Practical Exercise 8, 2014: Same data as Practical Ex. 7

```
norw.death=read.table("http://folk.uio.no/borgan/abg-2008/data/
                                        causes_death.txt", header=T)
norw.death=norw.death[norw.death$smkgr!=6,]
library(survival)
# a)
> fit.a=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
                              +sbp+bmi,data=norw.death,subset=sbp>0)
> fit.a
                     coef exp(coef) se(coef)     z       p
factor(sex)2    -0.50340   0.60447  0.09570 -5.26 1.4e-07
factor(smkgr)2   0.34677   1.41449  0.13602  2.55   0.011
factor(smkgr)3   0.92299   2.51681  0.14835  6.22 4.9e-10
factor(smkgr)4   0.93972   2.55927  0.13006  7.23 5.0e-13
factor(smkgr)5   1.16163   3.19515  0.15687  7.41 1.3e-13
sbp              0.01738   1.01753  0.00218  7.99 1.3e-15
bmi              0.01592   1.01604  0.01244  1.28   0.201

Likelihood ratio test=218  on 7 df, p=0
n= 3860, number of events= 548
   (65 observations deleted due to missingness)
```

Practical Exercise 8, 2014, b: Checking linearity

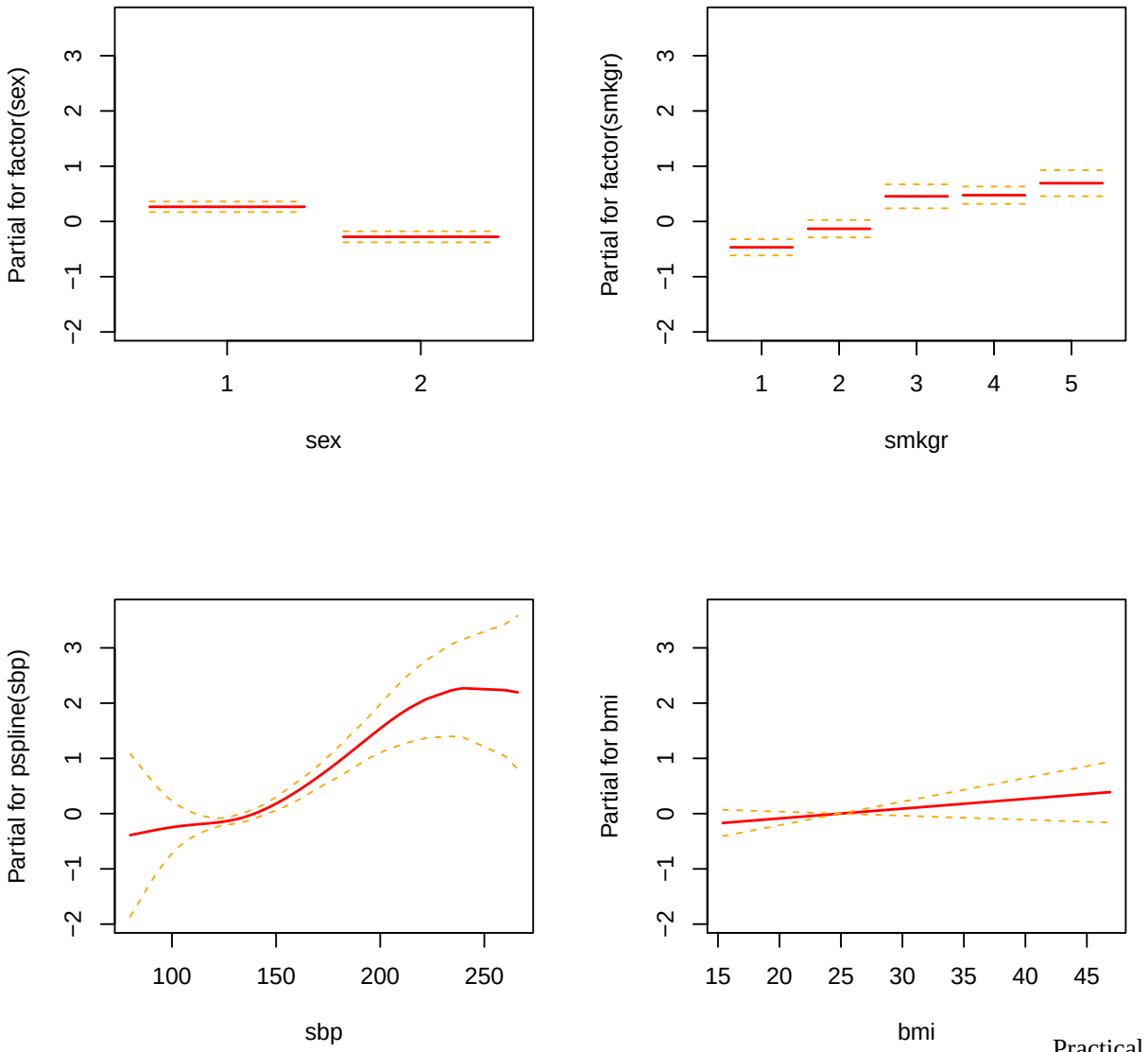
```
> fit.b=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
                    +pspline(sbp)+bmi, data=norw.death,subset=sbp>0)
> print(fit.b)

                          coef se(coef)      se2    Chisq DF       p
factor(sex)2          -0.54283  0.09750  0.09745 30.99532  1 2.6e-08
factor(smkgr)2         0.33570  0.13609  0.13609  6.08437  1   0.014
factor(smkgr)3         0.92273  0.14851  0.14847 38.60348  1 5.2e-10
factor(smkgr)4         0.94282  0.12998  0.12996 52.61252  1 4.1e-13
factor(smkgr)5         1.16029  0.15679  0.15677 54.76247  1 1.4e-13
pspline(sbp), linear   0.01757  0.00204  0.00204 74.29024  1 < 2e-16
pspline(sbp), nonlin                             8.06404  3   0.045
bmi                    0.01770  0.01253  0.01253  1.99489  1   0.158

Iterations: 5 outer, 18 Newton-Raphson
     Theta= 0.524
Degrees of freedom for terms= 1 4 4 1
Likelihood ratio test=227  on 10 df, p=0
  n=3860 (65 observations deleted due to missingness)
```

