

Вероятности и Статистика

Домашно 1

Владимир Ананиев, 81459

Зад. 1 Определете променливите съдържащи се в данните “survey” от пакета “MASS”. Вземете 50 случайно избрани наблюдения.

Hide

```
library('MASS')
indexes = sample(nrow(survey), 50)
students = survey[indexes,]
attach(students)
```

Hide

students

	Sex <fctr>	Wr.Hnd <dbl>	NW.Hnd <dbl>	W.Hnd <fctr>	Fold <fctr>	Pulse <int>	Clap <fctr>	Exer <fctr>	Smoke <fctr>	
143	Female	19.0	18.8	Right	R on L	65	Right	Freq	Never	
81	Male	19.5	19.5	Left	R on L	66	Left	Some	Never	
129	Female	17.5	17.0	Right	R on L	74	Right	Freq	Never	
67	Female	19.0	19.1	Right	L on R	NA	Neither	Freq	Never	
37	Female	16.0	16.5	Right	L on R	NA	Right	Some	Never	
29	Male	17.8	17.8	Right	L on R	76	Neither	Freq	Never	
98	Female	17.5	17.5	Right	R on L	60	Right	Freq	Never	
122	Male	22.5	22.5	Right	L on R	76	Right	Freq	Occas	
3	Male	18.0	13.3	Right	L on R	87	Neither	None	Occas	
222	Female	15.9	16.5	Right	R on L	70	Right	Freq	Never	

1-10 of 50 rows | 1-10 of 12 columns

Previous12345Next

Въз основа на тези наблюдения:

a) Отпечатайте пълните данни за най-младата жена.

Hide

```
women = students[Sex == 'Female',]
women_by_age = women[order(Age),]
women_by_age[1,]
```

	Sex <fctr>	Wr.Hnd <dbl>	NW.Hnd <dbl>	W.Hnd <fctr>	Fold <fctr>	Pulse <int>	Clap <fctr>	Exer <fctr>	Smoke <fctr>	
129	Female	17.5	17	Right	R on L	74	Right	Freq	Never	

1 row | 1-10 of 12 columns

б) Определете пола на 3-те най-високи студента.

Hide

```
students_by_height = students[order(-Height),]
students_by_height[1:3, 'Sex']
```

```
[1] Male Male Male
Levels: Female Male
```

в) Определете средната възраст на мъжете с пулс над 80.

Hide

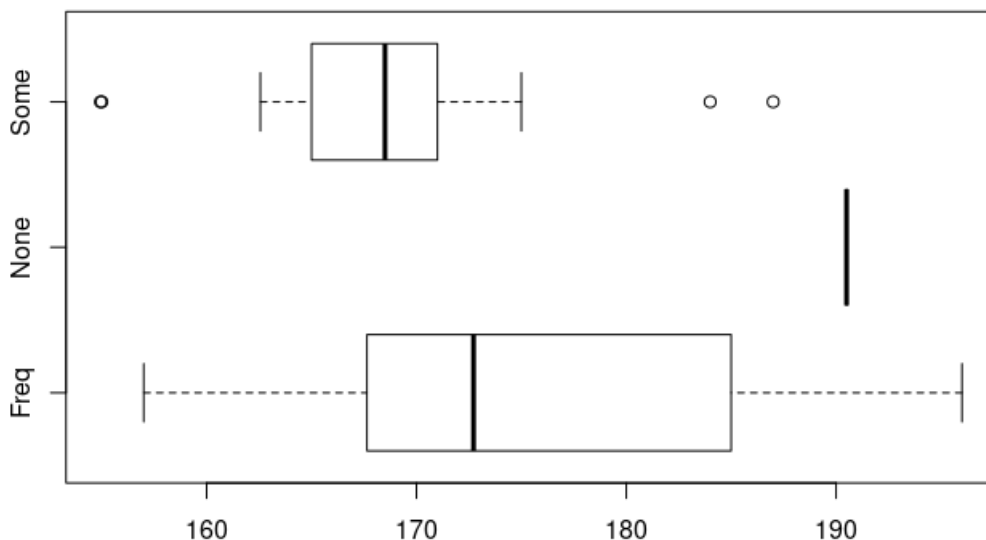
```
men_over_80_pulse = students[Sex == 'Male' & Pulse > 80,]  
mean(Age, na.rm = TRUE)
```

```
[1] 19.61502
```

г) Постройте боксплот на "Height" в зависимост от "Exer".

