**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

**Кафедра АМ**

отчет

**по ИДЗ №2**

**по дисциплине «Статистический анализ»**

Тема: Классические методы математической статистики.

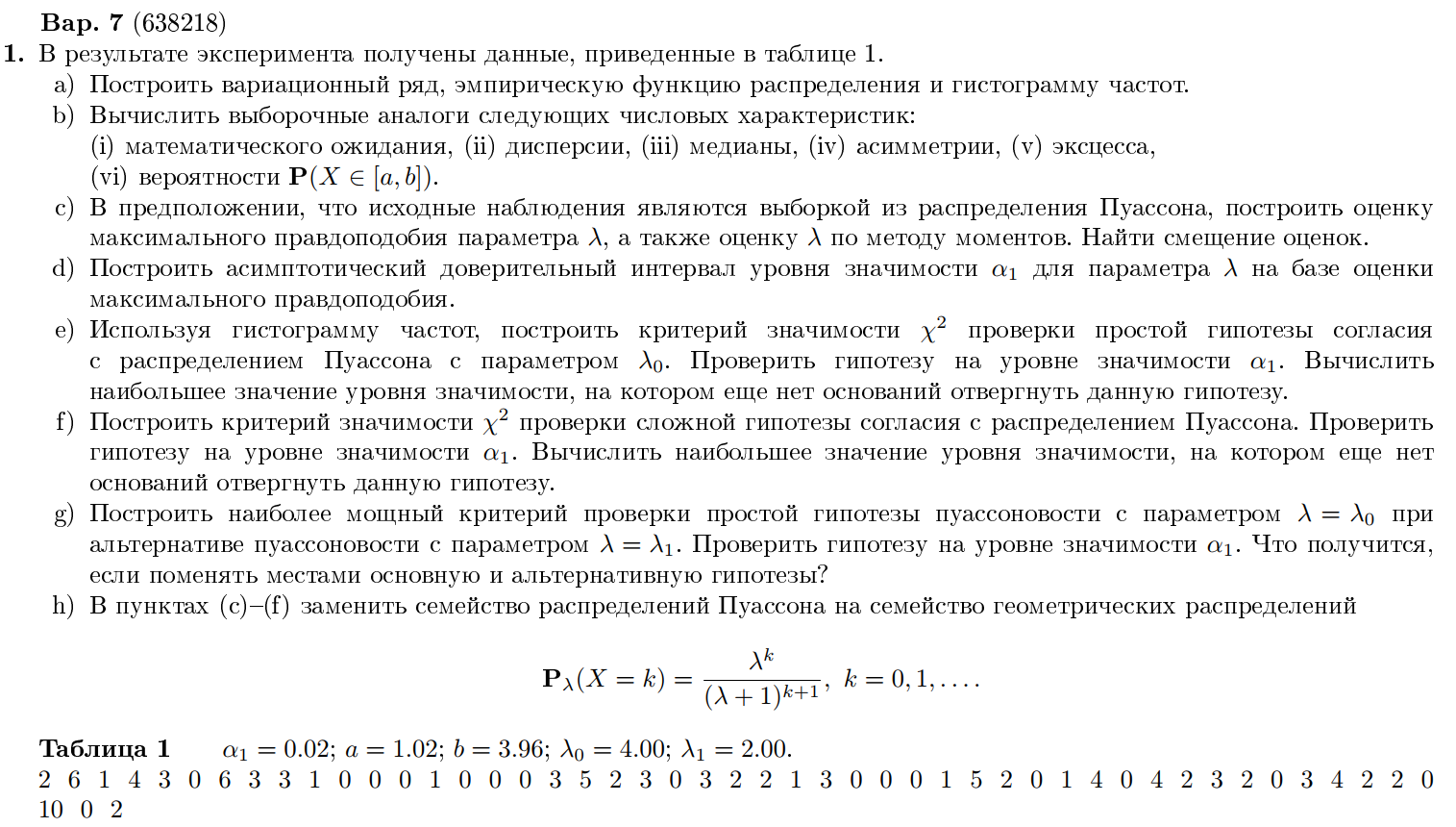
Вариант №12.

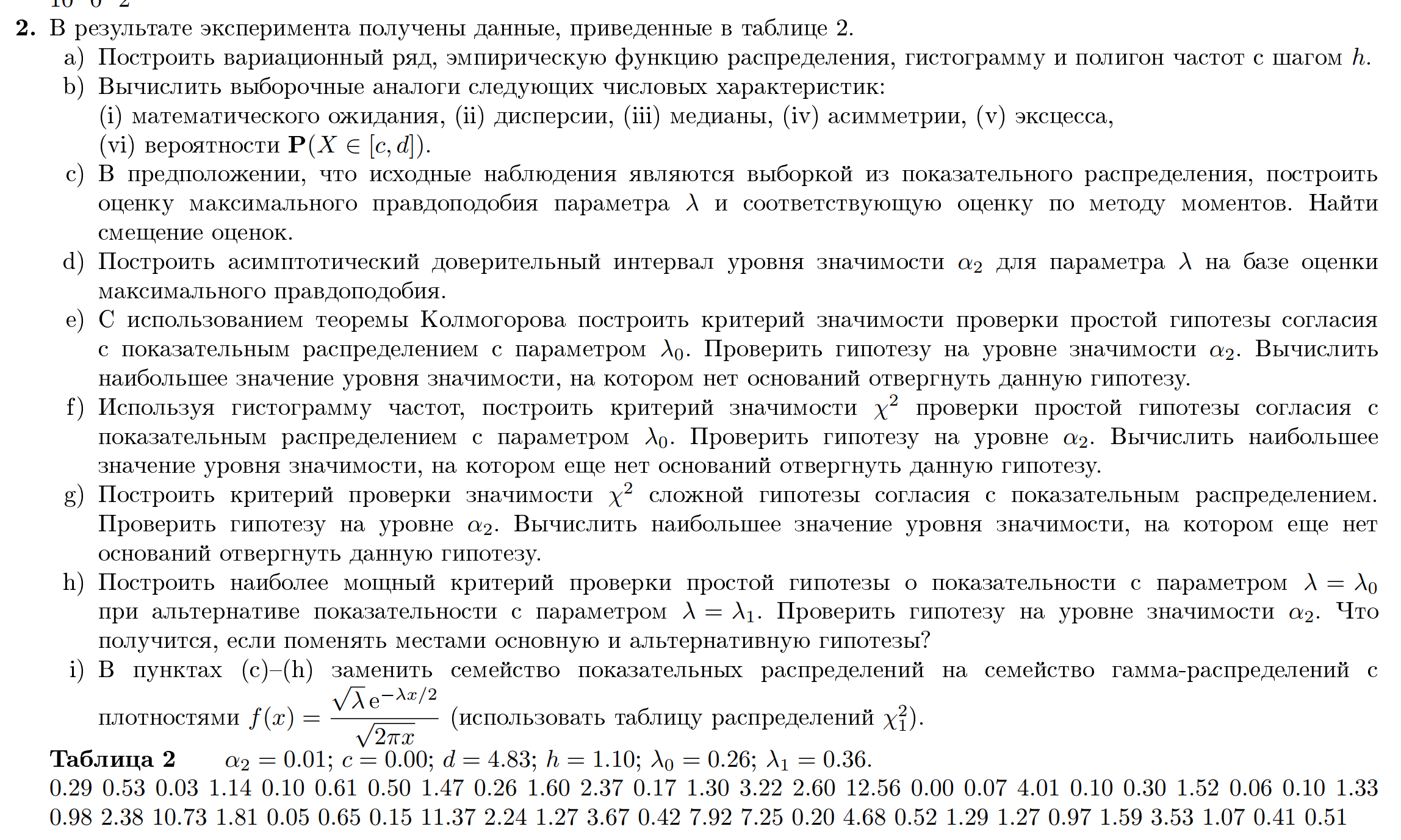
|  |  |  |
| --- | --- | --- |
| Студент гр. 6382 |  | Мартыненко П.П. |
| Преподаватель |  | Чирина А.В. |

