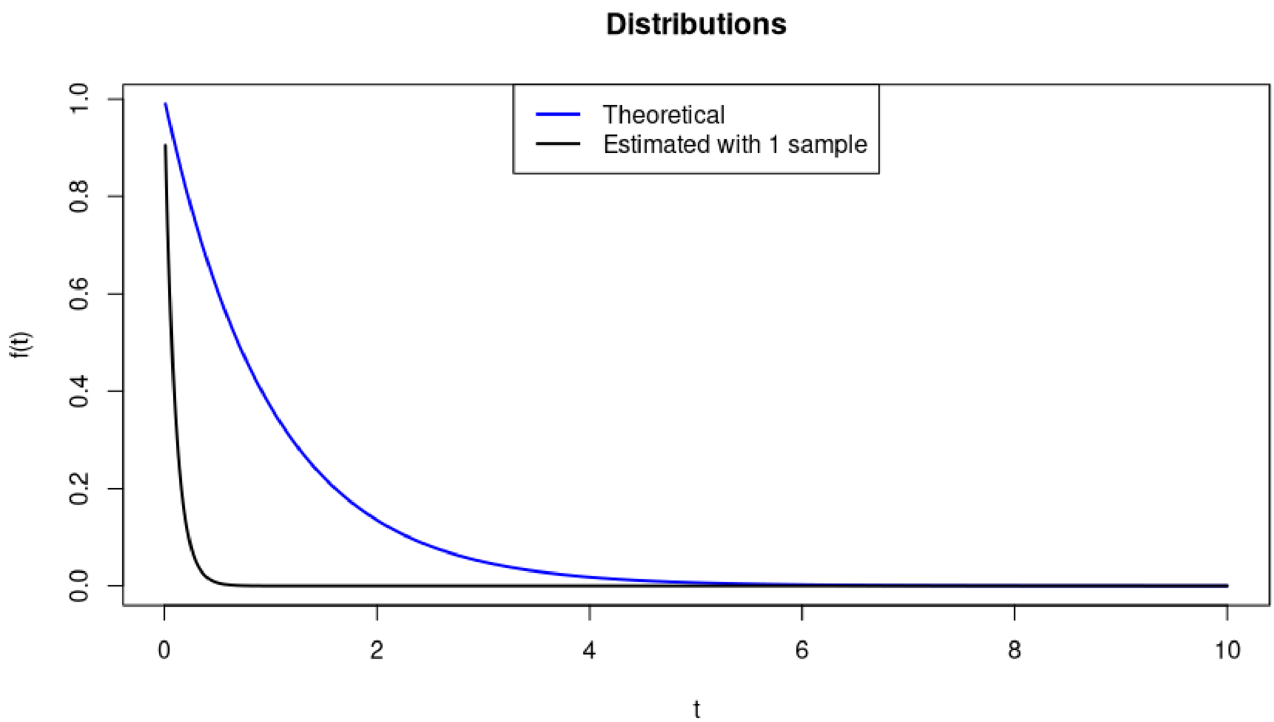
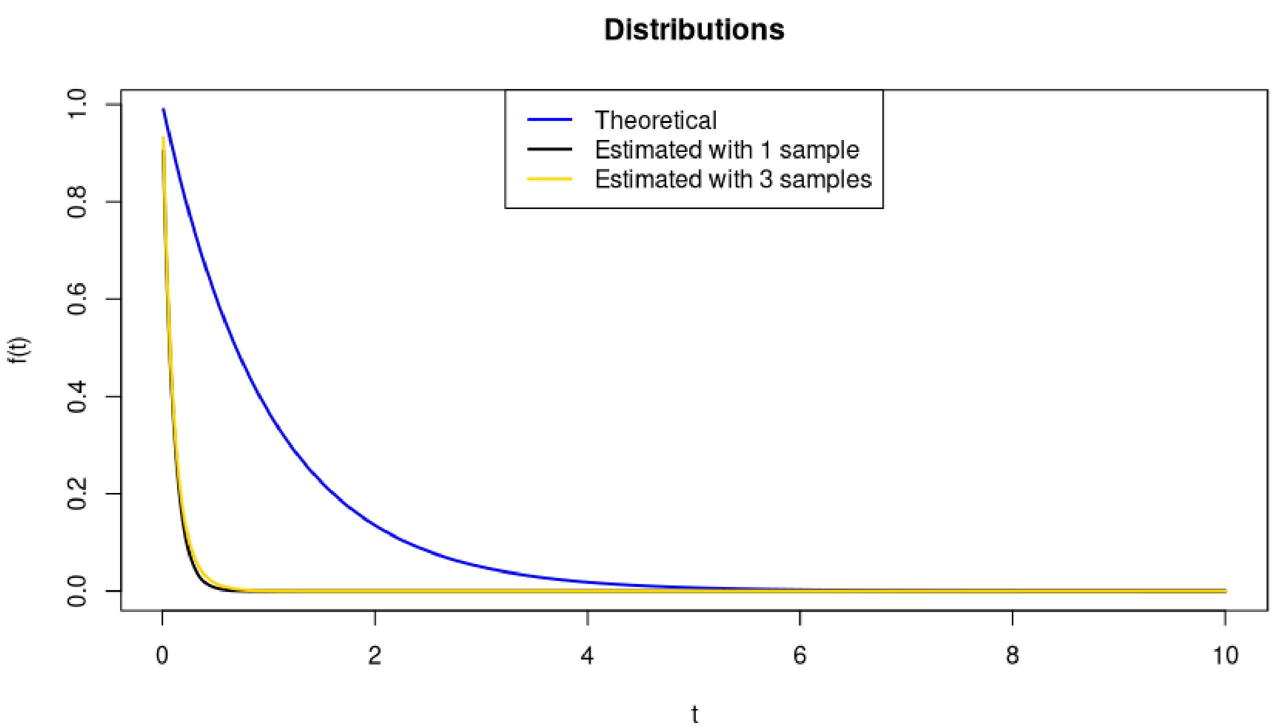
N1

```r
# task N1
# Parameters

theta = 1
n = 1000
sample1 = seq(1 / 100, 10, by = 0.01)
sample2 = seq(1 / 200, 12, by = 0.012)
sample3 = seq(1 / 150, 8, by = 0.008)

fnt1 = rep(0, n)
fnt2 = rep(0, n)
fnt3 = rep(0, n)

# theoretical function

f = function(t, theta){
  dexp(t, rate = theta, log = FALSE) / theta
}

# estimated function

fn = function(t, sum_sample){
  if (sum_sample > t){
    return((1 - t / sum_sample)^(n - 1))
  } else {
    return(0)
    }
}

for (i in 1:n){
  fnt1[i] = fn(sample1[i], sum(fn1))
  fnt2[i] = fn(sample2[i], sum(fn2))
  fnt3[i] = fn(sample3[i], sum(fn3))
}

# Plotting

plot(sample1, f(sample1, theta), type='l', col = 'blue', lwd = 2,
     main="Distributions", xlab="t", ylab="f(t)")
legend(x = 'top', col = c('blue', 'black'), lwd = 2,
       legend = c('Theoretical', 'Estimated with 1 sample'))
lines(sample1, fnt1, col = 'black', lwd = 2)

# Plotting mean of 3 samples

fnt_mean = rep(0, length(sample1))
for (i in 1:n){
  fnt_mean[i] = (fnt1[i] + fnt2[i] + fnt3[i]) / 3
}

plot(sample1, f(sample1, theta), type='l', col = 'blue', lwd = 2,
     main="Distributions", xlab="t", ylab="f(t)")
legend(x = 'top', col = c('blue', 'black', 'gold'), lwd = 2,
       legend = c('Theoretical',
                  'Estimated with 1 sample',
                  'Estimated with 3 samples'))
lines(sample1, fnt1, col = 'black', lwd = 2)
lines(sample1, fnt_mean, col = 'gold', lwd = 2)

# Mean value estimates theoretical distribution better than one-sample value
```


