PROBLEM 1 Q2

The resulting estimated functions are:

* y1(x) = x152.158 + -189.866
* y2(x) = x27.001 + x19.304 + -239.334
* y3(x) = x30.820+ x20.261+ -x10.0103+ - 175.277
* y4(x) = x40.005987 + x30.755 + x20.234 + x11.176 + -175.880
* y5(x) = x5 + x4 + x3 + x2 + x1 + x50.000853 + -x40.004698 + x30.7528 + x20.5260 + x10.9659 + -176.837

﻿[[52.158053801747215, -189.86610574097074], [7.001583333198915, 9.303864260428877, -239.3340329835961], [0.8201380988526162, 0.26176712315847694, -0.010327670472122463, -175.27713200891657], [0.005987963794383017, 0.7552180459273519, 0.23455985383724576, 1.1763636032250524, -175.88028826172473], [0.0008531198983952847, -0.004698036593498531, 0.7528113266207407, 0.5260849950247835, 0.9659162450400247, -176.83736889212605]]

PROBLEM 1 Q2

Sample size = 937

Sample mean = 0.743030411045

Standard error = 0.00415302728827

Standard cost = -1.6781948375

P value = 0.0933090692524

Is result at significance level 0.1? Yes

Is result at significance level 0.05? No

Based on the results gathered, we can conclude the hypothesis is:

significant at a = 0.1 (reject Ho),

not significant at a = 0.05 (cannot reject Ho)

PROBLEM 1 Q3

Largest standard error which the test will be significant at 0.05 is 0.00355597807416

minimum sample size is 1278.05931911 (1278)

PROBLEM 1 Q4

Null Hypothesis: mean engagement of students who become

knowledgeable in the material is the same with those who do not

Alternative Hypothesis: mean engagement of students who become

knowledgeable in the material is NOT the same with those who do not

Type of test can be used: two-sample-z-test

PROBLEM 1 Q5

This is the result for sample from eng1.txt:

Sample size = 1977

Sample mean = 0.639954507704

Standard error = 0.00571598958877

Standard cost = -19.2522205626

P value = 1.35240102861e-82

This is for sample from eng0.txt

Sample size: 937

Sample mean: 0.743030411045

Standard error: 0.00415302728827

Standard cost: -1.6781948375

P value: 0.0933090692524

Z-score: 14.58878454

P value for two-z-sample-test: 3.31043071683e-48

Based on the results gathered, we can conclude the hypothesis is significant at a = 0.1,0.01,0.05 (reject Ho), which means that the mean for samples in eng1.txt is different from the mean for samples in eng0.txt.

PROBLEM 2 Q1

I am using t-test because we don't know standard deviation and n < 30

Mean = 7.36363636364

Standard error = 5.07627767575

Standard statistic/t-value = 2.22813885196

Confidence interval for 95% is (3.9533466179911163, 10.773926109281611)

PROBLEM 2 Q2

I am using t-test because we don't know standard deviation and n < 30

Standard statistic/t-value: 1.81246112281

Confidence interval for 90% is (4.5895643725433333, 10.137708354729394)

The t-value is lower for 90% confidence interval and the confidence interval is also smaller.

PROBLEM 2 Q3

I am using z-test since we know the population standard deviation.

Mean: 7.36363636364

Standard error: 5.07624499731

Standard statistic/z-value: 1.95996398454

Confidence interval for 95% is (4.3638223960257854, 10.363450331246941)

The standard statistic and confidence interval are all different from the results in Problem 2 Q1 and Problem2 Q2.

PROBLEM 2 Q4

I can say with 82.24% confidence that the team is expected to win on average. This is based on the results calculated in the code:

T-value: 1.4505976296

p-value: 0.177524787237

Confidence value: 0.822475212763

Confidence interval: (5.143416462654649, 9.5838562646180776)