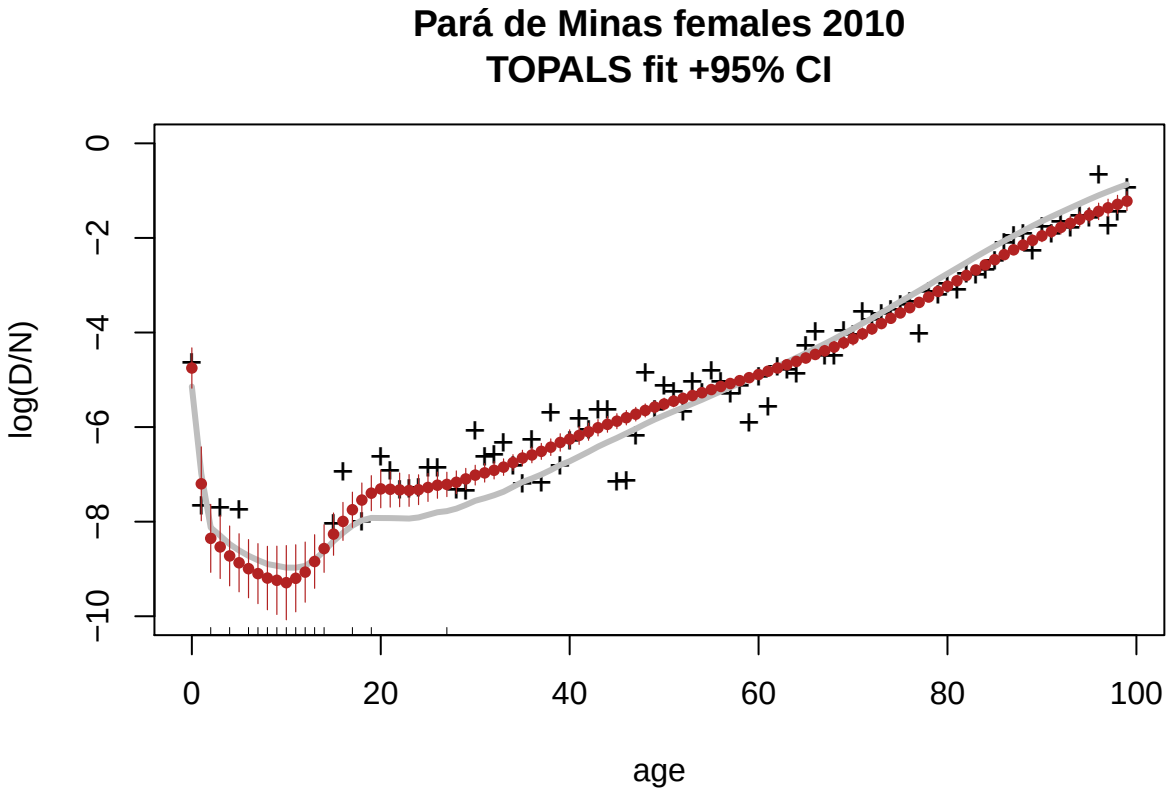
example_with_standard_error_calculation.R

Carl Schmertmann

Tue Aug 06 11:32:58 2019

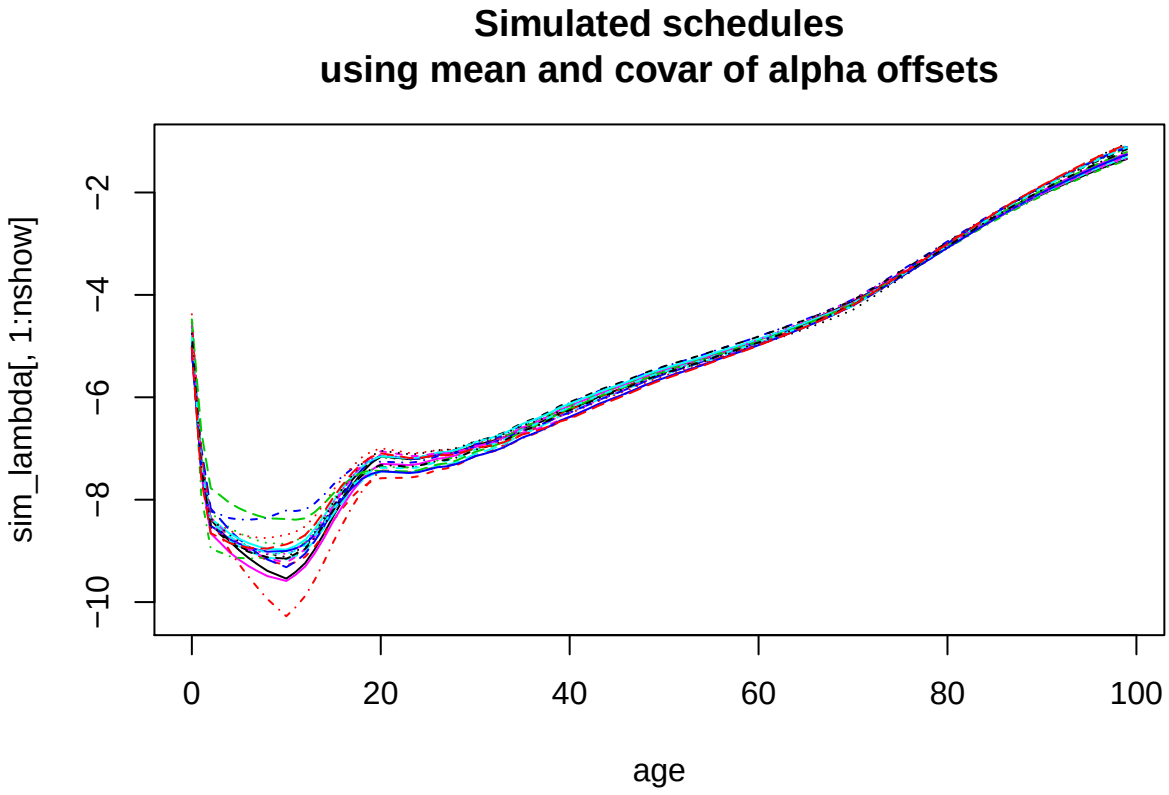
```
##############################################################################
# example using estimated standard errors of log mortality
# same data as Fig 1 in Gonzaga and Schmertmann 2016
##############################################################################

rm(list=ls())
graphics.off()
if (.Platform$OS.type == 'windows') windows(record=TRUE)

library(splines)

age = 0:99
B   = bs( 0:99, knots=c(0,1,10,20,40,70), degree=1 )

## Pará de Minas female data and HMD female standard

N = c(2289.1, 2278.6, 2321.8, 2367.8, 2431.4, 2473.6, 2530.7, 2661.3,
  2784.2, 2946.3, 2991.4, 3038, 3072.8, 3193.9, 3302.5, 3315.5,
  3272.8, 3182.7, 3165.3, 3183, 3232.3, 3259.5, 3221.7, 3178, 3077.6,
  3028.8, 3030.1, 3153.1, 3223.5, 3313.3, 3246.7, 3217.3, 3051.5,
  2991.8, 2887.3, 2810.4, 2770.1, 2791.6, 2851.9, 2878.9, 2877.1,
  2860.8, 2746, 2680.1, 2668.4, 2712.5, 2682.3, 2545.6, 2450.2,
  2360.2, 2308.6, 2222.3, 2154.8, 2118.1, 2057.1, 1953.8, 1809.8,
  1694, 1597.1, 1553.1, 1467.8, 1397.9, 1317.1, 1281.3, 1260.5,
  1214.1, 1147.3, 1045.1, 955.8, 886.1, 843.5, 815.5, 804.6, 785.4,