Distributions
Theoretical
Estimated with 1 sample
f(t)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
t


Distributions
Theoretical
Estimated with 1 sample
Estimated with 3 samples
f(t)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
t

N3

```r
# task N3

install.packages("stats4")
library(stats4)

# Parameters

alpha = 0.4
n = 1000

# Sample

X = rweibull(n, shape = 3, scale = alpha)

# Estimated alpha

LL = function(a){
  -sum(dweibull(X, shape = 3, scale = a, log=T))
}

mle(LL, start = 0.3, method = "L-BFGS-B", lower = c(0,0))


```

1:1 (Top Level)

Console Terminal Jobs

R 4.1.1 · ~/

```
> alpha = 0.4
> n = 1000
>
> # Sample
>
> X = rweibull(n, shape = 3, scale = alpha)
>
> # Estimated alpha
>
> LL = function(a){
+   -sum(dweibull(X, shape = 3, scale = a, log=T))
+ }
>
> mle(LL, start = 0.3, method = "L-BFGS-B", lower = c(0,0))

Call:
mle(minuslogl = LL, start = 0.3, method = "L-BFGS-B", lower = c(0,
    0))

Coefficients:
        a
0.4004409
>
```

N4

```r
# task N4

install.packages("tidyquant")
library(tidyquant)
library(ggplot2)
library(dplyr)
library(patchwork)
library(hrbrthemes)

# Downloading S&P500 prices from 2000 to 2020 years for every 2 years
# and calculating value at risk

values_at_risk = vector(length = 10, mode = 'list')
for (i in seq_along(values_at_risk)){
  from = sprintf("20%02d-01-01", (i - 1) * 2)
  to = sprintf("20%02d-01-01", (i - 1) * 2 + 2)
  getSymbols("^GSPC", from = from, to = to, warnings = FALSE)
  values_at_risk[[i]] = var(diff(log(drop(coredata(Ad(GSPC))))))
}

# Downloading S&P500 prices from 2000 to 2020 years

getSymbols("^GSPC", from = '2000-01-01',
           to = "2020-01-01",warnings = FALSE,
           auto.assign = TRUE)

GSPC = drop(coredata(Ad(GSPC)))[1:5030]

# Building dataframe

values_at_risk = unlist(rep(values_at_risk, each=503))

data = data.frame(
  day = as.Date("2000-01-01") + 0:5029,
  values_at_risk = values_at_risk,
  GSPC_price = GSPC
)

# Value used to transform the data

coeff = 6000000
```

```r
# Color constants

varsColor = "#66b2a3"
priceColor = rgb(0.19, 0.59, 0.89, 1)

# Plotting

ggplot(data, aes(x = day)) +
  geom_line(aes(y = GSPC_price), size = 0.3, color = priceColor) +
  geom_line(aes(y = values_at_risk * coeff), size = 1.5, color = varsColor) +
  scale_y_continuous(
    name = "S&P500 price",
    sec.axis = sec_axis(~./coeff, name = "Value at risk")
  ) +
  theme_ipsum() +
  theme(
    axis.title.y = element_text(color = priceColor, size = 13),
    axis.title.y.right = element_text(color = varsColor, size = 13)
  ) +
  ggtitle("S&P500 price and its value at risk from 2000 to 2020")
```

Name	Type	Value
values_at_risk	list [10]	List of length 10
[[1]]	double [1]	0.000190147
[[2]]	double [1]	0.0001922894
[[3]]	double [1]	4.538683e-05
[[4]]	double [1]	7.035973e-05
[[5]]	double [1]	0.000483249
[[6]]	double [1]	0.0001725132
[[7]]	double [1]	5.622454e-05
[[8]]	double [1]	7.331859e-05
[[9]]	double [1]	4.25569e-05
[[10]]	double [1]	8.918141e-05

S&P500 price and its value at risk from 2000 to 2020

T1

T1 $x_1, \dots, x_n \sim Exp(\theta) \sim \Gamma(1,\theta)$, $x \geq 0$

$F(x) = 1 - e^{-\theta x}$, $\theta > 0$

$\hat{\theta}_c = \frac{C(n)}{x_1 + \dots + x_n}$, $C(n) > 0$

• Найти $C(n)$ для несмещённой оценки $\hat{\theta}_c$

$x_i \sim \Gamma(1,\theta) \Rightarrow \sum_{i=1}^{n} x_i \sim \Gamma(\sum_{i=1}^{n} 1, \theta) = \Gamma(n,\theta)$

$p(x) = \frac{\theta^n}{\Gamma(n)} \cdot x^{n-1} \cdot e^{-\theta x}$

$E(\hat{\theta}_c) = \theta$

$E\left(\frac{C(n)}{x_1 + \dots + x_n}\right) = \theta$

Сделаем замену $Y = g(X)$, $y_i = \frac{1}{x_i}$, тогда

$$E(g(X)) = \int_{-\infty}^{+\infty} g(x) \cdot p(X) \cdot dx \, \mathbb{I}[x \geq 0] = \int_0^{+\infty} g(x) \cdot p(X) dx =$$

$$= \int_0^{+\infty} \frac{1}{x} \cdot \frac{\theta^n}{\Gamma(n)} \cdot x^{n-1} \cdot e^{-\theta x} \cdot dx = \frac{\theta^n}{\Gamma((n-1)+1)} \cdot \int_0^{+\infty} x^{n-2} \cdot e^{-\theta x} dx = \begin{bmatrix} \theta x = z \\ z = x/\theta \end{bmatrix}$$