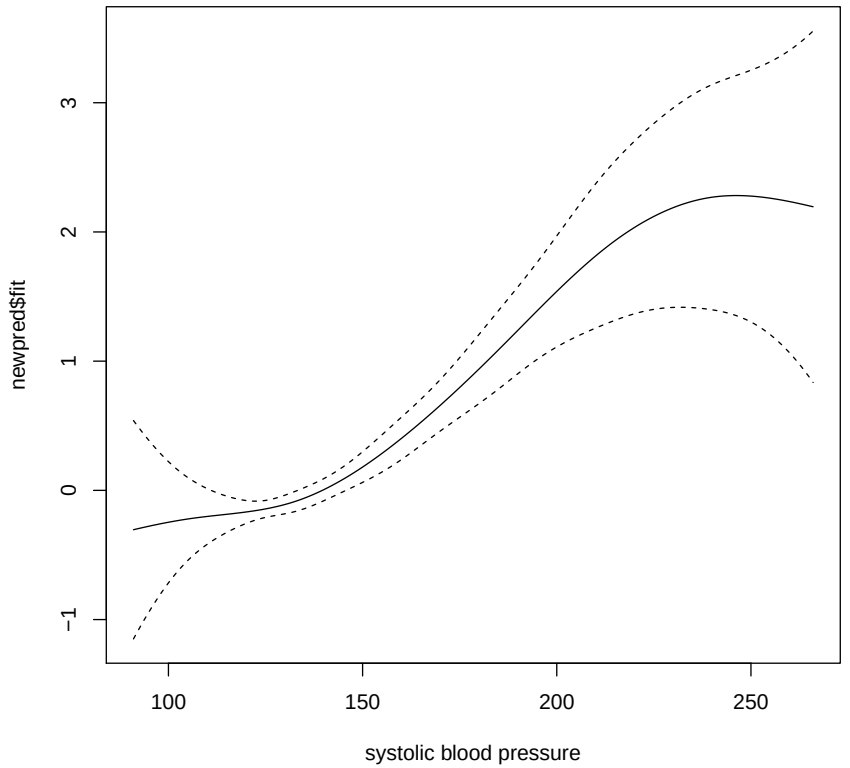
Practical Exercise 8, 2014, b, contd.

```
> par(mfrow=c(2,2))
> termplot(fit.b,se=T)
```



