Stat 6021: Guided Question Set 7

Tom Lever

10/16/22

Car drivers like to adjust the seat position for their own comfort. Car designers find it helpful to know where different drivers will position the seat. Researchers at HuMoSim laboratory at the University of Michigan collected data on 38 drivers. The response variable is *hipcenter*, the horizontal distance of the midpoint of the hips from a fixed location in the car in *mm*. They measured the following eight predictors:

- x_1: Age: Age in years
- x_2: Weight: Weight in pounds
- x_3: HtShoes: Height with shoes in *cm*.
- x_4: Ht: Height without shoes in *cm*.
- x_5: Seated: Seated height in *cm*.
- x_6: Arm: Arm length in *cm*.
- x_7: Thigh: Thigh length in *cm*.
- x_8: Leg: Lower leg length in *cm*.

The data are from the `faraway` package in R. After installing the faraway package, load the `seatpos` data set.

```
library(faraway)
head(seatpos, n = 3)
```

```
##   Age Weight HtShoes    Ht Seated  Arm Thigh  Leg hipcenter
## 1  46    180   187.2 184.9   95.2 36.1  45.3 41.3  -206.300
## 2  31    175   167.5 165.5   83.8 32.9  36.5 35.9  -178.210
## 3  23    100   153.6 152.2   82.9 26.0  36.6 31.0   -71.673
```

1. Fit the full model with all the predictors. Using the `summary` function, comment on the results of the `t` tests and ANOVA `F` test from the output.

   ```
   library(TomLeversRPackage)
   linear_model <- lm(hipcenter ~ ., data = seatpos)
   summarize_linear_model(linear_model)
   ```

   ```
   ##
   ## Call:
   ## lm(formula = hipcenter ~ ., data = seatpos)
   ##
   ## Residuals:
   ##     Min      1Q  Median      3Q     Max
   ## -73.827 -22.833  -3.678  25.017  62.337
   ##
   ## Coefficients:
   ##              Estimate Std. Error t value Pr(>|t|)
   ## (Intercept) 436.43213  166.57162   2.620   0.0138 *
   ## Age           0.77572    0.57033   1.360   0.1843
   ## Weight        0.02631    0.33097   0.080   0.9372
   ```

```
## HtShoes      -2.69241    9.75304  -0.276   0.7845
## Ht            0.60134   10.12987   0.059   0.9531
## Seated        0.53375    3.76189   0.142   0.8882
## Arm          -1.32807    3.90020  -0.341   0.7359
## Thigh        -1.14312    2.66002  -0.430   0.6706
## Leg          -6.43905    4.71386  -1.366   0.1824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.72 on 29 degrees of freedom
## Multiple R-squared:  0.6866, Adjusted R-squared:  0.6001
## F-statistic:  7.94 on 8 and 29 DF,  p-value: 1.306e-05
##
## E(y | x) =
##     B_0 +
##     B_Age * Age +
##     B_Weight * Weight +
##     B_HtShoes * HtShoes +
##     B_Ht * Ht +
##     B_Seated * Seated +
##     B_Arm * Arm +
##     B_Thigh * Thigh +
##     B_Leg * Leg
## E(y | x) =
##     436.43212822533 +
##     0.775716195411176 * Age +
##     0.0263130815934825 * Weight +
##     -2.69240773927674 * HtShoes +
##     0.601344580352112 * Ht +
##     0.533751697568726 * Seated +
##     -1.32806863757197 * Arm +
##     -1.14311887823954 * Thigh +
##     -6.43904626562725 * Leg
## Number of observations: 38
## Estimated variance of errors: 1422.82012070282
## Multiple R:  0.828585225565444   Adjusted R:  0.774651837544867
## Critical value t(alpha/2 = 0.05/2, DFRes = 29): 2.0452296421327
## Critical value F(alpha = 0.05, DFR = 8, DFRes = 29): 2.27825084905155
```

Since the above F statistic is greater than the above critical value F / the above p-value for an ANOVA F test is less than a significance level $\alpha = 0.05$, we reject a null hypothesis that all regression coefficients in our linear model are 0 / insignificant. We have sufficient evidence to support an alternate hypothesis that at least one regression coefficient in our linear model is not 0 / significant.

Since the magnitudes of the above t statistics for all predictors are less than the above critical value t, we have insufficient evidence to reject a null hypothesis that each regression coefficient is 0 / insignificant in the context of the multiple linear model and the other predictors. We have insufficient evidence to support an alternate hypothesis that each regression coefficient is not 0 / significant in the context of the multiple linear model and the other predictors.

2. Briefly explain why, based on your output from part 1, you suspect the model shows signs of multicollinearity.

   Per section 9.4.4: Multicollinearity: Multicollinearity Diagnostics: Other Diagnostics in *Introduction to Linear Regression Analysis* (Sixth Edition) by Douglas C. Montgomery et al., "if the overall F statistic is significant but the individual t statistics are all nonsignificant, multicollinearity is present".

Multicollinearity is present.

3. Provide the output of all the pairwise correlations among the predictors. Comment briefly on the pairwise correlations.

```
library(dplyr)
data_frame_of_predictor_values <- seatpos %>% select(-hipcenter)
correlation_matrix <- round(cor(data_frame_of_predictor_values), 3)
correlation_matrix
```

```
##             Age Weight HtShoes     Ht Seated   Arm Thigh    Leg
## Age       1.000  0.081  -0.079 -0.090 -0.170 0.360 0.091 -0.042
## Weight    0.081  1.000   0.828  0.829  0.776 0.698 0.573  0.784
## HtShoes -0.079  0.828   1.000  0.998  0.930 0.752 0.725  0.908
## Ht      -0.090  0.829   0.998  1.000  0.928 0.752 0.735  0.910
## Seated  -0.170  0.776   0.930  0.928  1.000 0.625 0.607  0.812
## Arm      0.360  0.698   0.752  0.752  0.625 1.000 0.671  0.754
## Thigh    0.091  0.573   0.725  0.735  0.607 0.671 1.000  0.650
## Leg     -0.042  0.784   0.908  0.910  0.812 0.754 0.650  1.000
```

```
analyze_correlation_matrix(correlation_matrix)
```

```
## Age
##     V:  Age
##     H:
##     M:
##     L:  Arm
##     N:  Weight, HtShoes, Ht, Seated, Thigh, Leg
## Weight
##     V:  Weight
##     H:  HtShoes, Ht, Seated, Leg
##     M:  Arm, Thigh
##     L:
##     N:  Age
## HtShoes
##     V:  HtShoes, Ht, Seated, Leg
##     H:  Weight, Arm, Thigh
##     M:
##     L:
##     N:  Age
## Ht
##     V:  HtShoes, Ht, Seated, Leg
##     H:  Weight, Arm, Thigh
##     M:
##     L:
##     N:  Age
## Seated
##     V:  HtShoes, Ht, Seated
##     H:  Weight, Leg
##     M:  Arm, Thigh
##     L:
##     N:  Age
## Arm
##     V:  Arm
##     H:  HtShoes, Ht, Leg
##     M:  Weight, Seated, Thigh
```

```
##     L:  Age
##     N:
## Thigh
##     V:  Thigh
##     H:  HtShoes, Ht
##     M:  Weight, Seated, Arm, Leg
##     L:
##     N:  Age
## Leg
##     V:  HtShoes, Ht, Leg
##     H:  Weight, Seated, Arm
##     M:  Thigh
##     L:
##     N:  Age
```