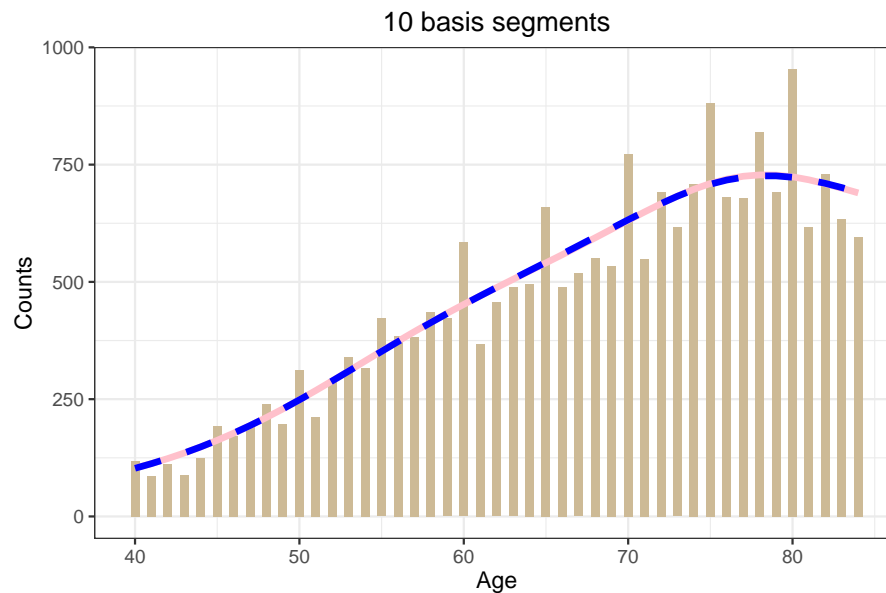# Quasi-likelihood fit of overdispersed (Greek death data)

10 basis segments

    Counts of male deaths in Greece, in 1960, versus age. P-spline fit with quasi-likelihood for 10 (broken blue curve) and 40 (pink curve) segments in the basis. The two curves are almost identical. R code in `f-Greece-QL.R`

```
# Quasi-likelihood fit of overdispersed counts with bases of different size (Greek death data)
# A graph in the book 'Practical Smoothing. The Joys of P-splines'
# Paul Eilers and Brian Marx, 2019

library(ggplot2)
library(gridExtra)
library(JOPS)

# Read the data (Greek number of deaths)
data(Greece_deaths)
Dat = subset(Greece_deaths, 40 <= Age & Age < 85)  # Get rid of grouped data above 84
sel = 'Males'
x = Dat$Age
y = Dat[, sel]
m = length(x)

nsegs = c(10, 40)

M = NULL
for (nseg in nsegs) {
  # Create basis and penalty matrix
  B = bbase(x, nseg = nseg)
  n = ncol(B)
  D = diff(diag(n), diff = 2)
  P = t(D) %*% D

# Compute P-spline Poisson fit
  eta = log(y + 1)
  a = rep(0, n)
  mon = T
  for (it in 1:40) {
    mu = exp(eta)
    z = y - mu + mu * eta
    G = t(B) %*% diag(c(mu)) %*% B
```

```r
    anew = solve(G +  lambda * P, t(B) %*% z)
    da = max(abs(anew - a))
    if (mon) cat(nseg, it, lambda, da, '\n')
    a = anew
    eta = B %*% a
    if (da < 1e-6) break
    H = solve(G + lambda * P, G)
    hd = diag(H)
    ed = sum(hd)
    tau2 = sum((D %*% a) ^2) / ed
    r = (y - mu) / sqrt(mu)
    sig2 = sum(r ^2) / (m - ed - 2)
    lanew = sig2 / tau2
    dla = (lanew - lambda) / lambda
    lambda = lanew
  }
  M = cbind(M, mu)
}

F2 = data.frame(x = x, y = y, mu1 = M[, 1], mu2 = M[, 2])

plt = ggplot(F2) +
  geom_bar(aes(x, y), stat = 'identity', fill = 'wheat3', width = 0.5) +
  geom_line(aes(x, mu2), col = 'pink', size = 1.5) +
  geom_line(aes(x, mu1), col = 'blue', size = 1.5, linetype = 2) +
  xlab('Age') + ylab('Counts') +
  ggtitle(paste(nsegs[1], 'basis segments')) +
  JOPS_theme()


# Make and save graph
plot(plt)
```