Санкт – Петербургский государственный электротехнический университет

«ЛЭТИ»

Отчет

по индивидуальному домашнему заданию № 2

по дисциплине «Вычислительная математика»

Вариант № 13

Выполнил: Сарсенов М.А.

факультет: КТИ

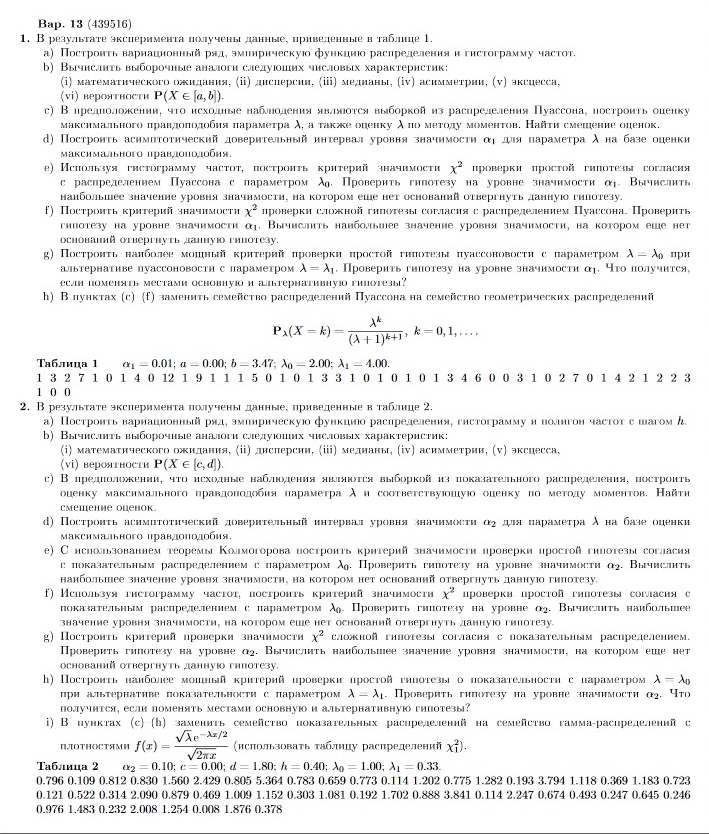
группа: 4395

Преподаватель:

Санкт - Петербург

2016

Постановка задания:



Решение

Задание 1

a) Построить вариационный ряд, эмпирическую функцию распределения гистограмму частот.

setwd("E:\\МатСтат") #Устанавливаем рабочую директорию

x<-scan ("Viborka.txt") #Вводим в программу исходные данные:

print (x)

1 3 2 7 1 0 1 4 0 12 1 9 1 1 1 5 0 1 0 1 3 3 1 0 1 0 1 0 1 3 4 6 0 0 3 1 0 2 7 0 1 4 2 1 2 2 3 1 0 0

sortVec<-sort(x); print(sortVec) #Получаем вариационный ряд:

0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 3 4 4 4 5 6 7 7 9 12

В дальнейшем может быть полезной информация о значениях элементов в выборке и их количестве. Воспользуемся командой

table(x)

x

0 1 2 3 4 5 6 7 9 12

13 17 5 6 3 1 1 2 1 1

Эмпирическая функция распределения имеет вид: 

