Modeling extreme values with a GEV mixture probability distributions

Standard Gumbel distribution

Pascal Alain Dkengne Sielenou

2023-10-28

```
# Load useful functions
```

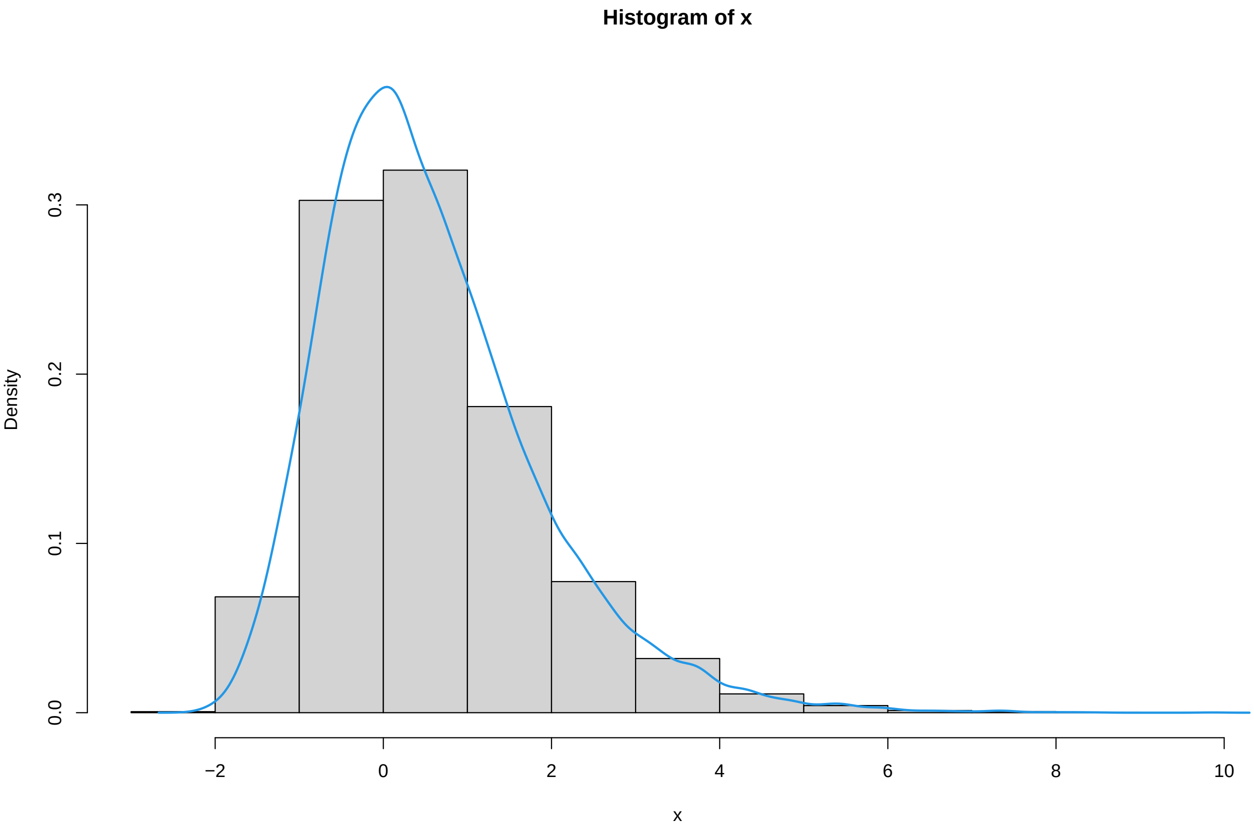
```
path <- ".."
