0.1 Estimation and reporting of linear and curved effects

In this exercise we will use the `testisDK` data from the `Epi` package:

1. First we load the diet data, inspect the dataset, and convert the dates to clandar years; a bit easier to work with:

```
library( Epi )
sessionInfo()
```

```
R version 3.1.0 (2014-04-10)
Platform: i386-w64-mingw32/i386 (32-bit)

locale:
[1] LC_COLLATE=Danish_Denmark.1252  LC_CTYPE=Danish_Denmark.1252
[3] LC_MONETARY=Danish_Denmark.1252 LC_NUMERIC=C
[5] LC_TIME=Danish_Denmark.1252

attached base packages:
[1] utils     datasets  graphics  grDevices stats     methods   base

other attached packages:
[1] Epi_1.1.65     foreign_0.8-61

loaded via a namespace (and not attached):
[1] tools_3.1.0
```

```
data( testisDK )
str( testisDK )
```

```
data.frame:        4860 obs. of  4 variables:
 $ A: num  0 1 2 3 4 5 6 7 8 9 ...
 $ P: num  1943 1943 1943 1943 1943 ...
 $ D: num  1 1 0 1 0 0 0 0 0 0 ...
 $ Y: num  39650 36943 34588 33267 32614 ...
```

```
head( testisDK )
```

```
  A    P D        Y
1 0 1943 1 39649.50
2 1 1943 1 36942.83
3 2 1943 0 34588.33
4 3 1943 1 33267.00
5 4 1943 0 32614.00
6 5 1943 0 32020.33
```

We can tabulate both failures and person-years using either `xtabs` or `stat.table`, the latter is a bit more versatile, because we can get rates too:

```
round( ftable( xtabs( cbind(D,PY=Y/1000) ~ I(floor(A/10)*10) +
                                           I(floor(P/10)*10),
                      data=testisDK ),
               row.vars=c(3,1) ), 1 )
```

```
                       I(floor(P/10) * 10)   1940   1950   1960   1970   1980   1990
   I(floor(A/10) * 10)
D  0                                         10.0    7.0   16.0   18.0    9.0   10.0
   10                                        13.0   27.0   37.0   72.0   97.0   75.0
```

```
   20                                        124.0  221.0  280.0  535.0  724.0  557.0
   30                                        149.0  288.0  377.0  624.0  771.0  744.0
   40                                         95.0  198.0  230.0  334.0  432.0  360.0
   50                                         40.0   79.0  140.0  151.0  193.0  155.0
   60                                         29.0   43.0   54.0   83.0   82.0   44.0
   70                                         18.0   26.0   35.0   41.0   40.0   32.0
   80                                          7.0    9.0   13.0   19.0   18.0   21.0
PY 0                                        2604.7 4037.3 3885.0 3820.9 3070.9 2165.5
   10                                       2135.7 3505.2 4004.1 3906.1 3847.4 2261.0
   20                                       2225.5 2923.2 3401.6 4028.6 3941.2 2824.6
   30                                       2195.2 3058.8 2856.2 3410.6 3968.8 2728.4
   40                                       1874.9 2980.1 2986.8 2823.1 3322.6 2757.7
   50                                       1442.8 2426.5 2796.6 2813.3 2635.0 2069.2
   60                                       1041.9 1711.8 2055.1 2358.1 2357.3 1565.0
   70                                        537.6  967.9 1136.1 1336.9 1538.0 1100.9
   80                                        133.6  261.6  346.3  423.5  504.2  414.6
```

```
stat.table( list(A=floor(A/10)*10,
                 P=floor(P/10)*10),
            list( D=sum(D),
                  Y=sum(Y/1000),
               rate=ratio(D,Y,10^5) ),
            margins=TRUE,
            data=testisDK )
```

```
 -------------------------------------------------------------------------
        --------------------------------P--------------------------------
 A            1940      1950      1960      1970      1980      1990     Total
 -------------------------------------------------------------------------
 0           10.00      7.00     16.00     18.00      9.00     10.00     70.00
           2604.66   4037.31   3884.97   3820.88   3070.87   2165.54  19584.22
              0.38      0.17      0.41      0.47      0.29      0.46      0.36

 10          13.00     27.00     37.00     72.00     97.00     75.00    321.00
           2135.73   3505.19   4004.13   3906.08   3847.40   2260.97  19659.48
              0.61      0.77      0.92      1.84      2.52      3.32      1.63

 20         124.00    221.00    280.00    535.00    724.00    557.00   2441.00
           2225.55   2923.22   3401.65   4028.57   3941.18   2824.58  19344.74
              5.57      7.56      8.23     13.28     18.37     19.72     12.62

 30         149.00    288.00    377.00    624.00    771.00    744.00   2953.00
           2195.23   3058.81   2856.20   3410.58   3968.81   2728.35  18217.97
              6.79      9.42     13.20     18.30     19.43     27.27     16.21

 40          95.00    198.00    230.00    334.00    432.00    360.00   1649.00
           1874.92   2980.15   2986.83   2823.11   3322.59   2757.72  16745.30
              5.07      6.64      7.70     11.83     13.00     13.05      9.85

 50          40.00     79.00    140.00    151.00    193.00    155.00    758.00
           1442.85   2426.54   2796.60   2813.32   2635.00   2069.18  14183.49
              2.77      3.26      5.01      5.37      7.32      7.49      5.34

 60          29.00     43.00     54.00     83.00     82.00     44.00    335.00
           1041.94   1711.79   2055.08   2358.05   2357.28   1564.98  11089.13
              2.78      2.51      2.63      3.52      3.48      2.81      3.02

 70          18.00     26.00     35.00     41.00     40.00     32.00    192.00
            537.62    967.88   1136.06   1336.95   1538.02   1100.86   6617.39
              3.35      2.69      3.08      3.07      2.60      2.91      2.90
```

```
80          7.00      9.00     13.00     19.00     18.00     21.00      87.00
          133.57    261.61    346.26    423.50    504.20    414.61    2083.75
            5.24      3.44      3.75      4.49      3.57      5.06       4.18


Total     485.00    898.00   1182.00   1877.00   2366.00   1998.00    8806.00
        14192.04  21872.50  23467.78  24921.03  25185.34  17886.80  127525.49
            3.42      4.11      5.04      7.53      9.39     11.17       6.91
-------------------------------------------------------------------------------
```