  749, 710, 659.8, 594.8, 557, 523.5, 497.1, 447.6, 400.2, 358.1,
  305.4, 253.5, 226.5, 200.2, 171.6, 134.6, 111.5, 94.1, 83.3,
  76.4, 54, 35.5, 16.6, 12, 13.4, 13.6)

D = c(21L, 1L, 0L, 1L, 0L, 1L, 0L, 0L, 0L, 0L, 0L, 0L, 0L, 0L, 0L,
      1L, 3L, 0L, 1L, 0L, 4L, 3L, 2L, 2L, 2L, 3L, 3L, 0L, 2L, 2L, 7L,
      4L, 4L, 5L, 3L, 2L, 5L, 2L, 9L, 3L, 5L, 8L, 6L, 9L, 9L, 2L, 2L,
      5L, 18L, 8L, 13L, 11L, 7L, 13L, 10L, 15L, 11L, 8L, 9L, 4L, 10L,
      5L, 11L, 10L, 9L, 16L, 20L, 11L, 10L, 16L, 14L, 22L, 18L, 20L,
      21L, 22L, 22L, 10L, 23L, 20L, 24L, 19L, 24L, 21L, 20L, 20L, 26L,
      27L, 24L, 13L, 18L, 13L, 15L, 12L, 11L, 7L, 8L, 2L, 3L, 5L)

HMD_female_std =
  c(-5.1434, -6.9847, -8.1263, -8.296, -8.475, -8.6084, -8.7217,
  -8.8148, -8.8967, -8.933, -8.9716, -8.9725, -8.9353, -8.8044,
  -8.6244, -8.4125, -8.2371, -8.084, -7.9726, -7.9208, -7.9235,
  -7.9242, -7.9292, -7.9346, -7.911, -7.8563, -7.8009, -7.7772,
  -7.7247, -7.6466, -7.5585, -7.5035, -7.4414, -7.3672, -7.2642,
  -7.1601, -7.0877, -7.0044, -6.908, -6.8007, -6.7229, -6.6227,
  -6.5213, -6.4139, -6.3197, -6.2298, -6.1335, -6.0364, -5.934,
  -5.8394, -5.7533, -5.6695, -5.5882, -5.5047, -5.422, -5.3325,
  -5.2444, -5.1596, -5.0766, -4.9908, -4.9012, -4.8086, -4.718,
