

Practicum 1 Analysis

Peter Claussen

Ben Derenge

Stephanie Liebl

9/16/2021

Analysis

Data Processing

```
metrics.dat <- read.csv('Practicum 1 Data.csv',header=TRUE)
metrics.dat <- metrics.dat[!is.na(metrics.dat$Snumber),]

metrics.dat$CalcTMM <- with(metrics.dat, 8*Vig.ex.Time + 4*Mod.ex.time + 3.3*Walk.ex.Time)

metrics.dat$shift[metrics.dat$shift==''] <- 'missing'
shift.levels <- c(paste(c(7:11),'am',sep=''),paste(c(12,1:2),'pm',sep=''),'other','missing')
metrics.dat$shift <- factor(metrics.dat$shift,shift.levels)
summary(metrics.dat$shift)
```

```
##      7am      8am      9am     10am     11am     12pm      1pm      2pm    other missing
##      31      115      56      50      44      14       8      15      15      4
```

```
metrics.dat$MissingLbs <- is.na(metrics.dat$pounds_gained)
table(metrics.dat$MissingLbs,metrics.dat$weightgain)
```

```
##
##           No Yes
##  FALSE    0   1 231
##   TRUE    4 110   6
```

We consider two subsets for analysis. First we create a data table that has appropriate values for `weightgain`. This will be the larger of the two data sets.

```
gained.dat <- metrics.dat[metrics.dat$weightgain %in% c('Yes','No'),]
gained.dat$WG <- FALSE
gained.dat$WG[gained.dat$weightgain=='Yes'] <- TRUE
gained.dat$pounds_gained[!gained.dat$WG] <- 0
dim(gained.dat)
```

```
## [1] 348  86
```

```
gained.dat$MissingLbs <- is.na(gained.dat$pounds_gained)
table(gained.dat$MissingLbs,gained.dat$weightgain)
```

```
##
##           No Yes
##  FALSE  111 231
##   TRUE    0   6
```

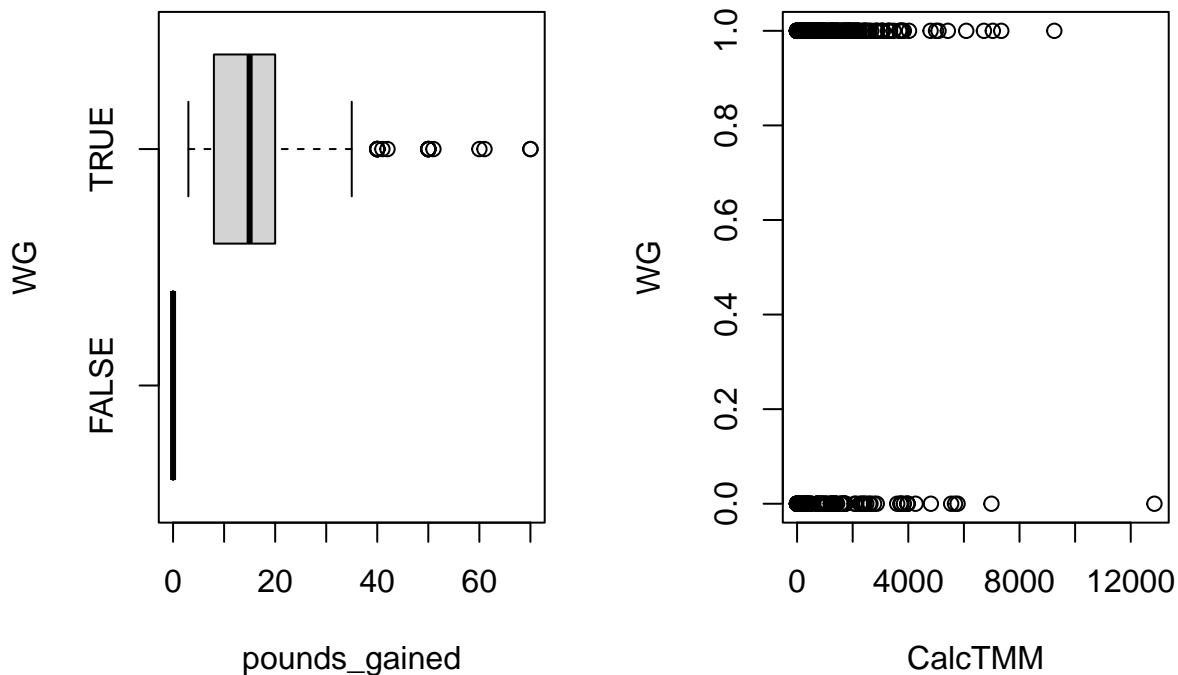
```
gained.dat <- gained.dat[!is.na(gained.dat$CalcTMM),]
dim(gained.dat)
```

```
## [1] 347 86
```

```
#gained.dat <- gained.dat[!gained.dat$MissingLbs,]
dim(gained.dat)
```

```
## [1] 347 86
```

```
par(mfrow=c(1,2))
boxplot(pounds_gained ~ WG,data=gained.dat,horizontal = TRUE)
plot(WG~CalcTMM,data=gained.dat)
```



#

Analysis of Binary Response (WG)

(SA1) Does *total metabolic minutes* have an effect on *weight gain*?

Simple logistic regression

```
SA1.model1 <- glm(WG ~ CalcTMM, data=gained.dat,family = binomial)
summary(SA1.model1)
```

```
##
```

```
## Call:
```

```
## glm(formula = WG ~ CalcTMM, family = binomial, data = gained.dat)
```

```
##
```

```
## Deviance Residuals:
```

```
##      Min       1Q   Median       3Q      Max
```

```
## -1.5702 -1.4777  0.8403  0.8697  1.1523
```

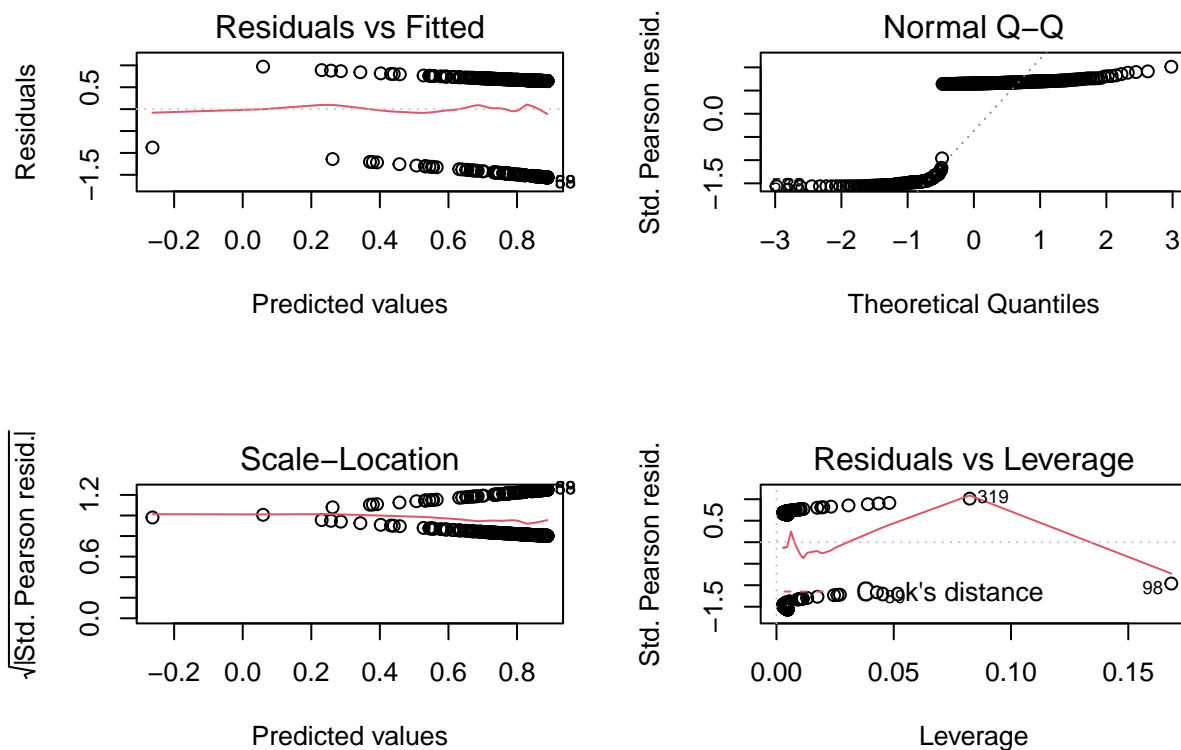
```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept) 8.882e-01 1.517e-01 5.855 4.78e-09 ***
## CalcTMM -8.959e-05 7.141e-05 -1.255 0.21
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 433.47 on 346 degrees of freedom
## Residual deviance: 431.91 on 345 degrees of freedom
## AIC: 435.91
##
## Number of Fisher Scoring iterations: 4
```

```
par(mfrow=c(2,2))
plot(SA1.model1)
```



