**МИНОБРНАУКИ РОССИИ**

**Санкт-Петербургский государственный**

**электротехнический университет**

**«ЛЭТИ» им. В.И. Ульянова (Ленина)**

**Кафедра ВМ-2**

отчет

**по индивидуальному заданию №2**

**по дисциплине «Статический анализ»**

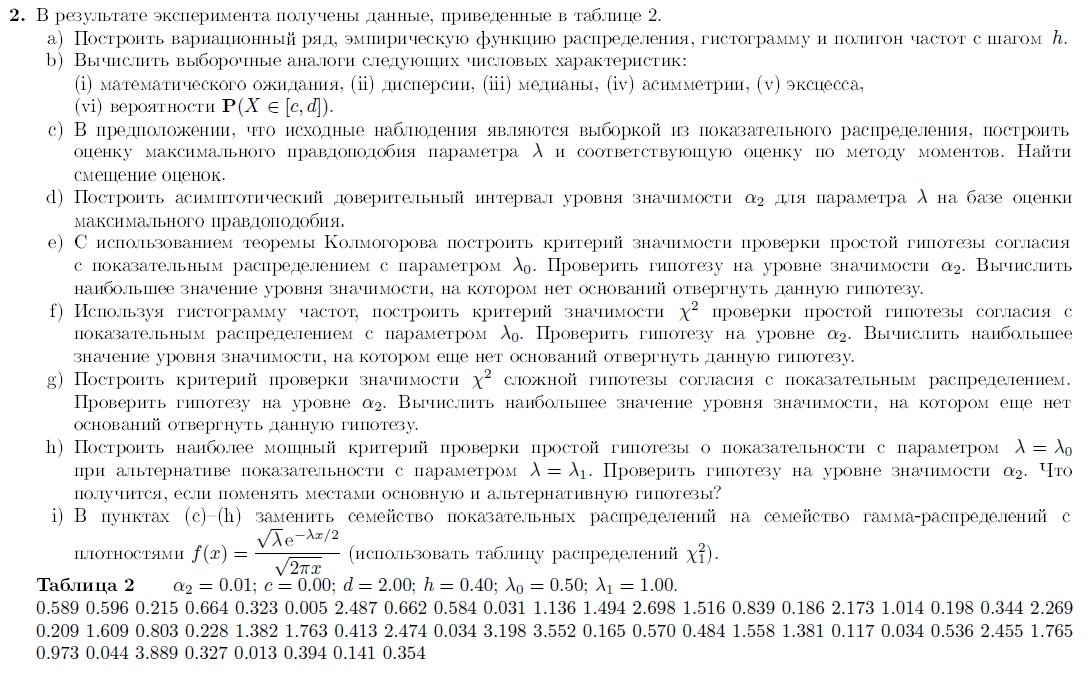
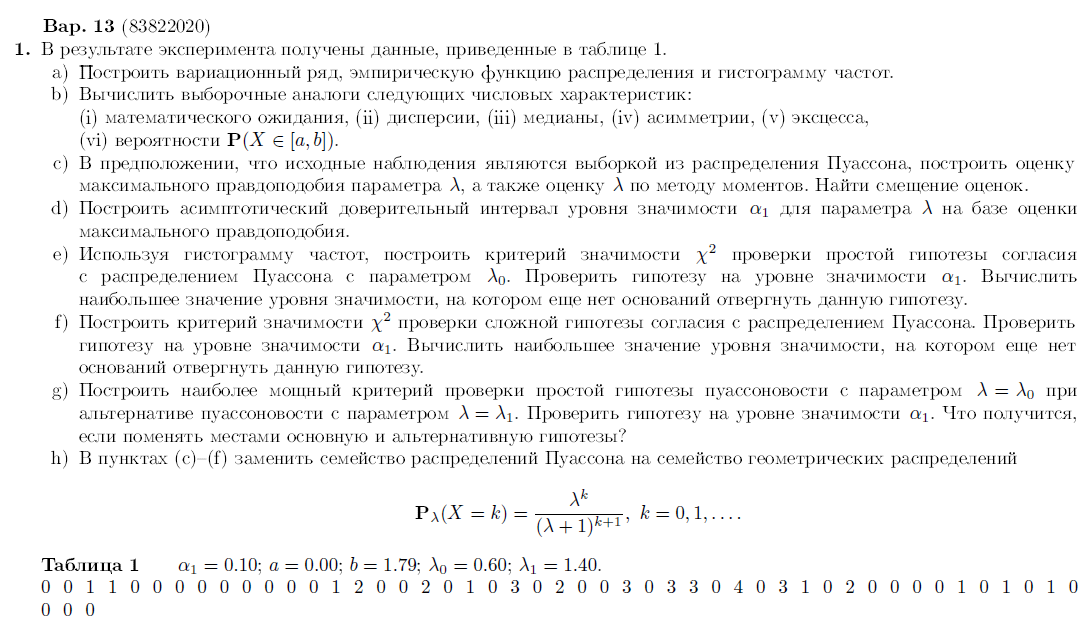
Вариант №13

|  |  |  |
| --- | --- | --- |
| Студент гр. 8382 |  | Мирончик П.Д. |
| Преподаватель |  | Малов С.В. |

Санкт-Петербург

2020

**Постановка задачи**

****

# Ход решения

# 1.

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

Построение ряда:

## x <- c(0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 2, 0, 0, 2, 0, 1, 0, 3, 0, 2, 0, 0, 3, 0, 3, 3, 0, 4, 0, 3, 1, 0, 2, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0)

## vararr <- sort(x)

Полученный вариационный ряд

## 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 2 2 2 2 3 3 3 3 3 4

Построение эмпирической ф.р.:

## F <- function(x, t) {

## z <- x[x<t]

## return(length(z) / length(x))

## }

Полученные значения

## 0, x <= 0

## 0.64, 0 < x <=1,

## 0.8, 1 < x <= 2,

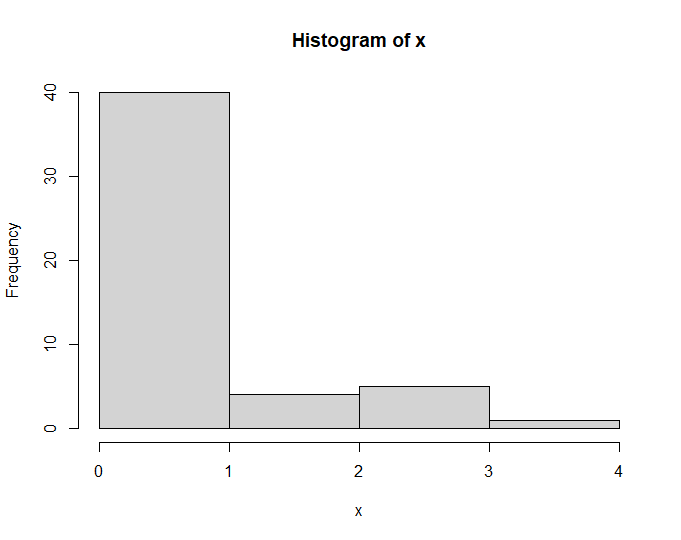
## 0.88, 2 < x <= 3,

## 0.98, 3 < x <= 4

## 1, x > 4

Гистограмма частот:

## hist(x, breaks = c(min(x):max(x)), right = TRUE)



**b) Вычислить выборочные аналоги следующих числовых характеристик:**

**(i) математического ожидания, (ii) дисперсии, (iii) медианы, (iv) асиметрии, (v) эксцесса, (vi) вероятности P(X in [a,b])**

(i) математическое ожидание

## meanX <- sum(x) / length(x)

получили 0.7

(ii) дисперсия

## variance <- sum(x^2) / length(x) - meanX^2

получили 1.21

(iii) медиана

## if (length(x) %% 2 != 0) {

## median <- vararr[(length(vararr) - 1) / 2]

## } else {

## median <- (vararr[length(vararr) / 2] + vararr[length(vararr) / 2 + 1]) / 2

## }

получили 0

(iv) ассиметрия

## assym <- sum((x - meanX)^3) / (length(x) \* variance^(3/2))

получили 1.424493

(v) эксцесс

## excess <- (sum((x - meanX)^4) / (length(x) \* variance^2)) - 3

получили 0.7932518

(vi) вероятность P(X in [a,b])

## Px <- function (a, b) {

## return(F(x, b) - F(x, a))

## }

a=0, b=1.79

получим 0.8

**c) 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.**

ОМП:

## library("maxLik")

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

## maxLL<-maxNR(LL,start=c(1))

## OMP<-maxLL$estimate

полученное ОМП: 0.7

оценка λ по методу моментов:

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

получили также 0.7

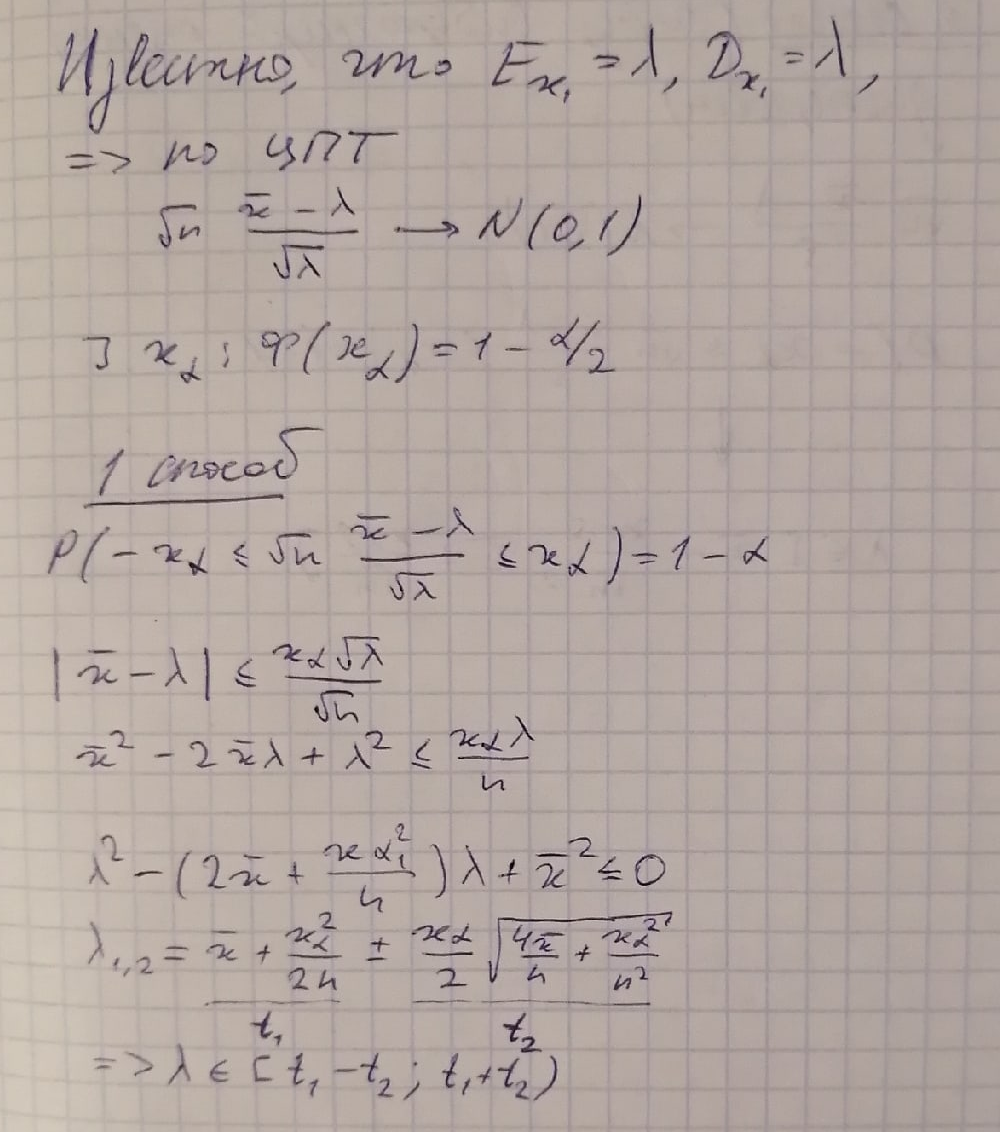
Плотность распределения Пуассона:

Метод моментов:

Значит оценка несмещенная

**d) Построить асимптотический доверительный интервал уровня значимости α1 для параметра λ на базе оценки максимального правдоподобия.**

Тремя способами.



## a1 = 0.1

## x\_mean = sum(x) / n

## x\_l = qnorm (1-a1/2)

## print(x\_mean + x\_l^2/2/n - x\_l/2\*sqrt(4\*x\_mean/n+x\_l^2/n^2))

## print(x\_mean + x\_l^2/2/n + x\_l/2\*sqrt(4\*x\_mean/n+x\_l^2/n^2))

Получим [0.5305622, 0.9235487]

2 способ

По методу максимального правдоподобия:

где - квантиль порядка xα стандартного нормального закона распределения.

Код:

## TI <- function(a1) {

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

## xa<-qnorm (1-a1/2)

## return(c(

## mean-xa\*sqrt(mean/length(x)),

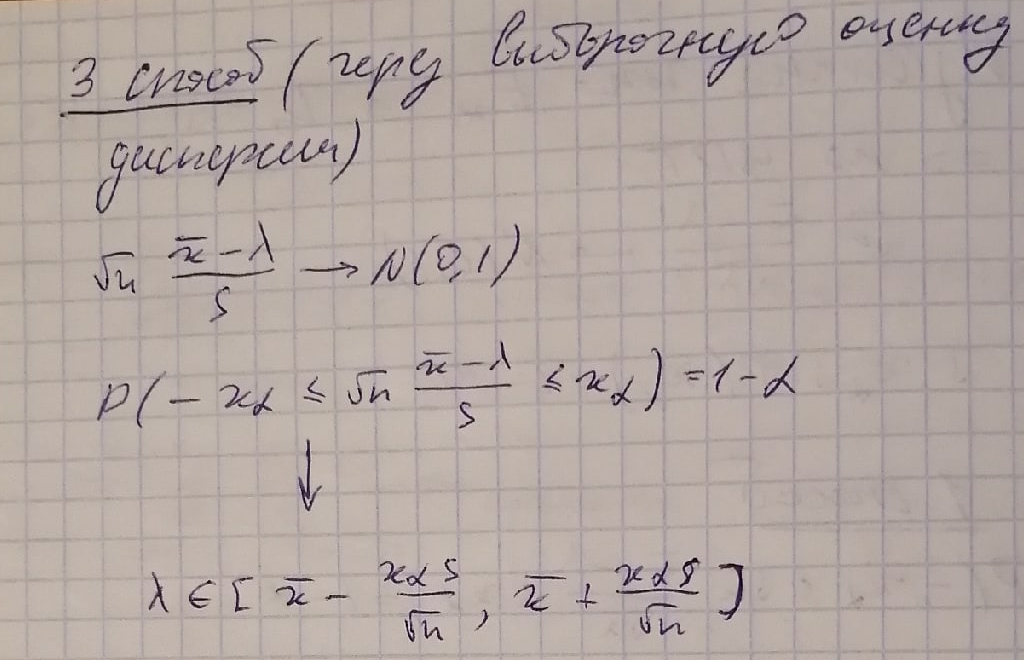
## mean+xa\*sqrt(mean/length(x))

## ))

## }

Результат для a1=0.1: [0.5053783, 0.8946217]

3 способ



## a1 = 0.1

## x\_mean = sum(x) / n

## x\_l = qnorm (1-a1/2)

## s = sqrt(sum((x-x\_mean)^2)/n)

## print(x\_mean - x\_l\*s/sqrt(n))

## print(x\_mean + x\_l\*s/sqrt(n))

Получим [0.4441208, 0.9558792]

**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 H0:

|  |  |  |
| --- | --- | --- |
| **x** | **ni** | **pi** |
| (-inf;0.5) | 32 | 0.54881164 |
| [0.5;1.5] | 8 | 0.32928698 |
| [1.5;2.5) | 4 | 0.09878609 |
| [2.5,inf) | 5 | 0.02311529 |
| sum | 50 | 1 |

## a1 = 0.1

## breaks = c(0.5, 1.5, 2.5)

## Fn = function(t) {

## z <- x[x < t]

## return(length(z))

## }

## ni = c()

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## ni = c(ni, Fn(breaks[i]))

## else if (i == length(breaks)+1)

## ni = c(ni, n - Fn(breaks[i-1]))

## else

## ni = c(ni, Fn(breaks[i]) - Fn(breaks[i-1]))

## }

## pi = c()

## Fv = function(t) {

## return(ppois(t, 0.6))

## }

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## pi = c(pi, Fv(breaks[i]))

## else if (i == length(breaks)+1)

## pi = c(pi, 1 - Fv(breaks[i-1]))

## else

## pi = c(pi, Fv(breaks[i]) - Fv(breaks[i-1]))

## }

## npi = pi\*n

## X2 = sum((ni-npi)^2/npi)

## xa = pchisq(X2, length(breaks),lower.tail = FALSE)

Получим статистику критерия , что больше границы критической области , значит гипотеза отклоняется

Полученное наибольшее значение уровня значимости (по таблице распределения это 26.11493)

**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.**

Нулевая гипотеза: “«Исходные наблюдения являются выборкой из некоторого распределения Пуассона”

Возьмем прежние интервалы (из задания e).