```

```
  -4.6262, -4.5331, -4.4375, -4.3396, -4.2397, -4.1368, -4.0272,
  -3.919, -3.8085, -3.6981, -3.5805, -3.4648, -3.3469, -3.2321,
  -3.1152, -2.9968, -2.8732, -2.7562, -2.6382, -2.5215, -2.4009,
  -2.2871, -2.1739, -2.0632, -1.9559, -1.8517, -1.7465, -1.6463,
  -1.5501, -1.4579, -1.3665, -1.2765, -1.1861, -1.1002, -1.02,
  -0.9429, -0.8695)
#####################################################

source('TOPALS_fit function.R')

fit = TOPALS_fit( N, D, HMD_female_std, details=TRUE)

fitted_logmx = HMD_female_std + B %*% fit$a

# standard errors for fitted logmx

se_logmx = sqrt( diag (B %*% fit$covar %*% t(B)) )


plot( age, log(D/N), pch='+',cex=1.2,ylim=c(-10,0),
      main='Pará de Minas females 2010\nTOPALS fit +95% CI')
rug(age[D==0], side=1, ticksize=.015)
lines(age, HMD_female_std, lty=1, lwd=3, col='grey')
points(age, fitted_logmx, cex=.80, pch=16, col='firebrick')

L = fitted_logmx - 1.96 * se_logmx
H = fitted_logmx + 1.96 * se_logmx

segments( age, L, age, H, col='firebrick', lwd=.60)
```




```
# uncertainty about alpha offsets
# means uncertainty about log mortality
# and about indices like e0
```

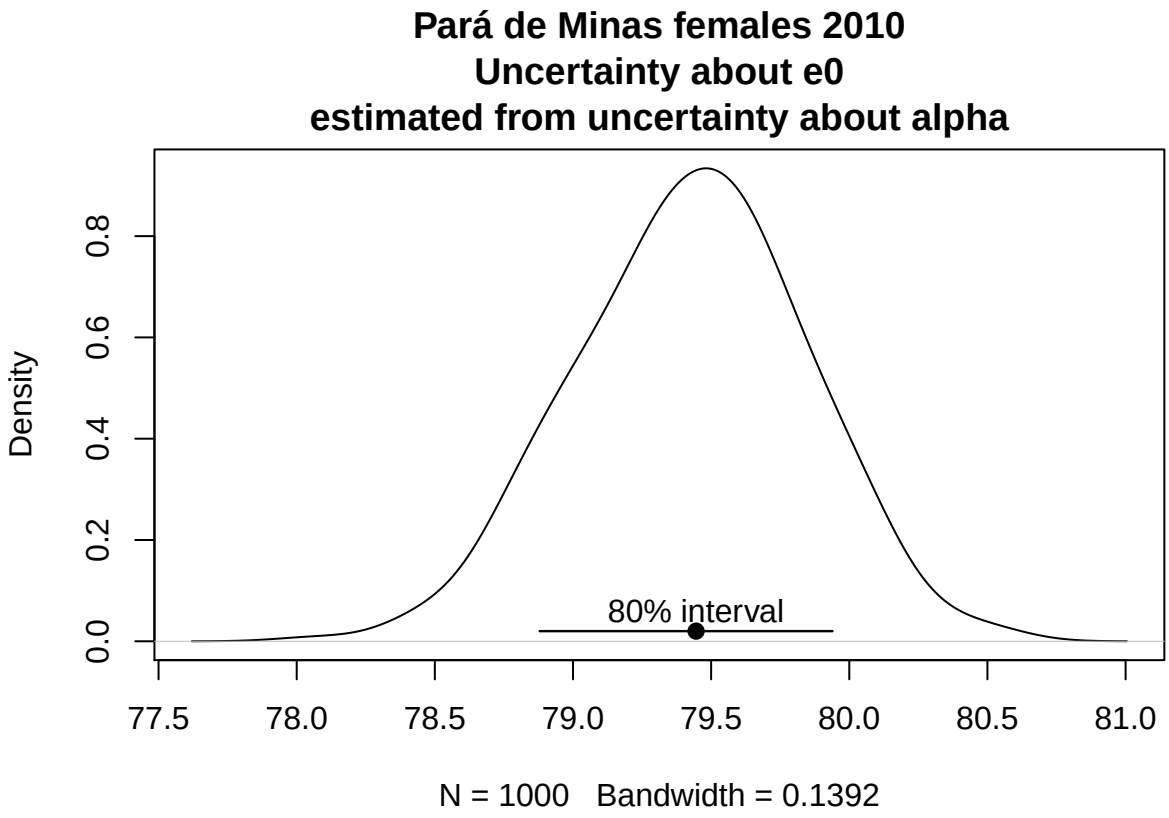
```
# Cholesky decomp of Var(alpha-hat)

CH = t( chol( fit$covar))