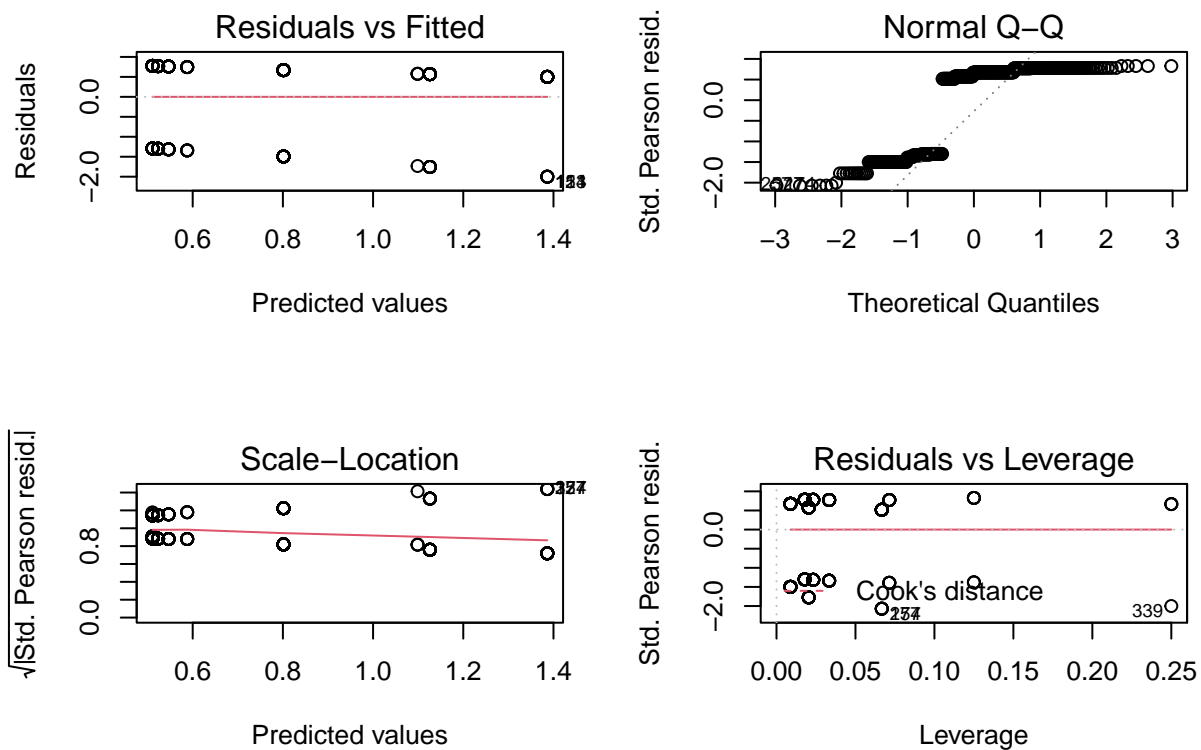
- (SA2) Does *shift* have an effect on *weight gain*?

```
SA2.model1 <- glm(WG ~ shift, data=gained.dat, family = binomial)
summary(SA2.model1)
```

```
##
## Call:
## glm(formula = WG ~ shift, family = binomial, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.7941  -1.4006   0.7585   0.9400   0.9695
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)    0.54654    0.37887    1.443    0.149
## shift8am       0.25482    0.43004    0.593    0.553
## shift9am      -0.03572    0.46875   -0.076    0.939
## shift10am      0.57947    0.50389    1.150    0.250
## shift11am     -0.02330    0.49303   -0.047    0.962
## shift12pm      0.04124    0.67428    0.061    0.951
## shift1pm      -0.03572    0.82272   -0.043    0.965
## shift2pm       0.83975    0.74847    1.122    0.262
## shifttother    0.83975    0.74847    1.122    0.262
## shiftmissing   0.55207    1.21527    0.454    0.650
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 433.47  on 346  degrees of freedom
## Residual deviance: 428.07  on 337  degrees of freedom
## AIC: 448.07
##
## Number of Fisher Scoring iterations: 4
```

```
par(mfrow=c(2,2))
plot(SA2.model1)
```



Model 2 Interactions

```
SA12.model2 <- glm(WG ~ shift+CalcTMM, data=gained.dat, family = binomial)
summary(SA12.model2)
```

```
##
```

```
## Call:
## glm(formula = WG ~ shift + CalcTMM, family = binomial, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8851  -1.4015   0.8104   0.9193   1.2372
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  6.964e-01  3.935e-01   1.770  0.0768 .
## shift8am     2.485e-01  4.312e-01   0.576  0.5645
## shift9am    -5.053e-02  4.702e-01  -0.107  0.9144
## shift10am    5.807e-01  5.049e-01   1.150  0.2501
## shift11am   -5.755e-02  4.947e-01  -0.116  0.9074
## shift12pm    2.385e-02  6.758e-01   0.035  0.9718
## shift1pm    -5.826e-02  8.241e-01  -0.071  0.9436
## shift2pm     8.839e-01  7.514e-01   1.176  0.2395
## shifttother  8.951e-01  7.540e-01   1.187  0.2352
## shiftmissing 5.872e-01  1.218e+00   0.482  0.6297
## CalcTMM     -1.069e-04  7.262e-05  -1.473  0.1408
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 433.47  on 346  degrees of freedom
## Residual deviance: 425.92  on 336  degrees of freedom
## AIC: 447.92
##
## Number of Fisher Scoring iterations: 4
```