Hide

```
boxplot(Height~Exer, horizontal = TRUE)
```



Зад. 2 Нека X и Y са последните две цифри на факултетния ви номер, т.е. $??XY$. Генерирайте 50 случайни наблюдения в равномерно разпределение в интервала $[0; X+1]$ и 5- случайни наблюдения експоненциално разпределени с параметър $\lambda = 1/(Y + 1)$

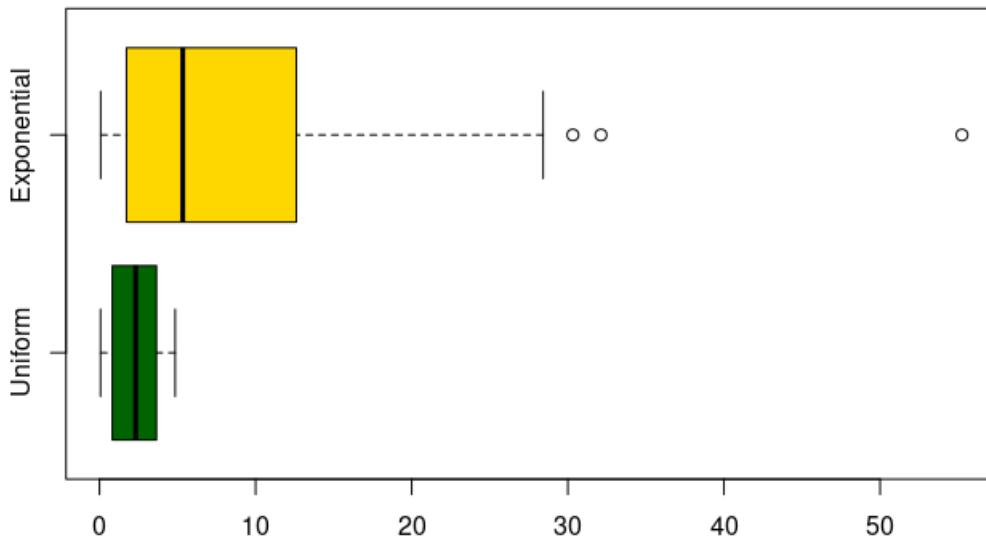
Hide

```
X = 5  
Y = 10  
uniform = runif(50, 0, X)  
exponential = rexp(50, rate=1/Y)
```

а) Постройте боксплот на двете извадки на една графика, направете извод.

Hide

```
boxplot(uniform, exponential, horizontal = TRUE, names = c('Uniform', 'Exponential'), col = c("darkgreen", "gold"))
```



