# Problem Set 4

*Wenjing Li*

*11/18/2019*

## Question 1

```
## New names:
## * `` -> ...1
```

Table 1: *Estimation for 4 different measurements.*

| ...1 | tot_dist | max_abs_dev | avg_abs_dev | auc |
|---|---|---|---|---|
| estimation | -0.161 | -0.510 | -0.651 | -0.406 |
| lwr | -0.234 | -0.731 | -0.906 | -0.596 |
| upr | -0.089 | -0.289 | -0.396 | -0.215 |

## Question 2

```
## Parsed with column specification:
## cols(
##   division = col_double(),
##   dispar_weight = col_double(),
##   lwr_dispar = col_double(),
##   upr_dispar = col_double()
## )
```

Table 2: *key data summaries for disparity in different census division.*

| division | dispar_weight | lwr_dispar | upr_dispar |
|---|---|---|---|
| 9 | 0.185 | 0.127 | 0.243 |
| 6 | 0.093 | 0.039 | 0.148 |
| 4 | 0.077 | 0.025 | 0.128 |
| 8 | 0.055 | -0.005 | 0.115 |
| 7 | 0.051 | 0.013 | 0.089 |
| 10 | 0.034 | -0.004 | 0.072 |
| 5 | 0.033 | -0.002 | 0.067 |
| 2 | 0.019 | -0.013 | 0.052 |
| 1 | 0.018 | 0.001 | 0.035 |
| 3 | 0.000 | -0.026 | 0.027 |

Note that the division 9 has the largest disparity. Division 9 is the mountain south.

## Question 3

```
## New names:
```

```
## * `` -> ...1
```

Table 3: *Result for logistic regression.*

| ...1 | weekday | winter | centered_age | centerd_age^2 | gender | PIR |
|---|---|---|---|---|---|---|
| ratio | 1.001 | 0.955 | 1.185 | 0.977 | 1.362 | 1.090 |
| lwr | 0.862 | 0.891 | 1.162 | 0.970 | 1.271 | 1.066 |
| upr | 1.164 | 1.023 | 1.208 | 0.983 | 1.459 | 1.115 |
| margin | 0.000 | -0.009 | 0.035 | 0.058 | 0.016 | NA |
| lwr_margin | -0.028 | -0.022 | 0.031 | 0.045 | 0.012 | NA |
| upr_margin | 0.029 | 0.004 | 0.040 | 0.072 | 0.021 | NA |

```
## New names:
## * `` -> ...1
```

Table 4: *Result for mixed logistic regression.*

| ...1 | weekday | winter | centered_age | centerd_age^2 | gender | PIR |
|---|---|---|---|---|---|---|
| ratio | 0.961 | 0.920 | 1.332 | 0.962 | 1.726 | 1.165 |
| lwr | 0.738 | 0.802 | 1.280 | 0.949 | 1.503 | 1.114 |
| upr | 1.252 | 1.056 | 1.386 | 0.975 | 1.983 | 1.219 |
| margin | -0.004 | -0.009 | 0.033 | 0.059 | 0.016 | NA |
| lwr_margin | -0.033 | -0.024 | 0.028 | 0.044 | 0.012 | NA |
| upr_margin | 0.024 | 0.006 | 0.038 | 0.074 | 0.021 | NA |

Note that two methods product similar result and the result from the first logistic model is much more precise than the other one.

Since all of the odd ratios in the model are around 1, that means each increasion in the predictors may not cause a difference in response. Thus, the statement that people in the US are more likely to drink water on a weekday than a weekend day is not sure.