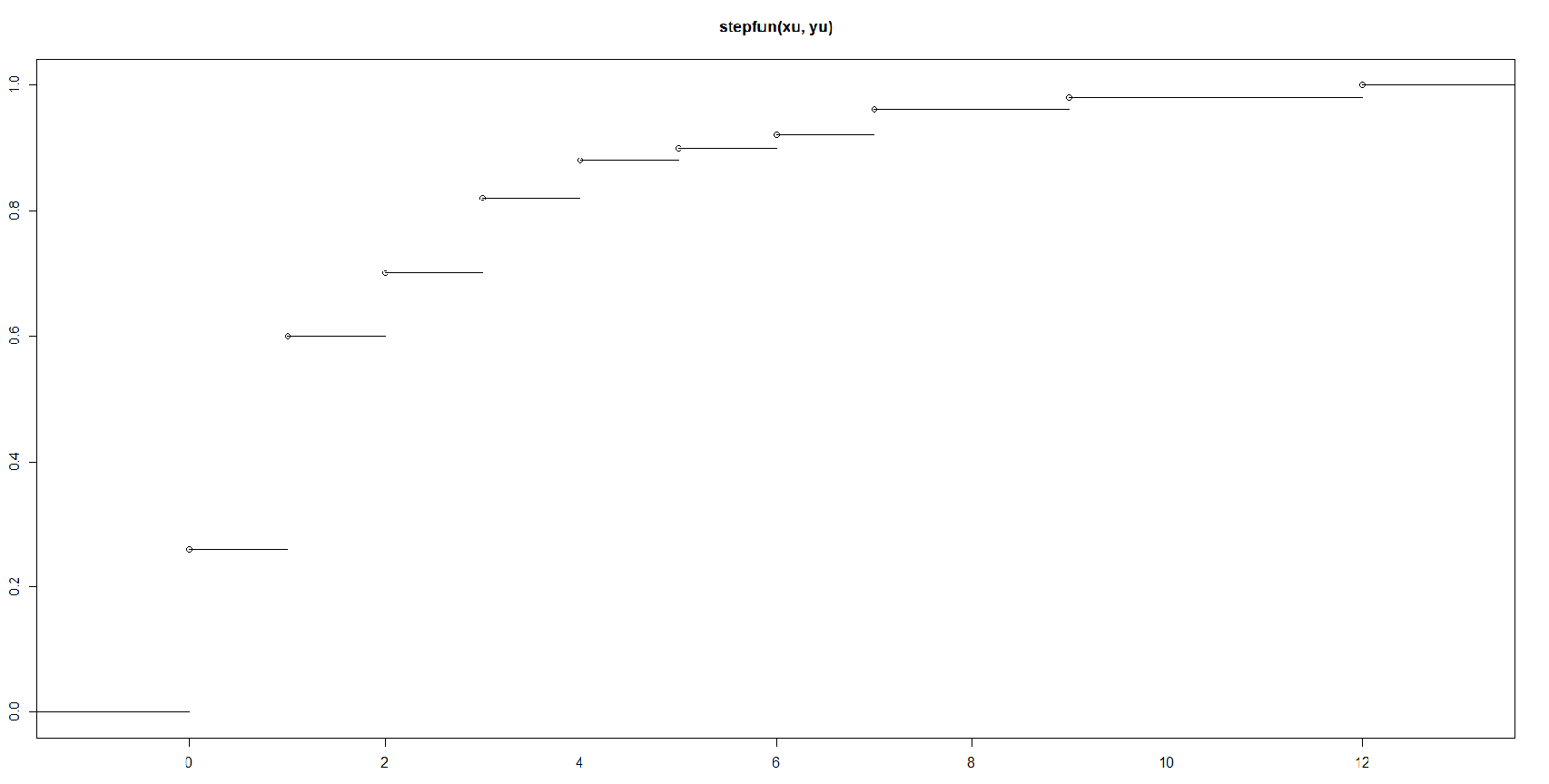


График эмпирической функции распределения

*Для построения эмпирической функции распределения:*

f<-function(x,t){z<-x[x<t]; length(z)/length(x)}

xu<-unique(sort(x))

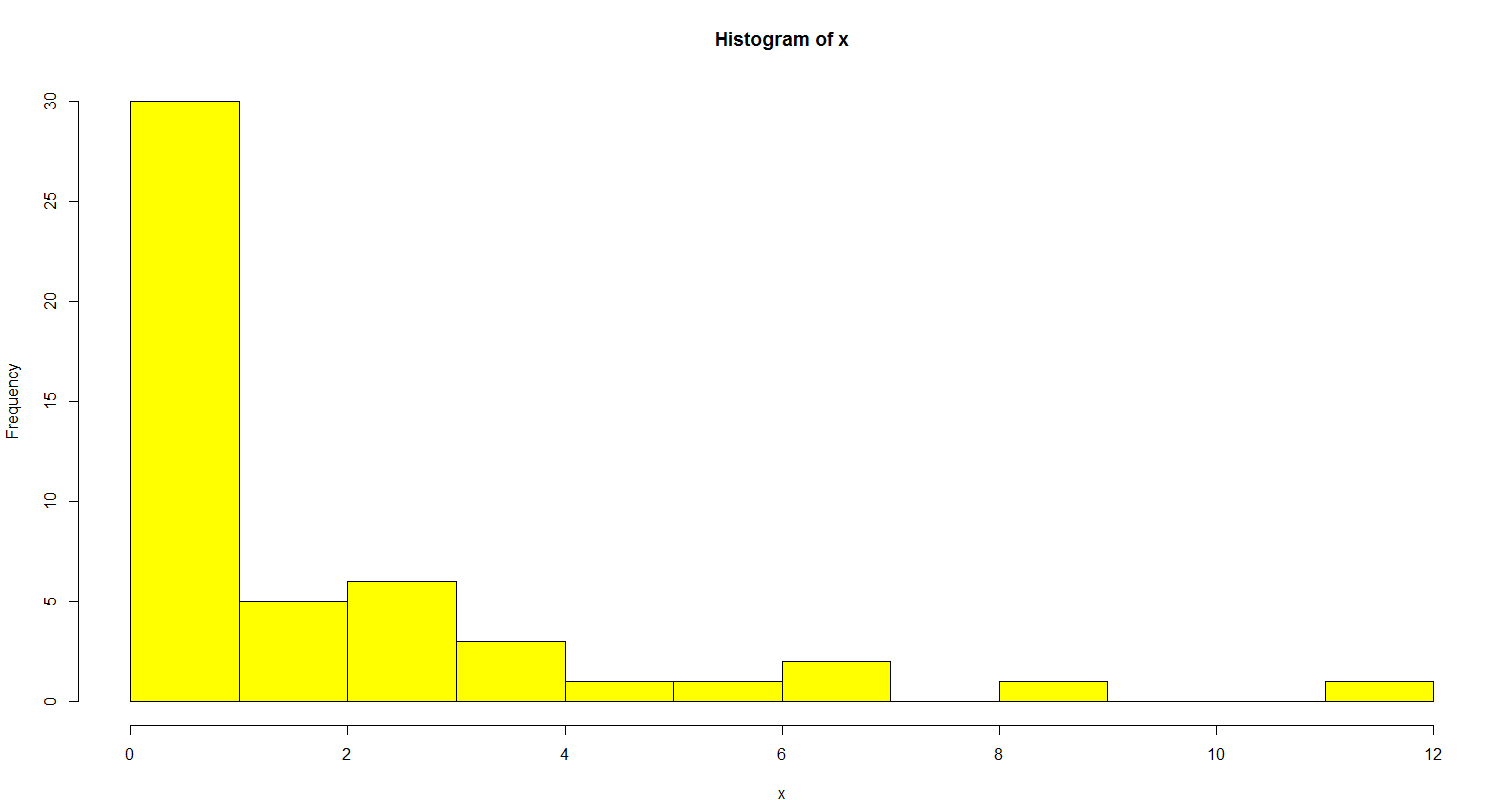
yu<-0; for(i in 1:length(xu)) yu[i]<-f(x,xu[i]); yu[length(xu)+1]<-1

z<-stepfun(xu,yu)

plot(z,verticals=FALSE)

*Гистограмма частот:*

hist(x,breaks=c(0:max(x)),col = "yellow",right = TRUE, freq = TRUE)