Note that for this type of cancer the peak age-specfic rates are in the 30es.

2. We then fit a Poisson-model for the mortality rates with a linear term for both age at entry:

```
ml <- glm( D ~ A, offset=log(Y), family=poisson, data=testisDK )
ci.exp( ml )
```

```
               exp(Est.)         2.5%        97.5%
(Intercept) 5.682883e-05 0.0000545697 0.0000591815
A           1.005499e+00 1.0045507062 1.0064479370
```

The parameter labelled `A` gives the annual increase in mortality by age (0.55%/year), but the intercept parameter is meaningless; it is the predicted mortality per person-year (because we used `Y` in the offset, and risk is in units of person-years), but for a 0 year old person.

3. But we can work out the predicted log-mortality rates for ages 25 to 45, say, by doing a hand-calculation based on the coefficients:

```
( cf <- coef( ml ) )
```

```
 (Intercept)            A
-9.775466746  0.005483811
```

We now have the intecept (the log-rate) and the slopes for age and calendar time, so to get the age-specific rates in ages 50 to 60 in the year 1970 we just plug in:

```
round( cbind( 25:45, exp( cf[1] + cf[2]*(25:45) )*10^5 ), 3 )
```

```
      [,1]  [,2]
 [1,]   25 6.518
 [2,]   26 6.554
 [3,]   27 6.590
 [4,]   28 6.626
 [5,]   29 6.662
 [6,]   30 6.699
 [7,]   31 6.736
 [8,]   32 6.773
 [9,]   33 6.810
[10,]   34 6.848
[11,]   35 6.885
[12,]   36 6.923
[13,]   37 6.961
[14,]   38 7.000
[15,]   39 7.038
[16,]   40 7.077
[17,]   41 7.116
[18,]   42 7.155
[19,]   43 7.194
[20,]   44 7.234
[21,]   45 7.273
```

4. But we do not have the standard errors of these mortality rates, and hence neither the confidence intervals. This is provided by `ci.exp`, if we provide the argument `ctr.mat=` as a matrix where each row corresponds to a prediction point and each column a parameter from the model.

