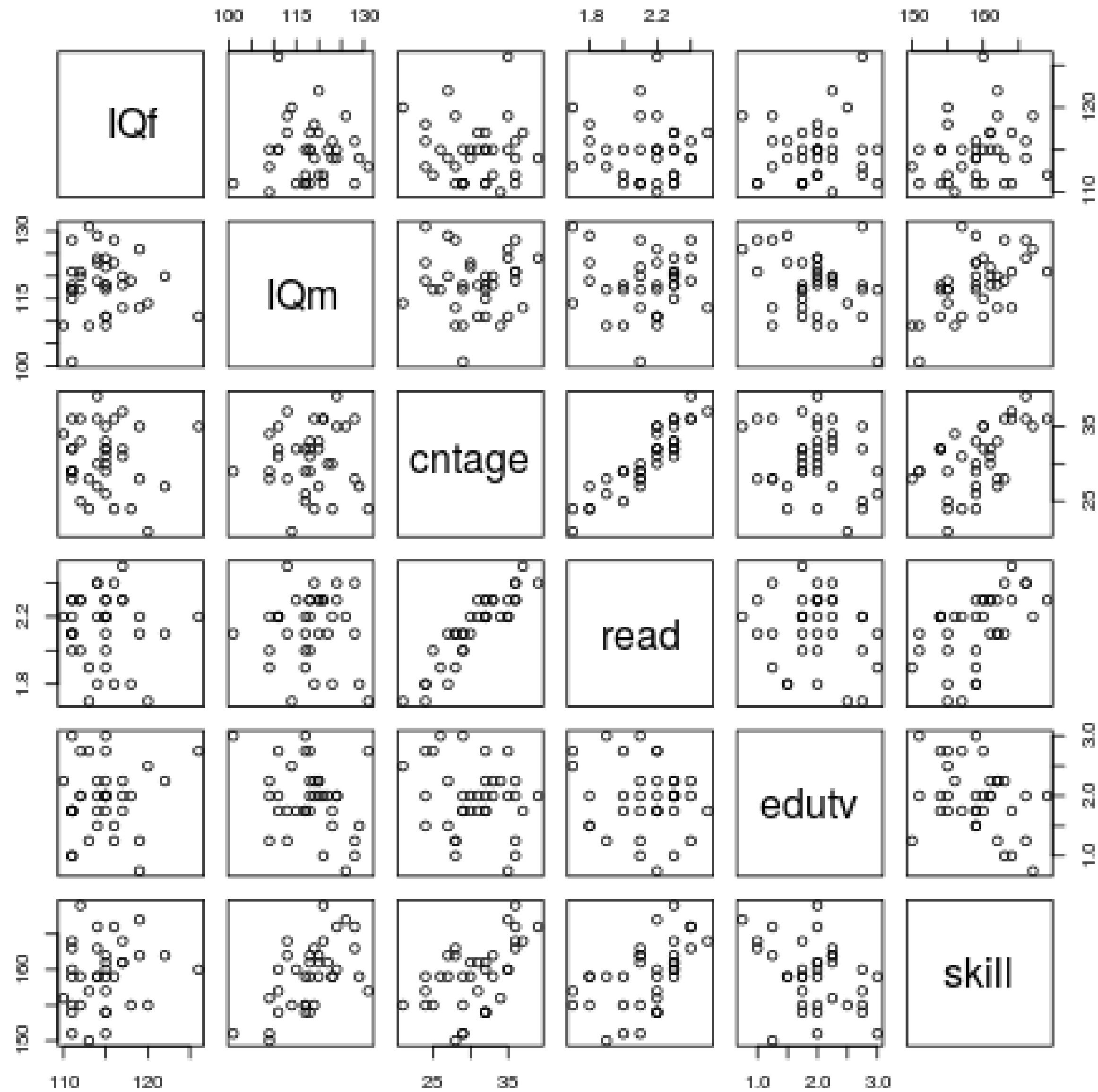
1

a




```
> summary(skills)
      IQf             IQm             cntage           read
 Min.   :110.0   Min.   :101.0   Min.   :21.00   Min.   :1.700
 1st Qu.:112.0   1st Qu.:113.8   1st Qu.:28.00   1st Qu.:2.000
 Median :115.0   Median :118.0   Median :31.00   Median :2.200
 Mean   :114.8   Mean   :118.2   Mean   :30.69   Mean   :2.136
 3rd Qu.:116.2   3rd Qu.:122.2   3rd Qu.:34.25   3rd Qu.:2.300
 Max.   :126.0   Max.   :131.0   Max.   :39.00   Max.   :2.500
```

```
     edutv            skill
 Min.   :0.750   Min.   :150.0
 1st Qu.:1.750   1st Qu.:155.0
 Median :2.000   Median :159.0
 Mean   :1.958   Mean   :159.1
 3rd Qu.:2.250   3rd Qu.:162.0
 Max.   :3.000   Max.   :169.0
```