$$= \frac{\theta^n}{(n-1)!} \cdot \int_0^{+\infty} \frac{1}{\theta} \cdot \frac{z^{((n-1)-1)}}{\theta^{n-2}} \cdot e^{-z} dz = \frac{\theta^n}{(n-1)!} \cdot \frac{1}{\theta^{n-1}} \cdot \Gamma((n-2)+1) =$$

$$= \frac{\theta}{(n-1)!} \cdot (n-2)! = \frac{\theta}{n-1}$$

$E(C(n) \cdot E(g(X))) = \theta$

$C(n) \cdot \frac{\theta}{(n-1)} = \theta$

$\boxed{C(n) = n-1}$ для несмещённой $\hat{\theta}_c$

• Найти $C(n)$ для состоятельной оценки $\hat{\theta}_c$

$\hat{\theta}_c \xrightarrow[n\to\infty]{P} \theta$

$\hat{\theta}_c = \frac{C(n)}{\sum_{i=1}^{n} x_i} = \frac{C(n)}{n} \cdot \frac{n}{\sum_{i=1}^{n} x_i}$

по ЗБЧ $\frac{1}{n} \cdot \sum_{i=1}^{n} x_i \xrightarrow[n\to\infty]{} E(X) = \frac{1}{\theta}$ (для $X \sim Exp(\theta)$)

Тогда $\frac{C(n)}{n} \cdot \frac{1}{1/\theta} = \theta \Rightarrow \frac{C(n)}{n} \cdot \theta = \theta \Rightarrow \boxed{C(n) = n}$

для состоятельной $\hat{\theta}_c$

T2

T2 $X_1, \dots, X_n \sim Unif([0,1])$

Плотность r-того члена задаётся β-распределением

$$P_{u(r)}(x) = \frac{n!}{(r-1)!(n-r)!} \cdot x^{r-1} \cdot (1-x)^{n-r}$$

I. Найдём матожидание

$$EX_{(r)} = \int_0^1 x \cdot \frac{n!}{(r-1)!(n-r)!} \cdot x^{r-1} \cdot (1-x)^{n-r}\, dx = \frac{n!}{(r-1)!(n-r)!} \cdot \int_0^1 x^{(r+1)-1} \cdot (1-x)^{(n-r+1)-1}\, dx =$$

$$= \frac{n!}{(r-1)!(n-r)!} \cdot B(r+1, n-r+1) = \frac{\cancel{n!} \cdot r! \cancel{(n-r)!}}{(r-1)! \cancel{(n-r)!} \cdot \cancel{(n+1)!}} = \frac{r}{n+1}$$

$$\boxed{EX_{(r)} = \frac{r}{n+1}}$$, ЧТД

II. Найдём экстремумы функции плотности

$$(p_{u(r)}(x))' = \frac{n!}{(r-1)!(n-r)!} \cdot \left((r-1)x^{r-2} \cdot (1-x)^{n-r} - (n-r)(1-x)^{n-r-1} \cdot x^{r-1}\right) = 0$$

$$(r-1)x^{r-2} \cdot (1-x)^{n-r} = (n-r)(1-x)^{n-r-1} \cdot x^{r-1}$$

$$\frac{r-1}{n-r} = \frac{(1-x)^{n-r-1} \cdot x^{r-1}}{(1-x)^{n-r} \cdot x^{r-2}}$$

$$\frac{r-1}{n-r} = \frac{x}{1-x}$$

$$(r-1)(1-x) - (n-r)x = 0$$

$$r - \cancel{rx} - 1 + x - nx + \cancel{rx} = 0$$

$$(r-1) - x(n-1) = 0$$

$$\boxed{x = \frac{r-1}{n-1}}$$ – экстремум $x_э$

Найдём вторую производную функции плотности и подставим $x_э = \frac{r-1}{n-1}$ и $1 - x_э = \frac{n-1}{n-1} - \frac{r-1}{n-1} = \frac{n-r}{n-1}$. Если при этом получим значение < 0, то $x_э$ является точкой максимума, а значит и модой величины $X_{(r)}$.

