5: Generalized Linear Models – Logistic and Poisson Regression

John H Maindonald

August 20, 2015

Ideas and issues illustrated by the graphs in this vignette

Generalized linear models (GLMs) are an extension of linear models. An important special case is models with a binary outcome. Poission regression models are another special case. Graphs shown here illustrate important aspects of the use of these models.

1 Code for the Figures

```
fig5.1 <-
function (){
    ylim <- range(bronchit$poll)+c(0,2.5)
    par(fig=c(0,.525, 0,1))
    plot(xlab="# cigarettes per day", ylab="Pollution", poll ~ cig,
        col=c(2,4)[r+1], pch=(3:2)[r+1], data=bronchit, ylim=ylim)
    legend(x="topleft", legend=c("Non-sufferer","Sufferer"), ncol=2,
        pch=c(3,2), col=c(2,4))
    mtext(side=3, line=1.0,
        expression("A: Untransformed "*italic(x)*"-scale"),
        cex=0.95, adj=0)
    par(fig=c(.475,1, 0,1), new=TRUE)
    plot(poll ~ log(cig+1), col=c(2,4)[r+1], pch=(3:2)[r+1],
        xlab="log(# cigarettes per day + 1)", ylab="",
        data=bronchit, ylim=ylim)
    xy1 <- with(subset(bronchit, r==0), cbind(x=log(cig+1), y=poll))
    xy2 <- with(subset(bronchit, r==1), cbind(x=log(cig+1), y=poll))
    est1 <- bkde2D(xy1, bandwidth=c(0.7, 3))
    est2 <- bkde2D(xy2, bandwidth=c(0.7, 3))
    lev <- pretty(c(est1$fhat, est2$fhat),4)
    contour(est1$x1, est1$x2, est1$fhat, levels=lev, add=TRUE, col=2)
    contour(est2$x1, est2$x2, est2$fhat, levels=lev, add=TRUE, col=4,
        lty=2)
```

```
    legend(x="topleft", legend=c("Non-sufferer","Sufferer"), ncol=2,
           lty=1:2, col=c(2,4), x.intersp=0.5)
    mtext(side=3, line=1.0,
          expression("B: Log-transformed "*italic(x)*"-scale"),
          cex=0.95, adj=0)
    par(fig=c(0,1,0,1))
}


fig5.2 <-
function (plotit=TRUE)
{
    par(mfrow=c(1,2))
    cig2.glm <- glm(r ~ log(cig+1) + poll, family=binomial,
                    data=bronchit)
    termplot(cig2.glm, se=TRUE, ylim=c(-2,4))
    par(mfrow=c(1,1))
}


fig5.3 <-
function ()
{
    nassnew <- subset(nassCDS,
                      !is.na(yearVeh) & yearVeh>=1986 & weight>0)
    nassnew.glm <- glm(dead ~ seatbelt + airbag + dvcat + yearVeh +
                       ageOFocc, weights=weight, family = quasibinomial,
                       data=nassnew)
    par(mfrow=c(1,2))
    termplot(nassnew.glm, terms=c("yearVeh","ageOFocc"),
             smooth=panel.smooth, se=TRUE)
    par(mfrow=c(1,1))
    par(fig=c(0,0.5,0,1), new=TRUE)
    mtext(side=3, line=1.0, "A", adj=0)
    par(fig=c(0.5,1,0,1), new=TRUE)
    mtext(side=3, line=1.0, "B", adj=0)
    par(fig=c(0,1,0,1))
}


fig5.4 <-
function (){
    qqnorm(rpois(30, 5), ylab="", main="")
    qqnorm(rpois(30, 5), ylab="", main="")
}
```

```
fig5.5 <-
function (){
    msg <- "As 'car::spm' is not available, cannot do plot."
    if(!require(car))return(msg)
    car::spm(~ . | habitat, data=moths, cex.labels=1.2,
        smooth=FALSE, reg.line=NA, diag="boxplot")
}
```

```
fig5.6 <-
function ()
{
    P.glm <- glm(P ~ habitat + log(meters), data=moths,
                    family=quasipoisson)
    par(mfrow=c(2,2))
    plot(P.glm, which=1:4)
    par(mfrow=c(1,1))
}
```