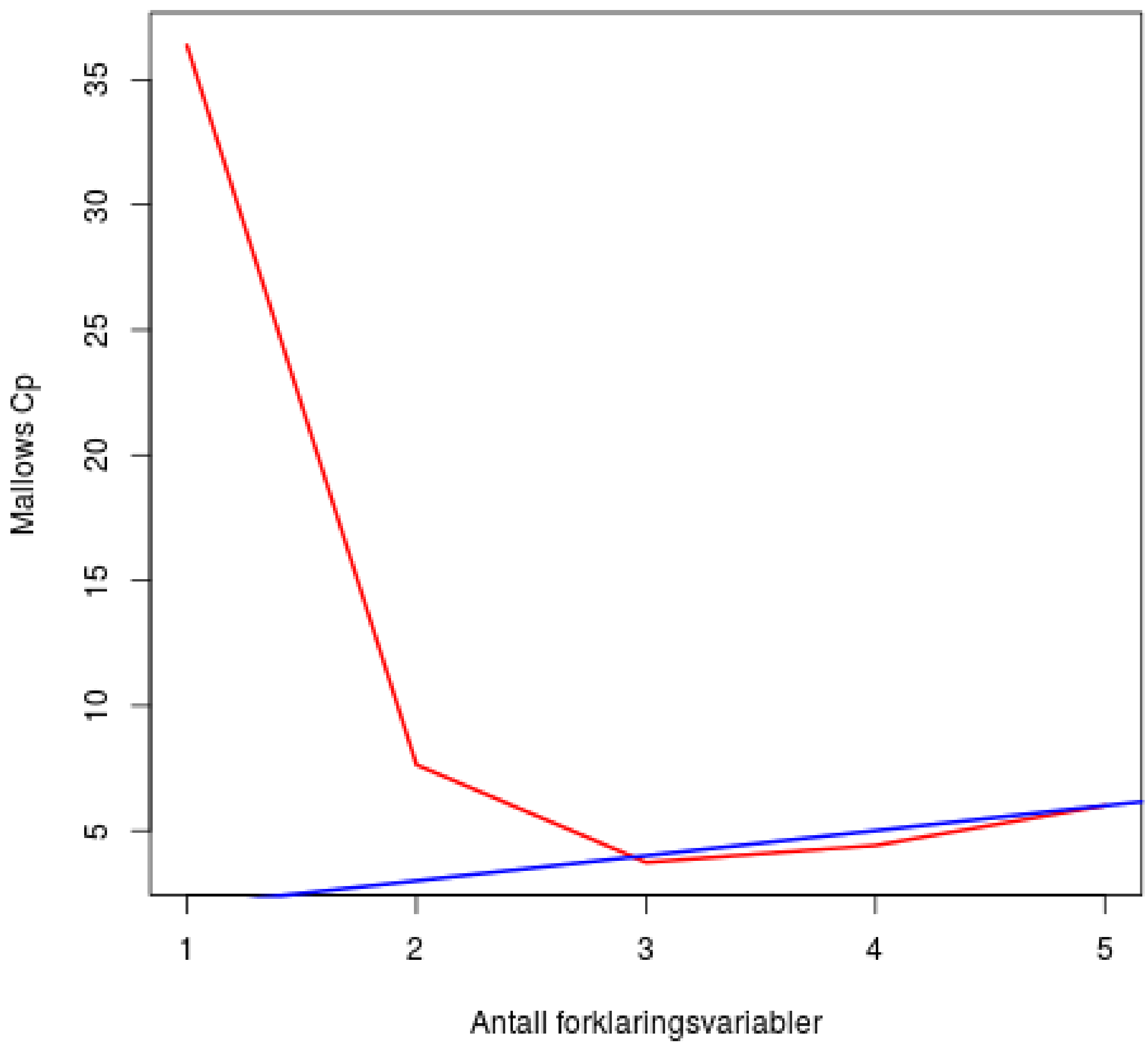
IQm,cntage and read seems highly correlated with skill. And cntage and read is also highly correlated.

