

Практика №1 Моделирование и описательная статистика

Долаева А.Р., г. 20.М04-мм

06/04/2021

Вариант 5. Гамма распределение $\Gamma(\alpha = 1, \lambda = 4)$

Параметры обозначенные с точкой вычислены по формулам: $\mu = \frac{\lambda}{\alpha}$; $\sigma^2 = \frac{\alpha}{\lambda^2}$; $\sigma = \frac{\alpha^{1/2}}{\lambda}$.

Парметры без точки вычисляются на основе статистических формул.

1. Моделирование распределения

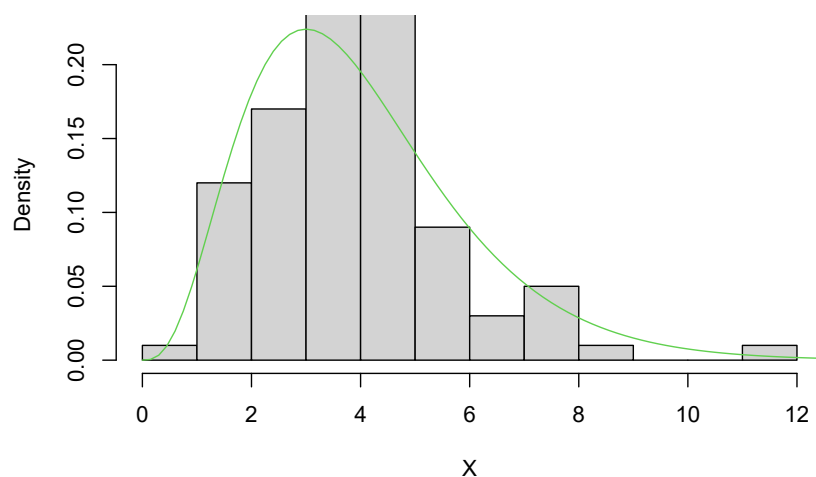
```
rate <- 1
shape <- 4
N <- 100
X <- rgamma(N, shape, rate); X[0:10]
```

```
## [1] 3.367415 3.032291 3.231879 1.086006 4.304567 3.456630 3.809718 4.636548
## [9] 2.348533 2.280803
```

```
peak = dgamma((shape-1)/rate, shape/rate)
```

2. Построение гистограммы плотности гамма распределения

```
hist(X, freq=FALSE, ylim=c(0, peak), main="")
f.<-function(x) dgamma(x, shape/rate)
curve(f., 0, 15, add=TRUE, col=3)
```



2. Нахождение математического ожидания

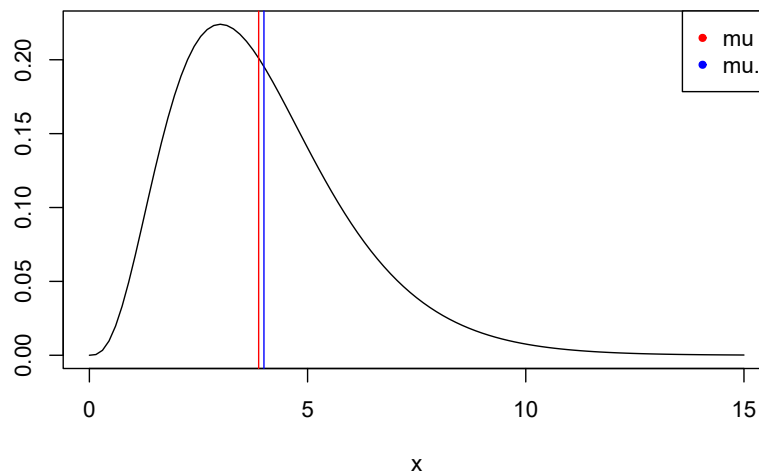
```

mu <-mean(X)
mu.<-shape/rate
c(mu.=mu.,mu=mu)

##      mu.      mu
## 4.000000 3.879038

plot(f.,0,15, type = "l", lty = 1, main="", ylab="")
abline(v = mu, col="red")
abline(v = mu., col="blue")
legend('topright',c("mu", "mu."),pch=20,col=c("red","blue"))

```



2. Дисперсии

```

sigma<-sd(X)
sigma2.<-shape/(rate^2)
c(sigma2.=sigma2.,sigma2=sigma^2)

```

```

##  sigma2.  sigma2
## 4.000000 3.150684

```

2. Стандартного отклонения

```

sigma.= sqrt(sigma2.)
c(sigma.=sigma.,sigma=sigma)