Model 3 SA1 and 2 plus anthropometric variables

```
SA12.model3a <- glm(WG ~ gender + Age + shift + CalcTMM, data=gained.dat,family = binomial)
summary(SA12.model3a)
```

```
##
## Call:
## glm(formula = WG ~ gender + Age + shift + CalcTMM, family = binomial,
##      data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9936  -1.3159   0.7661   0.8673   1.4101
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.897e-01  1.623e+00   0.117  0.907
## genderFemale  1.240e+00  1.472e+00   0.842  0.400
## genderMale    7.071e-01  1.486e+00   0.476  0.634
## Age          -6.189e-03  1.263e-02  -0.490  0.624
## shift8am     -1.491e-01  4.821e-01  -0.309  0.757
```

```

## shift9am      -3.842e-01  5.240e-01  -0.733    0.464
## shift10am     1.715e-01  5.524e-01   0.310    0.756
## shift11am     -1.932e-01  5.542e-01  -0.349    0.727
## shift12pm     -4.140e-01  7.192e-01  -0.576    0.565
## shift1pm      -4.538e-01  8.572e-01  -0.529    0.597
## shift2pm       1.331e+00  1.151e+00   1.157    0.247
## shifttother    5.339e-01  8.128e-01   0.657    0.511
## shiftmissing  1.374e+01  6.107e+02   0.022    0.982
## CalcTMM       -1.102e-04  7.636e-05  -1.444    0.149
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 393.55  on 318  degrees of freedom
## Residual deviance: 379.52  on 305  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 407.52
##
## Number of Fisher Scoring iterations: 13
SA12.model3b <- glm(WG ~ gender + Age + height +shift + CalcTMM, data=gained.dat,family = binomial)
summary(SA12.model3b)

##
## Call:
## glm(formula = WG ~ gender + Age + height + shift + CalcTMM, family = binomial,
## data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9392  -1.2928   0.7580   0.8566   1.3899
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.657e-01  3.381e+00  -0.049    0.961
## genderFemale  1.268e+00  1.489e+00   0.852    0.394
## genderMale    6.805e-01  1.486e+00   0.458    0.647
## Age          -2.496e-03  1.294e-02  -0.193    0.847
## height        3.124e-03  4.170e-02   0.075    0.940
## shift8am     -1.112e-01  4.834e-01  -0.230    0.818
## shift9am     -3.525e-01  5.278e-01  -0.668    0.504
## shift10am     2.318e-01  5.611e-01   0.413    0.680
## shift11am    -1.797e-01  5.549e-01  -0.324    0.746
## shift12pm    -4.107e-01  7.199e-01  -0.571    0.568
## shift1pm     -4.402e-01  8.578e-01  -0.513    0.608
## shift2pm      1.339e+00  1.153e+00   1.161    0.246
## shifttother   4.589e-01  8.196e-01   0.560    0.576
## shiftmissing  1.375e+01  6.104e+02   0.023    0.982
## CalcTMM      -1.052e-04  7.680e-05  -1.370    0.171
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 384.25  on 312  degrees of freedom
## Residual deviance: 370.32  on 298  degrees of freedom
## (34 observations deleted due to missingness)
## AIC: 400.32

```

```
##
## Number of Fisher Scoring iterations: 13
```

Model 4 Partition CalcTMM into components

```
SA12.model4a <- glm(WG ~ gender + Age + shift + Vig.ex.Time + Mod.ex.time + Walk.ex.Time, data=gained.dat)
summary(SA12.model4a)
```

```
##
## Call:
## glm(formula = WG ~ gender + Age + shift + Vig.ex.Time + Mod.ex.time +
##     Walk.ex.Time, family = binomial, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9919  -1.3054   0.7611   0.8657   1.4278
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   3.761e-01  1.662e+00   0.226   0.821
## genderFemale  1.046e+00  1.517e+00   0.690   0.490
## genderMale    5.222e-01  1.527e+00   0.342   0.732
## Age          -6.569e-03  1.269e-02  -0.518   0.605
## shift8am      -1.245e-01  4.846e-01  -0.257   0.797
## shift9am      -3.901e-01  5.260e-01  -0.742   0.458
## shift10am     2.003e-01  5.557e-01   0.360   0.719
## shift11am     -1.591e-01  5.583e-01  -0.285   0.776
## shift12pm     -4.125e-01  7.198e-01  -0.573   0.567
## shift1pm      -4.500e-01  8.575e-01  -0.525   0.600
## shift2pm      1.336e+00  1.160e+00   1.152   0.249
## shifttother    5.447e-01  8.137e-01   0.669   0.503
## shiftmissing  1.376e+01  6.103e+02   0.023   0.982
## Vig.ex.Time   -1.434e-03  1.231e-03  -1.165   0.244
## Mod.ex.time   -4.887e-05  1.167e-03  -0.042   0.967
## Walk.ex.Time  -2.474e-04  5.875e-04  -0.421   0.674
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 393.55  on 318  degrees of freedom
## Residual deviance: 379.25  on 303  degrees of freedom
## (28 observations deleted due to missingness)
## AIC: 411.25
##
## Number of Fisher Scoring iterations: 13
```

```
SA12.model4b <- glm(WG ~ gender + Age + height + shift + Vig.ex.Time + Mod.ex.time + Walk.ex.Time, data=gained.dat)
summary(SA12.model4b)
```

```
##
## Call:
## glm(formula = WG ~ gender + Age + height + shift + Vig.ex.Time +
##     Mod.ex.time + Walk.ex.Time, family = binomial, data = gained.dat)
##
```

```
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9366  -1.2871   0.7556   0.8564   1.3901
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -3.843e-03  3.414e+00  -0.001   0.999
## genderFemale  1.092e+00  1.540e+00   0.709   0.478
## genderMale    5.064e-01  1.531e+00   0.331   0.741
## Age          -2.712e-03  1.301e-02  -0.208   0.835
## height        3.138e-03  4.175e-02   0.075   0.940
## shift8am      -8.385e-02  4.860e-01  -0.173   0.863
## shift9am      -3.485e-01  5.298e-01  -0.658   0.511
## shift10am     2.608e-01  5.638e-01   0.463   0.644
## shift11am     -1.447e-01  5.592e-01  -0.259   0.796
## shift12pm     -4.137e-01  7.202e-01  -0.574   0.566
## shift1pm      -4.334e-01  8.579e-01  -0.505   0.613
## shift2pm       1.321e+00  1.163e+00   1.136   0.256
## shifttother    4.707e-01  8.202e-01   0.574   0.566
## shiftmissing  1.377e+01  6.105e+02   0.023   0.982
## Vig.ex.Time   -1.348e-03  1.259e-03  -1.071   0.284
## Mod.ex.time    1.043e-04  1.198e-03   0.087   0.931
## Walk.ex.Time  -3.354e-04  5.892e-04  -0.569   0.569
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 384.25  on 312  degrees of freedom
## Residual deviance: 370.06  on 296  degrees of freedom
## (34 observations deleted due to missingness)
## AIC: 404.06
##
## Number of Fisher Scoring iterations: 13
```