```

```
xfun::in_dir(dir = path, expr = source("./src/generate_gev_sample.R"))
xfun::in_dir(dir = path, expr = source("./src/calculate_gev_inverse_cdf.R"))
xfun::in_dir(dir = path, expr = source("./src/estimate_gev_mixture_model_parameters.R"))
xfun::in_dir(dir = path, expr = source("./src/plot_gev_mixture_model_pdf.R"))
xfun::in_dir(dir = path, expr = source("./src/plot_gev_mixture_model_cdf.R"))
xfun::in_dir(dir = path, expr = source("./src/estimate_gev_mixture_model_quantile.R"))


# Generate a random sample
```

```
n <- 20000
```

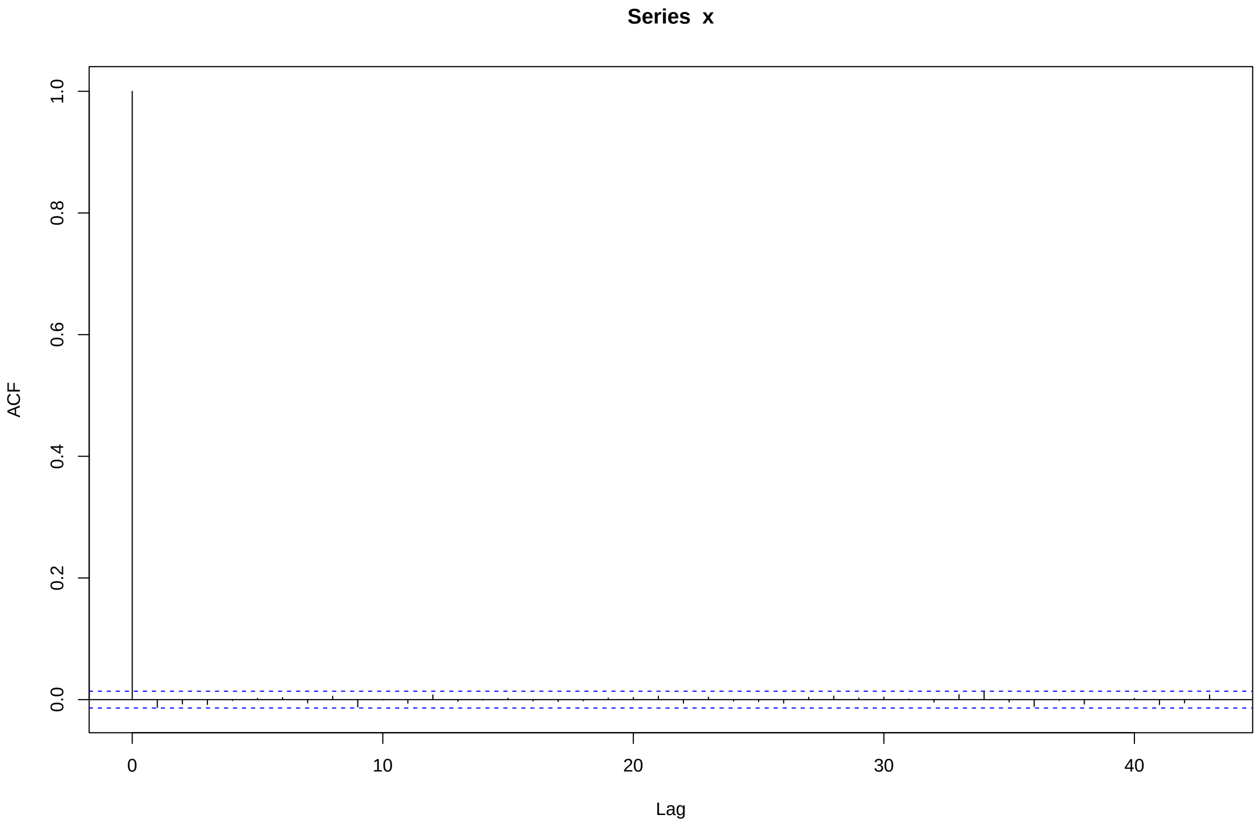
```
loc <- 0
scale <- 1
shape <- 0

set.seed(1122)
x <- generate_gev_sample(n = n, loc = loc, scale = scale, shape = shape)

# Histogram of all data
```

```
dens_x <- density(x)
hist(x, prob = TRUE, ylim = range(dens_x$y))
lines(dens_x, lwd = 2, col = 4)
```



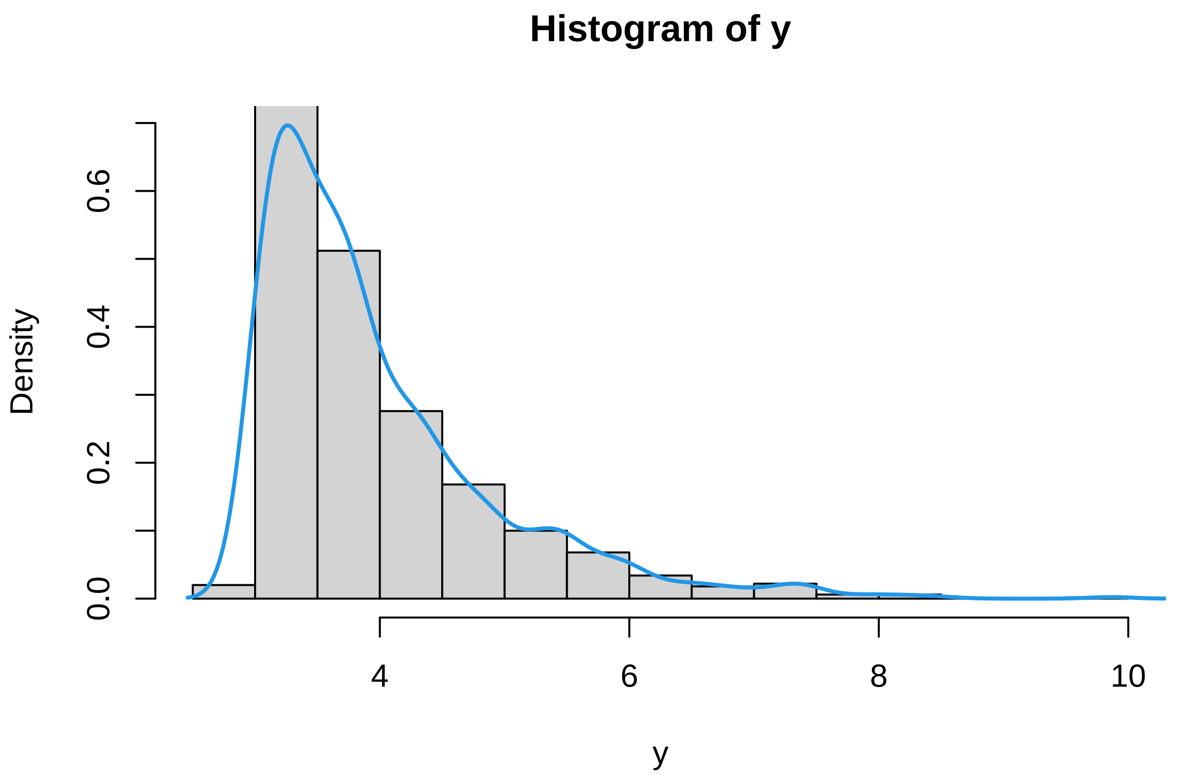
```
# Autocorrelation function of all data
```

```
acf(x)
```



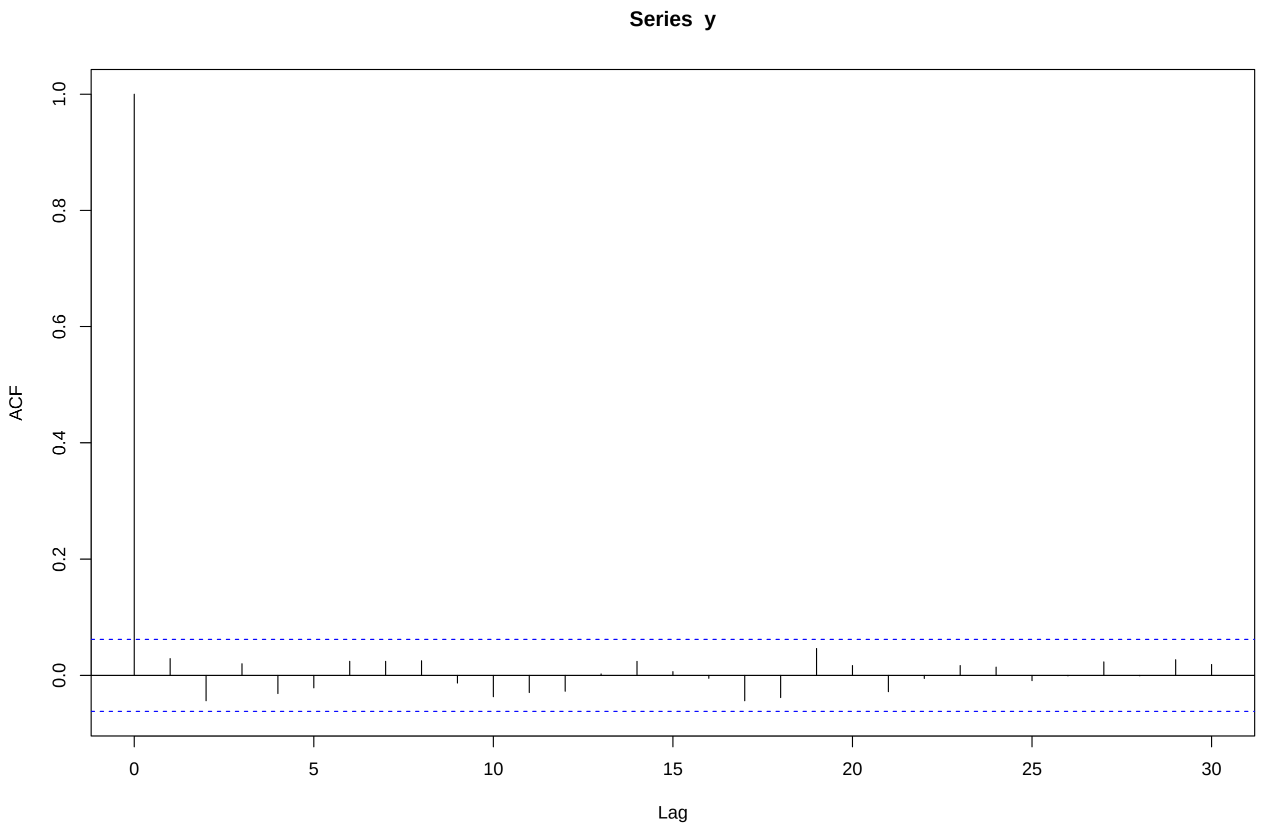
```
# Histogram of the largest data
```

```
nlargest <- 1000
y <- extract_nlargest_sample(x, n = nlargest)
dens_y <- density(y)
hist(y, prob = TRUE, ylim = range(dens_y$y))
lines(density(y), lwd = 2, col = 4)
```



```
# Autocorrelation function of the largest data
```

```
acf(y)
```



```
# Estimation of gev mixture models
```

```
gev_mixture_model <- estimate_gev_mixture_model_parameters(x = x,
                                                           block_sizes = 2:20,
                                                           minimum_nblocks = 50,
                                                           threshold = NULL,
                                                           nlargest = nlargest,
                                                           confidence_level = 0.95,
                                                           use_extremal_index = TRUE,
                                                           use_lower_threshold = FALSE,
                                                           maximum_iterations = 1500,