   Thus for each age we need the corresponding multipliers for the coefficients:

```
( CM <- cbind( 1, 25:45 ) )
```

```
      [,1] [,2]
 [1,]    1   25
 [2,]    1   26
 [3,]    1   27
 [4,]    1   28
 [5,]    1   29
 [6,]    1   30
 [7,]    1   31
 [8,]    1   32
 [9,]    1   33
[10,]    1   34
[11,]    1   35
[12,]    1   36
[13,]    1   37
[14,]    1   38
[15,]    1   39
[16,]    1   40
[17,]    1   41
[18,]    1   42
[19,]    1   43
[20,]    1   44
[21,]    1   45
```

```
round( ci.exp( ml, ctr.mat=CM )*10^5, 3 )
```

```
      exp(Est.)  2.5% 97.5%
 [1,]     6.518 6.365 6.674
 [2,]     6.554 6.403 6.708
 [3,]     6.590 6.441 6.742
 [4,]     6.626 6.479 6.777
 [5,]     6.662 6.516 6.812
 [6,]     6.699 6.554 6.847
 [7,]     6.736 6.592 6.883
 [8,]     6.773 6.630 6.919
 [9,]     6.810 6.667 6.956
[10,]     6.848 6.705 6.993
[11,]     6.885 6.743 7.031
[12,]     6.923 6.780 7.069
[13,]     6.961 6.817 7.108
[14,]     7.000 6.855 7.147
[15,]     7.038 6.892 7.187
[16,]     7.077 6.929 7.228
[17,]     7.116 6.966 7.268
[18,]     7.155 7.003 7.310
[19,]     7.194 7.040 7.352
[20,]     7.234 7.077 7.394
[21,]     7.273 7.113 7.437
```

5. We can now use this machinery to plot the mortality rates over the range from 15 to 65 years:

```
CM <- cbind( 1, 15:65 )
matplot( 15:65, ci.exp( ml, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

6. Now suppose we want to see if the mortality rates really are eksponentially increasing by age, we could add a quadratic term to the model:

```
mq <- glm( D ~ A + I(A^2), offset=log(Y), family=poisson, data=testisDK )
ci.exp( mq, Exp=F )
```

```
                Estimate         2.5%        97.5%
(Intercept) -12.365625166 -12.482504296 -12.248746037
A             0.180595889   0.174140158   0.187051619
I(A^2)       -0.002325937  -0.002410829  -0.002241045
```

We can the plot the estimated rates using the same machinery, but now with 3 columns in the matrix:

```
aa <- 15:65
CM <- cbind( 1, aa, aa^2 )
matplot( aa, ci.exp( mq, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
matlines( aa, ci.exp( ml, ctr.mat=CM[,1:2] )*10^5,
          type="l", lty=1, lwd=c(3,1,1), col="blue" )
```

Which indeed is dramatically different.

7. We could do the same using a 3rd degree polynomial:

```
mc <- glm( D ~ A + I(A^2) + I(A^3), offset=log(Y), family=poisson, data=testisDK )
CM <- cbind( 1, aa, aa^2, aa^3 )
matplot( aa, ci.exp( mc, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

8. Instead of continuing with higher powers of age we could use fractions of powers, or we could use splines, piecevise polynomial curves, that fit nicely together at join points (knots). This is implemented in the `splines` package, in the function `ns`, which returns a matrix. There is a wrapper `Ns` in the `Epi`-package that automatically designate the smalles and largest knots a *boundary knots*, beyond which the resulting curve is linear:

```
library( splines )
ms <- glm( D ~ Ns(A,knots=seq(15,65,10)), offset=log(Y), family=poisson, data=testisDK )
As <- Ns( aa, knots=seq(15,65,10) )
matplot( aa, ci.exp( ms, ctr.mat=cbind(1,As) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

9. Now in addition to this we would like to see how the dependence on calendar was, so we add a linear term to the model, and make a prediction for 1980, say:

```
msp <- glm( D ~ Ns(A,knots=seq(15,65,10)) + P, offset=log(Y), family=poisson, data=testisDK )
CM <- cbind( 1, Ns( aa, knots=seq(15,65,10) ), 1980 )
matplot( aa, ci.exp( msp, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

10. But we would like to see how the RR reltive to 1980 is, so we select only the period parameter, using the `subset` argument:

```
ci.exp( msp, subset="P" )
```

```
  exp(Est.)     2.5%    97.5%
P  1.024235 1.022769 1.025704
```

SO we have an increase of 2.4% per year.

To get the RR relative to 1980 for the years 1943 to 1996 we must multiply the log-RR for period with the distance form 1980, such as:

```
yy <- 1943:1996
Cp <- cbind( yy - 1980 )
matplot( yy, ci.exp( msp, ctr.mat=Cp, subset="P" ),
         log="y", xlab="Date", ylab="RR of Testis cancer",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
abline( h=1 )
```

11. As above we might like to see how it looks if we add a quadratic to the period effect:

```
msp <- glm( D ~ Ns(A,knots=seq(15,65,10)) + P + I(P^2), offset=log(Y), family=poisson, data=te
Cq <- cbind( yy, yy^2 ) - cbind( rep(1980,length(yy)),
                                        1980^2 )
matplot( yy, ci.exp( msp, ctr.mat=Cq, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence rate ratio",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
abline( h=1, v=1980 )
```