Model 5 - Model 4 plus BMI and initial body weight

For these models, we may include BMI, or just the anthropometric variables used to calculate BMI

```
gained.dat['initial_bweight'] <- gained.dat$bweight - gained.dat$pounds_gained
gained.dat['initial_BMI'] <- (gained.dat$initial_bweight / (gained.dat$height)^2)*703
```

```
SA12.model4a <- glm(WG ~ gender + Age + shift + Vig.ex.Time + Mod.ex.time + Walk.ex.Time + initial_BMI
summary(SA12.model4a)
```

```
##
## Call:
## glm(formula = WG ~ gender + Age + shift + Vig.ex.Time + Mod.ex.time +
##      Walk.ex.Time + initial_BMI, family = binomial, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.0146  -1.1930   0.7081   0.8825   1.3010
##
## Coefficients:
```



```

##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  5.545e-01  1.816e+00  0.305  0.7602
## genderFemale  1.475e+00  1.555e+00  0.949  0.3428
## genderMale    9.743e-01  1.570e+00  0.620  0.5349
## Age           1.258e-02  1.536e-02  0.819  0.4128
## shift8am      -2.794e-01  5.425e-01 -0.515  0.6065
## shift9am       1.927e-01  6.331e-01  0.304  0.7609
## shift10am      5.366e-01  6.421e-01  0.836  0.4034
## shift11am     -2.815e-01  6.246e-01 -0.451  0.6523
## shift12pm      3.545e-01  8.427e-01  0.421  0.6740
## shift1pm       6.780e-02  1.023e+00  0.066  0.9471
## shift2pm       1.590e+01  9.545e+02  0.017  0.9867
## shifttother    4.309e-01  8.647e-01  0.498  0.6182
## shiftmissing   1.613e+01  1.675e+03  0.010  0.9923
## Vig.ex.Time   -4.905e-04  1.427e-03 -0.344  0.7310
## Mod.ex.time    -5.897e-04  1.473e-03 -0.400  0.6889
## Walk.ex.Time   -3.366e-04  6.597e-04 -0.510  0.6098
## initial_BMI    -5.306e-02  2.664e-02 -1.992  0.0464 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 301.19  on 241  degrees of freedom
## Residual deviance: 280.85  on 225  degrees of freedom
## (105 observations deleted due to missingness)
## AIC: 314.85
##
## Number of Fisher Scoring iterations: 15
SA12.model4b <- glm(WG ~ gender + Age + height + shift + Vig.ex.Time + Mod.ex.time + Walk.ex.Time + initial_bweight, family = binomial, data = gained.dat)
summary(SA12.model4b)

##
## Call:
## glm(formula = WG ~ gender + Age + height + shift + Vig.ex.Time +
##      Mod.ex.time + Walk.ex.Time + initial_bweight, family = binomial,
##      data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9710  -1.2121   0.7130   0.8866   1.3026
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.567e+00  3.891e+00 -0.660  0.5095
## genderFemale  1.520e+00  1.581e+00  0.962  0.3362
## genderMale    9.834e-01  1.572e+00  0.626  0.5316
## Age           1.207e-02  1.539e-02  0.785  0.4326
## height        4.495e-02  5.134e-02  0.875  0.3813
## shift8am      -2.806e-01  5.430e-01 -0.517  0.6053
## shift9am       1.854e-01  6.329e-01  0.293  0.7696
## shift10am      5.229e-01  6.434e-01  0.813  0.4164
## shift11am     -2.799e-01  6.261e-01 -0.447  0.6549
## shift12pm      3.139e-01  8.410e-01  0.373  0.7090

```

```
## shift1pm      5.388e-02  1.018e+00  0.053  0.9578
## shift2pm      1.594e+01  9.521e+02  0.017  0.9866
## shifttother    4.353e-01  8.649e-01  0.503  0.6147
## shiftmissing   1.611e+01  1.670e+03  0.010  0.9923
## Vig.ex.Time    -4.464e-04  1.425e-03 -0.313  0.7541
## Mod.ex.time     -6.272e-04  1.471e-03 -0.426  0.6699
## Walk.ex.Time    -3.272e-04  6.621e-04 -0.494  0.6211
## initial_bweight -7.707e-03  4.215e-03 -1.828  0.0675 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 301.19 on 241 degrees of freedom
## Residual deviance: 281.36 on 224 degrees of freedom
## (105 observations deleted due to missingness)
## AIC: 317.36
##
## Number of Fisher Scoring iterations: 15
```

Pounds Gained analysis

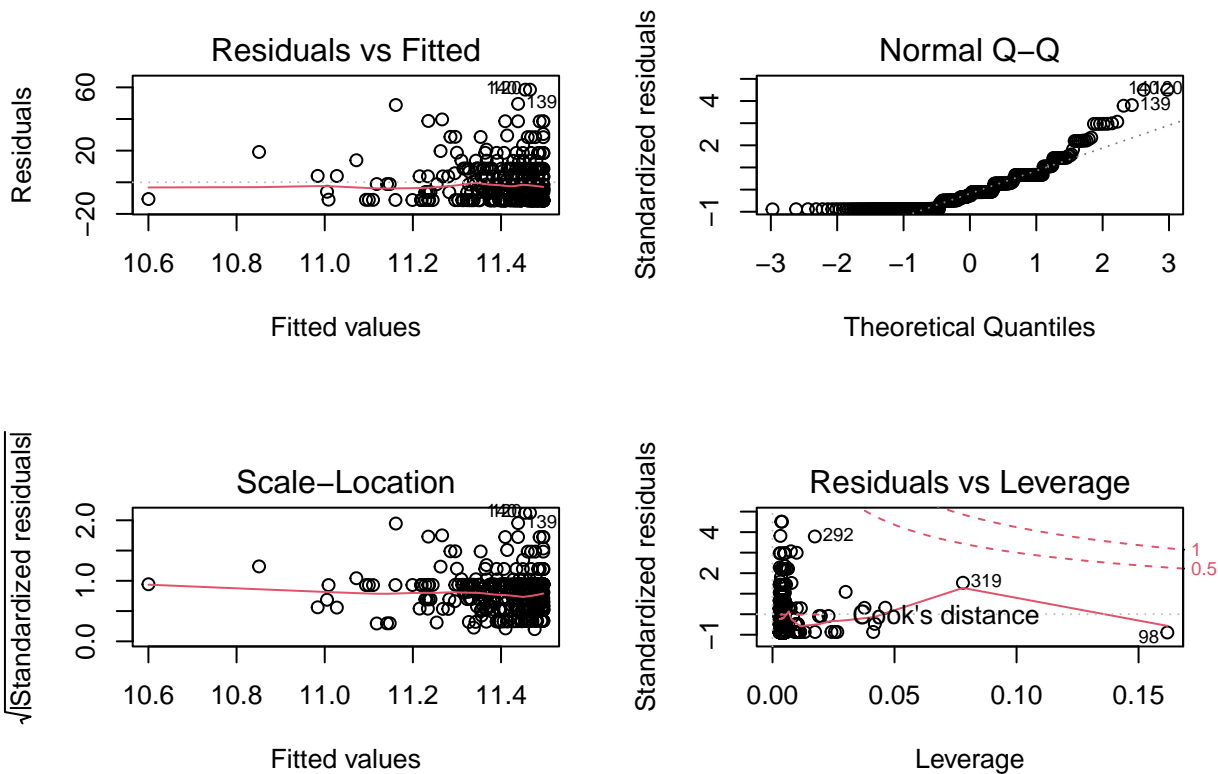
```
gained.dat$LBS <- round(gained.dat$pounds_gained)
```