Practical Exercise 8, 2014, b, contd.

```
> newnor=data.frame(cbind(rep(1,176),rep(2,176),91:266,rep(25,176)))
> names(newnor)=names(norw.death[c(8,13,10,11)])
> newpred<-predict(fit.b,newnor,type="terms",term=3,se=T)
> mi<-min(newpred$fit-1.96*newpred$se.fit)
> ma<-max(newpred$fit+1.96*newpred$se.fit)
> plot(newnor$sbp,newpred$fit,type="l",ylim=c(mi,ma),xlab="systolic blo
> lines(newnor$sbp,newpred$fit+1.96*newpred$se.fit,lty=2)
> lines(newnor$sbp,newpred$fit-1.96*newpred$se.fit,lty=2)
```



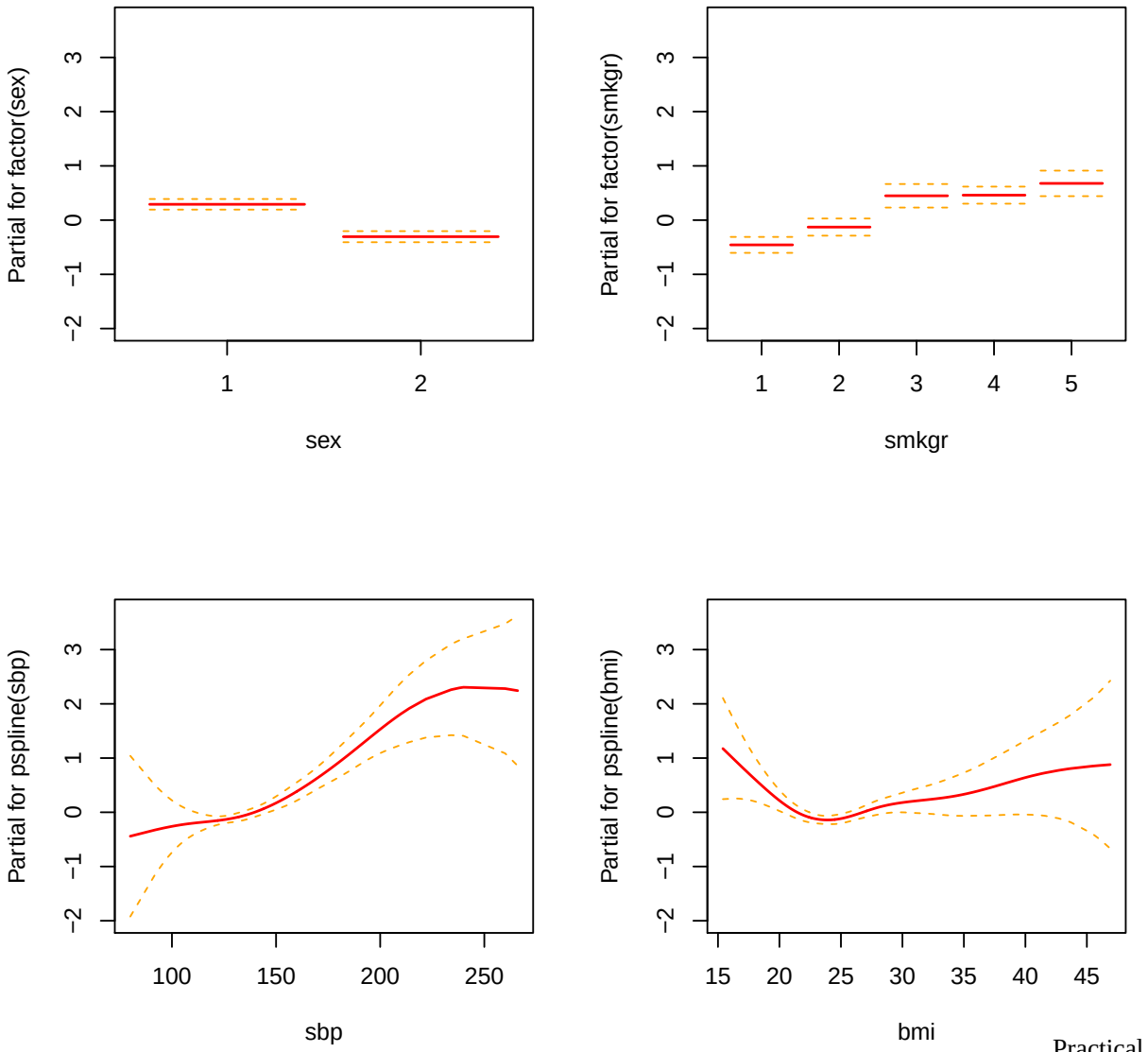
Practical Exercise 8, 2014, c, contd.

```
> #c)
> fit.c=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
             +pspline(sbp)+pspline(bmi), data=norw.death,subset=sbp>0)
> print(fit.c)
                          coef se(coef)      se2    Chisq   DF       p
factor(sex)2           -0.59709  0.09990  0.09976 35.72425 1.00 2.3e-09
factor(smkgr)2          0.32843  0.13624  0.13620  5.81152 1.00 0.01592
factor(smkgr)3          0.90533  0.14893  0.14886 36.95452 1.00 1.2e-09
factor(smkgr)4          0.91725  0.13028  0.13023 49.57166 1.00 1.9e-12
factor(smkgr)5          1.13405  0.15697  0.15691 52.19300 1.00 5.0e-13
pspline(sbp), linear    0.01731  0.00205  0.00205 71.41570 1.00 < 2e-16
pspline(sbp), nonlin                               8.11188 3.00 0.04372
pspline(bmi), linear    0.01541  0.01153  0.01152  1.78673 1.00 0.18132
pspline(bmi), nonlin                              16.70400 3.06 0.00087

Iterations: 5 outer, 19 Newton-Raphson
     Theta= 0.521
     Theta= 0.86
Degrees of freedom for terms= 1.0 4.0 4.0 4.1
Likelihood ratio test=245  on 13.1 df, p=0
  n=3860 (65 observations deleted due to missingness)
```