Санкт-Петербург

2018

# **Постановка задачи**



****

# **Порядок выполнения работы**

1. Выполним задания на основе данных из таблицы 1 (выборка из дискретного распределения).

Таблица 1.

|  |  |
| --- | --- |
|  |  |
|  | 2 6 1 4 3 0 6 3 3 1 0 0 0 1 0 0 0 3 5 2 3 0 3 2 2 1 3 0 0 0 1 5 2 0 1 4 0 4 2 3 2 0 3 4 2 2 0 10 0 2 |

1. Построим вариационный ряд, эмпирическую функцию распределения и гистограмму частот.

Получим вариационный ряд выборки , воспользовавшись командой sort(x).

Таблица 1.1.

|  |  |
| --- | --- |
|  | 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 4 4 4 4 5 5 6 6 10 |

Тогда эмпирическая функция распределения будет иметь вид

Построив её график, получим

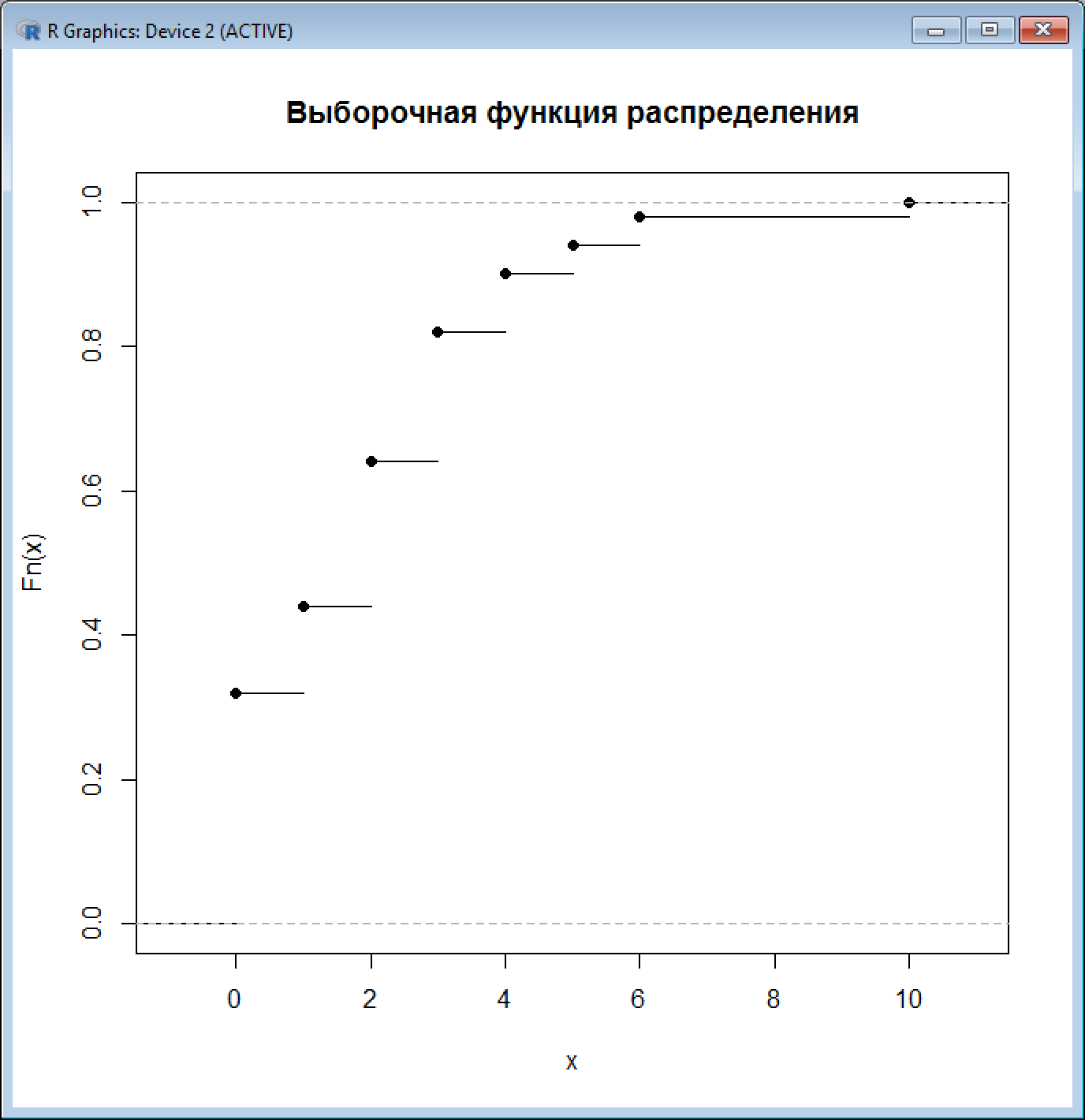


Рис.1.1 Выборочная функция распределения (п.1).