```

```

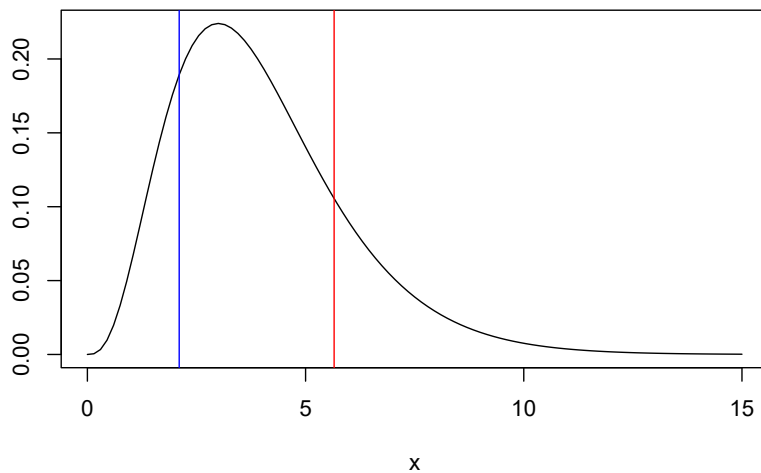
##  sigma.  sigma
## 2.000000 1.775017

```

```

plot(f.,0,15, type = "l", lty = 1, main="", ylab="")
abline(v = mu+sigma, col="red")
abline(v = mu-sigma, col="blue")

```



2. Ошибки среднего

```
c(me.=sigma./sqrt(N),me=sigma/sqrt(N))
```

```
##      me.      me
## 0.2000000 0.1775017
```

2. Медианы

```
c(mu.=mu., mu=mu, median=median(X))
```

```
##      mu.      mu  median
## 4.000000 3.879038 3.809138
```

2. Минимума

```
min(X)
```

```
## [1] 0.9892032
```

2. Максимума

```
max(X)
```

```
## [1] 11.39793
```

2. Размаха

```
max(X)-min(X)
```

```
## [1] 10.40873
```

2. Квартилей

```
q<-quantile(X, probs = seq(0, 1, 0.25));q
```

```
##      0%      25%      50%      75%     100%
## 0.9892032 2.6851094 3.8091377 4.6649450 11.3979338
```

```
median = q[50%]
```

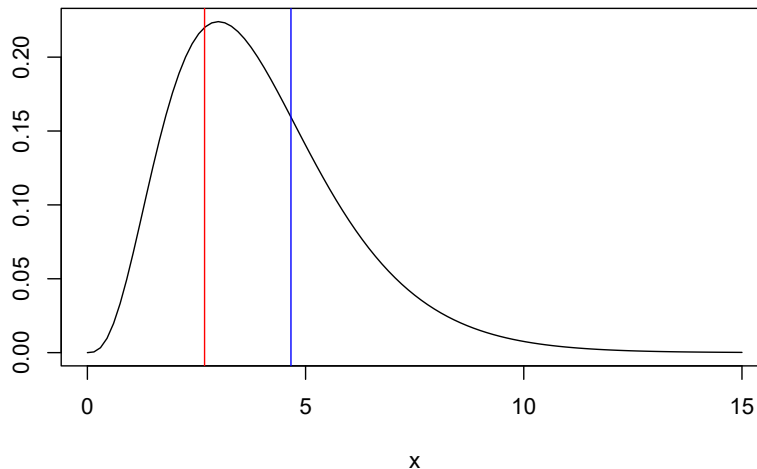
2. Интерквартильного размаха

```
Q<-q[4]-q[2];Q
```

```
##      75%
```

```
## 1.979836
```

```
plot(f.,0,15, type = "l", lty = 1, main="", ylab="")
abline(v = q[2], col="red")
abline(v = q[4], col="blue")
```



2. Ассиметрии (положительное значение - правый хвост длинее, правосторонняя асимметрия)

```
library(moments)
m3<-moment(X, order = 3, central = TRUE)
c(mean((X-mu.)^3), mean((X-mu)^3), m3) #
```

```
## [1] 4.908845 6.042518 6.042518
```

```
c(m3/sigma.^3, m3/sigma^3, skewness(X)) #
```

```
## [1] 0.7553148 1.0804649 1.0968768
```

2. Эксцесса (положительный показатель - “вытянутое” распределение)

```
c(mean((X-mean(X))^4),m4<-moment(X, order = 4, central = TRUE)) #
```

```
## [1] 52.75562 52.75562
```