                                                           log_mv = TRUE,
                                                           log_pw = TRUE,
                                                           trace = FALSE,
                                                           method = "MLE")
```

```
##   Successful convergence.
##   Successful convergence.
```

```
gev_mixture_model$extremal_indexes
```

```
##            2            3            4            5            6            7 
## 1.0000000000 1.0000000000 1.0000000000 0.9505128397 1.0000000000 0.9227004051 
##            8            9           10           11           12           13 
## 1.0000000000 0.8986079715 0.8814126530 0.8920296849 0.8986079715 0.8920296849 
##           14           15           16           17           18           19 
## 0.9412167819 0.9109104376 0.9663117674 1.0000000000 1.0000000000 0.9153737685 
##           20 
## 1.0000000000
```

```
gev_mixture_model$normalized_gev_parameters_object
```

```
##       loc_star   scale_star       shape_star
## 2  3.454605293 0.5832000521  0.2111565171035
## 3  3.451575296 0.6484150395  0.1386579009456
## 4  3.362602086 0.7558493950  0.0722082556615
## 5  3.269780332 0.8947437879 -0.0005846449204
## 6  3.232093972 0.8874985063  0.0027627645238
## 7  3.214480062 0.9326121778 -0.0185431111737
## 8  2.799019802 1.2378015517 -0.0938714661724
## 9  2.995567295 1.0957280392 -0.0640452270516
## 10 2.831895207 1.2313663539 -0.0959486532973
## 11 2.898313459 1.1318309330 -0.0700506158632
## 12 2.769811517 1.2153314029 -0.0910494470469
## 13 2.738730574 1.2972204786 -0.1109421448437
## 14 3.173500173 1.0537637674 -0.0660448518493
## 15 2.962237100 1.1338401059 -0.0795657776488
## 16 2.943030605 1.1978828474 -0.0948710316732
## 17 2.640938537 1.3340390618 -0.1177243518245
## 18 3.339054187 0.9284680982 -0.0294881585819
## 19 2.200619920 1.5943004373 -0.1536709731556
## 20 3.387462307 0.9167513605 -0.0282185982086
```

```
gev_mixture_model$full_normalized_gev_parameters_object