Построим таблицу частот для выборки

Таблица 1.2

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Частоты | 16 | 6 | 10 | 9 | 4 | 2 | 2 | 1 |
| Элементы выборки | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 10 |

Построим гистограмму частот, воспользовавшись следующими командами

l <- as.integer(min(x-1))

u <- as.integer(max(x+1))

hist (x, breaks = c(l: r), right = TRUE, freq = TRUE, main = "Гистограмма частот")

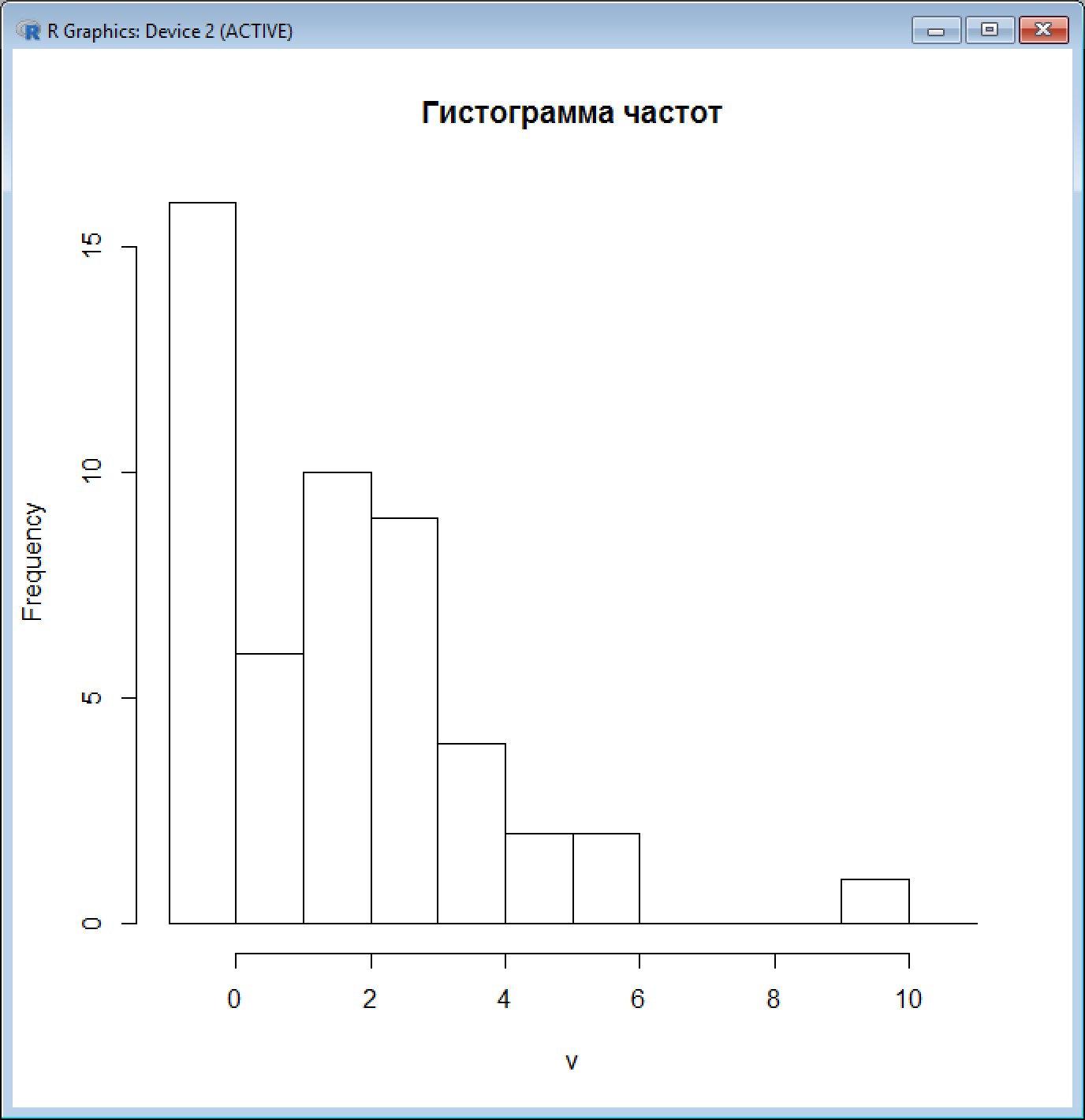


Рис.1.2. Гистограмма частот (п.1).

1. *Вычислим выборочные аналоги следующих числовых характеристик*
2. Математическое ожидание
3. Дисперсия
4. Медиана
5. Асимметрия
6. Эксцесс
7. Вероятность попадания в промежуток [a,b] при
8. В предположении, что исходные наблюдения являются выборкой из распределения Пуассона, построим оценку максимального правдоподобия параметра , а также оценку по методу моментов. Найдем смещение оценок.

Плотность распределения Пуассона имеет вид

* Метод максимального правдоподобия

Функция правдоподобия будет равна

И её логарифм

Дифференцируя по , получим

Приравняем производную к нулю

Проверим результаты, используя следующий код

LL<-function(t){sum(dpois(x,t[1],log=TRUE))}

ml<-maxNR(LL,start=c(1))

ml$estimate

[1] 2.02

ОМП была посчитана верно.

* Метод моментов

Математическое ожидание

Выборочный средний момент

И смещение оценки максимального правдоподобия

Значит, – несмещенная оценка.

