

HW 6

1. Model with *total* as the response and *takers*, *ratio*, *salary*, and *expend* as predictors from dataset *sat*.

a. Ordinary Least Squares, in this model **takers** seems to be the only significant (alpha=5%) predictor having a negative coefficient. All other predictors are insignificant with *ratio* having a negative coefficient and *salary* & *expend* having positive coefficients. The intercept is quite large and positive. The R^2 (0.82) is high & the RSE (32) is fairly low.

```
##
## Call:
## lm(formula = total ~ takers + ratio + salary + expend, data = sat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -90.531 -20.855  -1.746  15.979  66.571
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1045.9715    52.8698   19.784 < 2e-16 ***
## takers       -2.9045     0.2313  -12.559 2.61e-16 ***
## ratio        -3.6242     3.2154   -1.127  0.266
## salary        1.6379     2.3872    0.686  0.496
## expend        4.4626    10.5465    0.423  0.674
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 32.7 on 45 degrees of freedom
## Multiple R-squared:  0.8246, Adjusted R-squared:  0.809
## F-statistic: 52.88 on 4 and 45 DF,  p-value: < 2.2e-16
```

b. Least Absolute Deviations, in this model **takers** & **ratio** seem to be both significant (alpha=5%) predictors having both negative coefficients. All other predictors are insignificant with *expend* having a negative coefficient and *salary* having a positive coefficient. The intercept is quite large and positive. The SE for *expend* is quite high resonating its insignificance.

```
##
## Call: rq(formula = total ~ takers + ratio + salary + expend, data = sat)
##
## tau: [1] 0.5
##
## Coefficients:
##              Value      Std. Error t value    Pr(>|t|)
## (Intercept) 1090.89886    58.48207   18.65356  0.00000
## takers       -3.13961     0.26233  -11.96841  0.00000
## ratio        -7.26632     3.27271   -2.22028  0.03148
## salary        3.18313     2.05291    1.55054  0.12802
## expend       -0.79753     9.10816   -0.08756  0.93061
```

c. Huber's Robust Regression, in this model **takers** seems to be the only significant (alpha=5%) predictor having a negative coefficient. All other predictors are insignificant with *ratio* having a negative coefficient and *salary* & *expend* having positive coefficients. The intercept is quite large and positive. The RSE (25) is fairly low.

```
##
## Call: rlm(formula = total ~ takers + ratio + salary + expend, data = sat)
## Residuals:
##      Min       1Q   Median       3Q      Max
## -92.510 -17.701  -1.002  15.015  77.058
##
## Coefficients:
##              Value      Std. Error t value
## (Intercept) 1060.2074    49.8845    21.2533
## takers       -2.9778     0.2182   -13.6470
## ratio        -5.1254     3.0339    -1.6894
## salary        2.0933     2.2525     0.9293
## expend       3.9158     9.9510     0.3935
##
## Residual standard error: 25.58 on 45 degrees of freedom

##
## robust F-test (as if non-random weights)
##
## data:  from rlm(formula = total ~ takers + ratio + salary + expend, data = sat)
## F = 187.07, p-value < 2.2e-16
## alternative hypothesis: true takers is not equal to 0

##
## robust F-test (as if non-random weights)
##
## data:  from rlm(formula = total ~ takers + ratio + salary + expend, data = sat)
## F = 2.698, p-value = 0.1074
## alternative hypothesis: true ratio is not equal to 0

##
## robust F-test (as if non-random weights)
##
## data:  from rlm(formula = total ~ takers + ratio + salary + expend, data = sat)
## F = 0.88587, p-value = 0.3516
## alternative hypothesis: true salary is not equal to 0

##
## robust F-test (as if non-random weights)
##
## data:  from rlm(formula = total ~ takers + ratio + salary + expend, data = sat)
## F = 0.15966, p-value = 0.6914
## alternative hypothesis: true expend is not equal to 0
```

Model Comparison

Ordinary least squares and **Huber's Robust Regression** are closer to each other relative in their results. Both models have *takers* as the only significant predictor; both models share the same coefficients *signs* (*takers* (-), *ratio*(-), *salary*(+), *expand*(+)); both models have similar *intercept* (high & positive); both models share similar *RSE* (32 & 25); both models have similar *standard errors*. The models differ slightly in the *estimates of the coefficients*. The models p-values for **ratio** differ greatly- 26% vs 10%.

OLS/HRR and **Least Absolute Deviations** are a bit further apart from each other relative in their results. Both models have *takers* as the a significant predictor-but LAD considers *ratio* as a significant

predictor too; both models share the same coefficients *signs* (takers (-), ratio(-), salary(+))-but LAD considers *expend* negative in its sign (ols/HRR +); both models have similar *intercept* (high & positive); both models have similar *standard errors*. The models differ greatly in the *estimates of the coefficients*. The models p-values for **salary** differ greatly- on one side OLS/HRR (49%,35%) and on the other LAD with 12%.