```

```
##       loc_star   scale_star       shape_star
## 2  3.454605293 0.5832000521  0.2111565171035
## 3  3.451575296 0.6484150395  0.1386579009456
## 4  3.362602086 0.7558493950  0.0722082556615
## 5  3.224368182 0.8947703379 -0.0005846449204
## 6  3.232093972 0.8874985063  0.0027627645238
## 7  3.139394781 0.9340044925 -0.0185431111737
## 8  2.799019802 1.2378015517 -0.0938714661724
## 9  2.878022798 1.1032562032 -0.0640452270516
## 10 2.675515519 1.2463707743 -0.0959486532973
## 11 2.768476238 1.1409261103 -0.0700506158632
## 12 2.639247949 1.2272191436 -0.0910494470469
## 13 2.589572170 1.3137684319 -0.1109421448437
## 14 3.109533392 1.0579884440 -0.0660448518493
## 15 2.856043969 1.1422894450 -0.0795657776488
## 16 2.901913848 1.2017836365 -0.0948710316732
## 17 2.640938537 1.3340390618 -0.1177243518245
## 18 3.339054187 0.9284680982 -0.0294881585819
## 19 2.058685279 1.6161116717 -0.1536709731556
## 20 3.387462307 0.9167513605 -0.0282185982086
```

```
gev_mixture_model$automatic_weights_pw_shape
```

```
##               2               3               4               5               6
## 0.000000000e+00 0.000000000e+00 8.470329473e-22 2.183756817e-22 5.293955920e-23
##               7               8               9              10              11
## 2.117582368e-22 8.880514871e-02 5.148244031e-02 9.140432304e-02 5.899740946e-02
##              12              13              14              15              16
## 8.527398119e-02 1.101651068e-01 5.398471522e-02 7.090423010e-02 9.005591532e-02
##              17              18              19              20
## 1.186512494e-01 8.970553645e-03 1.636264762e-01 7.678450675e-03
```

```
gev_mixture_model$automatic_weights_pw_scale
```

```
##               2               3               4               5               6
## 0.0001143644717 0.0044736785713 0.0166985927128 0.0326824739968 0.0318410782290
##               7               8               9              10              11
## 0.0372289708539 0.0729358031798 0.0569578702091 0.0739924387575 