ФЕДЕРАЛЬНОЕ БЮДЖЕТНОЕ ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ

ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНЖЕНЕРНЫХ ТЕХНОЛОГИЙ

КАФЕДРА ИНФОРМАЦИОННЫХ И УПРАВЛЯЮЩИХ СИСТЕМ

ОПРЕДЕЛЕНИЕ ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ. СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ ИСПЫТАНИЙ

Методические указания для выполнения практической работы по дисциплине «Надежность информационных систем»

Для студентов, обучающихся по направлению 09.03.02 – «Информационные системы и технологии» очной формы обучения

Воронеж 2020

Некоторая информационная система сложной структуры функционирует на сервере. Система построена на архитектуре клиент-сервер. Обслуживание *n*-клиентов осуществляется *n*-параллельными потоками. При этом система постоянно загружает все n-потоков. В процессе эксплуатации по прошествии первых *t* минут выяснилось, что в интервале времени от 0 до ⍰*t*
произошел сбой (ошибка времени исполнения) у *n1* клиентов, в интервале от ⍰*t* до 2⍰- у *n2* клиентов и т.д. Требуется определить следующие показатели надежности информационной системы:

1) вероятность безотказной работы;

2) среднюю наработку до отказа (среднее время безотказной работы);

3) среднеквадратическое отклонение и дисперсию времени безотказной работы;

4) интенсивность отказов;

5) плотность распределения времени безотказной работы.

Графически отобразить найденные величины.

Теоретические сведения.

По результатам испытаний *N* невосстанавливаемых одинаковых объектов получена статистическая выборка – массив наработки (в любых единицах измерения) до отказа каждого из *N* испытывавшихся объектов. Выборка характеризует случайную величину наработки до отказа объекта *T* = *{t}*. Как известно случайная величина подчиняется некоторому закону распределения.

Подбор закона распределения осуществляется на основе аппроксимации экспериментальных данных о наработке до отказа, которые должны быть представлены в наиболее компактном графическом виде. Выбор той или иной аппроксимирующей функции носит характер гипотезы, которую выдвигает исследователь. Экспериментальные данные могут с большим или меньшим правдоподобием подтверждать или не подтверждать справедливость той или иной гипотезы.

Алгоритм обработки результатов и расчета показателей надежности

Формирование статистического ряда

При большом числе испытываемых объектов полученный массив наработок {…, t_i, …} случайной величины T. Поэтому для компактности и наглядности выборка представляется в графическом изображении статистического ряда – гистограмме наработки до отказа. Для этого необходимо:

- установить интервал наработки [t_{min}, t_{max}] и его длину $\zeta t = t_{max} - t_{min}$,

где $t_{min} \le МИН\{\ldots, t_i, \ldots\}$, $t_{max} \ge МАКС\{\ldots, t_i, \ldots\}$;

- разбить интервал наработки [t_{min}, t_{max}] на k интервалов равной ширины Δt – шаг гистограммы

$$\Delta t = \frac{\zeta t}{k}, \quad \Delta t = t_{i+1} - t_i = t_i - t_{i-1}.$$

- подсчитать частоты появления отказов во всех k интервалах

$$\hat{P}_i = \frac{\Delta n(t_i, t_i + \Delta t)}{N} = \frac{\Delta n(t_i, t_{i+1})}{N},$$

где $\Delta n(t_i, t_i + \Delta t)$ – число объектов, отказавших в интервале [t_i, $t_i + \Delta t$].

При этом выполняется условие

$$\sum_{1}^{k} \hat{P}_i = 1;$$

- полученный статистический ряд представляется в виде гистограммы, которая строится следующим образом. По оси абсцисс (t) откладываются интервалы Δt, на каждом из которых, как на основании, строится прямоугольник, высота которого пропорциональна соответствующей частоте $\hat{P}$. Возможный вид гистограммы приведен на рис. 1.