b

Forwards selection works by first finding the best model with one variable, then finding the best model found by including 1 more variable etc. Mallows Cp works by finding the number of parameters where Mallows Cp is closest to the number of parameters (m) + 1.

```
library(leaps)
skills = read.table(
    "http://www.uio.no/studier/emner/matnat/math/STK2120/v16/skills.txt",
    header=T)
summary(skills)
png("Oblig2-1summary.png")
plot(skills)
dev.off()

fit.forward = regsubsets(skill~.,data=skills, nvmax=10,method="forward")
summary.forward = summary(fit.forward)
print(summary.forward)
png("Oblig2-1mallows.png")

plot(1:5,summary.forward$cp,
     xlab="Antall forklaringsvariabler", ylab="Mallows Cp",
     type='l',col='red',lwd=2)
abline(1,1,lwd=2,col='blue')
dev.off()

fit.backward = regsubsets(skill~.,data=skills, nvmax=10,method="backward")
summary.backward = summary(fit.backward)
print(summary.backward)

fit.exhaustive = regsubsets(skill~.,data=skills, nvmax=10,method="exhaustive")
summary.exhaustive = summary(fit.exhaustive)
print(summary.exhaustive)
```

```
Selection Algorithm: forward
         IQf IQm cntage read edutv
1  ( 1 ) " " "*" " "    " "  " "
2  ( 1 ) " " "*" " "    "*"  " "
3  ( 1 ) "*" "*" " "    "*"  " "
4  ( 1 ) "*" "*" " "    "*"  "*"
5  ( 1 ) "*" "*" "*"    "*"  "*"
```



n C_p crosses $m+1$ at $m=3$, so we use a model with 3 parameters. By forward selection this is: IQf, IQm and read. Cntage is looks highly correlated with skill, but is not included since it is also correlated with read, so including both does not improve the model much

c

Backwards works like forwards, except starting with all variables, and iteratively removing 1 variable. Exhaustive works by trying every combination of variables.

```
Selection Algorithm: backward
          IQf IQm cntage read edutv
1  ( 1 )  " " "*" " "    " "  " "
2  ( 1 )  " " "*" " "    "*"  " "
3  ( 1 )  "*" "*" " "    "*"  " "
```

```
4  ( 1 ) "*" "*" "_"   "*"  "*"
5  ( 1 ) "*" "*" "*"   "*"  "*"

Selection Algorithm: exhaustive
         IQf IQm cntage read edutv
1  ( 1 ) "_" "*" "_"   "_"  "_"
2  ( 1 ) "_" "*" "_"   "*"  "_"
3  ( 1 ) "*" "*" "_"   "*"  "_"
4  ( 1 ) "*" "*" "_"   "*"  "*"
5  ( 1 ) "*" "*" "*"   "*"  "*"
```

All methods yields the same models.

2

a

$$L(\beta_0, \beta_1) = \prod_i \left[p(x_i)^{y_i}(1 - p(x_i))^{1-y_i}\right]$$

$$L(\beta_0, \beta_1) = \prod_i \left[\left(\frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right)^{y_i}\left(1 - \frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right)^{1-y_i}\right]$$

$$L(\beta_0, \beta_1) = \prod_i \left[\left(\frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right)^{y_i}\left(\frac{1}{1 + e^{\beta_0+\beta_1 x_i}}\right)^{1-y_i}\right]$$

$$L(\beta_0, \beta_1) = \prod_i \left[\left(\frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right)^{y_i}\frac{(1 + e^{\beta_0+\beta_1 x_i})^{y_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right]$$

$$L(\beta_0, \beta_1) = \prod_i \left[\frac{e^{\beta_0 y_i+\beta_1 x_i y_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right]$$

b

$$s_1(\beta_0, \beta_1) = \frac{\partial l(\beta_0, \beta_1)}{\partial \beta_0}$$

$$l(\beta_0, \beta_1) = \sum \log \frac{e^{\beta_0 y_i+\beta_1 x_i y_i}}{1 + e^{\beta_0+\beta_1 x_i}}$$

$$l(\beta_0, \beta_1) = \sum \left[\beta_0 y_i + \beta_1 x_i y_i - \log(1 + e^{\beta_0+\beta_1 x_i})\right]$$

$$s_1(\beta_0, \beta_1) = \frac{\partial}{\partial \beta_0}\sum \left[\beta_0 y_i + \beta_1 x_i y_i - \log(1 + e^{\beta_0+\beta_1 x_i})\right]$$

$$s_1(\beta_0, \beta_1) = \sum \left[y_i - \frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}}\right]$$