0.0613464962649
##              12              13              14              15              16
## 0.0716307070344 0.0822977573383 0.0516802111331 0.0615053188276 0.0685155052070
##              17              18              19              20
## 0.0847939057861 0.0365874313074 0.1194879376382 0.0352294584811
```

```
gev_mixture_model$automatic_weights_pw_loc
```

```
##             2             3             4             5             6
## 0.03151327008 0.03163672781 0.03525914818 0.04087908271 0.04056410954
##             7             8             9            10            11
## 0.04469694513 0.06036235829 0.05669398836 0.06608959778 0.06177961953
##            12            13            14            15            16
## 0.06776967013 0.07010121330 0.04604487507 0.05771487937 0.05558768577
##            17            18            19            20
## 0.06769141099 0.03621714162 0.09515080238 0.03424747394
```

```
gev_mixture_model$weighted_normalized_gev_parameters_object[3, ]
```

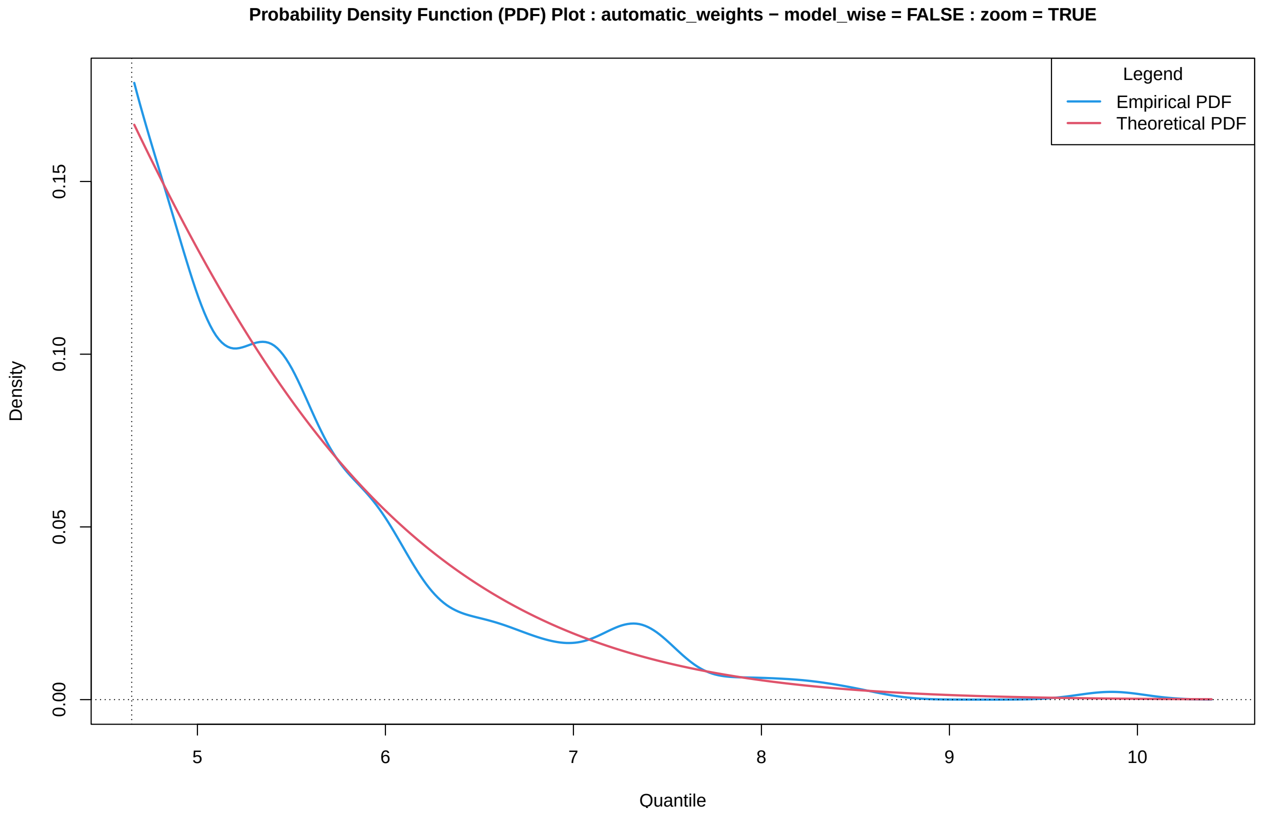
```
##                      loc_star  scale_star    shape_star
## automatic_weights 2.935561646 1.191014542 -0.1018671445
```

```r
gev_mixture_model$automatic_weights_mw
```

```
##                2                3                4                5                6 
## 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 0.000000000e+00 
##                7                8                9               10               11 
## 0.000000000e+00 2.646977960e-22 0.000000000e+00 4.528971983e-02 1.270549421e-21 
##               12               13               14               15               16 
## 9.235168642e-02 5.042293601e-02 3.069592430e-01 2.555299303e-01 2.239280883e-01 
##               17               18               19               20 
## 2.551839623e-02 0.000000000e+00 0.000000000e+00 0.000000000e+00
```

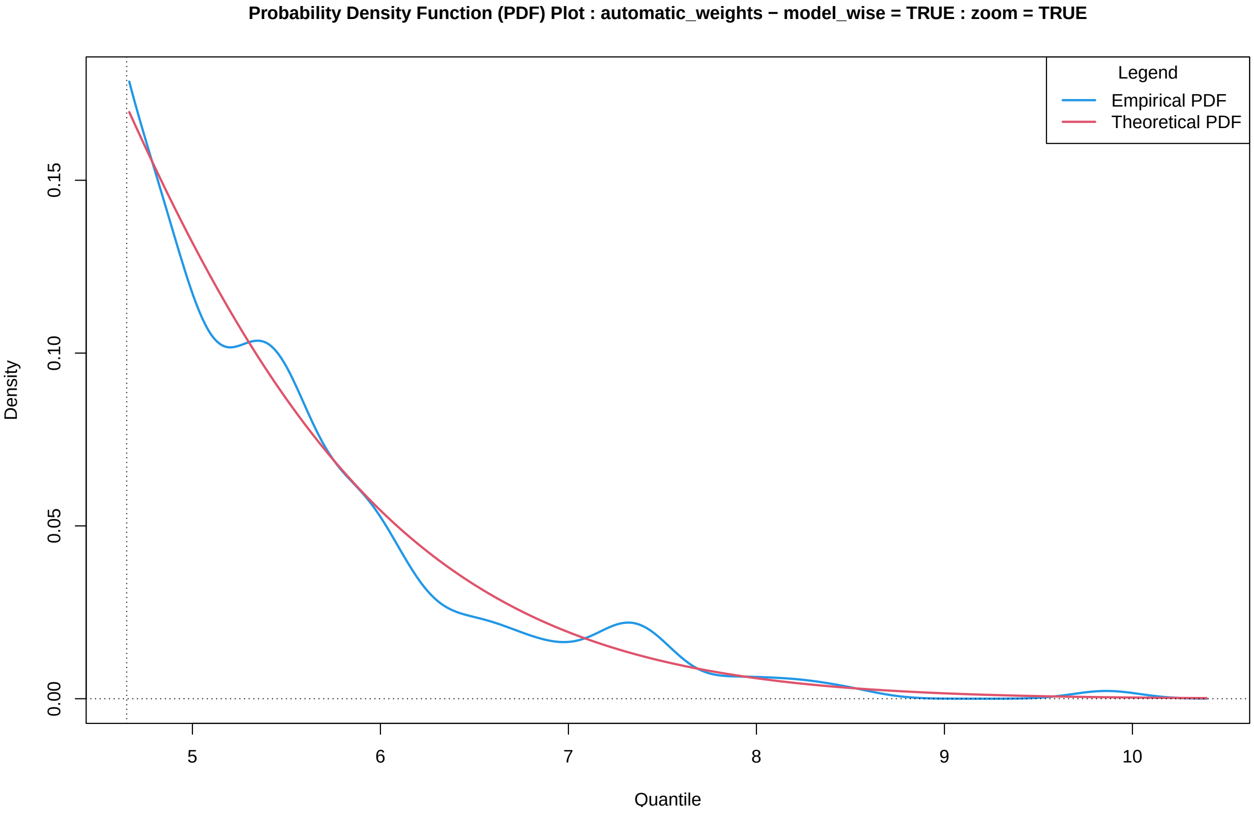
```r
# Model diagnostics

## GEV mixture model with respect to parameters
```

```r
plot_gev_mixture_model_pdf(gev_mixture_model,