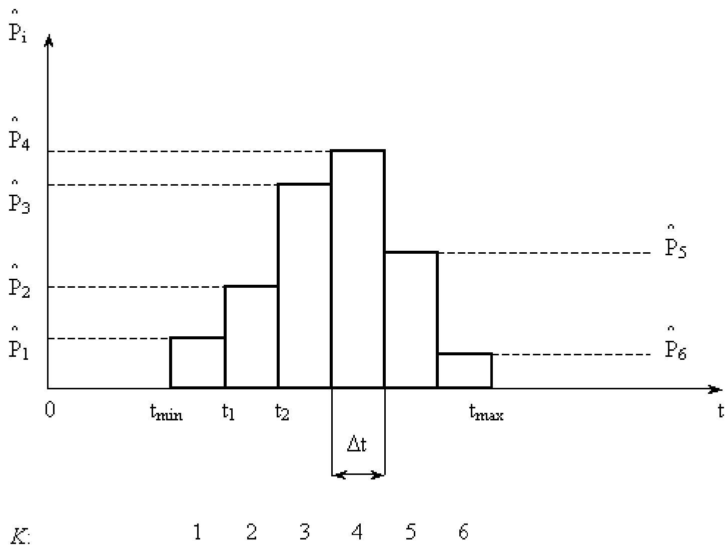


Рис. 1. Пример гистограммы.

Расчет эмпирических функций

Используя данные сформированного статистического ряда, определяются статистические оценки показателей надежности, т. е. эмпирические функции:

- функция распределения отказов (оценка ВО)

$$\hat{Q}(t_{min}) = \frac{n(t_{min})}{N} = 0$$

$$\hat{Q}(t_1) = \frac{n(t_1)}{N} = \frac{\Delta n(t_{min}, t_1)}{N} = \hat{P}_1 ;$$

$$\hat{Q}(t_2) = \frac{n(t_2)}{N} = \frac{\Delta n(t_{min}, t_1) + \Delta n(t_1, t_2)}{N} = \hat{P}_1 + \hat{P}_2 ;$$

… …

$$\hat{Q}(t_{max}) = \frac{n(t_{max})}{N} = \sum_{1}^{k} \hat{P}_i = 1 ;$$

- функции надежности *(оценка ВБР)*

$$\hat{P}(t_{min}) = 1 - \hat{Q}(t_{min}) = 1 ;$$

… … … … … … … … … … … …

$$\hat{P}(t_{max}) = 1 - \hat{Q}(t_{max}) = 0 ;$$

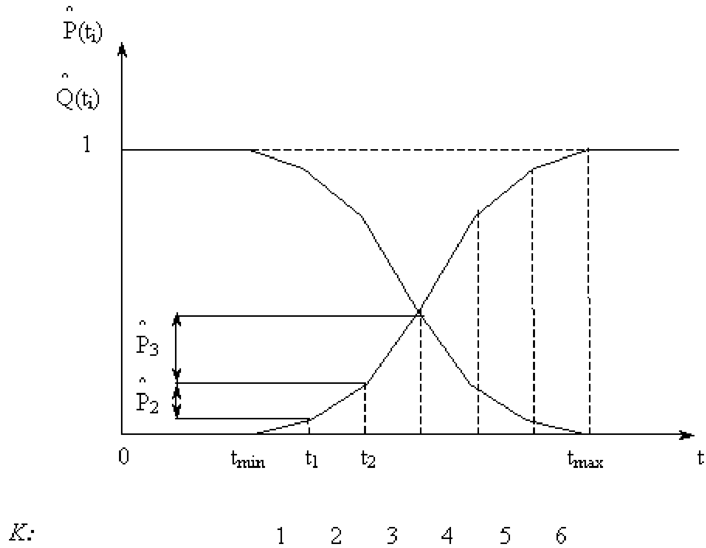


Рис.2. Пример функции надежности и распределения отказов.

- плотность распределения отказов (оценка ПРО)

$$\hat{f}(t_i) = \frac{\Delta n(t_i,\, t_{i+1})}{N \cdot \Delta t} = \frac{\hat{P}_i}{\Delta t};$$

- интенсивность отказов (оценка ИО)

$$\hat{\lambda}(t_i) = \frac{\Delta n(t_i,\, t_{i+1})}{N(t_i) \cdot \Delta t} = \frac{\Delta n(t_i,\, t_{i+1})}{[N - n(t_i)\,] \cdot \Delta t}.$$