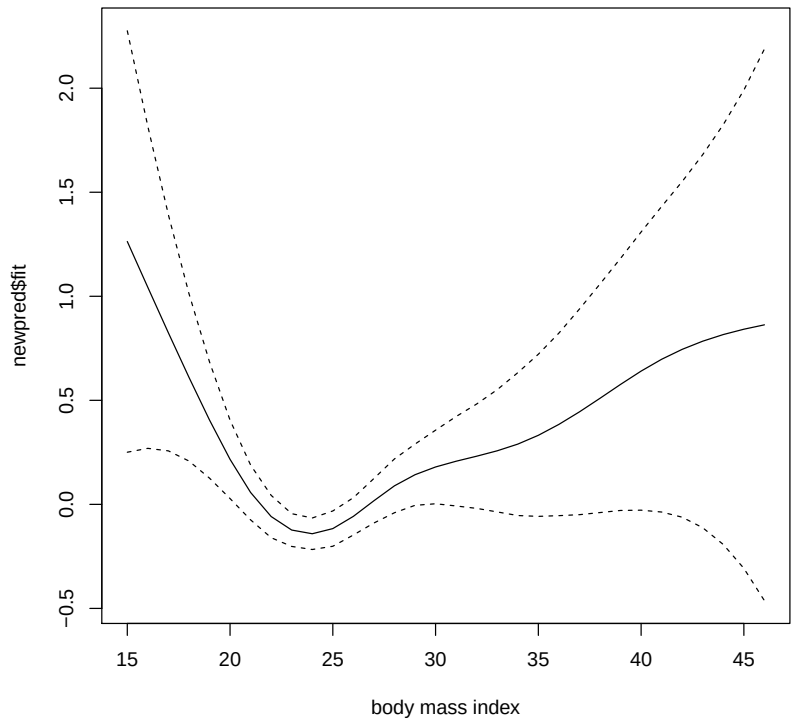
Practical Exercise 8, 2014, c, contd.

```
> par(mfrow=c(2,2))
> termplot(fit.c,se=T)
```



Practical Exercise 8, 2014, c, contd.

```
> newnor=data.frame(cbind(rep(1,32),rep(2,32),rep(120,32),15:46))
> names(newnor)=names(norw.death[c(8,13,10,11)])
> newpred<-predict(fit.c,newnor,type="terms",term=4,se=T)
> mi<-min(newpred$fit-1.96*newpred$se.fit)
> ma<-max(newpred$fit+1.96*newpred$se.fit)
> par(mfrow=c(1,1))
> plot(newnor$bmi,newpred$fit,type="l",ylim=c(mi,ma),xlab="body mass in
> lines(newnor$bmi,newpred$fit+1.96*newpred$se.fit,lty=2)
> lines(newnor$bmi,newpred$fit-1.96*newpred$se.fit,lty=2)
```