Simple gaussian model for pounds gained, with 'No' coded as 0

```
SA1.model1.lm <- lm(pounds_gained ~ CalcTMM, data=gained.dat)
summary(SA1.model1.lm)

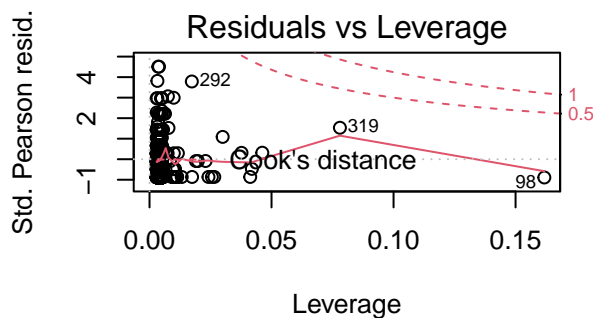
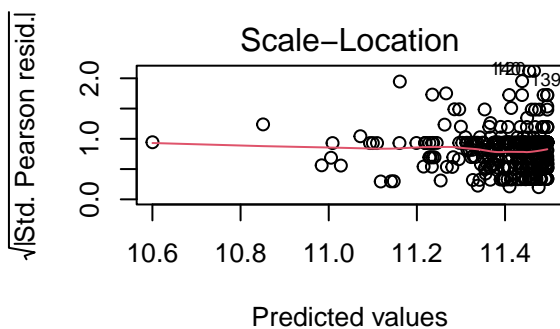
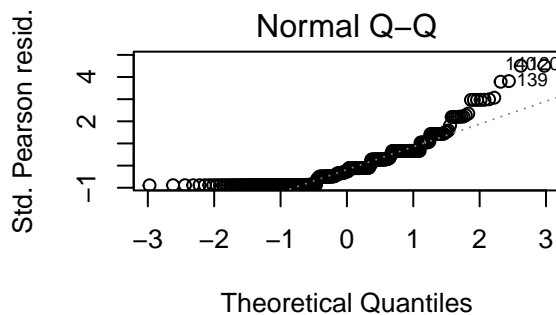
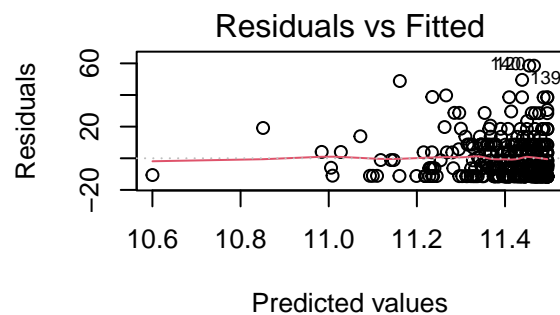
##
## Call:
## lm(formula = pounds_gained ~ CalcTMM, data = gained.dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.496 -11.348  -3.441   6.672  58.545
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.4958780  0.9241793  12.439  <2e-16 ***
## CalcTMM      -0.0000697  0.0004500  -0.155   0.877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 13.01 on 339 degrees of freedom
## (6 observations deleted due to missingness)
## Multiple R-squared:  7.076e-05, Adjusted R-squared:  -0.002879
## F-statistic: 0.02399 on 1 and 339 DF, p-value: 0.877

par(mfrow=c(2,2))
plot(SA1.model1.lm)
```



```
SA1.model1.gauss <- glm(pounds_gained ~ CalcTMM, data=gained.dat, family = gaussian)
summary(SA1.model1.gauss)
```

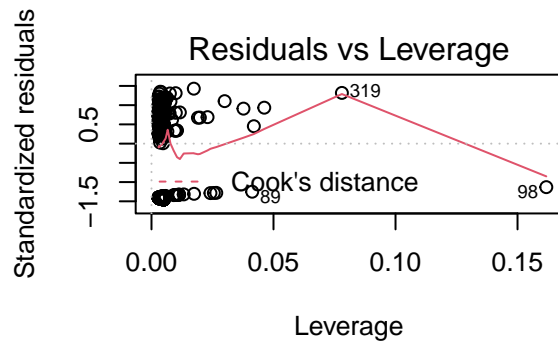
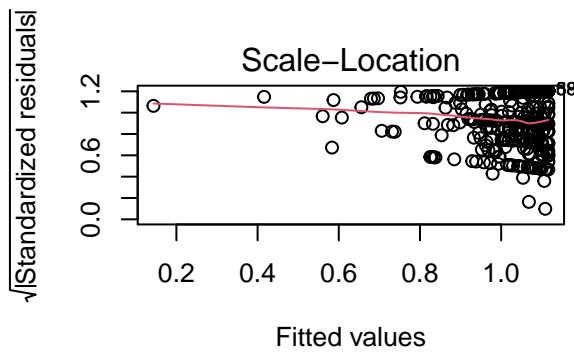
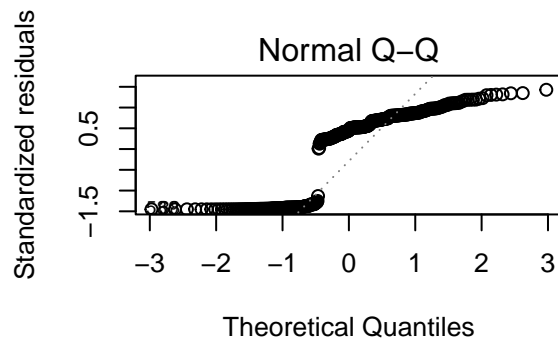
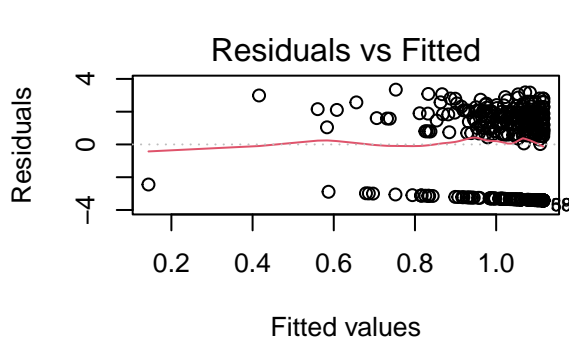
```
##
## Call:
## glm(formula = pounds_gained ~ CalcTMM, family = gaussian, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -11.496  -11.348   -3.441    6.672   58.545
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 11.4958780  0.9241793  12.439  <2e-16 ***
## CalcTMM      -0.0000697  0.0004500  -0.155    0.877
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 169.2323)
##
##      Null deviance: 57374  on 340  degrees of freedom
## Residual deviance: 57370  on 339  degrees of freedom
## (6 observations deleted due to missingness)
## AIC: 2721.5
##
## Number of Fisher Scoring iterations: 2
par(mfrow=c(2,2))
plot(SA1.model1.gauss)
```



```
#error in log(0)
SA1.model1.log <- lm(log(pounds_gained+0.1) ~ CalcTMM, data=gained.dat)
summary(SA1.model1.log)
```

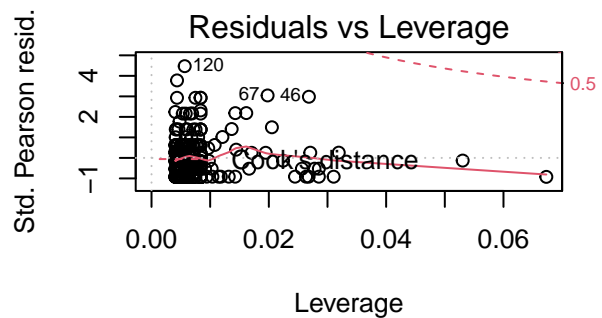
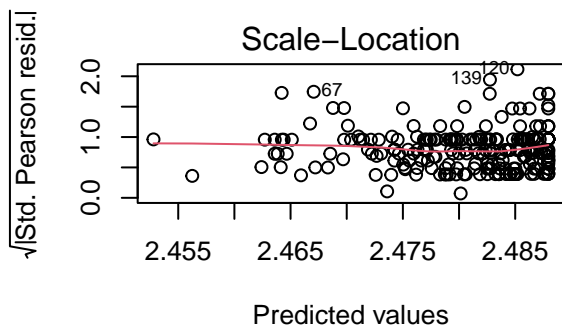
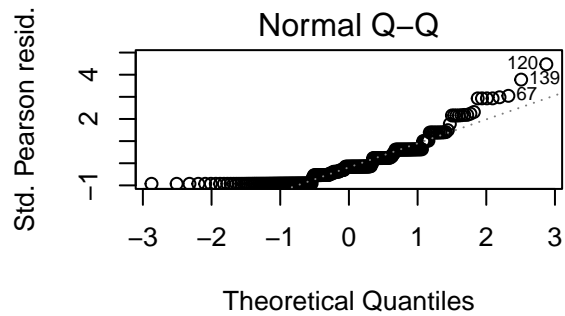
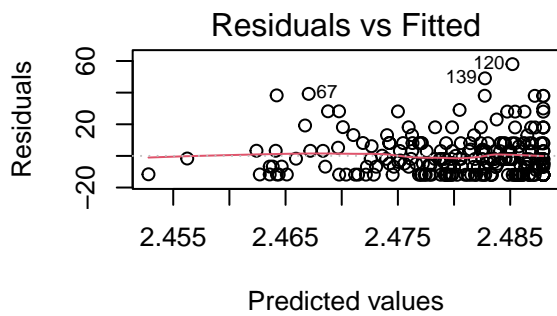
```
##
## Call:
## lm(formula = log(pounds_gained + 0.1) ~ CalcTMM, data = gained.dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.419 -3.259  1.046  1.892  3.343
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.116e+00  1.677e-01   6.654 1.15e-10 ***
## CalcTMM      -7.567e-05  8.167e-05  -0.927   0.355
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.361 on 339 degrees of freedom
## (6 observations deleted due to missingness)
## Multiple R-squared:  0.002526, Adjusted R-squared: -0.0004164
## F-statistic: 0.8585 on 1 and 339 DF, p-value: 0.3548
```

```
par(mfrow=c(2,2))
plot(SA1.model1.log)
```