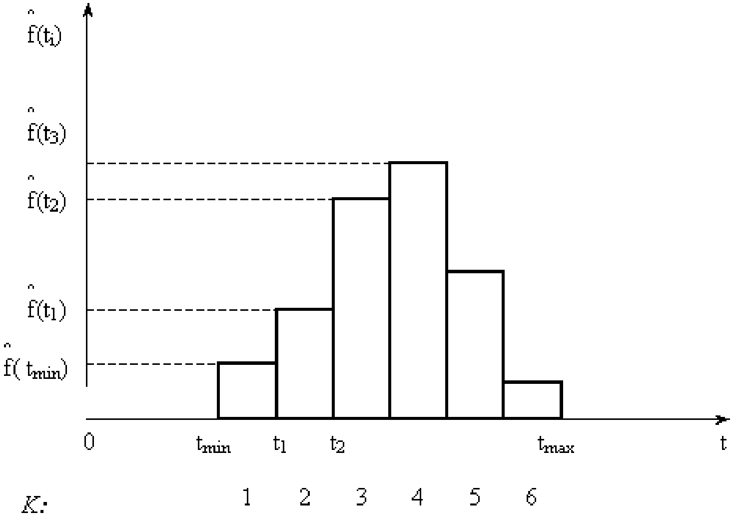


Рис. 3. Интенсивность отказов.

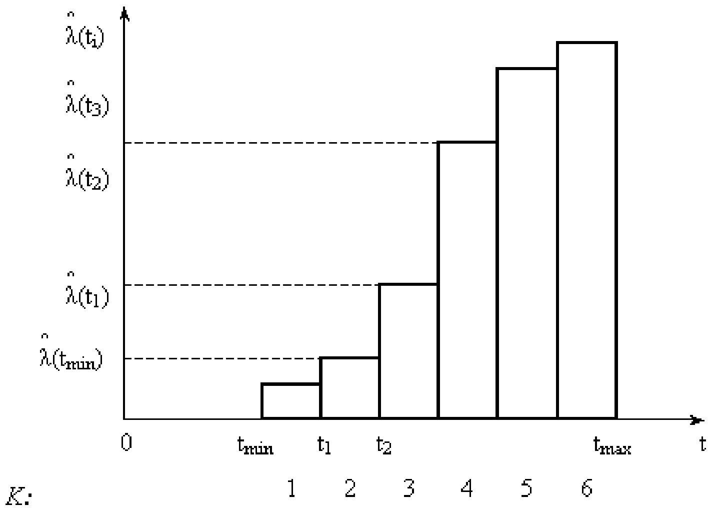


Рис. 4. Интенсивность отказов.

Расчет статистических оценок числовых характеристик

Для расчета статистических оценок числовых характеристик можно воспользоваться данными сформированного статистического ряда.

Оценки характеристик определяются:

- *оценка средней наработки до отказа (статистическое среднее наработки):*

$$\hat{T}_0 = \sum_{I}^{K} \tilde{t}_i \cdot \hat{P}_i ;$$

- *оценка дисперсии наработки до отказа (эмпирическая дисперсия наработки):*

$$\hat{D} = \sum_{I}^{K} (\tilde{t}_i - \hat{T}_0)^2 \cdot \hat{P}_i ,$$

где $\tilde{t}_i = t_i + \Delta t/2 = t_{i+1} - \Delta t/2$ – середина *i*-го интервала наработки, т. е. среднее значение наработки в интервале.

Оценка СКО $\hat{D} = \hat{S}^2$.

Выбор закона распределения

Выбор закона распределения состоит в подборе аналитической функции наилучшим образом аппроксимирующей эмпирические функции надежности.

Выбор, в значительной мере, процедура неопределенная и во многом субъективная, при этом многое зависит от априорных знаний об объекте и его свойствах, условиях работы, а также анализа вида графиков $\hat{P}(t)$, $\hat{f}(t)$, $\hat{\lambda}(t)$.

Очевидно, что выбор распределения будет зависеть, прежде всего, от вида эмпирической функции ПРО $\hat{f}(t)$, а также от вида - $\hat{\lambda}(t)$.

Предположим, что по тем или иным соображениям, выбран гипотетический закон распределения, заданный теоретической ПРО

$$f(t) = \Psi(t, a, b, c, \ldots) \, ,$$

где *a, b, c,* … - неизвестные параметры распределения.

Требуется подобрать эти параметры так, чтобы функция *f(t)* наилучшим образом сглаживала ступенчатый график $\hat{f}(t)$. При этом используется следующий прием: параметры *a, b, c,* … выбираются с таким расчетом, чтобы несколько важнейших числовых характеристик теоретического распределения были равны соответствующим статистическим оценкам.

На графике вместе с $\hat{f}(t)$ строится теоретическая ПРО *f(t)*, что позволяет визуально оценить результаты аппроксимации (расхождения между $\hat{f}(t)$ и *f(t)*. Поскольку эти расхождения неизбежны, то возникает вопрос: объясняются ли они случайными обстоятельствами, связанными с тем, что теоретическое распределение выбрано ошибочным? Ответ на этот вопрос дает расчет критерия согласия.