2   Show the Figures

```
pkgs <- c("DAAG","KernSmooth","car")
z <- sapply(pkgs, require, character.only=TRUE, warn.conflicts=FALSE)
if(any(!z)){
  notAvail <- paste(names(z)[!z], collapse=", ")
  print(paste("The following packages should be installed:", notAvail))
}
if(!exists("bronchit")){
  if(require("SMIR")) data("bronchit", package="SMIR") else
    print("Dataset 'bronchit' is not available")
}
```
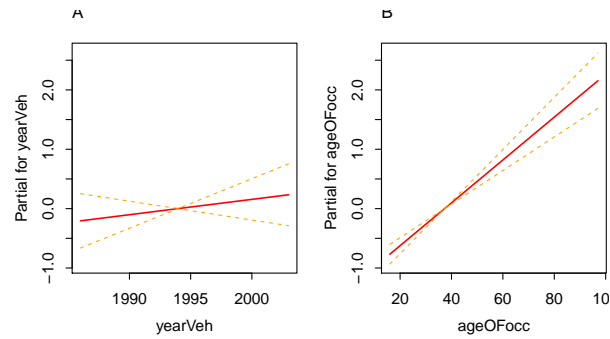
```
if(exists("bronchit"))fig5.1() else
  return("Cannot locate data set 'bronchit', get from 'SMIR'")
```
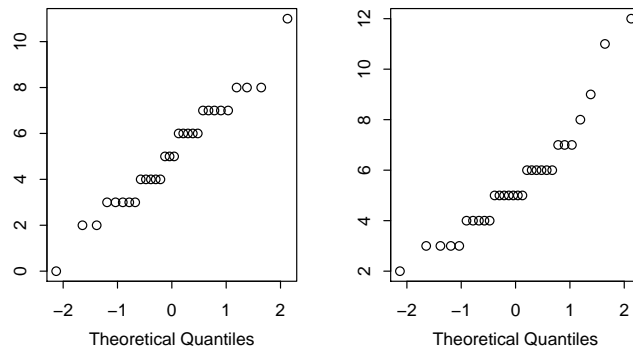
A: Untransformed *x*–scale   B: Log–transformed *x*–scale
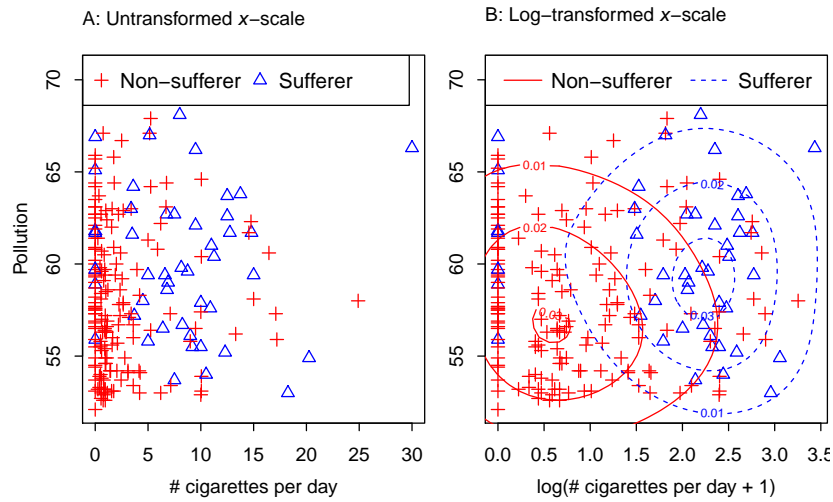

```r
if(exists("bronchit"))fig5.2() else
  return("Cannot locate data set 'bronchit', get from 'SMIR'")
```
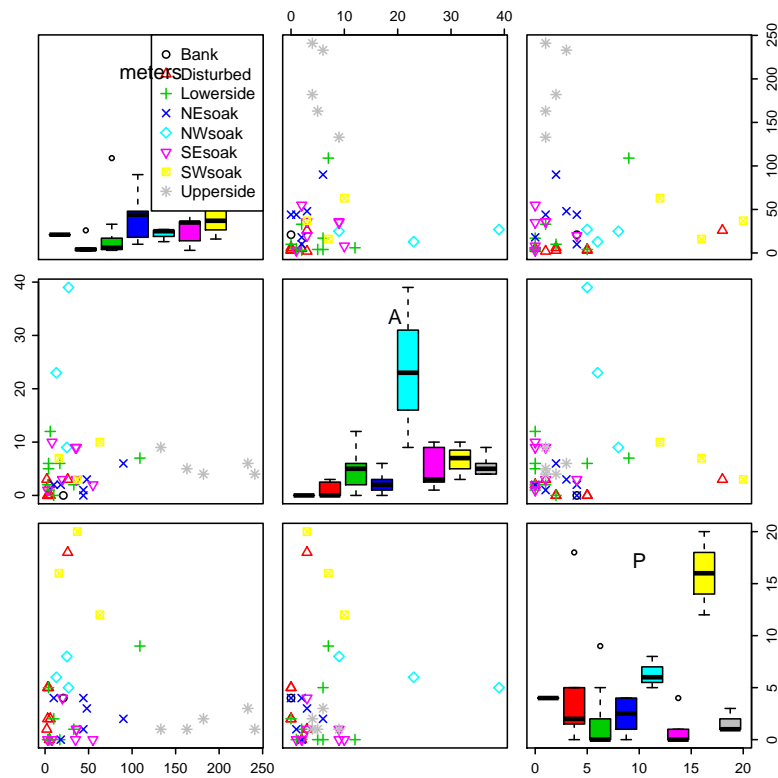


```r
fig5.3()
```


4
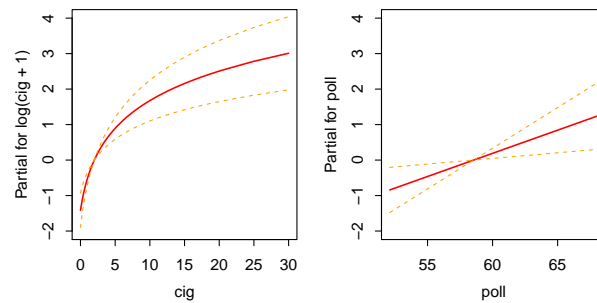
```
fig5.4()
```



```
if(require(DAAG)) fig5.5() else return("Dataset 'moths' is from 'DAAG', not available")
```

```
if(require(DAAG)) fig5.6() else return("Dataset 'moths' is from 'DAAG', not available")
```

Warning: not plotting observations with leverage one:
40