Используем код из пункта e, заменив

## Fv = function(t) {

## return(ppois(t, 0.6))

## }

## …

## xa = pchisq(X2, length(breaks),lower.tail = FALSE)

на

## Fv = function(t) {

## return(ppois(t, 0.7))

## }

## …

## xa = pchisq(X2, length(breaks)-1,lower.tail = FALSE)

Получим статистику критерия , что больше границы критической области 6.251389, значит гипотеза отклоняется.

Полученное наибольшее значение уровня значимости

**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?**

Основная гипотеза , альтернативная гипотеза .

Согласно лемме Неймана-Пирсона:

Функция правдоподобия для распределения Пуассона (была найдена ранее):

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение из которого можно найти и :

Из того, что , можно сделать вывод, что . Воспользуемся тем, что и найдем перебором :

## lambda<-0.6

## a1<-0.1

## c<-0

## a0<-ppois(c,lambda\*length(x))

## while (a0<a1)

## {

## c<-c+1;

## a0<-ppois(c,lambda\*length(x))

## }

## c<-c-1

## print(c)

## p<-(a1-ppois(c,lambda\*length(x)))/dpois(c,lambda\*length(x))

## print(p)

## print(a0)

Критерий примет вид:

Так как сумма , **принимаем основную гипотезу** .

Поменяем местами основную и альтернативную теории и проведём аналогичные вычисления:

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение:

Из того, что , можно сделать вывод, что . Воспользуемся тем, что и найдем перебором :

## lambda<-1.4

## a1<-0.1

## c<-0

## a0<-ppois(c,lambda\*length(x))

## while (a0<a1)

## {

## c<-c+1;

## a0<-ppois(c,lambda\*length(x))

## }

## c<-c-1

## print(c)

## p<-(a1-ppois(c,lambda\*length(x)))/dpois(c,lambda\*length(x))

## print(p)

## print(a0)

Критерий примет вид:

Так как сумма , **принимаем основную гипотезу** .

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

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

Нахождение оценки максимального правдоподобия параметра λ:

Плотность геометрического распределения:

Обозначим , и

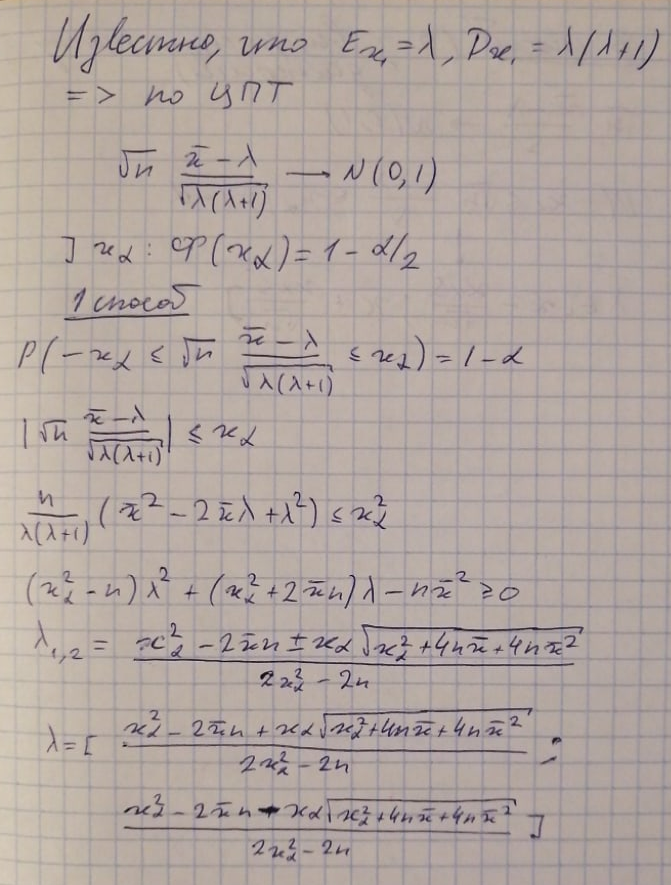
Метод моментов:

Плотность геометрического распределения:

Нахождение смещения оценок:

Следовательно, оценка несмещённая.

**h-d) Построить асимптотический доверительный интервал уровня значимости α1 для параметра λ на базе оценки максимального правдоподобия.**



## a = 0.1

## mean = sum(x) / length(x)

## xa = qnorm(1 - a/2)

## n = length(x)

## print((xa^2 - 2\*mean\*n + xa\*sqrt(xa^2 + 4\*n\*mean + 4\*n\*mean^2))/(2\*xa^2-2\*n))

## print((xa^2 - 2\*mean\*n - xa\*sqrt(xa^2 + 4\*n\*mean + 4\*n\*mean^2))/(2\*xa^2-2\*n))

Получим [0.4416486, 0.9812339]

2 способ

где - квантиль порядка xα стандартного нормального закона распределения.