## simulate alpha vectors using a multivar normal approx

nsim  = 1000
nshow = 20

sim_alpha = as.numeric(fit$a) +
             CH %*% matrix( rnorm(nsim*length(fit$a)),
                            ncol=nsim)

sim_lambda = HMD_female_std + B %*% sim_alpha

matplot( age, sim_lambda[,1:nshow], type='l',
         main='Simulated schedules\nusing mean and covar of alpha offsets')
```




```
# trapez approx of life expectancy from a logmx schedule over ages 0..99
e0 = function(logmx) {
  mx = exp(logmx)
  px = exp(-mx)
  lx = c(1,cumprod(px))
  return( sum(head(lx,-1) + tail(lx,-1)) / 2)
}

sim_e0 = apply(sim_lambda, 2, 'e0')

Q10 = quantile(sim_e0, .10)
Q50 = quantile(sim_e0, .50)
Q90 = quantile(sim_e0, .90)

round( head( sim_e0), 2)
```

```
## [1] 79.76 79.74 79.42 78.35 79.91 79.37
```

```
plot( density(sim_e0, adjust=1.5),
      main='Pará de Minas females 2010\nUncertainty about e0\nestimated from uncertainty about alpha')
points( Q50, .02, pch=16, cex=1.2)
segments(Q10, .02, Q90, .02, lwd=1.2)
text( Q50, .06, '80% interval')
```


Pará de Minas females 2010
Uncertainty about e0
estimated from uncertainty about alpha
Density
0.0
0.2
0.4
0.6
0.8
80% interval
77.5
78.0
78.5
79.0
79.5
80.0
80.5
81.0
N = 1000 Bandwidth = 0.1392