Assignment Exercise 5

R software package used, for easy of data entry, data named; d5 for first case and dt5 for second case.

a)

Pr[$resp_i$ =1]Pr[$resp_i$ =0]β$_i$ + Pr[$resp_i$ =0]Pr[$resp_i$ =1]β$_i$

=0

Also by adding the resulting respective coefficients:

0.06995   +   (-0.06995)   = 0

See output in (bi) below:

b)

Table showing that the transformation implies that the sign of all parameters change:

	Coefficients: First Model	Coefficients: Second Model
(Intercept)	-2.48836	2.48836
male	0.95369	-0.95369
activity	0.91375	-0.91375
age	0.06995	-0.06995
X.age.2..100	-0.06869	0.06869

Computer output is as below:

b)i

```
> # Description of data
>
> head(d5)
  response intercept male activity age X.age.2..100  X X.1 x.beta
1        1         1    0        0  58        33.64 NA  NA       0
2        1         1    1        0  50        25.00 NA  NA       0
3        1         1    1        0  40        16.00 NA  NA       0
4        1         1    1        0  36        12.96 NA  NA       0
5        1         1    1        0  28         7.84 NA  NA       0
6        1         1    1        0  70        49.00 NA  NA       0
  loglik.per.observation
1             -0.6931472
2             -0.6931472
3             -0.6931472
4             -0.6931472
5             -0.6931472
6             -0.6931472
>
> summary(d5)
    response        intercept       male          activity           age
 Min.   :0.0000   Min.   :1   Min.   :0.0000   Min.   :0.0000   Min.   :11.00
 1st Qu.:0.0000   1st Qu.:1   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:40.00
 Median :1.0000   Median :1   Median :1.0000   Median :0.0000   Median :50.00
 Mean   :0.5081   Mean   :1   Mean   :0.7254   Mean   :0.1881   Mean   :50.68
 3rd Qu.:1.0000   3rd Qu.:1   3rd Qu.:1.0000   3rd Qu.:0.0000   3rd Qu.:61.00
```

```
 Max.   :1.0000   Max.    :1   Max.   :1.0000   Max.    :1.0000   Max.    :93.00
   X.age.2..100       X                X.1               x.beta
loglik.per.observation
 Min.   : 1.21   Mode:logical    Mode:logical    Min.    :0   Min.    :-0.6931
 1st Qu.:16.00   NA's:925        NA's:925        1st Qu.:0   1st Qu.:-0.6931
 Median :25.00                                   Median :0   Median :-0.6931
 Mean   :27.48                                   Mean    :0   Mean    :-0.6931
 3rd Qu.:37.21                                   3rd Qu.:0   3rd Qu.:-0.6931
 Max.   :86.49                                   Max.    :0   Max.    :-0.6931
>
>
> mylogit <- glm(response ~ male+ activity +age +X.age.2..100, data = d5,
family = "binomial")
>
>
> summary(mylogit)

Call:
glm(formula = response ~ male + activity + age + X.age.2..100,
    family = "binomial", data = d5)

Deviance Residuals:
    Min      1Q   Median       3Q      Max
-1.6926  -1.2156   0.7389   1.0982   1.8473

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.48836    0.88999  -2.796  0.00517 **
male          0.95369    0.15818   6.029 1.65e-09 ***
activity      0.91375    0.18478   4.945 7.61e-07 ***
age           0.06995    0.03561   1.964  0.04948 *
X.age.2..100 -0.06869    0.03410  -2.015  0.04394 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1282.1  on 924  degrees of freedom
Residual deviance: 1203.7  on 920  degrees of freedom
AIC: 1213.7

Number of Fisher Scoring iterations: 4

> ## odds ratios only
> exp(coef(mylogit))
 (Intercept)         male      activity          age X.age.2..100
  0.08304624   2.59527990   2.49365102   1.07244947   0.93361419
```

b) ii

Description of data
```
>
> summary(dt5)
    response          new.resp         intercept        male            activity
 Min.   :0.0000   Min.    :0.0000   Min.    :1   Min.    :0.0000   Min.
:0.0000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1   1st Qu.:0.0000   1st
Qu.:0.0000
 Median :1.0000   Median :0.0000   Median :1   Median :1.0000   Median
:0.0000
```

```
 Mean    :0.5081    Mean    :0.4919    Mean    :1    Mean    :0.7254    Mean
:0.1881
 3rd Qu.:1.0000    3rd Qu.:1.0000    3rd Qu.:1    3rd Qu.:1.0000    3rd
Qu.:0.0000
 Max.    :1.0000    Max.    :1.0000    Max.    :1    Max.    :1.0000    Max.
:1.0000
      age              X.age.2..100        X              X.1              x.beta
 Min.   :11.00    Min.    : 1.21    Mode:logical    Mode:logical    Min.    :0
 1st Qu.:40.00    1st Qu.:16.00    NA's:925        NA's:925        1st Qu.:0
 Median :50.00    Median :25.00                                    Median :0
 Mean   :50.68    Mean    :27.48                                   Mean    :0
 3rd Qu.:61.00    3rd Qu.:37.21                                    3rd Qu.:0
 Max.   :93.00    Max.    :86.49                                   Max.    :0
 loglik.per.observation
 Min.   :-0.693
 1st Qu.:-0.693
 Median :-0.693
 Mean   :-0.693
 3rd Qu.:-0.693
 Max.   :-0.693
>
> names(dt5)
 [1] "response"                "new.resp"                "intercept"
 [4] "male"                    "activity"                "age"
 [7] "X.age.2..100"            "X"                       "X.1"
[10] "x.beta"                  "loglik.per.observation"
>
> head(dt5)
  response new.resp intercept male activity age X.age.2..100  X X.1 x.beta
1        1        0         1    1        0  58        33.64 NA  NA      0
2        1        0         1    1        0  50        25.00 NA  NA      0
3        1        0         1    1        0  40        16.00 NA  NA      0
4        1        0         1    1        0  36        12.96 NA  NA      0
5        1        0         1    1        0  28         7.84 NA  NA      0
6        1        0         1    1        0  70        49.00 NA  NA      0
  loglik.per.observation
1                 -0.693
2                 -0.693
3                 -0.693
4                 -0.693
5                 -0.693
6                 -0.693
>
>
> mylogit <- glm(new.resp ~ male+ activity +age +X.age.2..100, data = dt5,
family = "binomial")
>
> summary(mylogit)

Call:
glm(formula = new.resp ~ male + activity + age + X.age.2..100,
    family = "binomial", data = dt5)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8473  -1.0982  -0.7389   1.2156   1.6926

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.48836    0.88999   2.796  0.00517 **
male        -0.95369    0.15818  -6.029 1.65e-09 ***
activity    -0.91375    0.18478  -4.945 7.61e-07 ***
age         -0.06995    0.03561  -1.964  0.04948 *
```

```
X.age.2..100   0.06869     0.03410    2.015   0.04394 *
---
Signif. codes:   0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1282.1  on 924  degrees of freedom
Residual deviance: 1203.7  on 920  degrees of freedom
AIC: 1213.7

Number of Fisher Scoring iterations: 4

>
> ## odds ratios only
> exp(coef(mylogit))
 (Intercept)         male      activity             age X.age.2..100
  12.0414837    0.3853149     0.4010184      0.9324449     1.0711063
```

c)

This cut off point can be adjusted to reflect the percentage of observation equal 1 in the sample for both male and female.