Расчет критерия согласия

Критерий согласия – это критерий проверки гипотезы о том, что случайная величина T, представленная своей выборкой, имеет распределение предполагаемого типа.

Проверка состоит в следующем. Рассчитывается критерий, как некоторая мера расхождения теоретического и эмпирического распределений, причем эта мера является случайной величиной.

Чем больше мера расхождения, тем хуже согласованность эмпирического распределения с теоретическим, т. е. меньше мала, то гипотезу о выборе закона распределения следует отвергнуть, как мало правдоподобную.

В противном случае – экспериментальные данные не противоречат принятому распределению.

Из известных критериев наиболее применяемый критерий согласия χ^2 (хи-квадрат) Пирсона.

Проверка согласованности распределений по критерию χ^2 производится следующим образом:

- рассчитывается критерий χ^2 (мера расхождения)

$$\chi^2 = N \cdot \sum_{1}^{k} (\hat{P}_i - P_i)^2 / P_i ,$$

где $P_i = f(\tilde{t}_i)\,\Delta t$ – теоретическая частота (вероятность) попадания случайной величины в интервал [t_i, $t_i + \Delta t$];

- определяется число степеней свободы $R = k - L$,

где L – число независимых условий, наложенных на частоты $\hat{P}_i$, например:

а) условие $\sum \hat{P}_i = 1$;

б) условие совпадения $\sum \tilde{t}_i \cdot \hat{P}_i = T_0$;

в) условие совпадения $\sum (\tilde{t}_i - \hat{T}_0)^2 \cdot \hat{P}_i$ = D и т. д.

Чаще всего $L = 3$. Чем больше число степеней свободы, тем больше случайная величина χ^2 подчиняется распределению Пирсона;

- по рассчитанным χ^2 и R определяется вероятность P того, что величина, имеющая распределение Пирсона с R степенями свободы, превзойдет рассчитанное значение χ^2.

Ответ на вопрос: насколько мала должна быть вероятность P, чтобы отбросить гипотезу о выборе того или иного закона распределения – во многом неопределенный.

На практике, если *P < 0,1*, то рекомендуется подыскать другой закон распределения.

В целом, с помощью критерия согласия, можно опровергнуть выбранную гипотезу, если же *P* достаточно велика, то это не может служить доказательством правильности гипотезы, а указывает лишь на то, что гипотеза не противоречит данным эксперимента.

Ход выполнения работы.

Пусть имеет статистические данные работы системы:

N = 1000, t = 100, Δt = 10,

n1 = 150, n2 = 100, n3 = 50, n4 = 200, n5 = 100, n6 = 100, n7 = 80, n8 = 20, n9 = 90, n10 = 10;

где N – количество потоков обслуживания клиентов;

t – единицы наработки;

Δt – приращение наработки;

ni – количество отказавших потоков в i-м интервале;

K – количество интервалов.

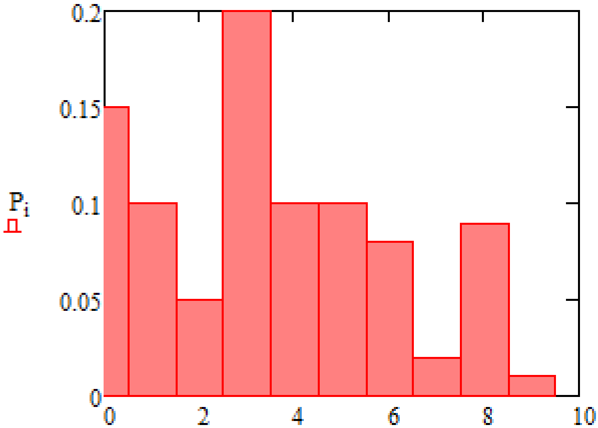
Сведем данные в таблицу:

k	1	2	3	4	5	6	7	8	9	10
ni	150	100	50	200	100	100	80	20	90	10

1. Подсчитаем частоты появления отказов по выражению

$$\hat{P}_i = \frac{\Delta n(t_i,\ t_i + \Delta t)}{N} = \frac{\Delta n(t_i,\ t_{i+1})}{N},$$

Далее в математическом пакете строим гистограмму:



2. Определим оценку вероятности отказов:

