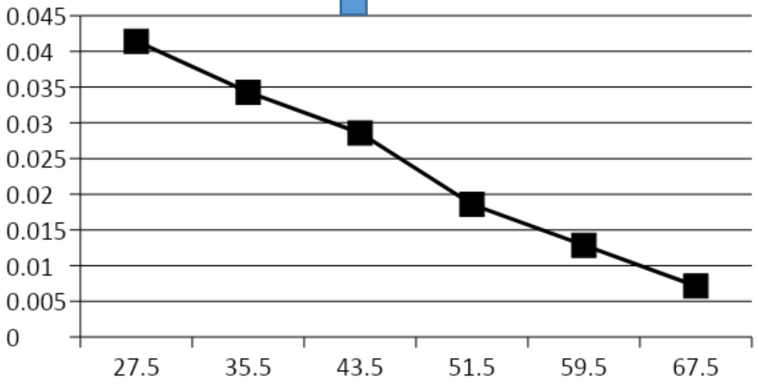
Экспоненциальный закон



Нормальный закон



Делаем вывод, что по очертаниям полученная статистическая плотность распределения соответствует экспоненциальному закону



$$f(t) = \lambda \cdot e^{-\lambda \cdot x} = \frac{1}{M(I)} \cdot e^{-\frac{x}{M(I)}}$$

Где λ=1/M(I) – среднее количество операций, выполняемых за единицу времени

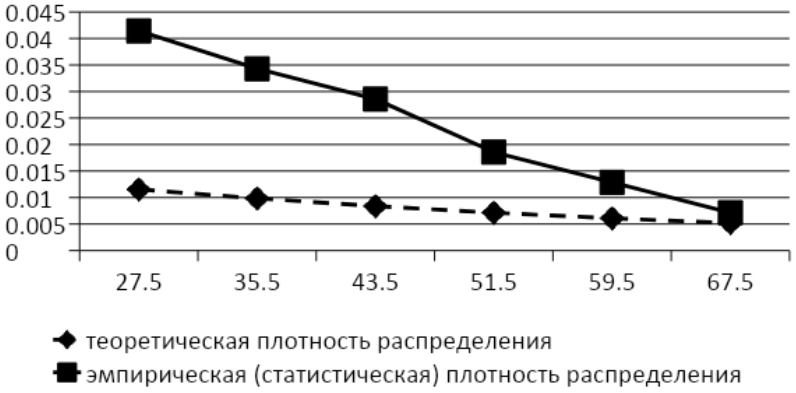
Теоретическая функция экспоненциального распределения - интеграл плотности распределения

$$F(t) = 1 - e^{-\lambda \cdot x} = 1 - e^{-\frac{x}{M(I)}}$$

4 Проверить гипотезу о соответствии выбранному теоретическому закону распределения случайной величины

1

Начальную оценку соответствия эмпирической функции распределения предполагаемому экспоненциальному закону распределения делаем путем визуального сравнения. На график наносим статистическую и теоретическую функцию



Вывод

экспоненциальное распределение не подходит для аппроксимации (выравнивания) распределения имеющихся статистических данных

2

Начальное предположение можно подтвердить посредством применения критерия согласия χ^2 (критерий хи-квадрат), основанный на измерении отклонений между эмпирическими ni и теоретическими частотами oi, соответствующими различным интервалам построенной гистограммы

Эмпирическая частота n_i, - известна

Теоретическая частота o_i, - рассчитывается $o_i = n\int_{t_{i-1}}^{t_i} f(t)dt = n(F(t_i) - F(t_{i-1})) = 100(e^{-\frac{x_{i-1}}{M(I)}} - e^{-\frac{x_i}{M(I)}})$

3

При заданных n_i и рассчитанных o_i для каждого интервала i гистограммы мера отклонения между эмпирическими и теоретическими частотами определяется следующей формулой

Расчетная формула χ^2

$$\chi^2 = \sum_{i=1}^{N} \frac{(n_i - o_i)^2}{o_i}$$

Начало интервала	Конец интервала	Наблюдаемая частота	Функция начала интервала	Функция конца интервала	Теоретическая частота	Критерий согласия
24	31	29	0,62	0,54	8,08	54,12
32	39	24	0,53	0,46	6,89	42,50
40	47	20	0,45	0,39	5,87	34,01
48	55	13	0,38	0,33	5,00	12,79
56	63	9	0,33	0,28	4,26	5,27
64	71	5	0,28	0,24	3,63	0,51
					Сумма	149,2