Выполним код для a1=0.1

## TIf <- function(a1) {

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

## xa<-qnorm (1-a1/2)

## return(c(

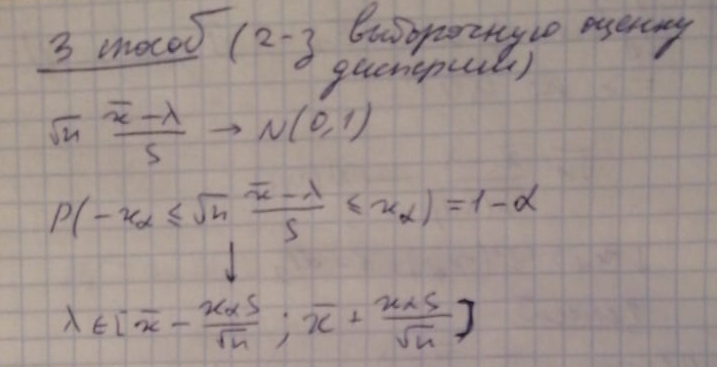
## mean-xa\*sqrt((mean\*(mean+1))/length(x)),

## mean+xa\*sqrt((mean\*(mean+1))/length(x))

## ))

## }

получим [0.4462443, 0.9537557]



## a1 = 0.1

## x\_mean = sum(x) / n

## x\_l = qnorm (1-a1/2)

## s = sqrt(sum((x-x\_mean)^2)/n)

## print(x\_mean - x\_l\*s/sqrt(n))

## print(x\_mean + x\_l\*s/sqrt(n))

Получили [0.4441208, 0.9558792]

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

Простая гипотеза H0:

Используем интервалы из пункта e.

Используем код

## a1 = 0.1

## breaks = c(0.5, 1.5, 2.5)

## Fn = function(t) {

## z <- x[x < t]

## return(length(z))

## }

## ni = c()

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## ni = c(ni, Fn(breaks[i]))

## else if (i == length(breaks)+1)

## ni = c(ni, n - Fn(breaks[i-1]))

## else

## ni = c(ni, Fn(breaks[i]) - Fn(breaks[i-1]))

## }

## pi = c()

## Fv = function(t) {

## l = 0.6

## prob = 1 / (l + 1)

## return(pgeom(t, prob))

## }

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## pi = c(pi, Fv(breaks[i]))

## else if (i == length(breaks)+1)

## pi = c(pi, 1 - Fv(breaks[i-1]))

## else

## pi = c(pi, Fv(breaks[i]) - Fv(breaks[i-1]))

## }

## npi = pi\*n

## X2 = sum((ni-npi)^2/npi)

## xa = pchisq(X2, length(breaks),lower.tail = FALSE)

Полученная статистика критерия . Граница критической области , значит гипотеза принимается.

Полученное наибольшее значение уровня значимости 0.137

**h-f) 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.**

Основная гипотеза H0: Х ~ Geom(λ).

Используем код из пункта h-e, заменив λ на 0.7 и уменьшив степень свободы на 1.

Полученная статистика критерия . Граница критической области , значит гипотеза отклоняется.

Полученное наибольшее значение уровня значимости 0.06317935

# 2.

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

Вариационный ряд:

## x = c(0.589, 0.596, 0.215, 0.664, 0.323, 0.005, 2.487, 0.662, 0.584, 0.031, 1.136, 1.494, 2.698, 1.516, 0.839, 0.186, 2.173, 1.014, 0.198, 0.344, 2.269, 0.209, 1.609, 0.803, 0.228, 1.382, 1.763, 0.413, 2.474, 0.034, 3.198, 3.552, 0.165, 0.570, 0.484, 1.558, 1.381, 0.117, 0.034, 0.536, 2.455, 1.765, 0.973, 0.044, 3.889, 0.327, 0.013, 0.394, 0.141, 0.354)

## vararr = sort(x)

Получим

## 0.005 0.013 0.031 0.034 0.034 0.044 0.117 0.141 0.165 0.186 0.198 0.209 0.215 0.228 0.323 0.327 0.344 0.354 0.394 0.413 0.484 0.536 0.570 .584 0.589 0.596 0.662 0.664 0.803 0.839 0.973 1.014 1.136 1.381 1.382 1.494 1.516 1.558 1.609 1.763 1.765 2.173 2.269 2.455 2.474 2.487 2.698 3.198 3.552 3.889

Эмпирическая функция распределения:

## x < 0.005: 0

## x < 0.013: 0.02

## x < 0.031: 0.04

## x < 0.034: 0.06

## x < 0.044: 0.1

## x < 0.117: 0.12

## x < 0.141: 0.14

## x < 0.165: 0.16

## x < 0.186: 0.18

## x < 0.198: 0.2

## x < 0.209: 0.22

## x < 0.215: 0.24

## x < 0.228: 0.26

## x < 0.323: 0.28

## x < 0.327: 0.3

## x < 0.344: 0.32

## x < 0.354: 0.34

## x < 0.394: 0.36

## x < 0.413: 0.38

## x < 0.484: 0.4

## x < 0.536: 0.42

## x < 0.57: 0.44

## x < 0.584: 0.46

## x < 0.589: 0.48

## x < 0.596: 0.5

## x < 0.662: 0.52

## x < 0.664: 0.54

## x < 0.803: 0.56

## x < 0.839: 0.58

## x < 0.973: 0.6

## x < 1.014: 0.62

## x < 1.136: 0.64

## x < 1.381: 0.66

## x < 1.382: 0.68

## x < 1.494: 0.7

## x < 1.516: 0.72

## x < 1.558: 0.74

## x < 1.609: 0.76

## x < 1.763: 0.78

## x < 1.765: 0.8

## x < 2.173: 0.82

## x < 2.269: 0.84

## x < 2.455: 0.86

## x < 2.474: 0.88

## x < 2.487: 0.9

## x < 2.698: 0.92

## x < 3.198: 0.94

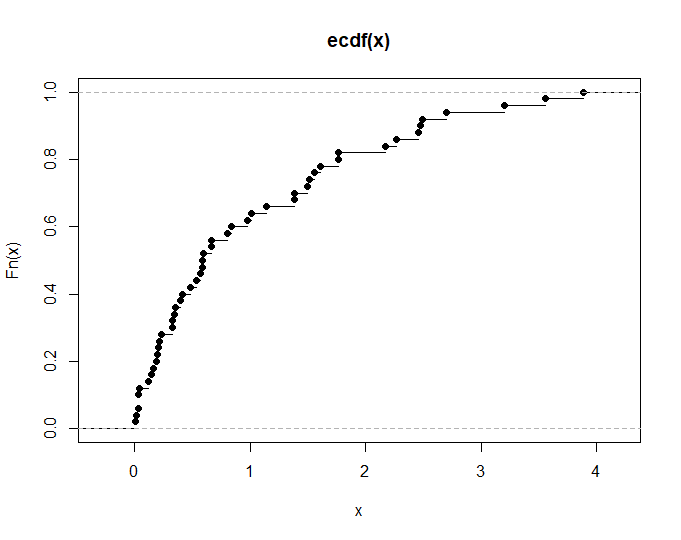
## x < 3.552: 0.96

## x < 3.889: 0.98

## x >= 3.889: 1

График:

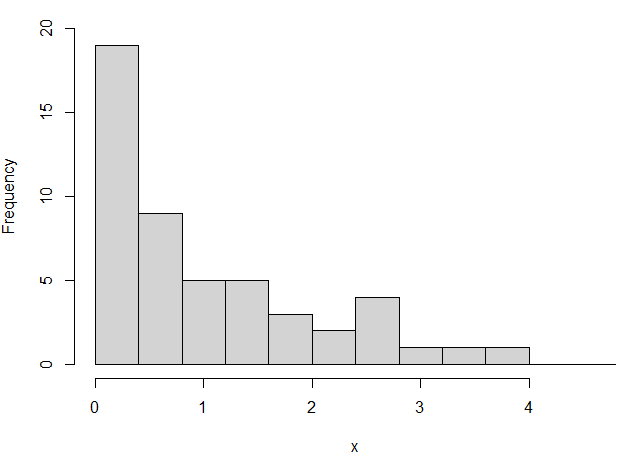
## plot(ecdf(x))

Гистограмма

Гистограмма частот:

## h = 0.4

## hist(x,breaks=seq(0, max(x)+1, by=h))

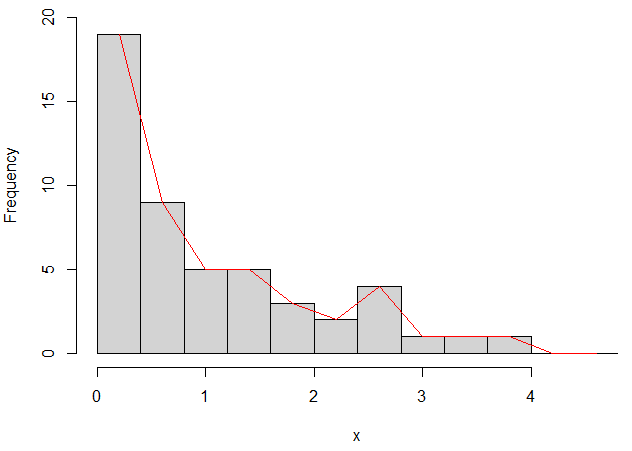


Построение полигона:

## h<-0.4

## pol<-hist(x,breaks=seq(0, max(x)+1, by=h))

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



**b) Вычислить выборочные аналоги следующих числовых характеристик: (i) математического ожидания, (ii) дисперсии, (iii) медианы, (iv) асимметрии, (v) эксцесса, (vi) вероятности P(X [c, d]).**

Математическое ожидание: 1.01776

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

Дисперсия: 0.9938275

## variance<-sum(x^2)/length(x)-mean^2

Медиана: 0.589

## if(length(vararr) %% 2 != 0){

## median<-vararr[(length(vararr)-1)/2]

## } else {

## median<-(vararr[(length(vararr))/2]+vararr[(length(vararr)+2)/2])/2

## }

Асимметрия: 1.137623

## assymmetry<-sum((x-mean)^3)/(length(x)\*variance^(3/2))

Эксцесс: 0.4342106

## excess<-(sum((x-mean)^4)/(length(x)\*variance^2))-3

Вероятность P(X [c,d]): 0.82

## c <- 0

## d <- 2

## P <- F(d)-F(c)

**c) 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.**

Нахождение оценки максимального правдоподобия параметра λ:

Плотность показательного распределения:

Метод моментов:

Плотность показательного распределения:

Нахождение смещения оценок:

Несмещённая оценка:

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

Найдем информацию Фишера:

Оценка максимального правдоподобия параметра λ:

 .

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

где - квантиль порядка xα стандартного нормального закона распределения.

## TI = function(a2) {

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

## xa<-qnorm (1-a2/2)

## return(c(

## 1/mean-xa/sqrt(mean\*mean\*length(x)),

## 1/mean+xa/sqrt(mean\*mean\*length(x))

## ))

## }

## print(TI(0.01))

Получим [0.6246293, 1.3404705]

**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.**

Гипотеза:

Согласно теореме Колмогорова, при справедливости гипотезы:

Обозначим

Согласно таблице распределения Колмогорова, α= 1,63.