1. Построим асимптотический доверительный интервал уровня значимости для параметра λ на базе оценки максимального правдоподобия.

Найдем вторую производную функции правдоподобия

Тогда наблюдаемая информация Фишера будет иметь вид

И выборочная дисперсия

Стандартная ошибка будет иметь вид

Значит, доверительный интервал с уровнем доверия будет иметь вид

Тогда, подставляя известные значения, получим доверительный интервал

1. Используя гистограмму частот, построим критерий значимости проверки простой гипотезы согласия с распределением Пуассона с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Тогда искомая вероятность примет вид

Построим таблицу оценки методом Хи-квадрат при

Таблица 1.3

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |
| Вероятности | 0,0183 | 0,0732 | 0,1465 | 0,1953 | 0,1953 | 0,1562 | 0,1041 | 0,1111 | 1 |
|  | 0,915 | 3,66 | 7,325 | 9,765 | 9,765 | 7,81 | 5,205 | 5,555 | 50 |
|  | 16 | 6 | 10 | 9 | 4 | 2 | 2 | 1 | 50 |
|  | -15,085 | -2,34 | -2,675 | 0,765 | 5,765 | 5,81 | 3,205 | 4,555 | 0 |
|  | 248,696 | 1,496 | 0,977 | 0,060 | 3,404 | 4,322 | 1,973 | 3,735 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит распределению Пуассона, и принимаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу, очень мало .

1. Построим критерий значимости проверки сложной гипотезы согласия с распределением Пуассона. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Сложная гипотеза имеет вид

Оценим неизвестный параметр как

Тогда искомая вероятность примет вид

Построим таблицу оценки методом Хи-квадрат при

Таблица 1.4

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |
| Вероятности | 0,1327 | 0,268 | 0,2706 | 0,1822 | 0,092 | 0,0372 | 0,0125 | 0,0048 | 1 |
|  | 6,6328 | 13,3982 | 13,5322 | 9,1117 | 4,6014 | 1,8590 | 0,6259 | 0,239 | 50 |
|  | 16 | 6 | 10 | 9 | 4 | 2 | 2 | 1 | 50 |
|  | -9,3672 | 7,3982 | 3,5322 | 0,1117 | 0,6014 | -0,141 | -1,374 | -0,761 | 0 |
|  | 13,229 | 4,0851 | 0,922 | 0,0014 | 0,0786 | 0,0107 | 3,0171 | 2,4237 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит распределению Пуассона, и принимаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу, мало .

1. Построим наиболее мощный критерий проверки простой гипотезы пуассоновости с параметром при альтернативе пуассоновости с параметром . Проверим гипотезу на уровне значимости . Что получится, если поменять местами основную и альтернативную гипотезы?

* Сформулируем гипотезы

Согласно лемме Неймана-Пирсона, наиболее мощный критерий проверки гипотезы при альтернативе имеет вид:

 , где 



Наиболее мощный критерий примет вид

Логарифмируем соотношение

Получим

И после преобразования

Введем обозначение

Тогда критерий примет вид

Вычислим и *p* из уравнения:



 , следовательно, 

Т.к. , то подбором (в цикле с помощью R) среди целых чисел можем найти такое наибольшее (а после и α0), что:



Тогда 

Проведём вычисления в R.  
> c<-0

> lambda0=4

> lambda1=2

> alpha1<-0.02

> alpha0<-1-ppois(c,lambda0\*n)-dpois(c, lambda0\*n)

> while (alpha0 > alpha1)

+ {

+ c<-c+1;

+ alpha0<-1-ppois(c,lambda0\*n)-dpois(c, lambda0\*n)

+ }

> p<-(alpha1-alpha0)/dpois(c,lambda0\*n)

> lche<- sum(x)

> lche>=c

[1] FALSE принимаем 

Тогда наиболее мощный критерий будет иметь вид



* Поменяем местами основную и альтернативную гипотезу.







Наиболее мощный критерий:

Логарифмируем соотношение

Получим

И после преобразования

Введем обозначение

Тогда критерий примет вид

Вычислим и *p* из уравнения



 , следовательно, 

Т.к.подбором (в цикле с помощью R) среди целых чисел можем найти наибольшее

Тогда с учётом уравнения выше получим



Проведём вычисления в R.

> c<-0; lambda1 <-2

> alpha0<-ppois(c,lambda1\*length(x))

while(alpha0<alpha1){

+ c=c+1

+ alpha0=ppois(c,lambda1\*length(x))}

> c= c-1

> c

[1] 79

> alpha0 = ppois(c,lambda1\*length(x))

> alpha0

[1] 0.01745132

> p = (alpha1 - alpha0)/dpois(c,lambda1\*length(x))

> alpha0

[1] 0.01745132

> p

[1] 0.6129158

> lche = sum(x)

> lche<=c

[1] FALSE отвергаем альтернативу

Тогда наиболее мощный критерий будет иметь вид



При замене основной и альтернативной гипотезы меняется также гипотеза, которую принимаем. Но т. к. изменение происходит со сменой гипотез местами, решение не меняется и принимается в пользу одной и той же гипотезы

1. Заменим семейство распределений Пуассона на семейство геометрических распределений и повторим пункты c)-f)
2. В предположении, что исходные наблюдения являются выборкой из геометрического распределения, построим оценку максимального правдоподобия параметра , а также оценку по методу моментов. Найдем смещение оценок.

Плотность геометрического распределения имеет вид

Тогда введем обозначения

* Метод максимального правдоподобия

Функция правдоподобия будет равна

И её логарифм

Дифференцируя по , получим

Приравняем производную к нулю

* Метод моментов

Математическое ожидание геометрического распределения имеет вид

Выборочный средний момент

И смещение оценки максимального правдоподобия

Значит, – несмещенная оценка.

1. Построим асимптотический доверительный интервал уровня значимости для параметра λ на базе оценки максимального правдоподобия.