12. But we would like to see if there were som non-linearity beyond the quadratic, with period as well, so we fit a spline for period (`P`) as well

```
mssp <- glm( D ~ Ns(A,knots=seq(15,65,10)) +
                 Ns(P,knots=seq(1950,1990,10)),
                 offset=log(Y), family=poisson, data=testisDK )
```

But as above we must compute the *difference* in the contribution from perid in year y and in the reference year, here 1970. So every row of the contrats matric must have the corresponding contribution from the reference year subtracted

```
Ps <- Ns(                 yy   , knots=seq(1950,1990,10) )
Pr <- Ns( rep(1970,length(yy)), knots=seq(1950,1990,10) )
matplot( yy, ci.exp( mssp, ctr.mat=Ps-Pr, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

13. But for this model we would also like to see the estimated age-specific rates in say 1980.

To this end we need a reference matrix with a number of rows equal to the numberof age-parameters:

```
Ar <- Ns( rep(1970,length(aa)), knots=seq(1950,1990,10) )
matplot( aa, ci.exp( mssp, ctr.mat=cbind(1,As,Ar) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

14. In order to do this in one go where we have overview of what we do, what is need is:

   - Where are the knots for age and period

- What are the prediction points for age and period

```
a.kn <- seq(15,65,20)
p.kn <- seq(1950,1990,10)
a.pt <- 10:65
p.pt <- 1945:1993
p.ref <- 1970
na <- length(a.pt)
np <- length(p.pt)
As <- Ns( a.pt, knots=a.kn )
Ps <- Ns( p.pt, knots=p.kn )
Pr <- Ns( rep(p.ref,np), knots=p.kn )
Ar <- Ns( rep(p.ref,na), knots=p.kn )
mAP <- glm( D ~ Ns(A,knots=a.kn) + Ns(P,knots=p.kn),
                offset=log(Y), family=poisson, data=testisDK )
par( mfrow=c(1,2) )
matplot( a.pt, ci.exp( mAP, ctr.mat=cbind(1,As,Ar) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20) )
matplot( p.pt, ci.exp( mAP, ctr.mat=Ps-Pr, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20)/5 )
abline( h=1, v=p.ref )
```