Вычислим величину Dn:

## lambda0 <- 0.5;

## Fy<-ecdf(x)

## Fexp <- pexp(x,lambda0)

## Diff <- array(dim=50)

## for(i in 1:length(x)){

## Diff[i]<-abs(Fy(x[i])-Fexp[i])

## }

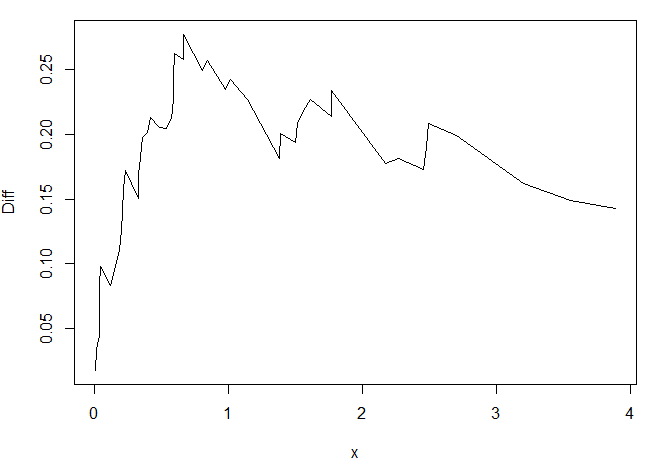
## D<-max(Diff)

## Dn<-D\*sqrt(length(x))

## print(Dn)

Получим . – значит, гипотеза H0 не принимается. Наибольшее значение уровня значимости, при котором мы все еще могли бы принять гипотезу 0,00092.

График достижения sup



x, при котором достигается sup, равен

Таблица sup

## x Fn Fexp Diff

## 1 0.005 0.02 0.0025 0.0175

## 2 0.013 0.04 0.0065 0.0335

## 3 0.031 0.06 0.0154 0.0446

## 4 0.034 0.10 0.0169 0.0831

## 5 0.034 0.10 0.0169 0.0831

## 6 0.044 0.12 0.0218 0.0982

## 7 0.117 0.14 0.0568 0.0832

## 8 0.141 0.16 0.0681 0.0919

## 9 0.165 0.18 0.0792 0.1008

## 10 0.186 0.20 0.0888 0.1112

## 11 0.198 0.22 0.0943 0.1257

## 12 0.209 0.24 0.0992 0.1408

## 13 0.215 0.26 0.1019 0.1581

## 14 0.228 0.28 0.1077 0.1723

## 15 0.323 0.30 0.1491 0.1509

## 16 0.327 0.32 0.1508 0.1692

## 17 0.344 0.34 0.1580 0.1820

## 18 0.354 0.36 0.1622 0.1978

## 19 0.394 0.38 0.1788 0.2012

## 20 0.413 0.40 0.1866 0.2134

## 21 0.484 0.42 0.2149 0.2051

## 22 0.536 0.44 0.2351 0.2049

## 23 0.570 0.46 0.2480 0.2120

## 24 0.584 0.48 0.2532 0.2268

## 25 0.589 0.50 0.2551 0.2449

## 26 0.596 0.52 0.2577 0.2623

## 27 0.662 0.54 0.2818 0.2582

## 28 0.664 0.56 0.2825 0.2775

## 29 0.803 0.58 0.3307 0.2493

## 30 0.839 0.60 0.3426 0.2574

## 31 0.973 0.62 0.3852 0.2348

## 32 1.014 0.64 0.3977 0.2423

## 33 1.136 0.66 0.4333 0.2267

## 34 1.381 0.68 0.4987 0.1813

## 35 1.382 0.70 0.4989 0.2011

## 36 1.494 0.72 0.5262 0.1938

## 37 1.516 0.74 0.5314 0.2086

## 38 1.558 0.76 0.5411 0.2189

## 39 1.609 0.78 0.5527 0.2273

## 40 1.763 0.80 0.5858 0.2142

## 41 1.765 0.82 0.5863 0.2337

## 42 2.173 0.84 0.6626 0.1774

## 43 2.269 0.86 0.6784 0.1816

## 44 2.455 0.88 0.7070 0.1730

## 45 2.474 0.90 0.7097 0.1903

## 46 2.487 0.92 0.7116 0.2084

## 47 2.698 0.94 0.7405 0.1995

## 48 3.198 0.96 0.7979 0.1621

## 49 3.552 0.98 0.8307 0.1493

## 50 3.889 1.00 0.8569 0.1431

код

## plot(x, Diff, type='l')

## print(

## data.frame(

## x = x,

## Fn = sapply(x, Fy),

## Fexp = round(Fexp, 4),

## Diff = round(Diff, 4)

## )

## )

**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 H0:

|  |  |  |
| --- | --- | --- |
| **Граница** | **ni** | **pi** |
| (-inf;0.4) | 19 | 0.18126925 |
| [0.4,0.8) | 9 | 0.14841071 |
| [0.8,1.2) | 5 | 0.12150841 |
| [0.8, 1.6) | 5 | 0.09948267 |
| [1.6, 2.4) | 5 | 0.14813475 |
| [2.4, inf) | 7 | 0.30119421 |
| sum | 50 | 1 |

Выполним код:

## a1 = 0.1

## breaks = c(0.4, 0.8, 1.2, 1.6, 2.4)

## Fn = function(t) {

## z <- x[x < t]

## return(length(z))

## }

## ni = c()

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## ni = c(ni, Fn(breaks[i]))

## else if (i == length(breaks)+1)

## ni = c(ni, n - Fn(breaks[i-1]))

## else

## ni = c(ni, Fn(breaks[i]) - Fn(breaks[i-1]))

## }

## pi = c()

## Fv = function(t) {

## l = 0.5

## return(pexp(t, l))

## }

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## pi = c(pi, Fv(breaks[i]))

## else if (i == length(breaks)+1)

## pi = c(pi, 1 - Fv(breaks[i-1]))

## else

## pi = c(pi, Fv(breaks[i]) - Fv(breaks[i-1]))

## }

## npi = pi\*n

## X2 = sum((ni-npi)^2/npi)

## xa = pchisq(X2, length(breaks),lower.tail = FALSE)

Получим статистику критерия , что больше границы критической области 15.08627, значит гипотеза отклоняется.

