Exam 2 - Question 6

Adam McQuistan

Tuesday, April 05, 2016

Problem 6 - Do problem 7.19 on page 291

- SPSS does part (a) and do not do extra computation to standardize the data
- Do part (c) unless you trust SPSS

Part A. Transform the variables by means of correlation transformation and fit the regression model.

*** Since I am using R I will transform with R ***

```
correlatesSd <- function(x){
  x_bar <- mean(x)
  x_sqr_diffs <- (x - x_bar)^2
  x_corr_sd <- sqrt(sum(x_sqr_diffs)/(length(x)-1))
  return(x_corr_sd)
}


correlatesTransform <- function(x, corrSd){
  corrTrans <- (1/sqrt(length(x)-1)) * ((x - mean(x))/corrSd)
  return(corrTrans)
}

setwd("C:\\Users\\AdamMcQuistan\\Documents\\ISQA 8340\\Exam 2")
df <- read.csv("data/6.18.csv")
names(df)[2:6] = c("Y", "X1", "X2", "X3", "X4")

sd_y <- correlatesSd(df$Y)
df$Y_trans <- correlatesTransform(df$Y, sd_y)

sd_x1 <- correlatesSd(df$X1)
df$X1_trans <- correlatesTransform(df$X1, sd_x1)

sd_x2 <- correlatesSd(df$X2)
df$X2_trans <- correlatesTransform(df$X2, sd_x2)

sd_x3 <- correlatesSd(df$X3)
df$X3_trans <- correlatesTransform(df$X3, sd_x3)

sd_x4 <- correlatesSd(df$X4)
df$X4_trans <- correlatesTransform(df$X4, sd_x4)

result <- lm(Y_trans ~ X1_trans + X2_trans + X3_trans + X4_trans, data=df)
summary(result)
```

```
##
## Call:
## lm(formula = Y_trans ~ X1_trans + X2_trans + X3_trans + X4_trans,
##     data = df)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.207223 -0.038429 -0.005914  0.036276  0.191422
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -5.161e-17  8.213e-03   0.000     1.00
## X1_trans    -5.479e-01  8.232e-02  -6.655 3.89e-09 ***
## X2_trans     4.236e-01  9.490e-02   4.464 2.75e-05 ***
## X3_trans     4.846e-02  8.504e-02   0.570     0.57
## X4_trans     5.028e-01  8.786e-02   5.722 1.98e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07392 on 76 degrees of freedom
## Multiple R-squared:  0.5847, Adjusted R-squared:  0.5629
## F-statistic: 26.76 on 4 and 76 DF,  p-value: 7.272e-14
```

```
anova(result)
```

```
## Analysis of Variance Table
##
## Response: Y_trans
##           Df  Sum Sq  Mean Sq F value    Pr(>F)
## X1_trans   1 0.06264 0.062642 11.4649  0.001125 **
## X2_trans   1 0.30776 0.307756 56.3262 9.699e-11 ***
## X3_trans   1 0.03543 0.035431  6.4846  0.012904 *
## X4_trans   1 0.17892 0.178920 32.7464 1.976e-07 ***
## Residuals 76 0.41525 0.005464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Part B. Interpret standardized regression coefficients for b2.

$H_0 : \beta_2 = 0$

$H_a : \beta_2 \neq 0$

Since the pvalue for for X2 (β_2) is less than 0.05 we reject the null hypthosis.