15. Finally with this in place we could do the same foir a model where we had replaced `P`, the data of follow-up by the the date of birth, `B=P-A`:

```
testisDK <- transform( testisDK, B = P-A )
with( testisDK, hist( rep(B,D), breaks=100, col="black" ) )
a.kn <- seq(15,65,5)
b.kn <- seq(1900,1970,5)
a.pt <- 10:65
b.pt <- 1890:1970
b.ref <- 1950
na <- length(a.pt)
nb <- length(b.pt)
As <- Ns( a.pt, knots=a.kn )
Bs <- Ns( b.pt, knots=b.kn )
Br <- Ns( rep(b.ref,nb), knots=b.kn )
Ar <- Ns( rep(b.ref,na), knots=b.kn )
mAB <- glm( D ~ Ns(A,knots=a.kn) + Ns(B,knots=b.kn),
                offset=log(Y), family=poisson, data=testisDK )
ci.exp( mAB, subset="B" )
```

```
                       exp(Est.)      2.5%    97.5%
Ns(B, knots = b.kn)1    1.233520 0.9474578 1.605953
Ns(B, knots = b.kn)2    1.370471 1.0987281 1.709424
Ns(B, knots = b.kn)3    1.180285 0.9292856 1.499080
Ns(B, knots = b.kn)4    1.860550 1.5035419 2.302327
Ns(B, knots = b.kn)5    1.937510 1.5646494 2.399225
Ns(B, knots = b.kn)6    2.306199 1.8844201 2.822382
Ns(B, knots = b.kn)7    2.769211 2.2752510 3.370411
Ns(B, knots = b.kn)8    1.839538 1.5195996 2.226837
Ns(B, knots = b.kn)9    3.437487 2.8485846 4.148137
Ns(B, knots = b.kn)10   3.460430 2.8670630 4.176599
Ns(B, knots = b.kn)11   5.056811 4.2019209 6.085629
Ns(B, knots = b.kn)12   4.554311 3.8422969 5.398268
Ns(B, knots = b.kn)13   4.793767 4.1597951 5.524359
Ns(B, knots = b.kn)14   5.078583 4.3433350 5.938296
```

```
par( mfrow=c(1,2) )
matplot( a.pt, ci.exp( mAB, ctr.mat=cbind(1,As,Ar) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20) )
matplot( b.pt, ci.exp( mAB, ctr.mat=Bs-Br, subset="B" ),
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20)/4 )
abline( h=1, v=b.ref )
```

```
 R 3.1.0
 ---------------------------------------------
 Program:  cont-eff.R
  Folder:  C:\Bendix\undervis\SPE\Repos\pracs
 Started:  fredag 23. maj 2014, 08:29:19
 ---------------------------------------------
> ### R code from vignette source cont-eff.rnw
>
> ###################################################
> ### code chunk number 1: cont-eff.rnw:3-7
> ###################################################
> options( width=90,
+          prompt=" ", continue=" ",
+          SweaveHooks=list( fig=function()
+          par(mar=c(3,3,1,1),mgp=c(3,1,0)/1.6) ) )


 ###################################################
 ### code chunk number 2: cont-eff.rnw:41-46
 ###################################################
 library( Epi )

Attaching package: Epi

The following object is masked from package:base:

    merge.data.frame

 sessionInfo()
R version 3.1.0 (2014-04-10)
Platform: i386-w64-mingw32/i386 (32-bit)

locale:
[1] LC_COLLATE=Danish_Denmark.1252  LC_CTYPE=Danish_Denmark.1252
[3] LC_MONETARY=Danish_Denmark.1252 LC_NUMERIC=C
[5] LC_TIME=Danish_Denmark.1252

attached base packages:
[1] utils     datasets  graphics  grDevices stats     methods   base

other attached packages:
[1] Epi_1.1.65     foreign_0.8-61
 data( testisDK )
 str( testisDK )
data.frame: 4860 obs. of  4 variables:
 $ A: num  0 1 2 3 4 5 6 7 8 9 ...
 $ P: num  1943 1943 1943 1943 1943 ...
 $ D: num  1 1 0 1 0 0 0 0 0 0 ...
 $ Y: num  39650 36943 34588 33267 32614 ...
 head( testisDK )
  A    P D        Y
1 0 1943 1 39649.50
2 1 1943 1 36942.83
3 2 1943 0 34588.33
4 3 1943 1 33267.00
5 4 1943 0 32614.00
6 5 1943 0 32020.33


 ###################################################
 ### code chunk number 3: cont-eff.rnw:51-63
 ###################################################
 round( ftable( xtabs( cbind(D,PY=Y/1000) ~ I(floor(A/10)*10) +
                                            I(floor(P/10)*10),
                      data=testisDK ),
               row.vars=c(3,1) ), 1 )
                      I(floor(P/10) * 10)   1940   1950   1960   1970   1980   1990
   I(floor(A/10) * 10)
D  0                                         10.0    7.0   16.0   18.0    9.0   10.0
   10                                        13.0   27.0   37.0   72.0   97.0   75.0
   20                                       124.0  221.0  280.0  535.0  724.0  557.0
```

```
   30                                          149.0  288.0  377.0  624.0  771.0  744.0
   40                                           95.0  198.0  230.0  334.0  432.0  360.0
   50                                           40.0   79.0  140.0  151.0  193.0  155.0
   60                                           29.0   43.0   54.0   83.0   82.0   44.0
   70                                           18.0   26.0   35.0   41.0   40.0   32.0