$$s_2(\beta_0, \beta_1) = \frac{\partial}{\partial \beta_1}\sum \left[\beta_0 y_i + \beta_1 x_i y_i - \log(1 + e^{\beta_0+\beta_1 x_i})\right]$$

$$s_2(\beta_0, \beta_1) = \sum \left[x_i y_i - \frac{e^{\beta_0+\beta_1 x_i} x_i}{1 + e^{\beta_0+\beta_1 x_i}}\right]$$

c

$$
\begin{aligned}
J_{11}(\beta_0,\beta_1) &= -\frac{\partial s_1(\beta_0,\beta_1)}{\partial\beta_0} \\
J_{11}(\beta_0,\beta_1) &= -\frac{\partial}{\partial\beta_0}\sum\left[y_i - \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right] \\
J_{11}(\beta_0,\beta_1) &= -\sum\left[-\frac{e^{\beta_0+\beta_1 x_i}(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i}e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2}\right] \\
&= \sum\left[\frac{e^{\beta_0+\beta_1 x_i}(1+e^{\beta_0+\beta_1 x_i} - e^{\beta_0+\beta_1 x_i})}{(1+e^{\beta_0+\beta_1 x_i})^2}\right] \\
&= \sum\frac{e^{\beta_0+\beta_1 x_i}}{(1+e^{\beta_0+\beta_1 x_i})^2}
\end{aligned}
$$

$$
\begin{aligned}
J_{22}(\beta_0,\beta_1) &= -\frac{\partial s_2(\beta_0,\beta_1)}{\partial\beta_1} \\
&= -\frac{\partial}{\partial\beta_1}\sum\left[x_i y_i - \frac{e^{\beta_0+\beta_1 x_i}x_i}{1+e^{\beta_0+\beta_1 x_i}}\right] \\
&= \sum\left[\frac{e^{\beta_0+\beta_1 x_i}x_i^2(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i}x_i e^{\beta_0+\beta_1 x_i}x_i}{(1+e^{\beta_0+\beta_1 x_i})^2}\right] \\
&= \sum\left[\frac{e^{\beta_0+\beta_1 x_i}x_i^2(1+e^{\beta_0+\beta_1 x_i} - e^{\beta_0+\beta_1 x_i})}{(1+e^{\beta_0+\beta_1 x_i})^2}\right] \\
&= \sum\frac{e^{\beta_0+\beta_1 x_i}x_i^2}{(1+e^{\beta_0+\beta_1 x_i})^2}
\end{aligned}
$$

$$
\begin{aligned}
J_{12}(\beta_0,\beta_1) = J_{21}(\beta_0,\beta_1) &= -\frac{\partial s_2(\beta_0,\beta_1)}{\partial\beta_0} \\
&= -\frac{\partial}{\partial\beta_0}\sum\left[x_i y_i - \frac{e^{\beta_0+\beta_1 x_i}x_i}{1+e^{\beta_0+\beta_1 x_i}}\right] \\
&= \sum\left[\frac{e^{\beta_0+\beta_1 x_i}x_i(1+e^{\beta_0+\beta_1 x_i}) - e^{\beta_0+\beta_1 x_i}x_i e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}}\right] \\
&= \sum\left[\frac{e^{\beta_0+\beta_1 x_i}x_i(1+e^{\beta_0+\beta_1 x_i} - e^{\beta_0+\beta_1 x_i})}{1+e^{\beta_0+\beta_1 x_i}}\right] \\
&= \sum\frac{e^{\beta_0+\beta_1 x_i}x_i}{1+e^{\beta_0+\beta_1 x_i}}
\end{aligned}
$$

d

$$\begin{aligned}
E(\hat{\beta_0} + \hat{\beta_1 x}) &= \beta_0 + \beta_1 x \\
V(\hat{\beta_0} + \hat{\beta_1 x}) &= I^{11}(\beta_0, \beta_1) + x^2 I^2(\beta_0, \beta_1) + 2x I^{12}(\beta_0, \beta_1) \\
&= \sum_i (\frac{(1+e^{\beta_0+\beta_1 x_i})^2}{e^{\beta_0+\beta_1 x_i}} + x^2 \sum_i \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} + 2x \sum_i \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i e^{\beta_0+\beta_1 x_i}} \\
&= \sum_i \left[\frac{(1+e^{\beta_0+\beta_1 x_i})^2}{e^{\beta_0+\beta_1 x_i}} + x^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} + 2x \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i e^{\beta_0+\beta_1 x_i}} \right] \\
&= \sum_i \left[(x^2 + 2x_i x + x_i^2) \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right] \\
&= \sum_i \left[(x + x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]
\end{aligned}$$

