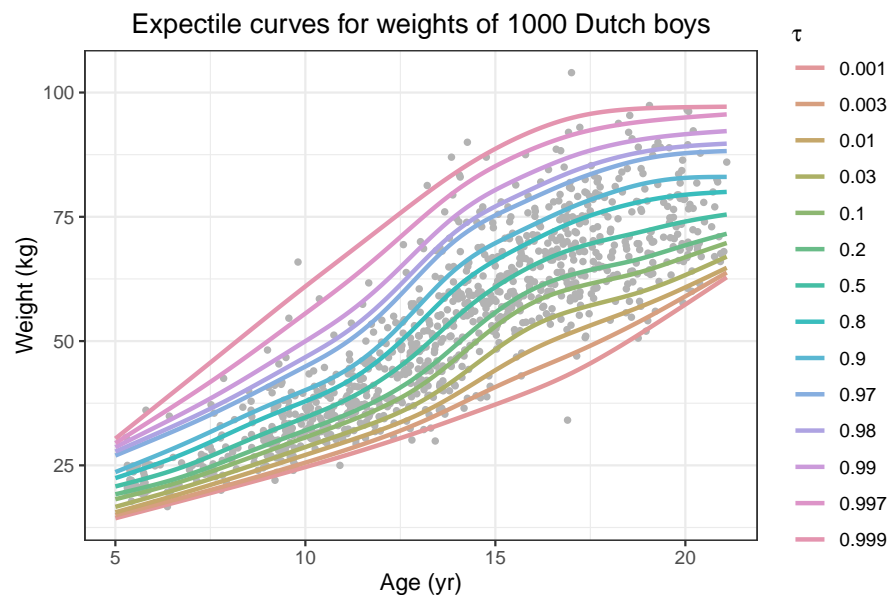# Expectile curves for body weight of 1000 boys (Dutch boys data)

Expectile curves for weights of 1000 Dutch boys

Expectile smoothing of weights of Dutch boys, with 20 cubic B-spline segments and $\lambda = 10$. R code in `f-exp-curves-1000.R`

```r
# Expectile curves for body weight of 1000 boys
# Paul Eilers and Brian Marx 2019
# A graph in "Practical Smoothing. The Joys of P-splines"

library(colorspace)
library(AGD)
library(ggplot2)
library(JOPS)

# Get the data
data(boys7482)
Dat = subset(boys7482, age > 5, select = c(age, wgt))
Dat = na.omit(Dat)
m0 = nrow(Dat)
m = 1000
set.seed(2018)
sel = sample(1:m0, m)
x = (Dat$age[sel])
y = Dat$wgt[sel]

# P-spline paremeters
xl = 5
xr = max(x)
nsegx = 20
bdeg = 3
pp = c(0.001, 0.003, 0.01, 0.03, 0.1, 0.2, 0.5, 0.8, 0.9, 0.97, 0.98, 0.99, 0.997, 0.999)
np= length(pp)

# Compute bases
B = bbase(x, xl, xr, nsegx, bdeg)
nbx = ncol(B)
xg = seq(xl, xr, length = 100)
Bg = bbase(xg, xl, xr, nsegx, bdeg)

# Compute penalty matrix
D = diff(diag(nbx), diff = 2)
```

```
lambdax = 10
P = lambdax * t(D) %*% D

# Fitting
Zg = NULL
for (p in pp) {
  b = 0.001
  dzmax = 1e-3 * max(y)
  z = 0
  for (it in 1:10) {
    r = y - z
    w = ifelse(r > 0, p, 1- p)
    W = diag(c(w))
    Q = t(B) %*% W %*% B
    a = solve(Q + P, t(B) %*% (w * y))
    znew = B %*% a
    dz = sum(abs(z - znew))
    if (dz < dzmax) break
    z = znew
  }
  Zg = cbind(Zg, Bg %*% a)
  cat(p, '\n')
}

cols = diverge_hcl(np, h = c(240, -30), c= 100, l = 60, power = 1)

# Dataframes for ggplot
DF1 = data.frame(x = x, y = y)
np = length(pp)
ng = length(xg)
DF2 = data.frame(x = rep(xg, np), y = as.vector(Zg),
                 p = as.factor(rep(pp, each = ng)))

# Plotting and saving graph
plt1 =  ggplot(DF1,  aes(x = x, y = y)) +
  geom_point(color = 'grey70', size = 1) +
  xlab('Age (yr)') + ylab('Weight (kg)') +
  ggtitle('Expectile curves for weights of 1000 Dutch boys') +
  geom_line(data = DF2, aes(x = x, y = y, group = p, color = p), size = 1) +
  JOPS_theme()  +
  scale_color_manual(values=rainbow_hcl(np, start = 10, end = 350)) +
  labs(color =  bquote(tau) )

plot(plt1)
```