Practical Exercise 8, 2014, c, bmi categorical

```
> # Using bmi as a categorical variable
> norw.death$bmicat=cut(norw.death$bmi,c(14,20,25,30,50))
> fit.c2=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
           +pspline(sbp)+factor(bmicat), data=norw.death,subset=sbp>0)
> fit.c2
                          coef se(coef)      se2    Chisq DF       p
factor(sex)2           -0.55704  0.09853  0.09848 31.96150  1 1.6e-08
factor(smkgr)2          0.34715  0.13609  0.13608  6.50752  1 0.01074
factor(smkgr)3          0.92783  0.14860  0.14855 38.98598  1 4.3e-10
factor(smkgr)4          0.93269  0.12999  0.12996 51.48629  1 7.2e-13
factor(smkgr)5          1.16023  0.15663  0.15661 54.87110  1 1.3e-13
pspline(sbp), linear    0.01740  0.00203  0.00203 73.21289  1 < 2e-16
pspline(sbp), nonlin                              8.25900  3 0.04091
factor(bmicat)(20,25] -0.71526  0.18704  0.18698 14.62465  1 0.00013
factor(bmicat)(25,30] -0.47241  0.18969  0.18966  6.20247  1 0.01276
factor(bmicat)(30,50] -0.35077  0.22996  0.22993  2.32677  1 0.12717

Iterations: 5 outer, 18 Newton-Raphson
     Theta= 0.525
Degrees of freedom for terms= 1 4 4 3
Likelihood ratio test=243  on 12 df, p=0
```

Practical Exercise 8, 2014, c, bmi categorical 2

```
> norw.death$bmicat2=as.numeric(norw.death$bmicat)
> norw.death$bmicat2[norw.death$bmicat2==2]=0
> fit.c3=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
          +pspline(sbp)+factor(bmicat2), data=norw.death,subset=sbp>0)
> fit.c3
                          coef se(coef)      se2    Chisq DF       p
factor(sex)2          -0.55704  0.09853  0.09848 31.96150  1 1.6e-08
factor(smkgr)2         0.34715  0.13609  0.13608  6.50752  1 0.01074
factor(smkgr)3         0.92783  0.14860  0.14855 38.98598  1 4.3e-10
factor(smkgr)4         0.93269  0.12999  0.12996 51.48629  1 7.2e-13
factor(smkgr)5         1.16023  0.15663  0.15661 54.87110  1 1.3e-13
pspline(sbp), linear   0.01740  0.00203  0.00203 73.21289  1 < 2e-16
pspline(sbp), nonlin                             8.25900  3 0.04091
factor(bmicat2)1       0.71526  0.18704  0.18698 14.62465  1 0.00013
factor(bmicat2)3       0.24285  0.09457  0.09454  6.59404  1 0.01023
factor(bmicat2)4       0.36449  0.16207  0.16200  5.05813  1 0.02451

Iterations: 5 outer, 18 Newton-Raphson
     Theta= 0.525
Degrees of freedom for terms= 1 4 4 3
Likelihood ratio test=243  on 12 df, p=0
```