$$(p_{u(r)}(x))'' = \frac{n!}{(r-1)!(n-r)!} \cdot \left((r-2)(r-1)x^{r-3} \cdot (1-x)^{n-r} - (r-1)(n-r)x^{r-2}(1-x)^{n-r-1} +\right.$$

$$\left. + (n-r)(n-r-1)x^{r-1}(1-x)^{n-r-2} - (r-1)(n-r)x^{r-2}(1-x)^{n-r-1}\right)$$

$$\frac{n!}{(r-1)!(n-r)!} \cdot \left((r-2)(r-1) \cdot \frac{(r-1)^{r-3}}{(n-1)^{r-3}} \cdot \frac{(n-r)^{n-r}}{(n-1)^{n-r}} - 2 \cdot (r-1)(n-r)\frac{(r-1)^{r-2}}{(n-1)^{r-2}} \cdot \frac{(n-r)^{n-r-1}}{(n-1)^{n-r-1}} + (n-r)(n-r-1) \cdot\right.$$

$$\left. \cdot \frac{(r-1)^{r-1}}{(n-1)^{r-1}} \cdot \frac{(n-r)^{n-r-2}}{(n-1)^{n-r-2}}\right) = \frac{n!}{(r-1)!(n-r)!} \cdot \left(\frac{((r-1)-1)(r-1)^{r-2}(n-r)^{n-r}}{(n-1)^{n-3}} - \frac{2 \cdot (r-1)^{r-1}(n-r)^{n-r}}{(n-1)^{n-3}} +\right.$$

$$\left. + \frac{(r-1)^{r-1} \cdot ((n-r)-1)(n-r)^{n-r-1}}{(n-1)^{n-3}}\right) = \frac{n!}{(r-1)!(n-r)!} \cdot \left(\frac{(r-1)^{r-1}(n-r)^{n-r} \cdot (1-2+1) - (r-1)^{r-2}(n-r)^{n-r}}{(n-1)^{n-3}} -\right.$$

$$\left. - \frac{(r-1)^{r-1}(n-r)^{n-r-1}}{(n-1)^{n-3}}\right) = \frac{n!}{(r-1)!(n-r)!} \cdot \left(-\frac{(r-1)^{r-2}(n-r)^{n-r-1}}{(n-1)^{n-3}}(n - \cancel{r} + \cancel{r} - 1)\right) = -\frac{n!}{(r-1)!(n-r)!}\left(\frac{(r-1)^{r-2}(n-r)^{n-r-1}}{(n-1)^{n-2}}\right)$$

Так как множитель с факториалами >0, и для $r > 1$ и $r < n \Rightarrow n > 2$, выражение $\frac{(r-1)^{r-2}(n-r)^{n-r-1}}{(n-1)^{n-2}}$ также будет >0, то учитывая минус перед множителями, всё выражение будет <0, значит, $x_э = \frac{r-1}{n-1}$ является модой, чтд.

T4

T4 | $P_\theta = \{ Law(\xi^2),\ \xi \sim N(0,\theta)\}$ => $\mu = 0$, $\sigma^2 = \theta$, $\sigma = \sqrt{\theta}$

Плотность вероятности

$$p(\xi) = \frac{1}{\sqrt{2\pi\theta}} \cdot e^{-\frac{\xi^2}{2\theta}}$$

Выразим плотность ξ^2

$$P\{\xi^2 \le \xi\} = P\{\xi \le \sqrt{\xi}\} \Rightarrow p_{\xi^2}(\xi) = \frac{1}{2\sqrt{\xi}} \cdot p(\sqrt{\xi})$$

Тогда плотность вероятности перепишем в виде $p(x,v) = g(x) \cdot e^{xv - d(}$

$$p(x,v) = \frac{1}{2\sqrt{\pi x}} \cdot e^{-\frac{x}{2\theta} - \frac{1}{2}\ln\theta}$$

Получаем $v = -\frac{1}{2\theta}$ => $\theta = -\frac{1}{2v}$, $d(v) = \frac{1}{2}\ln\left(-\frac{1}{2v}\right)$

Тогда а) семейство является экспоненциальным

б) $E(X) = d'(v) = \left(\frac{1}{2}\ln\left(-\frac{1}{2v}\right)\right)' = -\frac{1}{2v} = \theta$, $\boxed{E(X) = \theta}$

$Var(X) = d''(v) = \left(-\frac{1}{2v}\right)' = \frac{1}{2v^2} = 2\theta^2$, $\boxed{Var(X) = 2\theta^2}$

Оценки параметра θ ММП и ММ совпадают для экспоненциальных семейств, и будут равны

$$(d')^{-1} y = -\frac{1}{2y}$$

$\hat{v} = -\frac{1}{2\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)}$ для выборки $x_1, \dots, x_n$

$$\hat{\theta} = -\frac{1}{2\hat{v}} = -\frac{1}{2} \cdot \left(-2 \cdot \frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \boxed{\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$