Test Exercise 5: Answers to the Questions

- (a) Show that $\frac{\partial Pr[resp_i=1]}{\partial age_i} + \frac{\partial Pr[resp_i=0]}{\partial age_i} = 0$.

(respi is a binary variable, can have only 2 values 0 & 1)

$$Pr[resp_i = 1] + Pr[resp_i = 0] = 1 \quad \text{(Since mutually exclusive and exhaustive events)}$$

$$\Rightarrow \frac{\partial}{\partial age_i}(Pr[resp_i = 1] + Pr[resp_i = 0]) = \frac{\partial}{\partial age_i}(1)$$

$$\Rightarrow \frac{\partial Pr[resp_i = 1]}{\partial age_i} + \frac{\partial Pr[resp_i = 0]}{\partial age_i} = 0$$

- (b) Assume that you recode the dependent variable as follows: $resp_i^{new} = -resp_i + 1$. Hence, positive response is now defined to be equal to zero and negative response to be equal to 1. Use the *odds ratio* to show that this transformation implies that the *sign* of all parameters change.
    - $resp_i^{new} = -resp_i + 1 \Rightarrow resp_i = -resp_i^{new} + 1$. Now, we have,

$$\frac{Pr[resp_i = 1]}{Pr[resp_i = 0]} = exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2)$$

$$\Rightarrow \frac{Pr[-resp_i^{new} + 1 = 1]}{Pr[-resp_i^{new} + 1 = 0]} = exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2)$$

$$\Rightarrow \frac{Pr[resp_i^{new} = 0]}{Pr[resp_i^{new} = 1]} = exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2)$$

$$\Rightarrow \frac{Pr[resp_i^{new} = 1]}{Pr[resp_i^{new} = 0]} = \frac{1}{exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2)}$$

$$= exp(-(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2))$$

$$= exp(-\beta_0 - \beta_1 male_i - \beta_2 active_i - \beta_3 age_i - \beta_4 (age_i/10)^2))$$

- © Consider again the *odds ratio* positive response versus negative response:

$$\frac{Pr[resp_i=1]}{Pr[resp_i=0]} = exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2).$$

During lecture 5.5 you have seen that this odds ratio obtains its maximum value for age equal to 50 years for males as well as females. Suppose now that you want to extend the logit model and allow that this age value is possibly different for males than for females. Discuss how you can extend the logit specification.

- We consider the interaction term for gender and age and as shown below our model learnt is as follows:

$$\frac{Pr[resp_i = 1]}{Pr[resp_i = 0]} = exp(\beta_0 + \beta_1 male_i + \beta_2 active_i + \beta_3 age_i + \beta_4 (age_i/10)^2 + \beta_5 (male_i \times age_i))$$

$$\approx exp(-2.663741 + 1.171787 male_i + 0.912235 active_i + 0.073900 age_i - 0.069591 (age_i/10)^2 - 0.004308 (male_i \times age_i))$$

$$\approx 14.34987 \times 3.227755^{male_i} \times 2.489881^{active_i} \times exp(0.073900 age_i - 0.069591 (age_i/10)^2 - 0.004308 (male_i \times age_i))$$

So, when we have $male_i = 1$, i.e., **for males** the **first order condition** for the **highest odds ratio** becomes $[14.34987 \times 3.227755^1 \times 2.489881^{active_i} \times exp(0.073900 age_i - 0.069591 (age_i/10)^2 - 0.004308 (1 \times age_i))] \times (0.073900 - 0.069591 \times age_i/50 - 0.004308) = 0$. The solution to this first order condition **for males** is 50.00072 years.

Similarly, when we have $male_i = 0$, i.e., **for females** the **first order condition** for the **highest odds ratio** becomes $[14.34987 \times 3.227755^0 \times 2.489881^{active_i} \times exp(0.073900 age_i - 0.069591 (age_i/10)^2 - 0.004308 (0 \times age_i))] \times (0.073900 - 0.069591 \times age_i/50 - 0) = 0$. The solution to this first order condition **for females** is 53.09595 years.

```
##
## Call:
## glm(formula = response ~ . + I((age/10)^2) + male:age, family = "binomial",
##     data = df)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6925  -1.2071   0.7387   1.0973   1.8495
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -2.663741   1.009065  -2.640   0.0083 **
## male           1.171787   0.599319   1.955   0.0506 .
## activity       0.912235   0.184811   4.936 7.97e-07 ***
## age            0.073900   0.037236   1.985   0.0472 *
## I((age/10)^2) -0.069591   0.034220  -2.034   0.0420 *
## male:age      -0.004308   0.011400  -0.378   0.7055
## ---
## 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: 1203.6  on 919  degrees of freedom
##   (90 observations deleted due to missingness)
## AIC: 1215.6
##
## Number of Fisher Scoring iterations: 4
```