- *функция распределения отказов (оценка ВО)*

$$\hat{Q}(t_{min}) = \frac{n(t_{min})}{N} = 0$$

$$\hat{Q}(t_1) = \frac{n(t_1)}{N} = \frac{\Delta n(t_{min}, t_1)}{N} = \hat{P}_1 ;$$

$$\hat{Q}(t_2) = \frac{n(t_2)}{N} = \frac{\Delta n(t_{min}, t_1) + \Delta n(t_1, t_2)}{N} = \hat{P}_1 + \hat{P}_2 ;$$

… …

$$\hat{Q}(t_{max}) = \frac{n(t_{max})}{N} = \sum_{1}^{k} \hat{P}_i = 1 ;$$

3. Рассчитаем *функцию надежности (оценка ВБР)*

$$\hat{P}(t_{min}) = 1 - \hat{Q}(t_{min}) = 1 \ ;$$

$$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$$

$$\hat{P}(t_{max}) = 1 - \hat{Q}(t_{max}) = 0 \ ;$$

4. Рассчитаем плотность распределения отказов (оценка ПРО)

5. Определим интенсивность отказов (оценка ИО)

$$\hat{\lambda}(t_i) = \frac{\Delta n(t_i,\ t_{i+1})}{N(t_i) \cdot \Delta t} = \frac{\Delta n(t_i,\ t_{i+1})}{[N - n(t_i)\ ] \cdot \Delta t} .$$

6. Рассчитаем оценку средней наработки до отказа (статистическое среднее наработки):

7. Рассчитаем оценку дисперсии наработки до отказа (эмпирическая дисперсия наработки):

$$\hat{D} = \sum_{1}^{K} (\tilde{t}_i - \hat{T}_0)^2 \cdot \hat{P}_i ,$$

где $\tilde{t}_i = t_i + \Delta t/2 = t_{i+1} - \Delta t/2$ *– середина i-го интервала наработки, т. е. среднее значение наработки в интервале*

.

Содержание отчета.

1. Цель работы
2. Задание.
3. Основные результаты расчетов (численные массивы данных и графики).
4. Выводы.

Варианты заданий.

Вариант 1.
N = 1000, t = 100, Δt = 10, n1 = 500, n2 = 100, n3 = 50, n4 = 20, n5 = 10, n6 = 1, n7 = 8, n8 = 2, n9 = 9, n10 = 0
Вариант 2.
N = 500, t = 100, Δt = 10, n1 = 50, n2 = 10, n3 = 50, n4 = 20, n5 = 10, n6 = 10, n7 = 80, n8 = 20, n9 = 9, n10 = 10
Вариант 3.
N = 1000, t = 100, Δt = 10, n1 = 78, n2 = 101, n3 = 14, n4 = 26, n5 = 138, 9 n6 = 65, n7 = 8, n8 = 15, n9 = 73, n10 = 86
Вариант 4.
N = 300, t = 200, Δt = 20, n1 = 10, n2 = 10, n3 = 1, n4 = 7, n5 = 12, n6 = 19, n7 = 5, n8 = 14, n9 = 0, n10 = 10
Вариант 5.
N = 700, t = 100, Δt = 10, n1 = 32, n2 = 29, n3 = 1, n4 = 29, n5 = 1, n6 = 33, n7 = 7, n8 = 27, n9 = 34, n10 = 1
Вариант 6.

N = 657, t = 100, Δt = 10, n1 = 135, n2 = 42, n3 = 87, n4 = 4, n5 = 26, n6 = 17, n7 = 2, n8 = 105, n9 = 118, n10 = 121 Вариант 7.
N = 10000, t = 100, Δt = 10, n1 = 463, n2 = 476, n3 = 452, n4 = 359, n5 = 80, n6 = 296, n7 = 195, n8 = 316, n9 = 148, n10 = 434
Вариант 8.
N = 1200, t = 100, Δt = 10, n1 = 65, n2 = 22, n3 = 37, n4 = 31, n5 = 60, n6 = 43, n7 = 36, n8 = 5, n9 = 19, n10 = 0
Вариант 9.
N = 1000, t = 100, Δt = 10, n1 = 0, n2 = 9, n3 = 65, n4 = 44, n5 = 47, n6 = 28, n7 = 60, n8 = 97, n9 = 44, n10 = 81
Вариант 10.
N = 1200, t = 100, Δt = 10, N1=7, n2=90, n3=22, n4=13, n5=47, n6=79, n7=91, n8=30, n9=13, n10=49
Вариант 11.
N = 1200, t = 100, Δt = 10, n1=50,n2=31,n3=7,n4=1,n5=7,n6=100,n7=71,n8=45,n9= 12,n10=32
Вариант 12.
N = 1200, t = 100, Δt = 10, n1=84,n2=6,n3=7,n4=76,n5=27,n6=68,n7=55,n8=19,n9= 65,n10=15
Вариант 13.
N = 1200, t = 100, Δt = 10, n1=7,n2=28,n3=87,n4=76,n5=53,n6=52,n7=6,n8=82,n9= 17,n10=37