CI for $\beta_0 + \beta_1 x$:

$$\hat{\beta}_0 + \hat{\beta}_1 x \pm z_{.025} \sqrt{\sum_i \left[(x + x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]}$$

Since the logit function is monotone, the CI for $logit(\beta_0 + \beta_1 x)$ is just the logit of the CI for $\beta_0 + \beta_1 x$, so:

$$\left(\frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x - z_{.025} \sqrt{\sum_i \left[(x+x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]}}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x - z_{.025} \sqrt{\sum_i \left[(x+x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]}}}, \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x + z_{.025} \sqrt{\sum_i \left[(x+x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]}}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x + z_{.025} \sqrt{\sum_i \left[(x+x_i)^2 \frac{(1+e^{\beta_0+\beta_1 x_i})^2}{x_i^2 e^{\beta_0+\beta_1 x_i}} \right]}}} \right)$$

3

a

```
s1 <- function(sy, ebetas)
{
    return(sy - sum(ebetas/(1+ebetas)))
}

s2 <- function(sxy, ebetas, x)
{
    return(sxy - sum(x*ebetas/(1+ebetas)))
}

J11 <- function(ebetas)
{
    return(sum(ebetas/(1+ebetas)^2))
}
```

```
J12 <- function(x, ebetas)
{
    return(sum(x*ebetas/(1+ebetas)^2))
}

J22 <- function(x, ebetas)
{
    return(sum(x^2*ebetas/(1+ebetas)^2))
}

scorefunction <- function(x,sy,sxy,ebetas)
{
    return(c(s1(sy,ebetas), s2(sxy, ebetas, x)))
}

obsinfo <- function(ebetas,x)
{
    j12 = J12(x, ebetas)
    return (matrix(c(
               J11(ebetas),j12,
               j12,J22(x,ebetas)
               ), ncol=2))
}

logit <- function(x)
{
    return (exp(x)/(1+exp(x)))
}

logisticNewR <- function(x, y, theta0, eps=0.000001)
{
    n = length(x);
    sy = sum(y)
    sxy = sum(x*y)

    diff = 1; h = 0.0000001;
    theta = theta0
    while(diff>eps)
    {
        theta.old = theta
        ebetas = exp(theta[1] + theta[2]*x)
        s = scorefunction(x,sy, sxy, ebetas)
        Jbar = obsinfo(ebetas,x)
        theta = theta + solve(Jbar,s)
        diff = sum(abs(theta-theta.old))
    }
    return(list(theta=theta, J=Jbar))
}

exmp12.14=read.table(
    "http://www.uio.no/studier/emner/matnat/math/STK2120/v16/exmp12-14.txt",
    header=T)
```

```
x = exmp12.14$temp-70
y = exmp12.14$fail

res = logisticNewR(x, y,c(0,0))

beta0 = res$theta[1]
beta1 = res$theta[2]
J = res$J

Jinv = solve(J)
print(J)
print(Jinv)
SE0 = sqrt(Jinv[1,1])
SE1 = sqrt(Jinv[2,2])

pred = beta0 + (31-70)*beta1


print(beta0)
print(beta1)
print(varBeta0)
print(varBeta1)
```

b

```
exmp12.14=read.table(
+       "http://www.uio.no/studier/emner/matnat/math/STK2120/v16/exmp12-14.txt",
+       header=T)
>
> x = exmp12.14$temp-70
> y = exmp12.14$fail
>
> res = logisticNewR(x, y,c(0,0))
>
> beta0 = res$theta[1]
> beta1 = res$theta[2]
> Jinv = solve(res$J)
> SE0 = sqrt(Jinv[1,1])
> SE1 = sqrt(Jinv[2,2])
> beta0
[1] -1.20849
> beta1
[1] -0.2321627
> SE0
[1] 0.5952546
> SE1
[1] 0.1082365
```

c

```
> x = 31-70
> pred = beta0 + x*beta1
> phat = logit(pred)
> phat
[1] 0.9996088
```

```
> s = sqrt(Jinv[1,1]+t^2*Jinv[2,2] + 2*t* Jinv[1,2])
> z = qnorm(0.975)
> L = logit(pred-s*z)
> L
[1] 0.4816089
> U = logit(pred+s*z)
> U
[1] 0.9999999
```

So the 95prc CI for $p(x)$ is given by $(0.4816, 0.9999)$. The probability of error seems to be at least almost half, which is high. But we do not have any data points around $t = 31$ so our model might not be valid there.