Расчетное значение χ^2

$$\chi^2 = 149,2$$

4 Проверить гипотезу о соответствии выбранному теоретическому закону распределения случайной величины (продолжение)

4 По справочным таблицам найти предельно допустимое значение критерия согласия (т.е. расчетное значение должно быть меньше предельно-допустимого значения) с учетом:

- заданного уровня значимости критерия α
- имеющихся степеней свободы r = N-k-1

Число степеней свободы	Уровень значимости α					
	0,01	0,05	0,1	0,90	0,95	0,99
1	6,6	3,8	2,71	0,02	0,004	0,0002
2	9,2	6,0	4,61	0,21	0,1	0,02
3	11,3	7,8	6,25	0,58	0,35	0,12
4	13,3	9,5	7,78	1,06	0,71	0,30
5	15,1	11,1	9,24	1,61	1,15	0,55
6	16,8	12,6	10,6	2,20	1,64	0,87
7	18,5	14,1	12,0	2,83	2,17	1,24
8	20,1	15,5	13,4	3,49	2,73	1,65
9	21,7	16,9	14,7	4,17	3,33	2,09
10	23,2	18,3	16,0	4,87	3,94	2,56
11	24,7	19,7	17,3	5,58	4,57	3,05
12	26,2	21,0	18,5	6,30	5,23	3,57
13	27,7	22,4	19,8	7,04	5,89	4,11
14	29,1	23,7	21,1	7,79	6,57	4,66
15	30,6	25,0	22,3	8,5	7,26	5,23
16	32,0	26,3	23,5	9,31	7,98	5,81
17	33,4	27,6	24,8	10,1	8,67	6,41
18	34,8	28,9	26,0	10,9	9,39	7,01
19	36,2	30,1	27,2	11,7	10,1	7,63
20	37,6	31,4	28,4	12,4	10,9	8,26
21	38,9	32,7	29,6	13,2	11,6	8,90
22	40,3	33,9	30,6	14,0	12,63	9,54
23	41,6	35,2	32,0	14,8	13,1	10,2
24	43,0	36,4	33,2	15,7	13,8	10,9
25	44,3	37,7	34,4	16,5	14,6	11,5
26	45,6	38,9	35,6	17,3	15,4	12,2
27	47,0	40,1	36,7	18,1	16,2	12,9
28	48,3	41,3	37,9	18,9	16,9	13,6
29	49,6	42,6	39,1	19,8	17,7	14,3
30	50,9	43,8	40,3	20,6	18,5	15,0

k – число параметров, оцененных на основе исходной информации и использованных для определения теоретического распределения (для экспоненциального распределения в формуле используется математическое ожидание, т.е. k=1)

N – число интервалов гистограммы, если ni <5, то эти интервалы объединяются с предыдущими интервалами (для рассматриваемого примера N = 6)

Примем α = 0,05; r = 6-1-1=4

Тогда $\chi^2_{N-k-1,1-\alpha} = 9{,}5$

Вывод

Т.к. условие не выполняется

$$\chi^2 = 149{,}2 \quad > \quad \chi^2_{N-k-1,1-\alpha} = 9{,}5$$

Экспоненциальный закон не подходит для сглаживания статистической функции

5 Построить СМО железнодорожной станции

СМО (Система массового обслуживания) – один из методов математического моделирования

Цель моделирования

установление реальных простоев вагонов на станции с учетом неравномерности технологических операций, приводящих к колебанию продолжительности их выполнения и возникновению простоев в ожидании этих операций.

Основной принцип

Станцию представляют в виде совокупности подсистем с учетом фазового или линейного их взаимодействия между собой. Каждый элемент модели соответствует технологической операции, выполняемой каналом обслуживания.

Наличие элементов определяется наличием парков на станции

Объемы работы определяются размером движения и количеством подач на грузовые фронты

Таблица 4 – Поездопотоки, поступающие на станцию

Категория поезда	
	3
Транзитный с переработкой	8
Транзитный без переработки	14
Местный в адрес п/п №1	1 подача
Местный в адрес п/п №2	3подачи

Таблица 3 – Наличие парков на станции (1 вариант)

Парки	
Парк приема ПП	+
Сортировочно-отправочный парк СОП	+
Путь необщего пользования №1	+
Путь необщего пользования №2	+

Таблица 3 – Наличие парков на станции (2 вариант)

Парки	
Парк приема ПП	+
Транзитный парк ТРП	+
Сортировочно-отправочный парк СОП	+
Путь необщего пользования №1	+
Путь необщего пользования №2	+