                           type = "automatic_weights",
                           model_wise = FALSE,
                           zoom = TRUE,
                           xlab = "Quantile",
                           ylab = "Density",
                           main = "Probability Density Function (PDF) Plot")
```



```r
## GEV mixture model with respect to distribution functions
```

```r
plot_gev_mixture_model_pdf(gev_mixture_model,
                           type = "automatic_weights",
                           model_wise = TRUE,
                           zoom = TRUE,
                           xlab = "Quantile",
                           ylab = "Density",
                           main = "Probability Density Function (PDF) Plot")
```



```
# Estimation of an extreme quantile
```

```
estimator_types <- c("automatic_weights_mw",
                     "pessimistic_weights_mw",
                     "identic_weights_mw",
                     "automatic_weights_pw",
                     "pessimistic_weights_pw",
                     "identic_weights_pw",
                     "model_wise",
                     "parameter_wise",
                     "empirical")
```

```
alpha <- 10^(-14)

## Quantile from the true distribution
```

```
true_rl <- calculate_gev_inverse_cdf(p = 1 - alpha,
                                     loc = loc,
                                     scale = scale,
                                     shape = shape)

true_rl
```

```
## [1] 32.2369909
## Quantile from GEV mixture model with respect to parameters
```

```
rl_pw <- estimate_gev_mixture_model_quantile(gev_mixture_model,
                                             alpha = alpha,
                                             confidence_level = 0.95,
                                             do.ci = TRUE,
                                             estimator_type = estimator_types[4])

rl_pw
```

```
## [1] 14.03274952
```

```
## Quantile from GEV mixture model with respect to distribution functions
```

```