Полученное наибольшее значение уровня значимости 0.00551

**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.**

Основная гипотеза: Исходные наблюдения являются выборкой из показательного распределения

H0:

Интервалы берутся аналогично пункту f.

Выполним код предыдущего пункта, заменив λ на 0.9825499 и уменьшив степень свободы.

Получим статистику критерия , что меньше границы критической области 13.2767, значит гипотеза принимается.

Полученное наибольшее значение уровня значимости 0.598078

**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?**

Основная гипотеза , альтернативная гипотеза .

Согласно лемме Неймана-Пирсона:

Функция правдоподобия для показательного распределения (была найдена ранее):

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение:

, .

Найдём из уравнения:

- квантиль распределения уровня :

## lambda<-0.5

## a2<-0.01

## C<- qgamma(a2,50,lambda)

Критерий примет вид:

Так как сумма , принимаем альтернативную гипотезу.

Поменяем местами основную и альтернативную теории и проведём аналогичные вычисления:

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение:

, .

Найдём из уравнения:

- квантиль распределения уровня :

## lambda<-1

## a2<-0.01

## C<- qgamma(a2,50,lambda)

Критерий примет вид:

Так как сумма , принимаем альтернативную гипотезу .

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

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

Нахождение оценки максимального правдоподобия параметра λ:

Плотность гамма распределения:

Метод моментов:

Плотность гамма распределения:

Нахождение смещения оценок:

.

Несмещённая оценка:

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

Согласно центральной предельной теореме, при больших k {\displaystyle k} n гамма-распределение может быть приближено нормальным распределением:

*,* при k → ∞.

В нашем случае:

По лемме Фишера:

Следовательно:

Найдем область:

где - квантиль порядка xα стандартного нормального закона распределения.

Вычислим:

## TIi = function(a2) {

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

## xa<-qnorm (1-a2/2)

## return(c(

## 1/(mean+xa\*mean\*sqrt(2/length(x))),

## 1/(mean-xa\*mean\*sqrt(2/length(x)))

## ))

## }

## print(TIi(0.01))

Результат: [0.6484768, 2.0265692]

**i-e) 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.**

Гипотеза:

Согласно теореме Колмогорова, при справедливости гипотезы:

Обозначим

Согласно таблице распределения Колмогорова, α= 1,63.

Вычислим величину Dn:

## lambda0 = 0.5;

## Fy = ecdf(x)

## lambda = 0.5

## Fgamma = pgamma(x, 0.5, lambda / 2)

## Diff = array(dim=50)

## for(i in 1:length(x)){

## Diff[i] = abs(Fy(x[i])-Fgamma[i])

## }

## D = max(Diff)

## Dn<-D\*sqrt(length(x))

Полученное значение . – значит, гипотеза H0 принимается.Наибольший уровень значимости по таблице распределений Колмогорова будет примерно равен 0,064

**i-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.**

Простая гипотеза H0:

Интервалы берутся аналогично пункту f.

Выполним код

## a1 = 0.01

## breaks = c(0.4, 0.8, 1.2, 1.6, 2.4)

## Fn = function(t) {

## z <- x[x < t]

## return(length(z))

## }

## ni = c()

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## ni = c(ni, Fn(breaks[i]))

## else if (i == length(breaks)+1)

## ni = c(ni, n - Fn(breaks[i-1]))

## else

## ni = c(ni, Fn(breaks[i]) - Fn(breaks[i-1]))

## }

## pi = c()

## Fv = function(t) {

## l = 0.5

## return(pgamma(t, 0.5, l/2))

## }

## for (i in 1:(length(breaks)+1)) {

## if (i == 1)

## pi = c(pi, Fv(breaks[i]))

## else if (i == length(breaks)+1)

## pi = c(pi, 1 - Fv(breaks[i-1]))

## else

## pi = c(pi, Fv(breaks[i]) - Fv(breaks[i-1]))

## }

## npi = pi\*n

## X2 = sum((ni-npi)^2/npi)

## xa = pchisq(X2, length(breaks),lower.tail = FALSE)

## X2\_lim = qchisq(a1, length(breaks), lower.tail = FALSE)

Получим статистику критерия , что меньше границы критической области 15.08627, значит гипотеза принимается.

Полученное наибольшее значение уровня значимости 0.3734325

**i-g) 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.**

Простая гипотеза H0:

Интервалы берутся аналогично пункту f.

Выполним код предыдущего пункта, заменив λ на 0.9825499 и уменьшив степень свободы.

Получим статистику критерия , что меньше границы критической области 13.2767, значит гипотеза принимается.

Полученное наибольшее значение уровня значимости 0.7279941

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

Основная гипотеза , альтернативная гипотеза .

Согласно лемме Неймана-Пирсона:

Функция правдоподобия для гамма-распределения

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение:

Из того, что , можно сделать вывод, что .

Найдём из уравнения:

- квантиль распределения уровня :

## lambda<-0.5

## a2<-0.01

## C<-qnorm(a2,50/lambda,sqrt(2\*50)/lambda)

Критерий примет вид:

Так как сумма , принимаем основную гипотезу .

Поменяем местами основную и альтернативную теории и проведём аналогичные вычисления:

Найдём :

Прологарифмируем полученное выражение:

Обозначим . Тогда критерий примет вид:

Из этого следует уравнение:

Из того, что , можно сделать вывод, что .

Найдём из уравнения:

- квантиль распределения уровня :

## lambda<-1

## a2<-0.01

## C<-qnorm(a2,50/lambda,sqrt(2\*50)/lambda)

Критерий примет вид:

Так как сумма , принимаем альтернативную гипотезу .