6 Рассчитать расчетные характеристики для заданных технологических операций

Есть статистика выполнения исследуемой операции

Математическое ожидание продолжительности выполнения операций (основная расчетная характеристика) устанавливается с использованием метода обработки статистических данных

Нет статистики выполнения исследуемой операции

Корректировка технологических нормативов выполняется по формуле (1) для маневровых операций (за исключением интервала роспуска состава на горке) и (2) для остальных станционных технологических операций

$$M_{(t_{об})} = \frac{T_{тех}}{1-v_{t_{об}}^2} \qquad (1)$$

$$M_{(t_{об})} = \frac{T_{тех}}{1-v_{t_{об}}\sqrt{v_{t_{об}}}} \qquad (2)$$

Коэффициенты вариации продолжительности операций принимаются на основании среднесетевых значений, представленных в таблице.

Таблица - Приблизительные значения коэффициентов вариации

Процессы	Коэффициент вариации
1. Интервалы между моментами прибытия составов в парк прибытия	0,7 – 0,8
2. Длительность технического обслуживания в парке приема	0,2 – 0,3
3. Длительность горочного интервала	0,35 – 0,4
4. Интервалы между моментами окончания накопления составов в сортировочном парке: - при работе одного маневрового локомотива - двух локомотивов	1 0,8
5. Длительность окончания формирования составов	0,4 – 0,45
6. Длительность технического облуживания в парке отправления	0,3 – 0,4
7. Интервалы между моментами отправления поездов по графику: - на двухпутный участок - на однопутный участок	0,6 – 1 0,5 – 0,6

7 Рассчитать загрузку каналов элементов СМО

загрузка системы

$\rho = \lambda / \mu$

λ – интенсивность поступления поездов в каждую из систем (соответствует числу обрабатываемых составов/передач за сутки)

μ - интенсивность выхода поездов из рассматриваемой системы

Если канал, обслуживающий систему, выполняет только одну операцию

$$\mu^{\text{пп}} = \frac{1440\text{К}}{\text{М}(\text{Т}_{\text{об}})}$$

К – количество каналов

Если канал, обслуживающий систему, обслуживает и другие системы

$$\mu^{\text{пп}} = \frac{1440 \cdot K - \sum(N_i \cdot T_{\text{об}i}^{\text{др}})}{M\,(T_{\text{об}})}$$

Если канал обслуживает несколько систем, необходимо определить суммарную загрузку канала путем сложения всех временных затрат на выполнение операций рассматриваемым каналом, чтобы установить не превышение допустимых значений (0,9)

$$\rho_{\text{канал}} = \frac{\sum(N_i \cdot \text{М}(T_{\text{об}i}))}{1440\text{К}}$$

Если при расчете $\rho_{\text{канал}} > 1$, то необходимо рассмотреть меры по сокращению загрузки

8 Рассчитать число поездов в очереди в каждый элемент СМО станции и продолжительность простоя

Расчет числа поездов в очереди в каждую систему

Если система обслуживается одним каналом

$$L_s = \frac{\rho^2}{(1-\rho)}$$

Если система обслуживается несколькими каналами

$$L_s = p_n \frac{\rho_{\text{кан}}}{(1-\rho_{\text{кан}})^2}$$

p_n - вероятность того, что в системе находится n поездов для **многоканальной системы**

$$p_n = \frac{\frac{\rho_{\text{кан}}^k}{k!}}{\sum_{k-0}^{n} \frac{\rho_{\text{кан}}^k}{k!}}$$

k – число занятых сервисов (каналов) поездами

Пример расчета p_n при двух каналах

$$p_n = \frac{\frac{\rho^k}{k!}}{\sum_{k-0}^{n} \frac{p^k}{k!}} = \frac{\frac{0{,}27^2}{2!}}{1+0{,}27+\frac{0{,}27^2}{2!}} = 0,$$

Расчет числа поездов, находящихся в системе

включает как очередь, так и средства обслуживания

$$L_k = L_s + \rho_{\text{сист}}$$

Если канал выполняет несколько операций, то Ls рассчитывается для канала (одно общее число), а потом равномерно распределяется между всеми элементами, обслуживаемыми каналом

Расчет продолжительности пребывания поездов в очереди

$$W_s = \frac{24 \cdot L_s}{\lambda}$$

рассчитывается для канала (одно общее число), а потом равномерно распределяется между всеми элементами, обслуживаемыми каналом, исходя их предположения равнозначности всех операций

Расчет продолжительности пребывания поездов в системе

$$W_k = W_s + \frac{1}{\mu} = W_s + M(T_{\text{об}})$$

На основании полученных результатов определяется потребное число путей в парках