Найдем вторую производную функции правдоподобия

Тогда наблюдаемая информация Фишера будет иметь вид

И выборочная дисперсия

Стандартная ошибка будет иметь вид

Значит, доверительный интервал с уровнем доверия будет иметь вид

Тогда, подставляя известные значения, получим доверительный интервал

1. Используя гистограмму частот, построим критерий значимости проверки простой гипотезы согласия с геометрическим распределением с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Тогда искомая вероятность примет вид

Построим таблицу оценки методом Хи-квадрат при

Таблица 1.5

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |
| Вероятности | 0,2 | 0,16 | 0,128 | 0,1024 | 0,0819 | 0,0655 | 0,0524 | 0,2097 | 1 |
|  | 10 | 8 | 6,4 | 5,12 | 4,096 | 3,2768 | 2,6214 | 10,4858 | 50 |
|  | 16 | 6 | 10 | 9 | 4 | 2 | 2 | 1 | 50 |
|  | -6 | 2 | -3,6 | -3,88 | 0,096 | 1,2768 | 0,6214 | 9,4858 | 0 |
|  | 3,6 | 0,5 | 2,025 | 2,9403 | 0,0023 | 0,4975 | 0,1473 | 8,5811 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит геометрическому распределению, и принимаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу .

1. Построим критерий значимости проверки сложной гипотезы согласия с геометрическим распределением. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Сложная гипотеза имеет вид

Оценим неизвестный параметр как

Тогда искомая вероятность примет вид

Построим таблицу оценки методом Хи-квадрат при

Таблица 1.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |
| Вероятности | 0,331 | 0,221 | 0,148 | 0,099 | 0,066 | 0,044 | 0,029 | 0,059 | 1 |
|  | 16,556 | 11,074 | 7,407 | 4,954 | 3,313 | 2,216 | 1,482 | 2,994 | 50 |
|  | 16 | 6 | 10 | 9 | 4 | 2 | 2 | 1 | 50 |
|  | 0,556 | 5,074 | -2,592 | -4,045 | -0,686 | 0,2166 | -0,517 | 1,994 | 0 |
|  | 0,018 | 2,324 | 0,907 | 3,303 | 0,142 | 0,021 | 0,180 | 1,328 |  |

Итого

Следовательно, принимаем гипотезу , что выборка принадлежит геометрическому распределению, и отвергаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу

1. Выполним задания на основе данных из таблицы 2 (выборка из абсолютно непрерывного распределения).

Таблица 2.

|  |  |
| --- | --- |
|  | |
|  | 0.29 0.53 0.03 1.14 0.10 0.61 0.50 1.47 0.26 1.60 2.37 0.17 1.30 3.22 2.60 12.56 0.00 0.07 4.01 0.10 0.30 1.52 0.06 0.10 1.33  0.98 2.38 10.73 1.81 0.05 0.65 0.15 11.37 2.24 1.27 3.67 0.42 7.92 7.25 0.20 4.68 0.52 1.29 1.27 0.97 1.59 3.53 1.07 0.41 0.51 |

1. Построим вариационный ряд, эмпирическую функцию распределения и гистограмму частот.

Таблица 2.1.

|  |  |
| --- | --- |
|  | 0.00 0.03 0.05 0.06 0.07 0.10 0.10 0.10 0.15 0.17 0.20 0.26 0.29 0.30 0.41 0.42 0.50 0.51 0.52 0.53 0.61 0.65 0.97 0.98 1.07 1.14 1.27 1.27 1.29 1.30 1.33 1.47 1.52 1.59 1.60 1.81 2.24 2.37 2.38 2.60 3.22 3.53 3.67 4.01 4.68 7.25 7.92 10.73 11.37 12.56 |