#error in log(0)

```
SA1.model1.loggauss <- glm(pounds_gained+0.01 ~ Total_Met_Min, data=gained.dat, family = gaussian(link="log"))
par(mfrow=c(2,2))
plot(SA1.model1.loggauss)
```

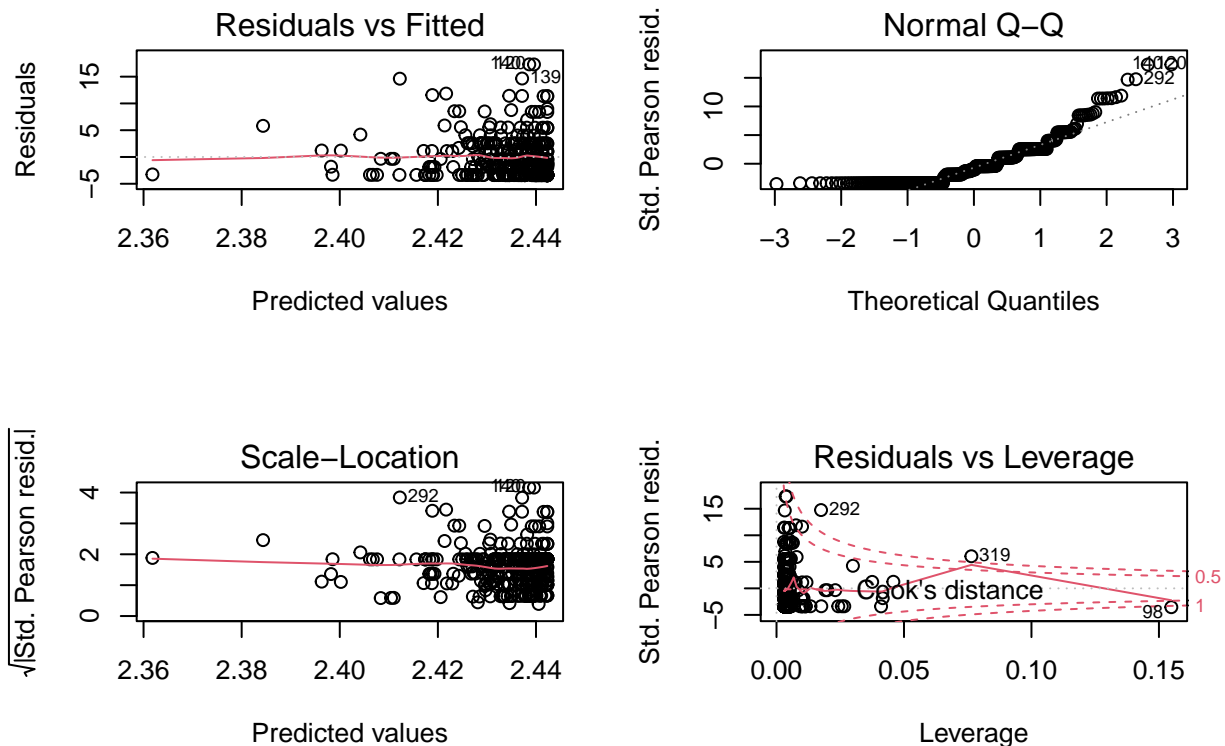


Poisson family

```
SA1.model1.LBS <- glm(LBS ~ CalcTMM, data=gained.dat, family = poisson)
summary(SA1.model1.LBS)
```

```
##
## Call:
## glm(formula = LBS ~ CalcTMM, family = poisson, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -4.796  -4.764  -1.077   1.824  11.675
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.442e+00  2.105e-02 116.024  <2e-16 ***
## CalcTMM      -6.262e-06  1.038e-05  -0.604    0.546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 4843.1  on 340  degrees of freedom
## Residual deviance: 4842.8  on 339  degrees of freedom
## (6 observations deleted due to missingness)
## AIC: 5872.4
##
## Number of Fisher Scoring iterations: 5
```

```
par(mfrow=c(2,2))
plot(SA1.model1.LBS)
```




```

library(pscl)

## Classes and Methods for R developed in the
## Political Science Computational Laboratory
## Department of Political Science
## Stanford University
## Simon Jackman
## hurdle and zeroinfl functions by Achim Zeileis

summary(zero.model1 <- zeroinfl(LBS ~ CalcTMM, data = gained.dat))

## Warning in sqrt(diag(object$vcov)): NaNs produced

##
## Call:
## zeroinfl(formula = LBS ~ CalcTMM, data = gained.dat)
##
## Pearson residuals:
##      Min      1Q  Median      3Q      Max
## -1.3941 -1.2980 -0.4077  0.7914  7.0383
##
## Count model coefficients (poisson with log link):
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  2.792e+00      NaN      NaN      NaN
## CalcTMM      2.426e-05      NaN      NaN      NaN
##
## Zero-inflation model coefficients (binomial with logit link):
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.509e-01  1.851e-01 -4.596 4.32e-06 ***
## CalcTMM      8.009e-05  1.241e-04  0.646  0.519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of iterations in BFGS optimization: 1
## Log-likelihood: -1631 on 4 Df

#SA1.model1.inv <- glm(LBS ~ Total_Met_Min, data=gained.dat,family = inverse.gaussian)
#summary(SA1.model1.inv)

SA1.model1.Gamma <- glm(LBS+0.01 ~ Total_Met_Min, data=gained.dat,family = Gamma)
summary(SA1.model1.Gamma)

##
## Call:
## glm(formula = LBS + 0.01 ~ Total_Met_Min, family = Gamma, data = gained.dat)
##
## Deviance Residuals:
##      Min      1Q  Median      3Q      Max
## -3.4913 -3.4866 -0.1781  0.5022  2.4768
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.304e-02  8.268e-03  10.043  <2e-16 ***
## Total_Met_Min 5.537e-07  5.323e-06  0.104   0.917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```



```
##
## (Dispersion parameter for Gamma family taken to be 1.184131)
##
## Null deviance: 992.41 on 246 degrees of freedom
## Residual deviance: 992.39 on 245 degrees of freedom
## (100 observations deleted due to missingness)
## AIC: 1418.8
##
## Number of Fisher Scoring iterations: 6

#SA1.model1.GammaId <- glm(LBS ~ Total_Met_Min, data=gained.dat,family = Gamma(link="identity"))
#summary(SA1.model1.GammaId)
#SA1.model1.GammaId <- glm(LBS ~ Total_Met_Min, data=gained.dat,family = Gamma(link="log"))
#summary(SA1.model1.GammaId)

par(mfrow=c(1,1))
hist(gained.dat$pounds_gained, freq = FALSE,ylim=c(0,max(dpois(0:70, mean(gained.dat$pounds_gained)))))

mean(gained.dat$pounds_gained)
sd(gained.dat$pounds_gained)
100*abs(sd(gained.dat$pounds_gained)-mean(gained.dat$pounds_gained))/mean(gained.dat$pounds_gained)
lines(0:70,dpois(0:70,mean(gained.dat$pounds_gained)),col='red')

nonzero <- gained.dat$pounds_gained[gained.dat$pounds_gained>2]

hist(nonzero, freq = FALSE,ylim=c(0,max(dpois(0:70, mean(nonzero)))))
lines(0:70,dpois(0:70,mean(nonzero)),col='red')
mean(nonzero)
sd(nonzero)
100*abs(sd(nonzero)-mean(nonzero))/mean(nonzero)

colors <- rep('red',dim(gained.dat)[1])
colors[gained.dat$pounds_gained==0] <- 'blue'
plot(pounds_gained~Total_Met_Min,gained.dat,col=colors)
abline(lm(pounds_gained~Total_Met_Min,gained.dat))
abline(lm(pounds_gained~Total_Met_Min,gained.dat[gained.dat$pounds_gained>0,]),col='red')