*b) Вычислить выборочные аналоги следующих числовых характеристик:*

1) математического ожидания: выборочное среднее 

mean<-sum(x)/length(x)

mean = 2.06

2) дисперсии: - выборочная дисперсия 

var<-sum(x^2)/length(x)-mean^2

var = 6.2164

3) медианы: - выборочная медиана, равная выборочной квантили порядка p: при p = 1/2

med<-sort(x)[trunc(length(x)\*1/2+1)]

med = 1

4) асимметрии – выборочная асимметрия: 

asm<-sum((x-mean)^3)/length(x)/var^(3/2)

asm = 1.983386

5) экцесса - выборочный эксцесс: 

exc<-sum((x-mean)^4)/length(x)/var^2-3

exc = 4.189488

6) вероятность попадания в заданный промежуток: P(X ∈ [a, b])

a<-0.00; b<-3.47

p<-f(x,b)-f(x,a)

p = 0.82

*с) В предположении, что исходные наблюдения являются выборкой из распределения Пуассона, построить оценку максимального правдоподобия параметра λ, а также оценку λ по методу моментов. Найти смещение оценок.*

library(maxLik) #Подключаем библиотеку

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

ml<-maxNR(LL,start=c(1)) #максимум функции правдоподобия

val<-ml$estimate; print (val) #оценка макс.правдоподобия

val = 2.06 # равно выборочному среднему (mean)

Аналогичный результат дает теория:  - плотность распределения Пуассона.

Метод максимального правдоподобия:

=>  => 

 => 

Метод моментов:

Математическое ожидание: **,** выборочный средний момент: **=**>

, значит - несмещенная оценка.

*d) Построить асимптотический доверительный интервал уровня значимости = 0.01 для параметра λ на базе оценки максимального правдоподобия.*

По методу максимального правдоподобия:

, 

Эксперимент регулярен, значит, подстановка ОМП вместо параметра в информацию Фишера не нарушает асимптотической нормальности.



;





где  - квантиль порядка стандартного закона распределения.

.



al<-0.01

xal<-qnorm(1-al/2)

T<-array(dim=2)

T[1]<-mean-xal\*sqrt(mean/length(x))

T[2]<-mean+xal\*sqrt(mean/length(x))

print(T) #Доверительный интервал

полученный результат: [1.537164, 2.582836]

*e) Используя гистограмму частот, построить критерий значимости χ2 проверки простой гипотезы согласия с распределением Пуассона с параметром λ0. Проверить гипотезу на уровне значимости α1. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.*

Простая гипотеза Hо : , *λ*o = 2.00

Таблица «значение-частота» имеет вид:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 12 |
| 13 | 17 | 5 | 6 | 3 | 1 | 1 | 2 | 1 | 1 |

Делим последовательность на r = 3 интервала и вводим вектор border границ интервалов, имеющий размерность r-1, который потребуется для получения значений частот из гистограммы: border<-c(0,2)

Нижние границы элементов: а1<-c(-Inf, 0, 1); Верхние границы элементов b1<-c(0, 1, Inf)

Число наблюдений, попавших в этот интервал: nu = 13, 17, 20 (берутся из гистограммы)



Критерий имеет вид:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № интервала |  |  |  |  |  |  |
| 1 | -Inf | 0 | 13 | 0.1353353 | 2.3961999 | 5.7417738 |
| 2 | 0 | 1 | 17 | 0.2706706 | 0.9422847 | 0.8879004 |
| r=3 | 1 | Inf | 20 | 0.5939942 | -1.7798465 | 3.1678536 |
|  | | | | =1 | = 9.797528 | |

n<-length(x); lambda0<-2.00 ; r<-3

a1<-c(-Inf, 0, 1); b1<-c(0, 1, Inf)

border <-c(0, 1) #общий массив границ интервалов

h<-hist(x,breaks=c(min(x),border,max(x)),plot=FALSE)

nu<-h$counts; print (nu) #частоты элементов

p1<-array(dim = r)

p1[1]<- ppois(border[1], lambda0)

p1[r] <- 1-ppois(border[r-1], lambda0)

p1[2:(r-1)]<-ppois(border[2:(r-1)],lambda0)-ppois(border[1:(r-2)],lambda0)

res <- array (dim = r)

res [1:r] <- (nu[1:r] - n\*p1[1:r])/sqrt(n\*p1[1:r])

res2 <- array (dim = r)

res2 [1:r]<- (res[1:r])^2

Xi2<-sum(res2)

xal<-qchisq(1-al, r-1)

Xi2>xal #TRUE

Итак, получили, что χ2 > xα, следовательно, нужно отвергнуть гипотезу H0

Для нахождения наибольшего значения уровня значимости, на котором ещё нет оснований отвергнуть данную гипотезу, вычисляем функцию распределения  в точке , и вычитаем полученное значение из единицы:

al2<-1-pchisq(Xi2,r-1); al2 #находим наибольший уровень значимости, при котором еще нет #оснований отвергнуть гипотезу:

Получили: al2 = 0.007455793

*f) Построить критерий значимости χ2 проверки сложной гипотезы согласия с распределением Пуассона. Проверить гипотезу на уровне значимости α1. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.*

Но – основная гипотеза: Х ~ Pois ()

Поделим область на *r*=3 интервалов, аналогично предыдущему пункту.