Тогда эмпирическая функция распределения будет иметь вид

Построив её график, получим

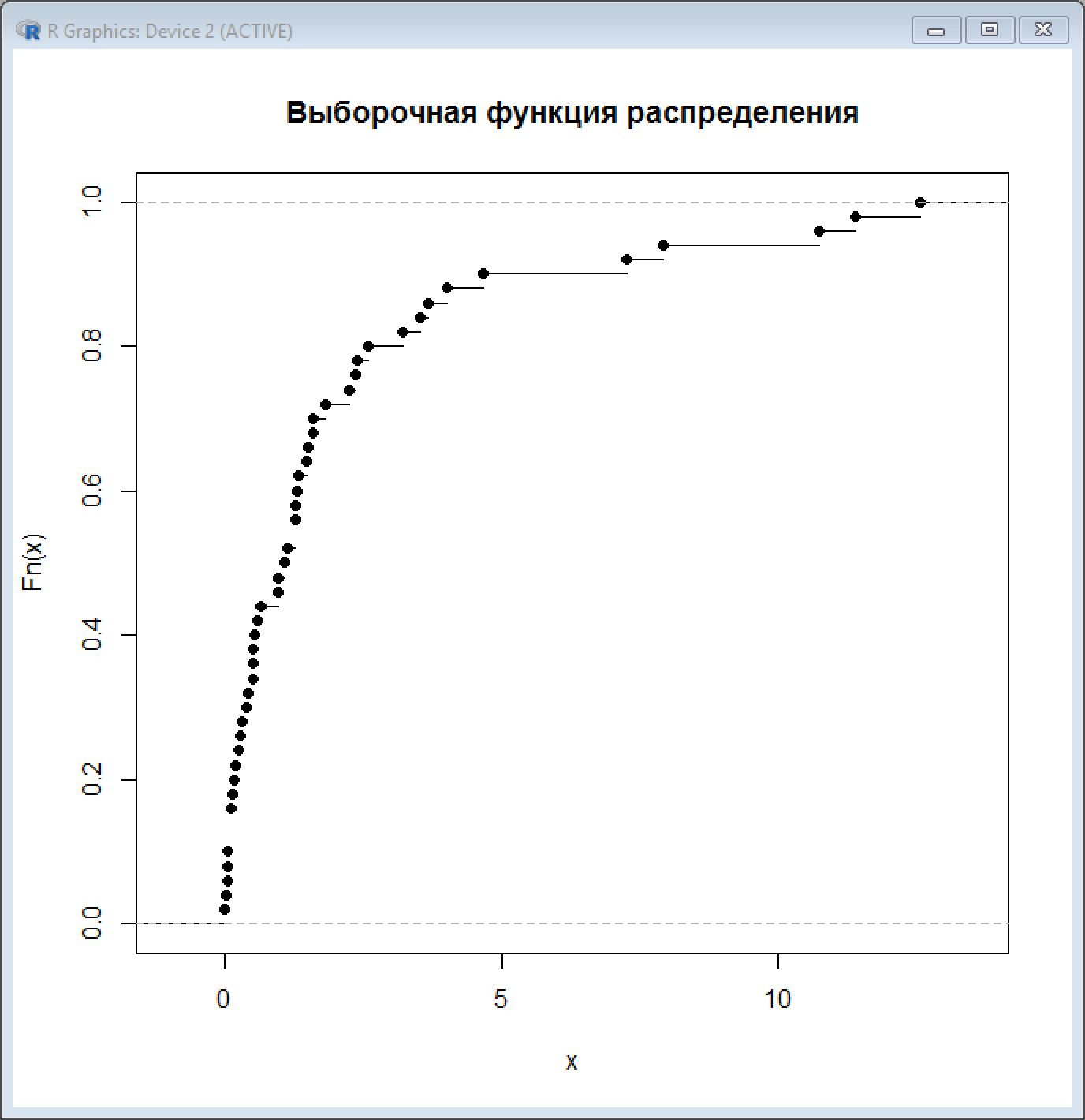


Рис.2.1 Выборочная функция распределения (п.2).

Построим гистограмму и полигон частот, воспользовавшись следующими командами

l <- as.integer(min(x))

u <- as.integer(round(max(x+1)))

p = array(dim = round ((u-l)/1.1))

for(i in 1:length(p)){ p[i] = l; l = l+1.1}

h1<-hist(x,breaks=p,freq=TRUE,right=TRUE, main = "Полигон частот с шагом h = 1.1")

lines(h1$counts ~ h1$mids, col="red")

Гистограмма и полигон частот изображены на рисунке 2.2.

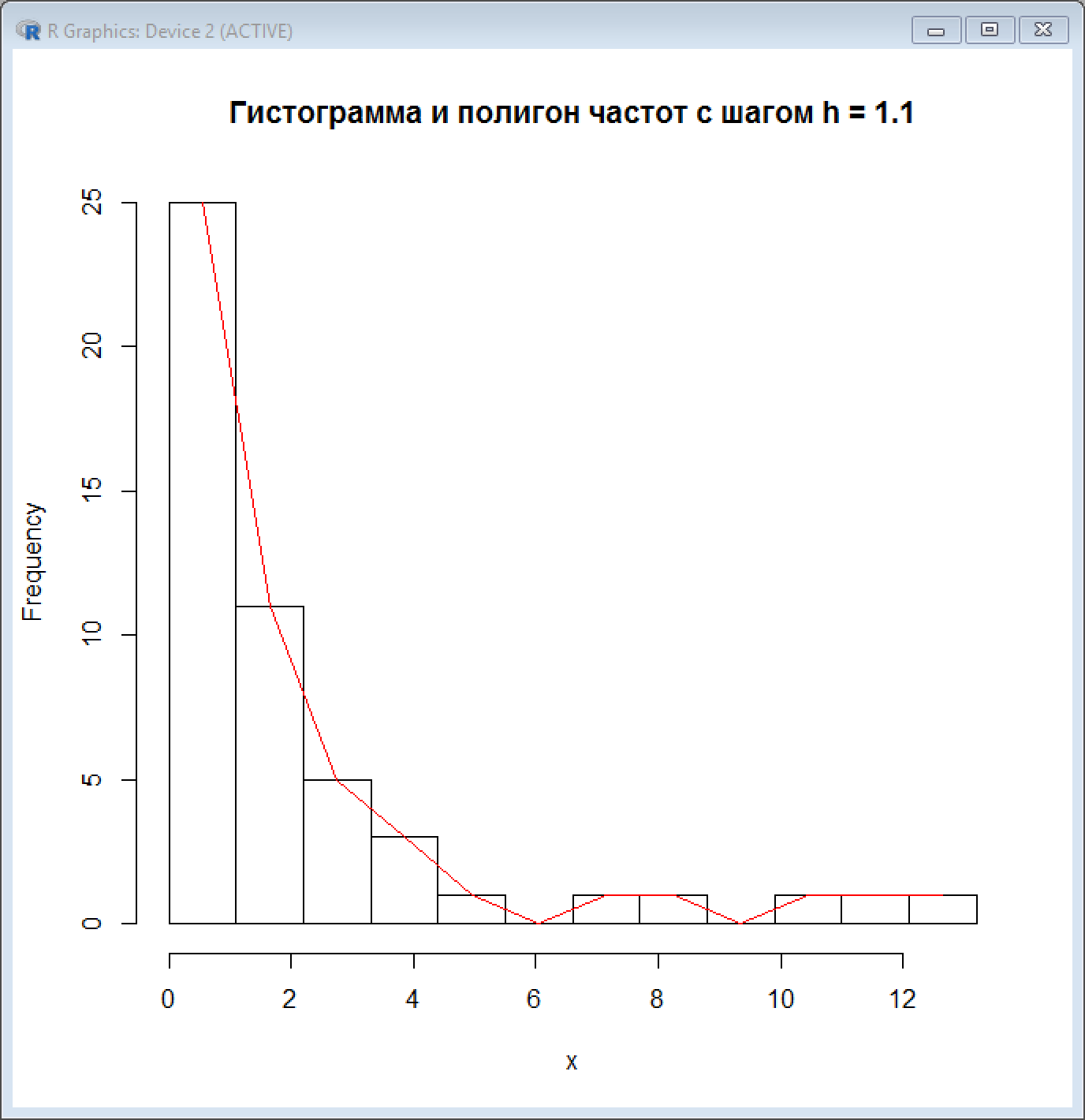


Рис.2.2. Гистограмма и полигон частот с шагом h=1.1 (п.2).

Построим таблицу частот для выборки , исходя из гистограммы.