rl_mw <- estimate_gev_mixture_model_quantile(gev_mixture_model,
                                             alpha = alpha,
                                             confidence_level = 0.95,
                                             do.ci = TRUE,
                                             estimator_type = estimator_types[1])

rl_mw
```

```
## [1] 16.61795698
```

```
## Quantiles from equivalent estimated distributions in GEV mixture model with respect to parameters
```

```
est_rl_pw <- suppressWarnings(estimate_gev_mixture_model_quantile(gev_mixture_model,
                                                                  alpha = alpha,
                                                                  confidence_level = 0.95,
                                                                  do.ci = TRUE,
                                                                  estimator_type = estimator_types[8]))

est_rl_pw
```

```
##                lower      quantile         upper
## 2  -1548.2494713716 1327.21569113 4202.68085363
## 3   -342.8653806363  268.33605006  879.53748075
## 4    -87.6415477026   79.36552830  246.37260430
## 5    -12.1025045934   29.16570240   70.43390939
## 6    -18.3180755733   30.26025014   78.83857585
## 7    -10.4455983330   24.22097690   58.88755214
## 8      2.7597142331   15.13787066   27.51602709
## 9     -1.6050240150   17.45605868   36.51714137
## 10     2.0165143819   14.88001775   27.74352113
## 11    -3.1652142635   16.95539829   37.07601084
## 12     0.9197729052   15.17725685   29.43474079
## 13     2.3468996728   13.96958275   25.59226582
## 14    -3.7474302756   16.80635608   37.36014243
## 15    -2.1419903358   15.81091531   33.76382096
## 16    -0.5952710422   14.77892345   30.15311794
## 17     2.3700405431   13.61030101   24.85056147
## 18   -23.8227373416   21.53142934   66.88559603
## 19     3.6014287009   12.45779173   21.31415476
## 20   -26.7220637170   21.63967153   70.00140678
```

```
## Comparison of estimated quantiles
```

```
est_rl_pw_range <- range(as.matrix(est_rl_pw))