   80                                            7.0    9.0   13.0   19.0   18.0   21.0
PY 0                                          2604.7 4037.3 3885.0 3820.9 3070.9 2165.5
   10                                         2135.7 3505.2 4004.1 3906.1 3847.4 2261.0
   20                                         2225.5 2923.2 3401.6 4028.6 3941.2 2824.6
   30                                         2195.2 3058.8 2856.2 3410.6 3968.8 2728.4
   40                                         1874.9 2980.1 2986.8 2823.1 3322.6 2757.7
   50                                         1442.8 2426.5 2796.6 2813.3 2635.0 2069.2
   60                                         1041.9 1711.8 2055.1 2358.1 2357.3 1565.0
   70                                          537.6  967.9 1136.1 1336.9 1538.0 1100.9
   80                                          133.6  261.6  346.3  423.5  504.2  414.6

 stat.table( list(A=floor(A/10)*10,
                  P=floor(P/10)*10),
             list( D=sum(D),
                   Y=sum(Y/1000),
                rate=ratio(D,Y,10^5) ),
             margins=TRUE,
             data=testisDK )
 ------------------------------------------------------------------------
        --------------------------------P-------------------------------
 A           1940      1950      1960      1970      1980      1990     Total
 ------------------------------------------------------------------------
 0          10.00      7.00     16.00     18.00      9.00     10.00     70.00
          2604.66   4037.31   3884.97   3820.88   3070.87   2165.54  19584.22
             0.38      0.17      0.41      0.47      0.29      0.46      0.36

 10         13.00     27.00     37.00     72.00     97.00     75.00    321.00
          2135.73   3505.19   4004.13   3906.08   3847.40   2260.97  19659.48
             0.61      0.77      0.92      1.84      2.52      3.32      1.63

 20        124.00    221.00    280.00    535.00    724.00    557.00   2441.00
          2225.55   2923.22   3401.65   4028.57   3941.18   2824.58  19344.74
             5.57      7.56      8.23     13.28     18.37     19.72     12.62

 30        149.00    288.00    377.00    624.00    771.00    744.00   2953.00
          2195.23   3058.81   2856.20   3410.58   3968.81   2728.35  18217.97
             6.79      9.42     13.20     18.30     19.43     27.27     16.21

 40         95.00    198.00    230.00    334.00    432.00    360.00   1649.00
          1874.92   2980.15   2986.83   2823.11   3322.59   2757.72  16745.30
             5.07      6.64      7.70     11.83     13.00     13.05      9.85

 50         40.00     79.00    140.00    151.00    193.00    155.00    758.00
          1442.85   2426.54   2796.60   2813.32   2635.00   2069.18  14183.49
             2.77      3.26      5.01      5.37      7.32      7.49      5.34

 60         29.00     43.00     54.00     83.00     82.00     44.00    335.00
          1041.94   1711.79   2055.08   2358.05   2357.28   1564.98  11089.13
             2.78      2.51      2.63      3.52      3.48      2.81      3.02

 70         18.00     26.00     35.00     41.00     40.00     32.00    192.00
           537.62    967.88   1136.06   1336.95   1538.02   1100.86   6617.39
             3.35      2.69      3.08      3.07      2.60      2.91      2.90

 80          7.00      9.00     13.00     19.00     18.00     21.00     87.00
           133.57    261.61    346.26    423.50    504.20    414.61   2083.75
             5.24      3.44      3.75      4.49      3.57      5.06      4.18


 Total     485.00    898.00   1182.00   1877.00   2366.00   1998.00   8806.00
         14192.04  21872.50  23467.78  24921.03  25185.34  17886.80 127525.49
             3.42      4.11      5.04      7.53      9.39     11.17      6.91
 ------------------------------------------------------------------------


 ###################################################
 ### code chunk number 4: cont-eff.rnw:78-80
 ###################################################
 ml <- glm( D ~ A, offset=log(Y), family=poisson, data=testisDK )
 ci.exp( ml )
               exp(Est.)         2.5%        97.5%
(Intercept) 5.682883e-05 0.0000545697 0.0000591815
A           1.005499e+00 1.0045507062 1.0064479370


 ###################################################
 ### code chunk number 5: cont-eff.rnw:91-92
 ###################################################
 ( cf <- coef( ml ) )
 (Intercept)            A
-9.775466746  0.005483811
```

```
###################################################
### code chunk number 6: cont-eff.rnw:97-98
###################################################