Х2()=- зависит от , т.к. величины *pi* не фиксированы. Известно, что в случае регулярности эксперимента статистика  сходится по распределению к **

Критерий имеет вид:

Далее проведем вычисления в R

#Данные, которые были определены ранее:

#a1<-c(-Inf, 0, 1); b1<-c(0, 1, Inf); nu<-c(13,17,20)

#al<-0.01; r<-3; n<-50

csq<-function (t){ #функция для χ2

p<-pnorm(b1,0,t) - pnorm (a1,0,t);

f<-sum((nu-n\*p)^2/(n\*p));

print (f)

}

X2<-nlm(csq,p=mean(x)) #стандартный минимизатор

xal1<-qchisq (1-al, r-2)

X2$minimum<=xal1 #производим сравнение

[1] FALSE

Итак, , и гипотеза отвергается

Наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу:

alpha2<-1-pchisq(X2$minimum,r-2)

print (alpha2)

Получаем: alpha2 = 0.0006885139

*g) Построить наиболее мощный критерий проверки простой гипотезы пуассоновости с параметром λ = λ0 при альтернативе пуассоновости с параметром λ = λ1. Проверить гипотезу на уровне значимости α1. Что получится, если поменять местами основную и альтернативную гипотезы?*



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

 , где 





Наиболее мощный критерий:

Логарифмируем соотношение

Получим:

Обозначим через:

Тогда критерий:

Отыщем 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 

Наиболее мощный критерий:

Проведём вычисления в R.  
c<-0

alpha1<-0.01

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)

}

с

[1] 123

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

p

[1] 0.5939768

alpha0

[1] 0.008180814

sum(x)

[1] 103

lche<- sum(x)

lche>=c

[1] FALSE #принимаем 

Критерий:



Теперь поменяем местами основную и альтернативную гипотезу.









Наиболее мощный критерий:

Логарифмируем соотношение

Получим:

Обозначим через:

Тогда критерий:

Отыщем и *p* из уравнения:



 , следовательно, 

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

Тогда с учётом уравнения выше 

Проведём вычисления в R.

c<-0; lambda1 <-4

alpha0<-ppois(c,lambda1\*n)

while(alpha0<alpha1){

c<-c+1;

alpha0<-ppois(c,lambda1\*n)

}

c<-c-1

c

[1] 167

alpha0<-ppois(c,lambda1\*n)

p<-(alpha1-alpha0)/(dpois(c,lambda1\*n))

alpha0

[1] 0.009315414

P

[1] 0.3976056

sum(x)

[1] 103

lche<- sum(x)

lche<=c

[1] TRUE #принимаем альтернативу

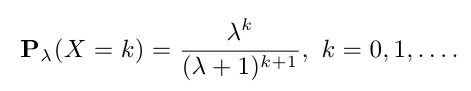
Критерий построен:



При замене основной и альтернативной гипотезы меняется также гипотеза, которую принимаем. Но т.к изменение происходит от 

на , то решение принимается в пользу одной и той же гипотезы: 

***h) В пунктах (c)-(f) заменить семейство распределений Пуассона на семейство геометрических распределений***





Обозначим 

Найдём оценку максимального правдоподобия:



Для геометрического распределения математическое ожидание: ****, выборочный средний момент: ** =**> .

, значит  - несмещенная оценка.

library(maxLik)

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

ml<-maxNR(LL,start=c(1)) #максимум функции правдоподобия

val<-ml$estimate; print (val) #оценка макс.правдоподобия

[1] 2.06

***Построить асимптотический доверительный интервал уровня значимости* *для параметра λ на базе оценки максимального правдоподобия.***



Найдём информацию Фишера.



ОМП параметра :





Эксперимент регулярен, значит, подстановка ОМП вместо параметра в информацию Фишера не нарушает асимптотической нормальности.







Где *,* т.е. *b* - квантиль стандартного нормального распределения.





alpha<-0.01

T<-array(dim=2)

xal<-qnorm (1-alpha/2)

T[1]<-mean-xal\*sqrt((mean\*(mean+1))/length(x)) #левая граница Д.И.

T[2]<-mean+xal\*sqrt((mean\*(mean+1))/length(x)) #правая граница Д.И.

[1] 1.14541 2.97459

Ответ: асимптотический доверительный интервал для параметра  уровня доверия 

[1.14541, 2.97459]

*Используя гистограмму частот, построить критерий значимости  проверки простой гипотезы согласия с геометрическим распределением с параметром. Проверить гипотезу на уровне значимости . Вычислить наибольшее значение уровня значимости, на котором ещё нет основания отвергнуть данную гипотезу.*

Простая гипотеза Hо: , *λ*o=2.00

Таблица «значение-частота» имеет вид:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 9 | 12 |
| 13 | 17 | 5 | 6 | 3 | 1 | 1 | 2 | 1 | 1 |

Делим последовательность на r = 3 интервала. Границы интервалов берем такие же, как

Число наблюдений, попавших в этот интервал: nu = 13, 17, 20 (берутся из гистограммы)



Критерий имеет вид

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № интервала |  |  |  |  |  |  |
| 1 | -Inf | 0 | 13 | 0.3333333 | -0.8981462 | 0.8066667 |
| 2 | 0 | 1 | 17 | 0.2222222 | 1.7666667 | 3.1211111 |
| r=3 | 1 | Inf | 20 | 0.4444444 | -1.341762 | 0.2222222 |
|  | | | | =1 | = 4.15 | |

r<-3 #количество интервалов

a<-2.00 #lambda0

b<-array(dim=r-1) #вектор границ

b[1]<-0; b[2]<-1;

h<-hist(x,breaks=c(min(x),b,max(x)),plot=FALSE) #построение гистограммы

p<-array(dim=3) #вектор теоретических вероятностей

p[1]<-pgeom(b[1],1/(a+1))

p[2]<-pgeom(b[2],1/(a+1))-pgeom(b[1],1/(a+1))

p[3]<-1-pgeom(b[2],1/(a+1))

nu<-h$counts #получение вектора частот

v1<-(nu-n\*p)^2/(n\*p) #вектор слагаемых величины X2

X2<-sum(v1) #вычисление величины X2

xa<-qchisq(1-alpha,2) #вычисление квантиля

X2>xa

[1] FALSE

alpha2<-1-pchisq(X2,2) #находим наибольший уровень значимости, при

alpha2 #котором нет оснований отвергнуть гипотезу

[1] 0.1255564

Итак, получили, что χ2 < xα, следовательно, нужно принять гипотезу Hо.