Таблица 2.2

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Частоты | 25 | 11 | 5 | 3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 |
| Элементы выборки | 0-1.1 | 1.1-2.2 | 2.2-3.3 | 3.3-4.4 | 4.4-5.5 | 5.5-6.6 | 6.6-7.7 | 7.7-8.8 | 8.8-9.9 | 9.9-11 | 11-12.1 | 12.1-13.2 |

1. Вычислим выборочные аналоги следующих числовых характеристик
2. Математическое ожидание
3. Дисперсия
4. Медиана
5. Асимметрия
6. Эксцесс
7. Вероятность попадания в промежуток [c,d] при
8. В предположении, что исходные наблюдения являются выборкой показательного распределения, построим оценку максимального правдоподобия параметра , а также оценку по методу моментов. Найдем смещение оценок.

Плотность показательного распределения с параметром имеет вид

* Метод максимального правдоподобия

Функция правдоподобия будет равна

И её логарифм

Дифференцируя по , получим

Приравняем производную к нулю

Проверим результаты, используя следующий код

LL<-function(t){sum(dexp(x,t[1],log=TRUE))}

ml<-maxNR(LL,start=c(1))

ml$estimate

[1] 0.484637

ОМП была посчитана верно.

* По методу моментов

Математическое ожидание

Выборочный средний момент

И смещение оценки максимального правдоподобия

Значит, – несмещенная оценка.

1. Построим асимптотический доверительный интервал уровня доверия для параметра на базе оценки максимального правдоподобия.

Найдем вторую производную функции правдоподобия

Тогда наблюдаемая информация Фишера будет иметь вид

И выборочная дисперсия

Стандартная ошибка будет иметь вид

Значит, доверительный интервал с уровнем доверия будет иметь вид

Тогда, подставляя известные значения, получим доверительный интервал

1. С использованием теоремы Колмогорова построим критерий значимости проверки простой гипотезы согласия с показательным распределением с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Вычислим статистику , которая будет наблюдаемым значением критерия .

Найдем , используя следующие команды

v1<-sort(x); v2<-c(0:(n-1))/n; v3<-c(1:n)/n

v4<-abs(pexp(v1,0.26)-v2); v5<-abs(pexp(v1,0.26)-v3)

D<-max(v4,v5)

K <- sqrt(n)\*D

Итого

При помощи таблицы распределения Колмогорова (табл. 2.2) определим, согласуется ли гипотеза с выборкой

Распределение Колмогорова

Таблица 2.3.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *р* | 0,6 | 0,7 | 0,9 | 0,95 | 0,975 | 0,995 | 0,999 | 0,9995 |
|  | 0,89 | 0,97 | 1,22 | 1,36 | 1,48 | 1,73 | 1,95 | 2,03 |

Необходимое значение вероятности .

Тогда получим

Следовательно, отвергаем гипотезу , что выборка принадлежит показательному распределению, и принимаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу очень мало .

1. Используя гистограмму частот, построим критерий значимости проверки простой гипотезы согласия с показательным распределением с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Тогда искомая вероятность примет вид

Найдем вероятности , используя следующий код в R

a1=0.01

n=50

r=12

p=array(dim=r)

nu = h1$counts

b = h1$breaks

l0=0.26

p[1] = pexp(b[2],l0)

p[2:(r-1)]=pexp(b[3:(r)],l0) - pexp(b[2:(r-1)],l0)

p[r] = 1 - sum(p[1: (r-1)])

v1 = (nu-p)/n/p

v2 = v1^2

Tn = sum(v2)

xa = qchisq(1-a1,r-1)

pv = pchisq(xa,r-1)

Tn;xa

[1] 9.833166

[1] 24.72497

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.2487378 | 0.18686710 | 0.14038626 | 0.10546695 | 0.07923338 | 0.05952508 | 0.04471896 | 0.03359569 | 0.02523918 | 0.01896125 | 0.01424488 | 0.04302388 | 1 |
|  | 12.4368692 | 9.3433549 | 7.0193132 | 5.2733476 | 3.9616689 | 2.9762538 | 2.2359482 | 1.6797843 | 1.2619591 | 0.9480627 | 0.7122441 | 2.1511939 | 50 |
|  | 25 | 11 | 5 | 3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 50 |
|  | 3.96070578 | 1.33936021 | 0.47930753 | 0.30128966 | 0.05401853 | 0.00040000 | 0.18253193 | 0.33098681 | 0.00040000 | 1.07077489 | 1.91549218 | 0.19789876 |  |

Итого

Следовательно, принимаем гипотезу , что выборка принадлежит показательному распределению с параметром .

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу .

1. Построим критерий значимости проверки сложной гипотезы согласия с показательным распределением. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Оценим неизвестный параметр как

Тогда искомая вероятность примет вид

Найдем вероятности , используя следующий код в R

a1=0.01

n=50

r=12

p=array(dim=r)

nu = h1$counts

b = h1$breaks

l0=0.49

p[1] = pexp(b[2],l0)

p[2:(r-1)]=pexp(b[3:(r)],l0) - pexp(b[2:(r-1)],l0)

p[r] = 1 - sum(p[1: (r-1)])

v1 = (nu-p)/n/p

v2 = v1^2

Tn = sum(v2)

xa = qchisq(1-a1,r-2)

pv = pchisq(xa,r-2)

Tn;xa

[1] 212.9093

[1] 23.20925

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.416668708 | 0.243055896 | 0.141782110 | 0.082705941 | 0.048244964 | 0.028142797 | 0.016416574 | 0.009576301 | 0.005586156 | 0.003258580 | 0.001900832 | 0.002661142 | 1 |
|  | 20.83343540 | 12.15279479 | 7.08910549 | 4.13529706 | 2.41224818 | 1.40713985 | 0.82082871 | 0.47881507 | 0.27930781 | 0.16292899 | 0.09504158 | 0.13305709 | 50 |
|  | 25 | 11 | 5 | 3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 50 |
|  | 1.3923861 | 0.7834756 | 0.4696465 | 0.4976763 | 0.1556705 | 0.0004000 | 1.4358773 | 4.2786467 | 0.0004000 | 37.4255589 | 110.2859331 | 56.1835838 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит показательному распределению, и принимаем альтернативу.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу крайне мало .