SA1.model1.quasi <- update(SA1.model1.LBS, family=quasipoisson)
summary(SA1.model1.quasi)

##
## Call:
## glm(formula = LBS ~ CalcTMM, family = quasipoisson, data = gained.dat)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -4.796 -4.764 -1.077 1.824 11.675
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.442e+00 8.110e-02 30.113 <2e-16 ***
## CalcTMM -6.262e-06 3.998e-05 -0.157 0.876
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for quasipoisson family taken to be 14.84488)
##
## Null deviance: 4843.1 on 340 degrees of freedom
## Residual deviance: 4842.8 on 339 degrees of freedom
## (6 observations deleted due to missingness)
## AIC: NA
##
## Number of Fisher Scoring iterations: 5
SA1.model2.gauss <- glm(pounds_gained ~ gender + Age + height + shift + CalcTMM, data=gained.dat,family
summary(SA1.model2.gauss)

##
## Call:
## glm(formula = pounds_gained ~ gender + Age + height + shift +
## CalcTMM, family = gaussian, data = gained.dat)
##
## Deviance Residuals:
## Min 1Q Median 3Q Max
## -17.704 -9.060 -2.504 5.747 55.559
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -4.7214382 20.4034341 -0.231 0.8172
## genderFemale 10.5196761 9.4719500 1.111 0.2676
## genderMale 8.6750343 9.4708836 0.916 0.3604
## Age -0.0366611 0.0757426 -0.484 0.6287
## height 0.1487949 0.2484185 0.599 0.5497
## shift8am 0.1892888 2.9469080 0.064 0.9488
## shift9am -2.8688532 3.2366154 -0.886 0.3761
## shift10am -2.9749370 3.3189699 -0.896 0.3708
## shift11am -6.0891531 3.3929321 -1.795 0.0737 .
## shift12pm -6.4094602 4.4514239 -1.440 0.1510
## shift1pm -9.2096214 5.2985354 -1.738 0.0832 .
## shift2pm 6.1186000 4.9404482 1.238 0.2165
## shifttother 1.0638677 4.4409485 0.240 0.8108
## shiftmissing 9.4630511 9.5155676 0.994 0.3208
## CalcTMM -0.0003234 0.0004686 -0.690 0.4906
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 166.5436)
##
## Null deviance: 52011 on 306 degrees of freedom
## Residual deviance: 48631 on 292 degrees of freedom
## (40 observations deleted due to missingness)
## AIC: 2458.2
##
## Number of Fisher Scoring iterations: 2
SA1.model3.gauss <- glm(pounds_gained ~ gender + Age + height + shift + Vig.ex.Time + Mod.ex.time + Wa
summary(SA1.model3.gauss)

##
```

```

## Call:
## glm(formula = pounds_gained ~ gender + Age + height + shift +
##     Vig.ex.Time + Mod.ex.time + Walk.ex.Time, family = gaussian,
##     data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -16.195   -8.917   -2.476    5.548   55.156
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.6484710 20.6125163  -0.080  0.9363
## genderFemale   8.2393567  9.7347609   0.846  0.3980
## genderMale    6.7219128  9.6917153   0.694  0.4885
## Age          -0.0449858  0.0762564  -0.590  0.5557
## height         0.1382322  0.2492902   0.555  0.5797
## shift8am       0.3453983  2.9615961   0.117  0.9072
## shift9am      -3.1369097  3.2541094  -0.964  0.3359
## shift10am     -2.8279949  3.3343971  -0.848  0.3971
## shift11am     -5.8159231  3.4134749  -1.704  0.0895 .
## shift12pm     -6.3172972  4.4592872  -1.417  0.1577
## shift1pm      -9.2493994  5.3061598  -1.743  0.0824 .
## shift2pm       6.6333157  5.0253488   1.320  0.1879
## shifttother    1.2332641  4.4496642   0.277  0.7819
## shiftmissing   9.6613770  9.5296220   1.014  0.3115
## Vig.ex.Time   -0.0088004  0.0072517  -1.214  0.2259
## Mod.ex.time   -0.0004069  0.0058075  -0.070  0.9442
## Walk.ex.Time   0.0021408  0.0036303   0.590  0.5559
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 166.9633)
##
##      Null deviance: 52011  on 306  degrees of freedom
## Residual deviance: 48419  on 290  degrees of freedom
## (40 observations deleted due to missingness)
## AIC: 2460.9
##
## Number of Fisher Scoring iterations: 2
SA1.model2.LBS <- glm(LBS ~ gender + Age + height+ initial_BMI + shift + CalcTMM, data=gained.dat,family=poisson)
summary(SA1.model2.LBS)

##
## Call:
## glm(formula = LBS ~ gender + Age + height + initial_BMI + shift +
##     CalcTMM, family = poisson, data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -5.8674  -3.9583  -0.6924   1.5600  10.8477
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   1.910e-01  6.391e-01   0.299  0.765106

```

```

## genderFemale  1.676e+00  4.510e-01   3.715 0.000203 ***
## genderMale    1.558e+00  4.513e-01   3.451 0.000558 ***
## Age           -4.806e-03  2.050e-03  -2.344 0.019057 *
## height        1.310e-02  6.203e-03   2.112 0.034671 *
## initial_BMI   2.589e-03  3.489e-03   0.742 0.458101
## shift8am      -2.530e-02  7.026e-02  -0.360 0.718748
## shift9am       2.463e-02  7.985e-02   0.308 0.757781
## shift10am     -1.617e-01  8.124e-02  -1.991 0.046526 *
## shift11am     -7.909e-01  9.733e-02  -8.127 4.42e-16 ***
## shift12pm     -4.657e-01  1.225e-01  -3.802 0.000144 ***
## shift1pm      -9.808e-01  1.936e-01  -5.067 4.04e-07 ***
## shift2pm       6.080e-01  1.076e-01   5.650 1.60e-08 ***
## shifttother    1.547e-01  9.868e-02   1.568 0.116928
## shiftmissing  5.609e-01  1.629e-01   3.443 0.000575 ***
## CalcTMM       -2.957e-05  1.147e-05  -2.577 0.009970 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##    Null deviance: 3529.0  on 241  degrees of freedom
## Residual deviance: 3229.5  on 226  degrees of freedom
##    (105 observations deleted due to missingness)
## AIC: 4000
##
## Number of Fisher Scoring iterations: 5
SA1.model3.LBS <- glm(LBS ~ gender + Age + height+ initial_BMI + shift + Vig.ex.Time + Mod.ex.time + W
summary(SA1.model3.LBS)