round( cbind( 25:45, exp( cf[1] + cf[2]*(25:45) )*10^5 ), 3 )
      [,1]  [,2]
 [1,]   25 6.518
 [2,]   26 6.554
 [3,]   27 6.590
 [4,]   28 6.626
 [5,]   29 6.662
 [6,]   30 6.699
 [7,]   31 6.736
 [8,]   32 6.773
 [9,]   33 6.810
[10,]   34 6.848
[11,]   35 6.885
[12,]   36 6.923
[13,]   37 6.961
[14,]   38 7.000
[15,]   39 7.038
[16,]   40 7.077
[17,]   41 7.116
[18,]   42 7.155
[19,]   43 7.194
[20,]   44 7.234
[21,]   45 7.273


###################################################
### code chunk number 7: cont-eff.rnw:108-110
###################################################
( CM <- cbind( 1, 25:45 ) )
      [,1] [,2]
 [1,]    1   25
 [2,]    1   26
 [3,]    1   27
 [4,]    1   28
 [5,]    1   29
 [6,]    1   30
 [7,]    1   31
 [8,]    1   32
 [9,]    1   33
[10,]    1   34
[11,]    1   35
[12,]    1   36
[13,]    1   37
[14,]    1   38
[15,]    1   39
[16,]    1   40
[17,]    1   41
[18,]    1   42
[19,]    1   43
[20,]    1   44
[21,]    1   45
round( ci.exp( ml, ctr.mat=CM )*10^5, 3 )
      exp(Est.)  2.5% 97.5%
 [1,]     6.518 6.365 6.674
 [2,]     6.554 6.403 6.708
 [3,]     6.590 6.441 6.742
 [4,]     6.626 6.479 6.777
 [5,]     6.662 6.516 6.812
 [6,]     6.699 6.554 6.847
 [7,]     6.736 6.592 6.883
 [8,]     6.773 6.630 6.919
 [9,]     6.810 6.667 6.956
[10,]     6.848 6.705 6.993
[11,]     6.885 6.743 7.031
[12,]     6.923 6.780 7.069
[13,]     6.961 6.817 7.108
[14,]     7.000 6.855 7.147
[15,]     7.038 6.892 7.187
[16,]     7.077 6.929 7.228
[17,]     7.116 6.966 7.268
[18,]     7.155 7.003 7.310
[19,]     7.194 7.040 7.352
[20,]     7.234 7.077 7.394
[21,]     7.273 7.113 7.437


###################################################
### code chunk number 8: mort-lin
###################################################
CM <- cbind( 1, 15:65 )
matplot( 15:65, ci.exp( ml, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
```

```
###################################################
### code chunk number 9: cont-eff.rnw:125-127
###################################################
mq <- glm( D ~ A + I(A^2), offset=log(Y), family=poisson, data=testisDK )
ci.exp( mq, Exp=F )
                Estimate          2.5%        97.5%
(Intercept) -12.365625166 -12.482504296 -12.248746037
A             0.180595889   0.174140158   0.187051619
I(A^2)       -0.002325937  -0.002410829  -0.002241045


###################################################
### code chunk number 10: mort-qdr
###################################################
aa <- 15:65
CM <- cbind( 1, aa, aa^2 )
matplot( aa, ci.exp( mq, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
matlines( aa, ci.exp( ml, ctr.mat=CM[,1:2] )*10^5,
          type="l", lty=1, lwd=c(3,1,1), col="blue" )


###################################################
### code chunk number 11: mort-qdr
###################################################
mc <- glm( D ~ A + I(A^2) + I(A^3), offset=log(Y), family=poisson, data=testisDK )
CM <- cbind( 1, aa, aa^2, aa^3 )
matplot( aa, ci.exp( mc, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )


###################################################
### code chunk number 12: mort-spl
###################################################
library( splines )
ms <- glm( D ~ Ns(A,knots=seq(15,65,10)), offset=log(Y), family=poisson, data=testisDK )
As <- Ns( aa, knots=seq(15,65,10) )
matplot( aa, ci.exp( ms, ctr.mat=cbind(1,As) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )


###################################################
### code chunk number 13: mort-spl-P
###################################################
msp <- glm( D ~ Ns(A,knots=seq(15,65,10)) + P, offset=log(Y), family=poisson, data=testisDK )
CM <- cbind( 1, Ns( aa, knots=seq(15,65,10) ), 1980 )
matplot( aa, ci.exp( msp, ctr.mat=CM )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence rate per 100,000 PY",
         type="l", lty=1, lwd=c(3,1,1), col="black" )


###################################################
### code chunk number 14: cont-eff.rnw:182-183
###################################################
ci.exp( msp, subset="P" )
  exp(Est.)     2.5%    97.5%
P  1.024235 1.022769 1.025704


###################################################
### code chunk number 15: cont-eff.rnw:189-195
###################################################
yy <- 1943:1996
Cp <- cbind( yy - 1980 )
matplot( yy, ci.exp( msp, ctr.mat=Cp, subset="P" ),
         log="y", xlab="Date", ylab="RR of Testis cancer",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
abline( h=1 )


###################################################