Для нахождения наибольшего значения уровня значимости, на котором ещё нет оснований отвергнуть данную гипотезу, вычисляем функцию распределения  в точке , и вычитаем полученное значение из единицы:

alpha2<-1-pchisq(X2,2); alpha2 #находим наибольший уровень значимости, при котором еще нет #оснований отвергнуть гипотезу:

Получили: al2 = 0.007455793

***Построить критерий значимости χ2 проверки сложной гипотезы согласия с геометрическим распределением. Проверить гипотезу на уровне значимости α1. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.***



Сложная гипотеза согласия: Но – основная гипотеза: Х ~ Geom (1/(+1))

Поделим область на *r*=3 интервалов, аналогично предыдущему пункту. *X*2- зависит от , т.к. величины *pi* не фиксированы. Известно, что в случае регулярности эксперимента статистика  сходится по распределению к **.

Критерий имеет вид:

Далее проведем вычисления в R:

> P<-function(a){

+ p[1]<-pgeom(b[1],1/(a+1))

+ i<-2

+ while(i<r){

+ p[i]<-pgeom(b[i],1/(a+1))-pgeom(b[i-1],1/(a+1));

+ i<-i+1;

+ }

+ p[r]<-1-pgeom(b[r-1],1/(a+1))

+ p;}

> X2<-function(a){g<-n\*P(a);f<-(nu-g)^2/g;sum(f)}

> XM<-nlm(X2,p=mean) #проводим минимизацию,

> xb<-qchisq(1-al,r-2) #вычисляем квантиль

> XM$minimum<xb # гипотезу принимаем на заданном уровне знач.

[1] TRUE

> alpha2<-1-pchisq(XM$minimum,r-2) #наибольший уровень значимости, на котором еще нет

оснований отвергнуть данную гипотезу

> alpha2

[1] 0.04568443

Следовательно, нужно принять гипотезу Hо.

**Задание 2**

1. *Построить вариационный ряд, эмпирическую функцию распределения, гистограмму и полигон частот с шагом h.*