## Quantiles from equivalent estimated GEV distributions in GEV mixture model respect to distribution f
```

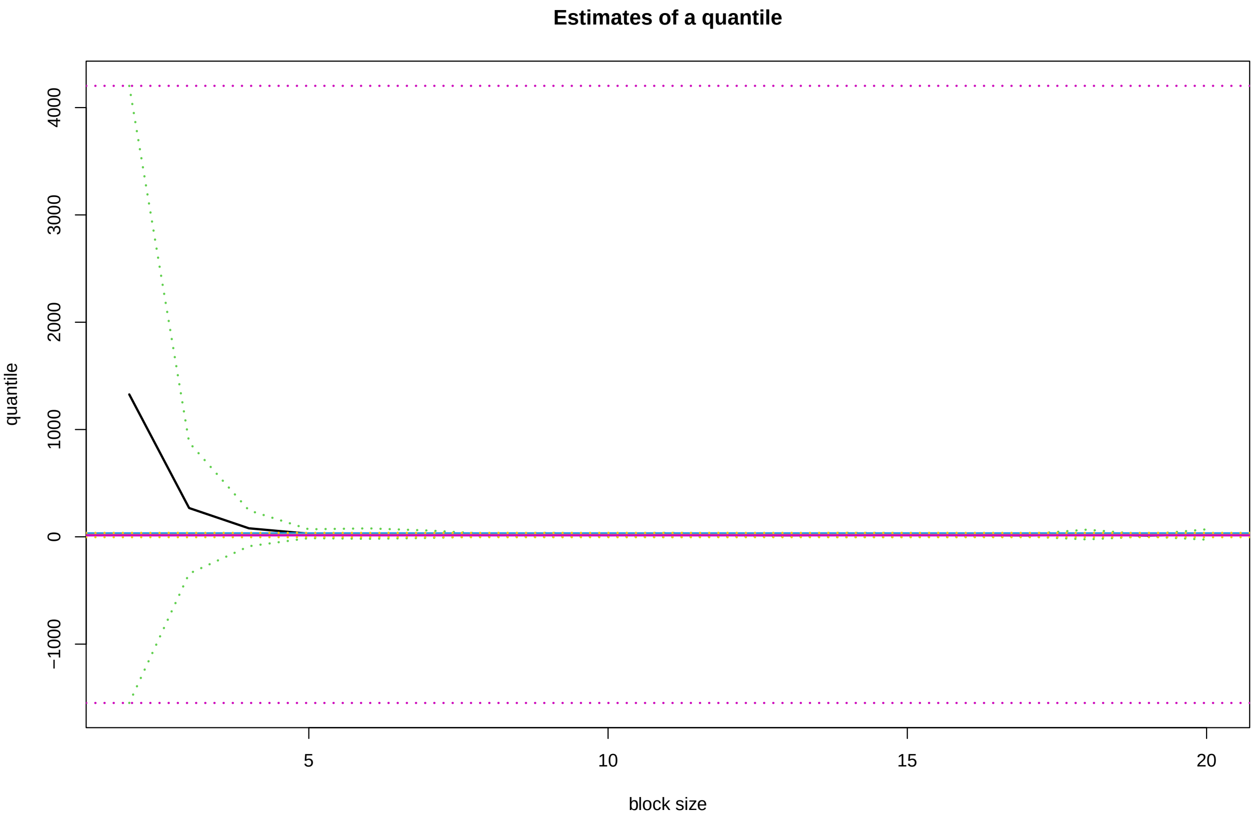
```
est_rl_mw <- suppressWarnings(estimate_gev_mixture_model_quantile(gev_mixture_model,
                                                                   alpha = alpha,
                                                                   confidence_level = 0.95,
                                                                   do.ci = TRUE,
                                                                   estimator_type = estimator_types[7]))
est_rl_mw
```

```
##              lower    quantile       upper
## 8   2.7597142331 15.13787066 27.51602709
## 10  2.0165143819 14.88001775 27.74352113
## 11 -3.1652142635 16.95539829 37.07601084
## 12  0.9197729052 15.17725685 29.43474079
## 13  2.3468996728 13.96958275 25.59226582
## 14 -3.7474302756 16.80635608 37.36014243
## 15 -2.1419903358 15.81091531 33.76382096
## 16 -0.5952710422 14.77892345 30.15311794
## 17  2.3700405431 13.61030101 24.85056147
```

```
est_rl_mw_range <- range(as.matrix(est_rl_mw))

est_rl_mw_range
```

```
## [1] -3.747430276 37.360142435
```

```
matplot(x = rownames(est_rl_pw),
        y = est_rl_pw,
        xlab = "block size",
        ylab = "quantile",
        main = "Estimates of a quantile",
        ylim = range(c(est_rl_pw_range, true_rl, rl_pw)),
        cex = 1,
        cex.lab = 1,
        cex.axis = 1,
        type = "l",
        lty = c("dotted", "solid", "dotted"),
        lwd = c(2,2,2),
        col = c(3, 1, 3))

abline(h = true_rl, col = 4, lwd = 2)
abline(h = rl_mw, col = 7, lwd = 2)
abline(h = rl_pw, col = 6, lwd = 2)
abline(h = est_rl_pw_range, col = 6, lty = "dotted", lwd = 2)
abline(h = est_rl_mw_range, col = 7, lty = "dotted", lwd = 2)
```



```
# Legend:
# blue: Quantile from the true distribution
# yellow: Quantile from GEV mixture model with respect to distribution functions
# pink: Quantile from GEV mixture model with respect to parameters
```