Вариант 14.
N = 1200, t = 100, Δt = 10, n1=59,n2=74,n3=76,n4=94,n5=69,n6=88,n7=78,n8=57,n9=98,n10=16
Вариант 15.
N = 1200, t = 100, Δt = 10, n1=6,n2=82,n3=40,n4=4,n5=93,n6=16,n7=84,n8=95,n9=67,n10=70
Вариант 16.
N = 1200, t = 100, Δt = 10, n1=30,n2=47,n3=10,n4=22,n5=89,n6=77,n7=44,n8=13,n9=82,n10=9
Вариант 17.
N = 1200, t = 100, Δt = 10, n1=11,n2=14,n3=45,n4=58,n5=29,n6=80,n7=11,n8=18,n9=88,n10=75
Вариант 18.
N = 1200, t = 100, Δt = 10, n1=47,n2=9,n3=79,n4=82,n5=94,n6=49,n7=25,n8=7,n9=47,n10=85
Вариант 19.
N = 1200, t = 100, Δt = 10, n1=97,n2=9,n3=33,n4=75,n5=22,n6=49,n7=37,n8=41,n9=15,n10=44
Вариант 20.
N = 1200, t = 100, Δt = 10, n1=35,n2=75,n3=78,n4=76,n5=44,n6=42,n7=7,n8=36,n9=23,n10=77
Вариант 21.
N = 1200, t = 100, Δt = 10, n1=13,n2=76,n3=9,n4=32,n5=97,n6=9,n7=22,n8=61,n9=54,n10=40
Вариант 22.
N = 1200, t = 100, Δt = 10,

n1=43,n2=51,n3=2,n4=28,n5=77,n6=61,n7=36,n8=10,n9=73,n10=58
Вариант 23.
N = 1200, t = 100, Δt = 10, n1=10,n2=94,n3=30,n4=65,n5=9,n6=86,n7=21,n8=1,n9=92,n10=15
Вариант 24.
N = 1200, t = 100, Δt = 10, n1=25,n2=26,n3=85,n4=20,n5=85,n6=63,n7=53,n8=57,n9=75,n10=13
Вариант 25.
N = 1200, t = 100, Δt = 10, n1=83,n2=92,n3=4,n4=2,n5=81,n6=61,n7=92,n8=37,n9=70,n10=95
Вариант 26.
N = 1200, t = 100, Δt = 10, n1=53,n2=24,n3=48,n4=11,n5=50,n6=90,n7=47,n8=65,n9=37,n10=9
Вариант 27.
N = 1200, t = 100, Δt = 10, n1=76,n2=29,n3=43,n4=10,n5=9,n6=43,n7=69,n8=21,n9=32,n10=15
Вариант 28.
N = 1200, t = 100, Δt = 10, n1=93,n2=41,n3=62,n4=94,n5=20,n6=41,n7=89,n8=89,n9=72,n10=79
Вариант 29.
N = 1200, t = 100, Δt = 10, n1=39,n2=16,n3=83,n4=62,n5=95,n6=95,n7=41,n8=64,n9=28,n10=77
Вариант 30.
N = 1200, t = 100, Δt = 10, n1=80,n2=47,n3=48,n4=51,n5=49,n6=29,n7=39,n8=7,n9=76,n10=33

Контрольные вопросы:

1. Что представляет математическая модель, и для каких целей она используется в задачах надежности?
2. Из каких условий выбирается закон распределения наработки до отказа объекта?
3. В чем заключается постановка задачи при испытаниях объектов на надежность?
4. Что представляет собой процедура формирования статистического ряда по результатам испытаний?
5. Какие эмпирические функции рассчитываются при обработке результатов испытаний?
6. В чем заключается выбор закона распределения наработки до отказа по результатам испытаний?
7. Что представляет собой критерий согласия?