> x<-scan ("VIBORKA2.txt"); print (x) #прочитать и вывести данные из файла

Read 50 items

[1] 0.796 0.109 0.812 0.830 1.560 2.429 0.805 5.364 0.783 0.659 0.773 0.114

[13] 1.202 0.775 1.282 0.193 3.794 1.118 0.369 1.183 0.723 0.121 0.522 0.314

[25] 2.090 0.879 0.469 1.009 1.152 0.303 1.081 0.192 1.702 0.888 3.841 0.114

[37] 2.247 0.674 0.493 0.247 0.645 0.246 0.976 1.483 0.232 2.008 1.254 0.008

[49] 1.876 0.378

> sortVec<-sort(x); print (sortVec) #вариационный ряд

[1] 0.008 0.109 0.114 0.114 0.121 0.192 0.193 0.232 0.246 0.247 0.303 0.314

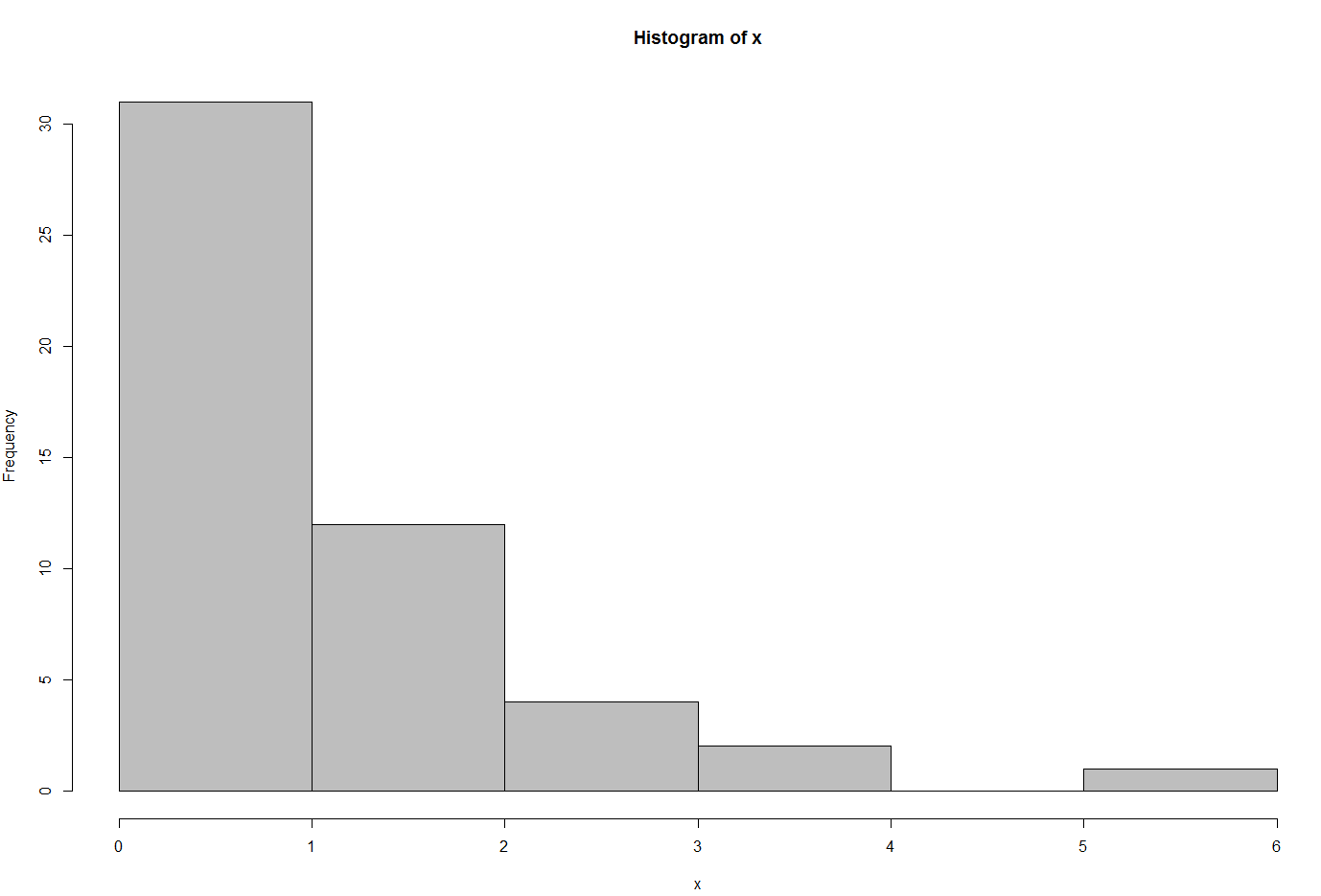
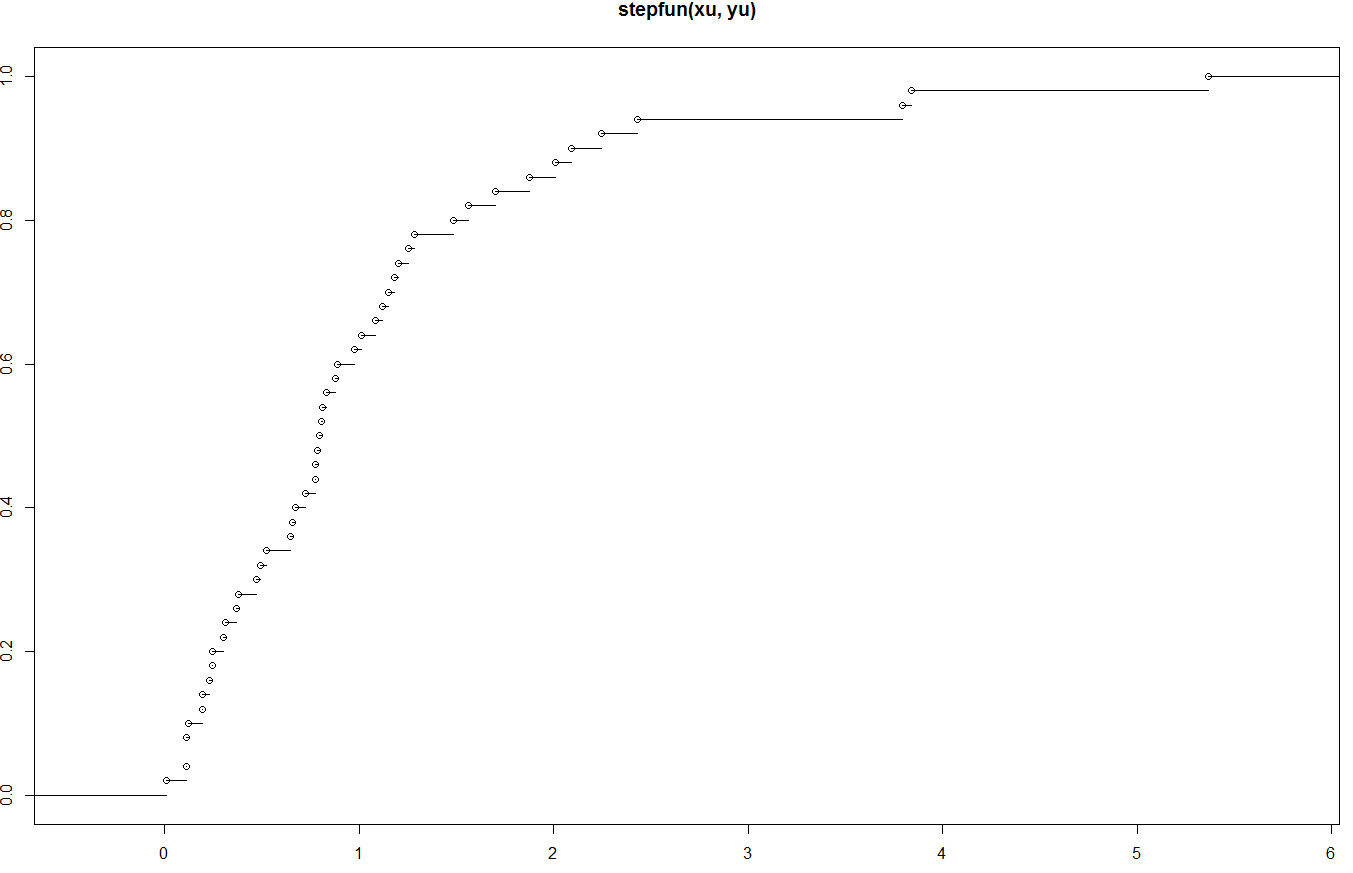
[13] 0.369 0.378 0.469 0.493 0.522 0.645 0.659 0.674 0.723 0.773 0.775 0.783

[25] 0.796 0.805 0.812 0.830 0.879 0.888 0.976 1.009 1.081 1.118 1.152 1.183

[37] 1.202 1.254 1.282 1.483 1.560 1.702 1.876 2.008 2.090 2.247 2.429 3.794

[49] 3.841 5.364

Для построения эмпирической функции распределения и гистограммы частот используем код:



*b) Вычислить выборочные аналоги следующих числовых характеристик:*

mean<-sum(x)/length(x); print (mean) #выборочное среднее

[1] 1.06234

var<-sum(x^2)/length(x)-mean^2; print (var) #выборочная дисперсия

[1] 1.060067

med<-sort(x)[trunc(length(x)\*1/2+1)]; print (med) #медиана - выборочная квантиль порядка 1/2

[1] 0.805

asm<-sum((x-mean)^3)/length(x)/var^(3/2); print (asm) #выборочная асимметрия

[1] 2.173395

exc<-sum((x-mean)^4)/length(x)/var^2-3; print (exc) #выборочный эксцесс

[1] 5.441274

c<-0.00; d<-1.80

> p<-f(x,d)-f(x,c); print (p) #вероятность попадания в промежуток

[1] 0.84

***с) В предположении, что исходные наблюдения являются выборкой из показательного распределения, построить оценку максимального правдоподобия параметра λ и соответствующую оценку по методу моментов. Найти смещение оценок.***

Оценка по методу моментов:

Рассмотрим:

Несмещенная оценка:

***d) Построить асимптотический доверительный интервал уровня значимости α2 для параметра λ на базе оценки максимального правдоподобия.*** Найдём информацию Фишера.