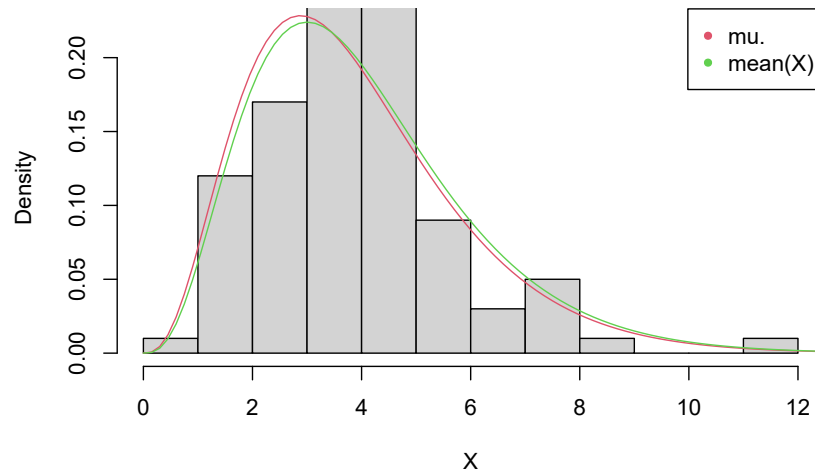
```
m2<-moment(X, order = 2, central = TRUE) #
c(kurtosis(X)-3,m4/m2^2-3) #
```

```
## [1] 2.422363 2.422363
```

3. Оценка параметров распределения по методу моментов

```
hist(X,freq=FALSE,ylim=c(0,peak), main="")
f.<-function(x) dgamma(x,shape/rate)
f<-function(x) dgamma(x,mean(X))
curve(f,0,15,add=TRUE,col=2)
curve(f.,0,15,add=TRUE,col=3)
```

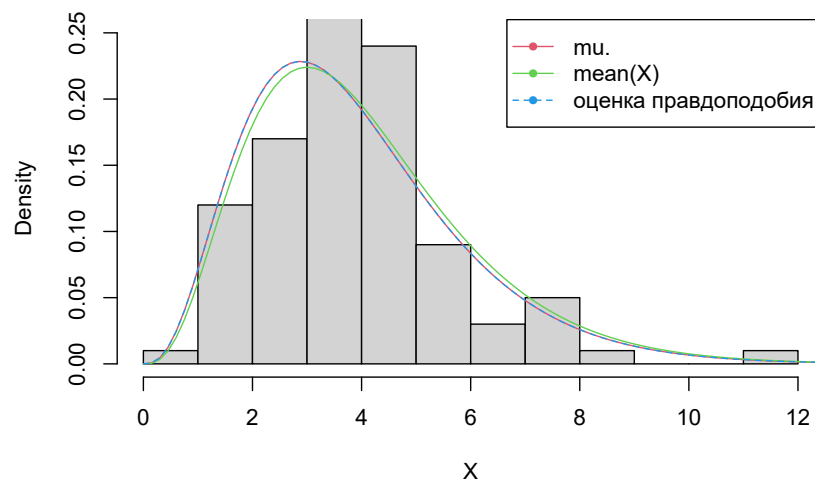
```
legend('topright',c("mu.", "mean(X)"),pch=20,col=c(2,3))
```



по методу максимального прав-

доподобия

```
Func.prob.log <- function(x) -sum(dnorm(X, mean = x[1], log = TRUE))
res<-optim(c(mu.,sigma.),Func.prob.log)
mu.<-res$par[1];
f.<-function(x)dgamma(x,mu..)
hist(X,freq=FALSE, ylim=c(0,0.25),main="")
curve(f,0,15,add=TRUE,col=2)
curve(f.,0,15,add=TRUE,col=3)
curve(f..,0,15,add=TRUE,col=4,lty=2)
legend('topright',c("mu.", "mean(X)", "оценка правдоподобия"),pch=20,col=c(2,3,4),lty=c(1,1,2))
```