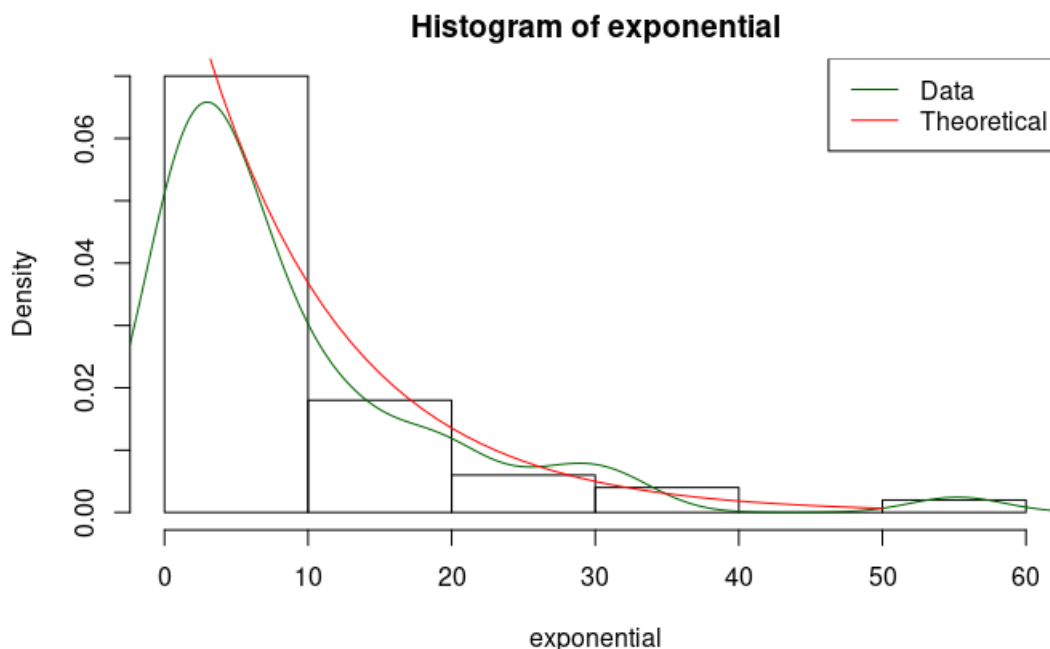
б) Постройте хистограма за експоненциалните данни. На същата графика начертайте плътността построена по данните и теоритичната плътност.

Hide

```
hist(exponential, probability = TRUE)
lines(density(exponential), col='darkgreen')
```

Hide

```
z = seq(0, 50, 1)
y = dexp(z, 1/Y)
lines(z, y, col='red')
legend('topright', legend=c('Data', 'Theoretical'), col=c('darkgreen', 'red'), lty=1:1)
```



Зад. 3 Нека X_1, X_2, X_3 са случайно избрани различни цели числа в интервала $[1, T]$. Напишете функция $f(k, T, n)$, която по направени n на брой санлюдения над X_1, X_2, X_3 пресмята емпиричната вероятност да е изпълнено $X_1 + 2X_2 + 3X_3 = k$, където k е цяло число. Генерирайте случайно T в интервала $[1, 10]$ и случайно k в интервала $[10, 30]$ и пресметнете търсената вероятност.

Hide

```
f <- function(k, T, n){
  counter = 0
  for(i in 1:n){
    X = sample(1:T, 3, replace = TRUE)
    sum = X[1] + 2*X[2] + 3*X[3]
    if (sum == k) {
      counter = counter+1
    }
  }
  return (counter / n)
}
T = sample(1:10, 1)
k = sample(10:30, 1)
for(n in c(100,1000,10000,100000,1000000)) {
  prob = f(k, T, n)
  res = sprintf("f(%d, %d, %d) = %f", k, T, n, prob)
  print(res)
}
```

```
[1] "f(18, 5, 100) = 0.070000"
[1] "f(18, 5, 1000) = 0.064000"
[1] "f(18, 5, 10000) = 0.068800"
[1] "f(18, 5, 100000) = 0.071220"
[1] "f(18, 5, 1000000) = 0.072021"
```

Зад. 4 Нека случайната величина **X** е гамма разпределена с параметри 2 и 0.5.

Hide

```
X = rgamma(1000, 2, 0.5)
```

Определете: **а)** $P(X > 2)$

Hide

```
1 - pgamma(2, 2, 0.5)
```

```
[1] 0.7357589
```

б) с, така че $P(X > c) = 0.35$

Hide

```
qgamma(0.65, 2, 0.5)
```

```
[1] 4.437689
```

в) Q1, M, Q3.

Hide

```
summary(X)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.2134  1.8033  3.1249  3.7456  5.0702 18.3539
```