ОМП параметра :



Эксперимент регулярен, значит, подстановка ОМП вместо параметра в информацию Фишера не нарушает асимптотической нормальности.





, где

*,* т.е. *b* - квантиль стандартного нормального распределения.



Итак, 

al<-0.10

n<-length(x)

xal<-qnorm (1-al/2)

T<-array(dim=2)

T[1]<-1/mean-xal\*(1/mean)/(sqrt(n))

T[2]<-1/mean+xal\*(1/mean)/(sqrt(n))

T

[1] 0.3178645 0.6821455

Итак, получили интервал: [0.7223512, 1.1602852]

***e) С использованием теоремы Колмогорова построить критерий значимости проверки простой гипотезы согласия с показательным распределением с параметром λ0. Проверить гипотезу на уровне значимости α2. Вычислить наибольшее значение уровня значимости, на котором нет оснований отвергнуть данную гипотезу.***

Ho: ,

Критерий:

К> Xα, значит, отвергаем гипотезу

> lambda0<- 1.00

> y<-sort(x)

> v2<-c(0:(n-1))/n

> v3<-c(1:n)/n

> v4<-pexp(y,lambda0)

> v5<-abs(v2-v4)

> v6<-abs(v3-v4)

> v7<-pmax(v5,v6)

> D<-max(v7)

> print(D)

[1] 0.1353375

> K<-sqrt(n)\*D

> print(K)

[1] 0.9569803

***f) Используя гистограмму частот, построить критерий значимости χ2 проверки простой гипотезы согласия с показательным распределением с параметром λ0. Проверить гипотезу на уровне α2. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.***

Простая гипотеза Hо: *λ*o=1.00

x

0.008 0.109 0.114 0.121 0.192 0.193 0.232 0.246 0.247 0.303 0.314 0.369 0.378

1 1 2 1 1 1 1 1 1 1 1 1 1

0.469 0.493 0.522 0.645 0.659 0.674 0.723 0.773 0.775 0.783 0.796 0.805 0.812

1 1 1 1 1 1 1 1 1 1 1 1 1

0.83 0.879 0.888 0.976 1.009 1.081 1.118 1.152 1.183 1.202 1.254 1.282 1.483

1 1 1 1 1 1 1 1 1 1 1 1 1

1.56 1.702 1.876 2.008 2.09 2.247 2.429 3.794 3.841 5.364

1 1 1 1 1 1 1 1 1 1

Делим последовательность на r = 5 интервала.

Нижние границы элементов: a1<- c(-Inf, 0.493, 0.976);

Верхние границы элементов: b1<-c(0.469, 0.888, Inf)

Число наблюдений, попавших в этот интервал: nu = (берутся из гистограммы)



Критерий имеет вид:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № интервала |  |  |  |  |  |  |
| 1 | -Inf | 0.247 | 10 | 0.21885931 | -0.2850547 | 0.08125618 |
| 2 | 0.247 | 0.674 | 10 | 0.27147486 | -0.9700040 | 0.94090772 |
| 3 | 0.674 | 0.888 | 10 | 0.5939942 | 2.2974979 | 5.27849644 |
| 4 | 0.888 | 1.483 | 10 | 0.09818795 | 0.2547846 | 0.06491520 |
| r=5 | 1.483 | Inf | 10 | 0.18452209 | -0.4000982 | 0.16007855 |
|  | | | | =1 | = 6.525654 | |

n<-length(x); lambda0<-1; r<-5; alpha2<-0.1

border <-c(0.247, 0.674, 0.888, 1.483) #общий массив границ интервалов

h<-hist(x,breaks=c(min(x),border,max(x)),plot=FALSE)

nu<-h$counts; print (nu) #частоты элементов

p1<-array(dim = r)

p1[1]<- pexp(border[1], lambda0)

p1[r] <- 1-pexp(border[r-1], lambda0)

p1[2:(r-3)]<-pexp(border[2:(r-3)],lambda0)-pexp(border[1:(r-4)],lambda0)

p1[3:(r-2)]<-pexp(border[3:(r-2)],lambda0)-pexp(border[2:(r-3)],lambda0)

p1[4:(r-1)]<-pexp(border[4:(r-1)],lambda0)-pexp(border[3:(r-2)],lambda0)

res <- array (dim = r)

res [1:r] <- (nu[1:r] - n\*p1[1:r])/sqrt(n\*p1[1:r])

res2 <- array (dim = r)

res2 [1:r]<- (res[1:r])^2

Xi2<-sum(res2)

xal<-qchisq(1-alpha2, r-1)

Xi2>xal #FALSE

#находим наибольший уровень значимости, при котором нет оснований отвергнуть гипотезу:

al2<-1-pchisq(Xi2,r-1); al2

Итак, получили, что χ2 > xα, следовательно, нужно принять гипотезу Hо.

наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу

al2<-1-pchisq(Xi2,r-1); al2

Получили: 0.1631811

***g) Построить критерий проверки значимости χ2 сложной гипотезы согласия с показательным распределением. Проверить гипотезу на уровне α2. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.***

Сложная гипотеза согласия: Но – основная гипотеза:

Поделим область на *r*=5 интервалов, аналогично предыдущему пункту. *X*2- зависит от , т.к. величины *pi* не фиксированы. Известно, что в случае регулярности эксперимента статистика  сходится по распределению к **.