### code chunk number 16: mort-spl-P
###################################################
msp <- glm( D ~ Ns(A,knots=seq(15,65,10)) + P + I(P^2), offset=log(Y), family=poisson, data=testisDK )
Cq <- cbind( yy, yy^2 ) - cbind( rep(1980,length(yy)),
                                 1980^2 )
matplot( yy, ci.exp( msp, ctr.mat=Cq, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence rate ratio",
         type="l", lty=1, lwd=c(3,1,1), col="black" )
abline( h=1, v=1980 )
```

```
###################################################
### code chunk number 17: cont-eff.rnw:213-216
###################################################
mssp <- glm( D ~ Ns(A,knots=seq(15,65,10)) +
                 Ns(P,knots=seq(1950,1990,10)),
                 offset=log(Y), family=poisson, data=testisDK )


###################################################
### code chunk number 18: mort-spl-splP
###################################################
Ps <- Ns(                  yy  , knots=seq(1950,1990,10) )
Pr <- Ns( rep(1970,length(yy)), knots=seq(1950,1990,10) )
matplot( yy, ci.exp( mssp, ctr.mat=Ps-Pr, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black" )


###################################################
### code chunk number 19: mort-spl-splP
###################################################
Ar <- Ns( rep(1970,length(aa)), knots=seq(1950,1990,10) )
matplot( aa, ci.exp( mssp, ctr.mat=cbind(1,As,Ar) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black" )


###################################################
### code chunk number 20: cont-eff.rnw:248-271
###################################################
a.kn <- seq(15,65,20)
p.kn <- seq(1950,1990,10)
a.pt <- 10:65
p.pt <- 1945:1993
p.ref <- 1970
na <- length(a.pt)
np <- length(p.pt)
As <- Ns( a.pt, knots=a.kn )
Ps <- Ns( p.pt, knots=p.kn )
Pr <- Ns( rep(p.ref,np), knots=p.kn )
Ar <- Ns( rep(p.ref,na), knots=p.kn )
mAP <- glm( D ~ Ns(A,knots=a.kn) + Ns(P,knots=p.kn),
                offset=log(Y), family=poisson, data=testisDK )
par( mfrow=c(1,2) )
matplot( a.pt, ci.exp( mAP, ctr.mat=cbind(1,As,Ar) )*10^5,
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20) )
matplot( p.pt, ci.exp( mAP, ctr.mat=Ps-Pr, subset="P" ),
         log="y", xlab="Age", ylab="Testis cancer incidence RR",
         type="l", lty=1, lwd=c(3,1,1), col="black",
         ylim=c(1,20)/5 )
abline( h=1, v=p.ref )


###################################################
### code chunk number 21: cont-eff.rnw:277-303
###################################################
testisDK <- transform( testisDK, B = P-A )
with( testisDK, hist( rep(B,D), breaks=100, col="black" ) )
a.kn <- seq(15,65,5)
b.kn <- seq(1900,1970,5)
a.pt <- 10:65
b.pt <- 1890:1970
b.ref <- 1950
na <- length(a.pt)
nb <- length(b.pt)
As <- Ns( a.pt, knots=a.kn )
Bs <- Ns( b.pt, knots=b.kn )
Br <- Ns( rep(b.ref,nb), knots=b.kn )
Ar <- Ns( rep(b.ref,na), knots=b.kn )
mAB <- glm( D ~ Ns(A,knots=a.kn) + Ns(B,knots=b.kn),
                offset=log(Y), family=poisson, data=testisDK )
ci.exp( mAB, subset="B" )
                    exp(Est.)      2.5%    97.5%
Ns(B, knots = b.kn)1   1.233520 0.9474578 1.605953
Ns(B, knots = b.kn)2   1.370471 1.0987281 1.709424
Ns(B, knots = b.kn)3   1.180285 0.9292856 1.499080
Ns(B, knots = b.kn)4   1.860550 1.5035419 2.302327
Ns(B, knots = b.kn)5   1.937510 1.5646494 2.399225
Ns(B, knots = b.kn)6   2.306199 1.8844201 2.822382
Ns(B, knots = b.kn)7   2.769211 2.2752510 3.370411
Ns(B, knots = b.kn)8   1.839538 1.5195996 2.226837
Ns(B, knots = b.kn)9   3.437487 2.8485846 4.148137
Ns(B, knots = b.kn)10  3.460430 2.8670630 4.176599
Ns(B, knots = b.kn)11  5.056811 4.2019209 6.085629
Ns(B, knots = b.kn)12  4.554311 3.8422969 5.398268
```

```
Ns(B, knots = b.kn)13  4.793767 4.1597951 5.524359
Ns(B, knots = b.kn)14  5.078583 4.3433350 5.938296
 par( mfrow=c(1,2) )
 matplot( a.pt, ci.exp( mAB, ctr.mat=cbind(1,As,Ar) )*10^5,
          log="y", xlab="Age", ylab="Testis cancer incidence RR",
          type="l", lty=1, lwd=c(3,1,1), col="black",
          ylim=c(1,20) )
 matplot( b.pt, ci.exp( mAB, ctr.mat=Bs-Br, subset="B" ),
          log="y", xlab="Age", ylab="Testis cancer incidence RR",
          type="l", lty=1, lwd=c(3,1,1), col="black",
          ylim=c(1,20)/4 )
 abline( h=1, v=b.ref )



 ---------------------------------------------
 Program: cont-eff.R
  Folder: C:\Bendix\undervis\SPE\Repos\pracs
   Ended: fredag 23. maj 2014, 08:29:22
 Elapsed: 00:00:02
 ---------------------------------------------
> proc.time()
   user  system elapsed
   2.29    0.28    2.98
```