1. Построим наиболее мощный критерий проверки простой гипотезы показательности с параметром при альтернативе показательности с параметром . Проверим гипотезу на уровне значимости . Что получится, если поменять местами основную и альтернативную гипотезы?
2. В пунктах (c)-(g) заменим семейство показательных распределений на семейство гамма-распределений.
3. В предположении, что исходные наблюдения являются выборкой из гамма-распределения, построим оценку максимального правдоподобия параметра , а также оценку по методу моментов. Найдем смещение оценок.

Плотность гамма-распределения с параметром имеет вид

* Метод максимального правдоподобия

Функция правдоподобия будет равна

И её логарифм

Дифференцируя по , получим

Приравняем производную к нулю

* По методу моментов

Математическое ожидание

Выборочный средний момент

И смещение оценки максимального правдоподобия

Значит, – несмещенная оценка.

1. Построим асимптотический доверительный интервал уровня доверия для параметра на базе оценки максимального правдоподобия.

Найдем вторую производную функции правдоподобия

Тогда наблюдаемая информация Фишера будет иметь вид

И выборочная дисперсия

Стандартная ошибка будет иметь вид

Значит, доверительный интервал с уровнем доверия будет иметь вид

Тогда, подставляя известные значения, получим доверительный интервал

1. С использованием теоремы Колмогорова построим критерий значимости проверки простой гипотезы согласия с показательным распределением с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Вычислим статистику , которая будет наблюдаемым значением критерия .

Найдем , используя следующие команды

v1<-sort(x); v2<-c(0:(n-1))/n; v3<-c(1:n)/n

v4<-abs(pchisq(v1,1,0.26)-v2); v5<-abs(pchisq(v1,1,0.26)-v3)D<-max(v4,v5)

K <- sqrt(n)\*D

Итого

При помощи таблицы распределения Колмогорова (табл 2.2) определим, согласуется ли гипотеза с выборкой

Распределение Колмогорова

Таблица 2.2.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *р* | 0,6 | 0,7 | 0,9 | 0,95 | 0,975 | 0,995 | 0,999 | 0,9995 |
|  | 0,89 | 0,97 | 1,22 | 1,36 | 1,48 | 1,73 | 1,95 | 2,03 |

Необходимое значение вероятности .

Тогда получим

Следовательно, принимаем гипотезу , что выборка принадлежит гамма-распределению.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу .

1. Используя гистограмму частот, построим критерий значимости проверки простой гипотезы согласия с гамма-распределением с параметром. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Тогда искомая вероятность примет вид

Найдем вероятности , используя следующий код R

a1=0.01

n=50

r=12

p=array(dim=r)

nu = h1$counts

b = h1$breaks

l0=0.26

p[1] = pchisq(b[2],1,l0)

p[2:(r-1)]= pchisq(b[3:(r)], 1,l0) - pchisq(b[2:(r-1)], 1,l0)

p[r] = 1 - sum(p[1: (r-1)])

v1 = (nu-p)/n/p

v2 = v1^2

Tn = sum(v2)

xa = qchisq(1-a1,r-1)

pv = pchisq(xa,r-1)

Tn;xa

[1] 698.4431

[1] 24.72497

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.645492029 | 0.166192268 | 0.082660245 | 0.044920101 | 0.025354073 | 0.014602425 | 0.008511779 | 0.004999587 | 0.002951560 | 0.001748537 | 0.001038355 | 0.001529041 | 1 |
|  | 32.27460143 | 8.30961340 | 4.13301224 | 2.24600506 | 1.26770367 | 0.73012125 | 0.42558895 | 0.24997936 | 0.14757798 | 0.08742685 | 0.05191776 | 0.07645205 | 50 |
|  | 25 | 11 | 5 | 3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 50 |
|  | 0.5694255 | 1.6998109 | 1.4155560 | 1.7310794 | 0.5910963 | 0.0004000 | 5.4274325 | 15.8430295 | 0.0004000 | 130.3737802 | 370.2250115 | 170.5660583 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит гамма-распределению с параметром .

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу крайне мало .

1. Построим критерий значимости проверки сложной гипотезы согласия с гамма-распределением. Проверим гипотезу на уровне значимости. Вычислим наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.

Простая гипотеза имеет вид

Оценим неизвестный параметр как

Тогда искомая вероятность примет вид

Найдем вероятности , используя следующий код в R

a1=0.01

n=50

r=12

p=array(dim=r)

nu = h1$counts

b = h1$breaks

l0=0.3

p[1] = pchisq(b[2],1,l0)

p[2:(r-1)]= pchisq(b[3:(r)], 1,l0) - pchisq(b[2:(r-1)], 1,l0)

p[r] = 1 - sum(p[1: (r-1)])

v1 = (nu-p)/n/p

v2 = v1^2

Tn = sum(v2)

xa = qchisq(1-a1,r-2)

pv = pchisq(xa,r-2)

Tn;xa

[1] 252.1112

[1] 23.20925

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Группы |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 0.636659703 | 0.167449958 | 0.084614426 | 0.046612129 | 0.026626152 | 0.015500288 | 0.009123316 | 0.005406665 | 0.003218229 | 0.001921165 | 0.001149084 | 0.001718884 | 1 |
|  | 31.83298514 | 8.37249791 | 4.23072132 | 2.33060645 | 1.33130758 | 0.77501441 | 0.45616582 | 0.27033327 | 0.16091143 | 0.09605824 | 0.05745422 | 0.08594422 | 50 |
|  | 25 | 11 | 5 | 3 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 50 |
|  | 0.2928795 | 0.8369920 | 0.6749263 | 0.8029215 | 0.2672837 | 0.0002000 | 2.3591944 | 6.7680273 | 0.0002000 | 53.9797002 | 151.1219309 | 67.4593801 |  |

Итого

Следовательно, отвергаем гипотезу , что выборка принадлежит гамма-распределению с параметром.

Наибольшее значение уровня значимости , при котором ещё можно принять гипотезу крайне мало .