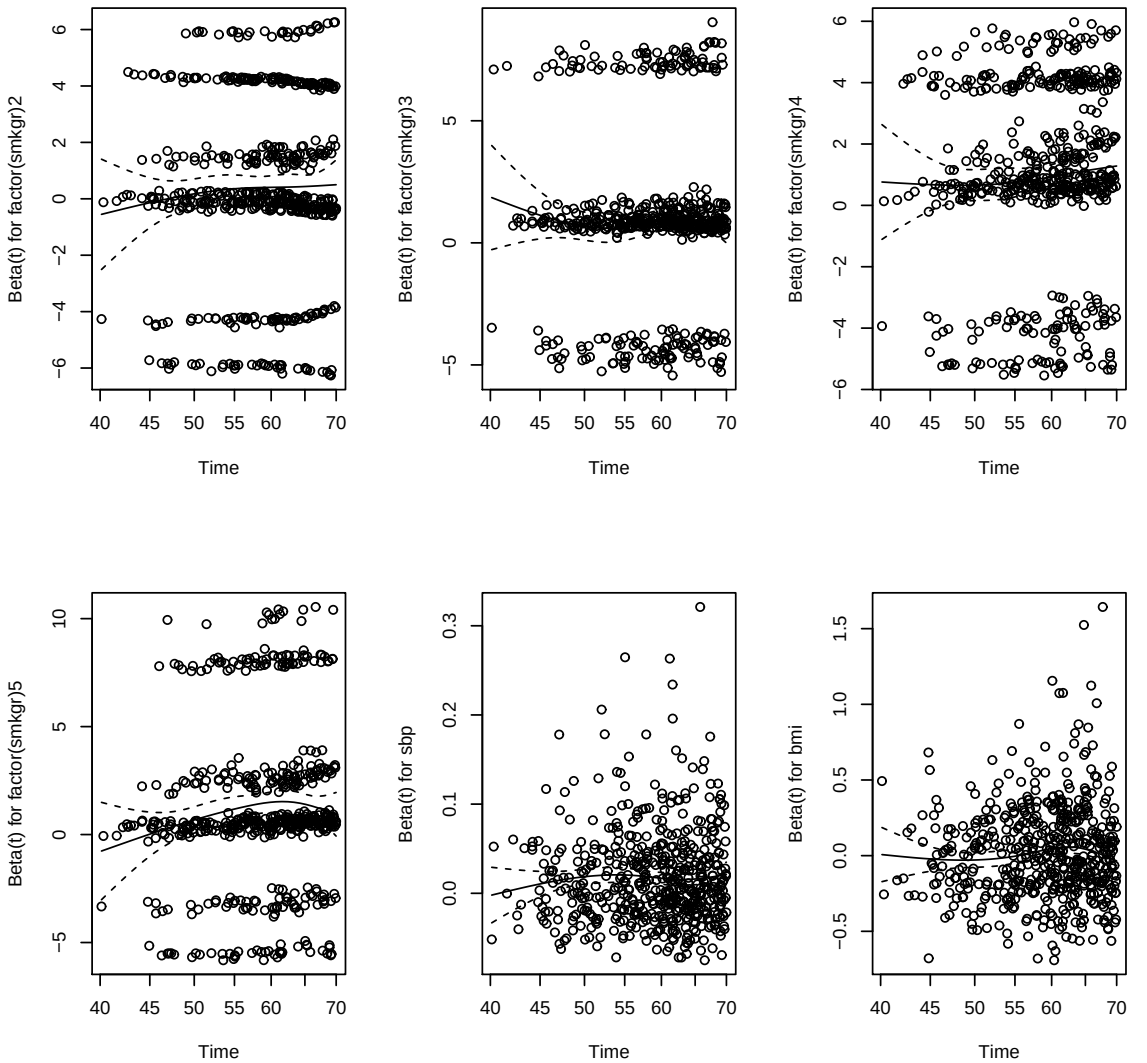
Practical Exercise 8, 2014, d, prop.haz. check

```
> #d)
> fit.d=coxph(Surv(agestart,agestop,dead)~strata(sex)+factor(smkgr)
                                          +sbp+bmi, data=norw.death)

> cox.zph(fit.d,transform='log')
                    rho chisq      p
factor(smkgr)2 0.0475 1.237 0.2660
factor(smkgr)3 0.0209 0.235 0.6278
factor(smkgr)4 0.0640 2.178 0.1400
factor(smkgr)5 0.0840 3.886 0.0487
sbp            0.0163 0.160 0.6890
bmi            0.0502 1.616 0.2037
GLOBAL             NA 6.307 0.3897
```

Practical Exercise 8, 2014, d, prop.haz. check - plot

```
> par(mfrow=c(2,3))
> plot(cox.zph(fit.d,transform='log'))
```