##
## Call:
## glm(formula = LBS ~ gender + Age + height + initial_BMI + shift +
##     Vig.ex.Time + Mod.ex.time + Walk.ex.Time, family = poisson,
##     data = gained.dat)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -5.8706  -3.8976  -0.6743   1.5308  10.7743
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  4.184e-01  6.463e-01   0.647 0.517388
## genderFemale  1.565e+00  4.545e-01   3.444 0.000574 ***
## genderMale    1.471e+00  4.547e-01   3.234 0.001219 **
## Age          -5.238e-03  2.056e-03  -2.547 0.010854 *
## height       1.148e-02  6.328e-03   1.814 0.069702 .
## initial_BMI  2.272e-03  3.501e-03   0.649 0.516344
## shift8am     -1.716e-02  7.065e-02  -0.243 0.808056
## shift9am      7.521e-03  8.020e-02   0.094 0.925282
## shift10am    -1.490e-01  8.156e-02  -1.827 0.067698 .
## shift11am    -7.752e-01  9.775e-02  -7.930 2.20e-15 ***
## shift12pm    -4.521e-01  1.227e-01  -3.686 0.000228 ***
## shift1pm     -9.657e-01  1.936e-01  -4.987 6.13e-07 ***
## shift2pm      6.561e-01  1.122e-01   5.850 4.91e-09 ***

```

```
## shiftother      1.616e-01  9.875e-02   1.636 0.101747
## shiftmissing    5.755e-01  1.631e-01   3.529 0.000417 ***
## Vig.ex.Time     -5.121e-04  1.859e-04  -2.754 0.005882 **
## Mod.ex.time     -2.149e-04  1.261e-04  -1.704 0.088442 .
## Walk.ex.Time    1.206e-04  8.244e-05   1.463 0.143572
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##      Null deviance: 3529.0  on 241  degrees of freedom
## Residual deviance: 3221.3  on 224  degrees of freedom
##      (105 observations deleted due to missingness)
## AIC: 3995.7
##
## Number of Fisher Scoring iterations: 6
best.model.quasi <- update(best.model.LBS,family=quasipoisson)

#par(mfrow=c(2,2))
#plot(best.model.quasi)
summary(best.model.quasi)
```

Zero-Inflated Poisson

```
# We can't use initial BMI, no 0 values
summary(zero.model2 <- zeroinfl(LBS ~ gender + + Age + + height + shift + CalcTMM, data = gained.dat))

## Warning in value[[3L]](cond): system is computationally singular: reciprocal
## condition number = 7.50822e-19FALSE
##
## Call:
## zeroinfl(formula = LBS ~ gender + +Age + +height + shift + CalcTMM, data = gained.dat)
##
## Pearson residuals:
##      Min      1Q  Median      3Q      Max
## -2.0144 -1.1409 -0.3485  0.6161  6.8557
##
## Count model coefficients (poisson with log link):
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.075e+00      NA      NA      NA
## genderFemale  1.351e+00      NA      NA      NA
## genderMale    1.409e+00      NA      NA      NA
## Age          -1.749e-03      NA      NA      NA
## height        8.416e-03      NA      NA      NA
## shift8am       4.330e-02      NA      NA      NA
## shift9am      -1.301e-01      NA      NA      NA
## shift10am     -3.223e-01      NA      NA      NA
## shift11am     -5.355e-01      NA      NA      NA
## shift12pm     -5.178e-01      NA      NA      NA
## shift1pm      -1.019e+00      NA      NA      NA
## shift2pm       1.189e-01      NA      NA      NA
```

```

## shiftother -1.322e-02 NA NA NA
## shiftmissing 1.562e-01 NA NA NA
## CalcTMM 1.195e-05 NA NA NA
##
## Zero-inflation model coefficients (binomial with logit link):
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 2.897e-01 NA NA NA
## genderFemale -1.256e+00 NA NA NA
## genderMale -6.889e-01 NA NA NA
## Age 1.621e-03 NA NA NA
## height -2.970e-03 NA NA NA
## shift8am 2.542e-02 NA NA NA
## shift9am 2.790e-01 NA NA NA
## shift10am -2.922e-01 NA NA NA
## shift11am 1.153e-01 NA NA NA
## shift12pm 3.131e-01 NA NA NA
## shift1pm 3.379e-01 NA NA NA
## shift2pm -1.431e+00 NA NA NA
## shiftother -5.515e-01 NA NA NA
## shiftmissing -1.385e+01 NA NA NA
## CalcTMM 9.579e-05 NA NA NA
##
## Number of iterations in BFGS optimization: 5
## Log-likelihood: -1370 on 30 Df
summary(zero.model3 <- zeroinfl(LBS ~ gender + Age + height + shift + Vig.ex.Time + Mod.ex.time + Walk.
##
## Call:
## zeroinfl(formula = LBS ~ gender + Age + height + shift + Vig.ex.Time +
## Mod.ex.time + Walk.ex.Time, data = gained.dat)
##
## Pearson residuals:
## Min 1Q Median 3Q Max
## -1.9982 -1.1472 -0.3542 0.6322 6.6198
##
## Count model coefficients (poisson with log link):
## Estimate Std. Error z value Pr(>|z|)
## (Intercept) 1.364e+00 6.210e-01 2.196 0.02812 *
## genderFemale 1.174e+00 4.719e-01 2.488 0.01286 *
## genderMale 1.264e+00 4.749e-01 2.662 0.00777 **
## Age -2.328e-03 1.704e-03 -1.366 0.17202
## height 7.091e-03 6.103e-03 1.162 0.24523
## shift8am 3.516e-02 6.364e-02 0.553 0.58059
## shift9am -1.812e-01 7.463e-02 -2.428 0.01518 *
## shift10am -3.270e-01 7.417e-02 -4.409 1.04e-05 ***
## shift11am -5.329e-01 8.265e-02 -6.449 1.13e-10 ***
## shift12pm -5.346e-01 1.198e-01 -4.463 8.10e-06 ***
## shift1pm -1.016e+00 1.790e-01 -5.677 1.37e-08 ***
## shift2pm 1.535e-01 9.674e-02 1.587 0.11250
## shiftother -1.222e-02 9.393e-02 -0.130 0.89651
## shiftmissing 1.538e-01 1.602e-01 0.960 0.33711
## Vig.ex.Time -2.916e-04 1.791e-04 -1.628 0.10348
## Mod.ex.time -3.933e-05 1.018e-04 -0.386 0.69915
## Walk.ex.Time 2.863e-04 6.931e-05 4.130 3.62e-05 ***

```

```
##
## Zero-inflation model coefficients (binomial with logit link):
##      Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.190e-01  3.427e+00  0.035  0.972
## genderFemale -1.077e+00  1.565e+00 -0.688  0.492
## genderMale   -5.136e-01  1.558e+00 -0.330  0.742
## Age          1.856e-03  1.302e-02  0.142  0.887
## height       -2.894e-03  4.186e-02 -0.069  0.945
## shift8am      -2.045e-03  4.942e-01 -0.004  0.997
## shift9am      2.764e-01  5.379e-01  0.514  0.607
## shift10am     -3.218e-01  5.727e-01 -0.562  0.574
## shift11am      7.960e-02  5.677e-01  0.140  0.888
## shift12pm      3.149e-01  7.252e-01  0.434  0.664
## shift1pm       3.309e-01  8.635e-01  0.383  0.702
## shift2pm      -1.419e+00  1.166e+00 -1.216  0.224
## shifttother   -5.647e-01  8.249e-01 -0.685  0.494
## shiftmissing -1.387e+01  1.008e+03 -0.014  0.989
## Vig.ex.Time   1.281e-03  1.259e-03  1.017  0.309
## Mod.ex.time   -1.215e-04  1.175e-03 -0.103  0.918
## Walk.ex.Time  2.873e-04  6.166e-04  0.466  0.641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Number of iterations in BFGS optimization: 1
## Log-likelihood: -1364 on 34 Df
```

Bootstrap the coefficients

```
boots.zero.model1.tbl$Estimate <- coef(zero.model1)
boots.zero.model1.tbl

confint(zero.model1)
```