Критерий имеет вид:

Далее проведем вычисления в R:

al<-0.1

r<-5

P<-function(a){

p[1]<-pexp(b[1],a)

i<-2

while(i<r){

p[i]<-pexp(b[i],a)-pexp(b[i-1],a);

i<-i+1;

}

p[r]<-1-pexp(b[r-1],a)

p;}

X2<-function(a){g<-n\*P(a);f<-(nu-g)^2/g;sum(f)}

XM<-nlm(X2,p=mean) #проводим минимизацию,

xb<-qchisq(1-al,r-2) #вычисляем квантиль

XM$minimum<xb

[1] FALSE

То есть нужно отвергнуть гипотезу Но

Наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу:

alpha2<-1-pchisq(XM$minimum,r-2) #наибольший уровень значимости, на котором

alpha2 #нет оснований отвергнуть гипотезу

[1] 0

**&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&**

***h) Построить наиболее мощный критерий проверки простой гипотезы о показательности с параметром λ = λ0 при альтернативе показательности с параметром λ = λ1. Проверить гипотезу на уровне значимости α2. Что получится, если поменять местами основную и альтернативную гипотезы?***

Построение наиболее мощного критерия.

H0: λ=λ0=1.00

HA: λ=λ1=0.33

Критерий:

**&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&&**

***i) В пунктах (c)-(h) заменить семейство показательных распределений на семейство гамма-распределений с плотностями***

Оценка по методу моментов:

Смещение оценок:

*Построить асимптотический доверительный интервал уровня значимости α2 для параметра λ на базе оценки максимального правдоподобия.*

> n<-length(x)

> xal<-qnorm (1-al/2)

>T<-array(dim=2)

> T[1]<-(mean+xal\*mean\*sqrt(2)/sqrt(n))^(-1)

> T[2]<-(mean-xal\*mean\*sqrt(2)/sqrt(n))^(-1)

> T

[1] 0.7083062 1.4027975

Итак, д.и равен: [0.7083062, 1.4027975]

***С использованием теоремы Колмогорова построить критерий значимости проверки простой гипотезы согласия с гамма распределением с параметром λ0. Проверить гипотезу на уровне значимости α2. Вычислить наибольшее значение уровня значимости, на котором нет оснований отвергнуть данную гипотезу.***

Ho: ,

Критерий:

К> Xα, значит, принимаем гипотезу

lambda0<- 1.00

y<-sort(x)

v2<-c(0:(n-1))/n

v3<-c(1:n)/n

v4<-pgamma(y,1/2,2/lambda0)

v5<-abs(v2-v4)

v6<-abs(v3-v4)

v7<-pmax(v5,v6)

D<-max(v7)

[1] 0.5517783

K<-sqrt(n)\*D

[1] 3.901662

***f) Используя гистограмму частот, построить критерий значимости χ2 проверки простой гипотезы согласия с гамма распределением с параметром λ0. Проверить гипотезу на уровне α2. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.***

Гипотеза: H0: λ=λ0, λ0 = 1,00; X

Делим последовательность на r = 5 интервала.

Нижние границы элементов: a1<-c(-Inf, 0.247, 0.674, 0.888, 1. 483)

Верхние границы элементов: b1<-c(0.247, 0.674, 0.888, 1. 483, Inf)

Число наблюдений, попавших в этот интервал: nu (берутся из гистограммы)



Критерий имеет вид:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| № интервала |  |  |  |  |  |  |
| 1 | -Inf | 0.247 | 10 | 0.51785160 | --3.123250 | 9.7546903 |
| 2 | 0.247 | 0.674 | 10 | 0.23652034 | -0.530989 | 0.2819493 |
| 3 | 0.674 | 0.888 | 10 | 0.06298679 | 3.860315 | 14.9020281 |
| 4 | 0.888 | 1.483 | 10 | 0.09760940 | 2.317391 | 5.3702991 |
| r=5 | 1.483 | Inf | 10 | 0.08503187 | 2.787864 | 7.7721871 |
|  | | | | =1 | = 38.08115 | |

n<-length(x); lambda0<-0.17; r<-5; alpha2<-0.02

> border <-c(0.02, 0.62, 1.47, 3,25) #общий массив границ интервалов

> nu<-c(10,10,10,10,10);

> p1<-array(dim = r)

> p1[1]<- pgamma(border[1], shape = 1/2)

> p1[r] <- 1-pgamma(border[r-1],shape = 1/2)

> p1[2:(r-1)]<-pgamma(border[2:(r-1)], shape = 1/2)-pgamma(border[1:(r-2)], shape = 1/2)

> res <- array (dim = r)

> res [1:r] <- (nu[1:r] - n\*p1[1:r])/sqrt(n\*p1[1:r])

> res2 <- array (dim = r)

> res2 [1:r]<- (res[1:r])^2

> Xi2<-sum(res2)

> xal<-qchisq(1-alpha2, r-1)

> Xi2>xal

[1] TRUE

Итак, получили, что χ2 > xα, следовательно, нужно отвергнуть гипотезу Hо.

***Построить критерий проверки значимости χ2 сложной гипотезы согласия с гамма- распределением. Проверить гипотезу на уровне α2. Вычислить наибольшее значение уровня значимости, на котором еще нет оснований отвергнуть данную гипотезу.***

Проверка сложной гипотезы согласия.

H0: λЄ[0,∞)

Статистика критерия:

>al<-0.1

> r<-5

> P<-function(a){

+ p[1]<-pgamma(b[1], shape=1/2)

+ i<-2

+ while(i<r){

+ p[i]<-pgamma(b[i], shape=1/2)-pgamma(b[i-1], shape=1/2);

+ i<-i+1;

+ }

+ p[r]<-1-pgamma(b[r-1], shape=1/2)

+ p;}

> X2<-function(a){g<-n\*P(a);f<-(nu-g)^2/g;sum(f)}

> XM<-nlm(X2,p=mean) #проводим минимизацию,

> xb<-qchisq(1-al,r-2) #вычисляем квантиль

> XM$minimum<xb

[1] FALSE #значит, отвергаем Но

alpha2<-1-pchisq(XM$minimum,r-2) #наибольший уровень значимости, на котором

alpha2

[1] 0