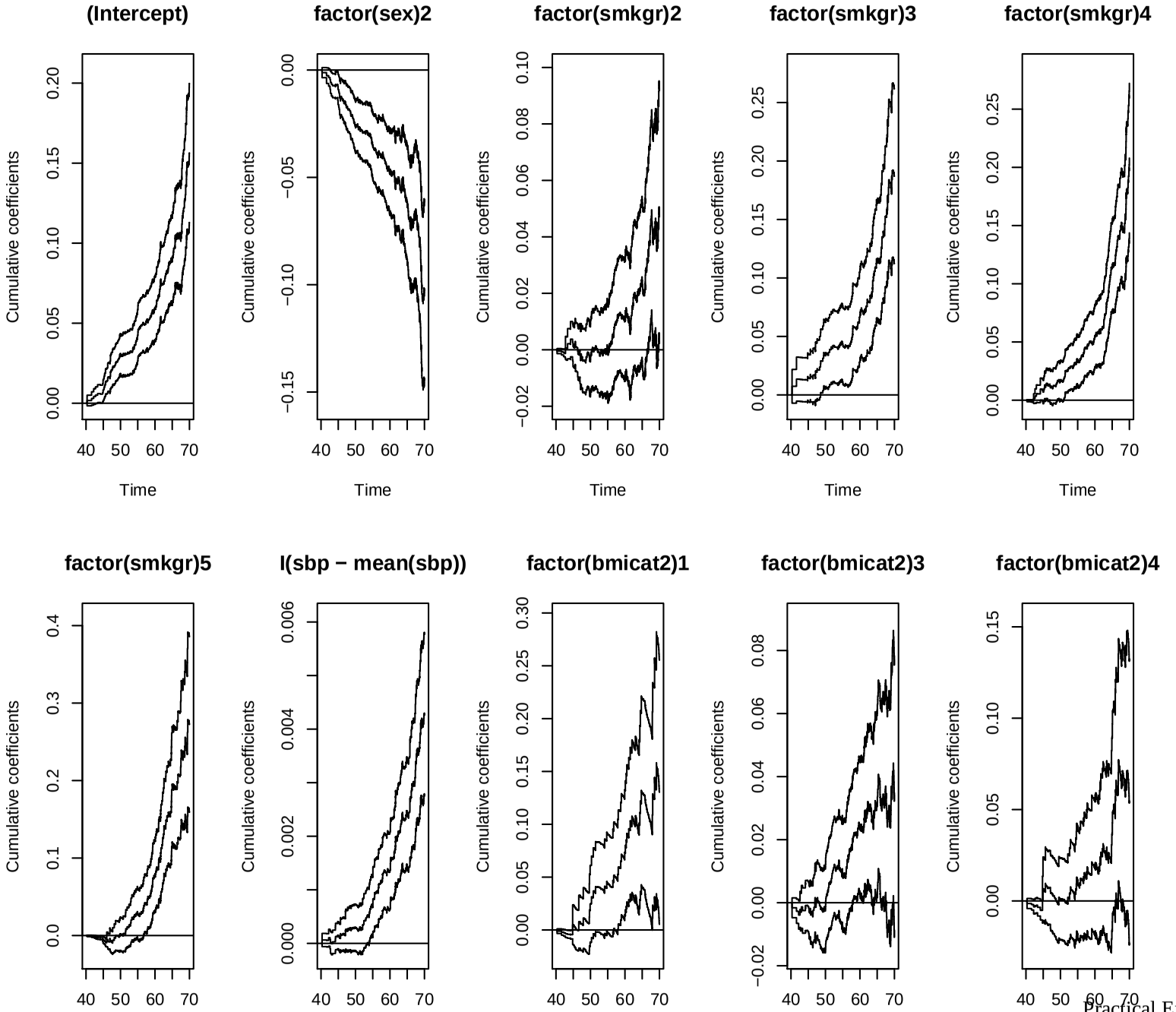


Aalen additive on the Norwegian mortality data

```
> plot(fit.aareg,xlim=c(40,80))
> norw.death=read.table("http://folk.uio.no/borgan/abg-2008/data/
                                                causes_death.txt", header=T)
> norw.death=norw.death[norw.death$smkgr!=6,]
> norw.death=norw.death[norw.death$sbp>0&!is.na(norw.death$bmi),]
> norw.death$bmicat=cut(norw.death$bmi,c(14,20,25,30,50))
> norw.death$bmicat2=as.numeric(norw.death$bmicat)
> norw.death$bmicat2[norw.death$bmicat2==2]=0
> library(survival)
> library(timereg)

> fit.aalen=aalen(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
                      +I(sbp-mean(sbp))+factor(bmicat2), data=norw.death)
> par(mfrow=c(2,5))
> plot(fit.aalen,start.time=40,stop.time=80)
```

Aalen additive, Norwegian mortality data, plots



Reminder: Cox-regression on the Norwegian mortality

```
> coxfit=coxph(Surv(agestart,agestop,dead)~factor(sex)+factor(smkgr)
                  +I(sbp-mean(sbp))+factor(bmicat2),data=norw.death)
> coxfit

                        coef exp(coef) se(coef)     z       p
factor(sex)2       -0.51695   0.59634  0.09678 -5.34 9.2e-08
factor(smkgr)2      0.35862   1.43135  0.13604  2.64 0.00839
factor(smkgr)3      0.92807   2.52962  0.14837  6.26 4.0e-10
factor(smkgr)4      0.92908   2.53218  0.13004  7.14 9.0e-13
factor(smkgr)5      1.16145   3.19458  0.15667  7.41 1.2e-13
I(sbp - mean(sbp))  0.01718   1.01733  0.00218  7.87 3.4e-15
factor(bmicat2)1    0.72140   2.05732  0.18684  3.86 0.00011
factor(bmicat2)3    0.23456   1.26435  0.09429  2.49 0.01286
factor(bmicat2)4    0.33206   1.39384  0.16066  2.07 0.03874

Likelihood ratio test=234  on 9 df, p=0
n= 3860, number of events
```