4. Проверка согласия эмпирического и теоретического распределения по критерию хи-квадрат Пирсона

```

h<-hist(X,plot=FALSE)
#
n.i<-sapply(seq(length(h$breaks)-1)+1,function(i)
length(X[X<h$breaks[i] & X>=h$breaks[i-1]]))
p.i<-sapply(seq(length(h$breaks)-1)+1,function(i)
pgamma(h$breaks[i],mu..)-pgamma(h$breaks[i-1],mu..))
sum(p.i)

## [1] 0.9980524

#
p.i[1]<-pgamma(h$breaks[2],mu..)
p.i[length(p.i)]<-1-pgamma(h$breaks[length(h$breaks)-1],mu..)
sum(p.i)

## [1] 1

доверительный уровень вероятности

#
n*p_i>5
tab<-cbind(h$counts,p.i*length(X))
tab

##      [,1]      [,2]
## [1,]    1  2.2881746
## [2,]   12 13.6483523
## [3,]   17 21.8274325
## [4,]   27 21.3581928
## [5,]   24 16.3276795
## [6,]    9 10.7677399
## [7,]    3  6.4369232
## [8,]    5  3.5879719
## [9,]    1  1.8979345
## [10,]   0  0.9639603
## [11,]   0  0.4739522
## [12,]   1  0.4216863

t1<-cumsum(tab[,2]);T1<-min(which(t1>5))
t1

## [1]  2.288175 15.936527 37.763959 59.122152 75.449832 86.217572
## [7] 92.654495 96.242467 98.140401 99.104362 99.578314 100.000000

t2<-cumsum(tab[seq(nrow(tab),1,-1),2]);t2;T2<-min(which(t1>5))

## [1]  0.4216863  0.8956385  1.8595988  3.7575333  7.3455052 13.7824284
## [7] 24.5501684 40.8778479 62.2360406 84.0634732 97.7118254 100.0000000

T2<-nrow(tab)-T2
T2

## [1] 10

Tab<-apply(tab,2,function(x) c(sum(x[seq(T1)]),x[c((T1+1):T2)],sum(x[c((T2+1):nrow(tab)]))))
Tab

##      [,1]      [,2]
## [1,]   13 15.9365268
## [2,]   17 21.8274325

```

```
## [3,] 27 21.3581928
## [4,] 24 16.3276795
## [5,] 9 10.7677399
## [6,] 3 6.4369232
## [7,] 5 3.5879719
## [8,] 1 1.8979345
## [9,] 0 0.9639603
## [10,] 1 0.8956385
```

```
chi2<-sum(apply(Tab,1,function(x) (x[1]-x[2])^2/x[2])) # -
p.value<-1-pchisq(chi2,nrow(Tab)-3);
print(paste("p.value",round(p.value,4),sep="="))
```

```
## [1] "p.value=0.1482"
```

Вероятность получения случайного отклонения. Ошибка в

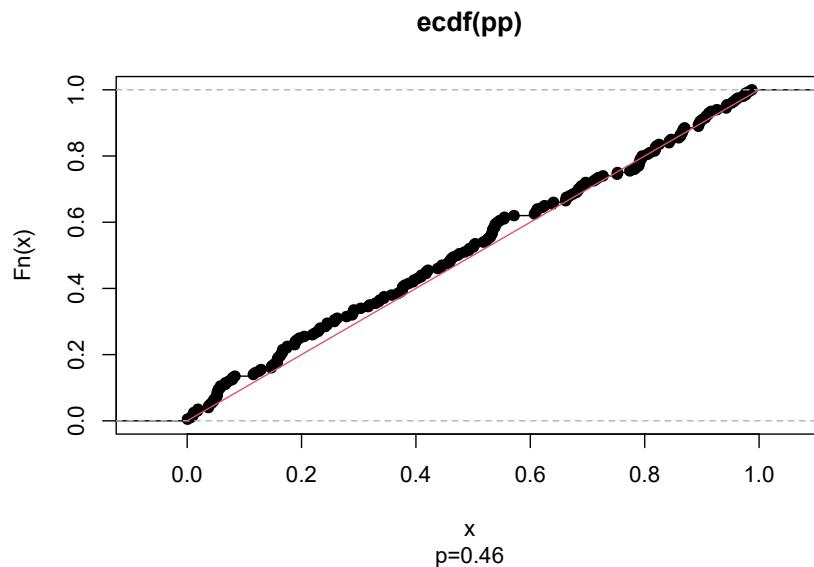
```
round(p.value,4)*100
```

```
## [1] 14.82
```

% случаев

5. Моделирование выборки с нормальным законом распределения и применить критерий согласия Шапиро-Уилка

```
pp<-sapply(seq(200),function(x)shapiro.test(rnorm(100,4,12))$p.value)
plot(ecdf(pp))
lines(c(0,1),c(0,1),type="l",col=2)
title(sub=paste("p",format(ks.test(pp, "punif")$p.value,3,2),sep="=" ))
```



Ошибка в

```
round(ks.test(pp, "punif")$p.value*100,2)
```

```
## [1] 46.42
```

% случаев