Accelerated failure time on the Norwegian mortality

```
> # Survreg, will not take left-truncated data. Use agestop-agestart
> survregfit=survreg(Surv(agestop-agestart,dead)~factor(sex)+factor(smk
     +I(sbp-mean(sbp))+factor(bmicat2),dist="exponential",data=norw.dea
> summary(survregfit)
                     Value Std. Error     z        p
(Intercept)         5.6138    0.12490 44.95 0.00e+00
factor(sex)2        0.4901    0.09647  5.08 3.77e-07
factor(smkgr)2     -0.3843    0.13563 -2.83 4.60e-03
factor(smkgr)3     -0.9393    0.14832 -6.33 2.41e-10
factor(smkgr)4     -0.8931    0.12966 -6.89 5.64e-12
factor(smkgr)5     -1.0944    0.15614 -7.01 2.40e-12
I(sbp - mean(sbp)) -0.0182    0.00215 -8.46 2.65e-17
factor(bmicat2)1   -0.7075    0.18666 -3.79 1.50e-04
factor(bmicat2)3   -0.2736    0.09423 -2.90 3.69e-03
factor(bmicat2)4   -0.3556    0.16067 -2.21 2.69e-02

Scale fixed at 1
Exponential distribution
Loglik(model)= -3193.6   Loglik(intercept only)= -3310.2
        Chisq= 233.15 on 9 degrees of freedom, p= 0
n= 3860
```

Accelerated failure time on the Norwegian mortality

```
> # Survreg, May be a good idea to include agestart as covariate
> survregfit=survreg(Surv(agestop-agestart,dead)~factor(sex)+factor(smk
+I(sbp-mean(sbp))+factor(bmicat2)+agestart,dist="exponential",data=norw
> summary(survregfit)
(Intercept)          8.9908    0.58518 15.36 2.85e-53
factor(sex)2         0.5048    0.09687  5.21 1.87e-07
factor(smkgr)2      -0.3646    0.13608 -2.68 7.38e-03
factor(smkgr)3      -0.9029    0.14851 -6.08 1.20e-09
factor(smkgr)4      -0.9136    0.13014 -7.02 2.21e-12
factor(smkgr)5      -1.1232    0.15680 -7.16 7.89e-13
I(sbp - mean(sbp)) -0.0162    0.00218 -7.45 9.40e-14
factor(bmicat2)1    -0.6834    0.18681 -3.66 2.54e-04
factor(bmicat2)3    -0.2369    0.09447 -2.51 1.21e-02
factor(bmicat2)4    -0.3334    0.16080 -2.07 3.81e-02
agestart            -0.0779    0.01305 -5.97 2.40e-09

Scale fixed at 1
Exponential distribution
Loglik(model)= -3176.1   Loglik(intercept only)= -3310.2
        Chisq= 268.28 on 10 degrees of freedom, p= 0
n= 3860
```

Accelerated failure time on the Norwegian mortality

```
> # Should also allow for non-constant hazard (Weibull)
> survregfit=survreg(Surv(agestop-agestart,dead)~factor(sex)+factor(smk
> summary(survregfit)
                     Value Std. Error      z        p
(Intercept)         6.64290    0.35738  18.59 4.03e-77
factor(sex)2        0.29109    0.05581   5.22 1.83e-07
factor(smkgr)2     -0.20384    0.07719  -2.64 8.28e-03
factor(smkgr)3     -0.51968    0.08622  -6.03 1.67e-09
factor(smkgr)4     -0.52366    0.07630  -6.86 6.73e-12
factor(smkgr)5     -0.65106    0.09198  -7.08 1.46e-12
I(sbp - mean(sbp)) -0.00954    0.00129  -7.41 1.30e-13
factor(bmicat2)1   -0.40310    0.10645  -3.79 1.53e-04
factor(bmicat2)3   -0.13403    0.05353  -2.50 1.23e-02
factor(bmicat2)4   -0.18860    0.09094  -2.07 3.81e-02
agestart           -0.04874    0.00762  -6.40 1.60e-10
Log(scale)         -0.57300    0.04127 -13.89 7.77e-44

Scale= 0.564 (Weibull distribution)
Loglik(model)= -3096.5   Loglik(intercept only)= -3239
        Chisq= 285.07 on 10 degrees of freedom, p= 0
n= 3860
```