**IE208 Applied Statistics**

**Fall 2021**

**Final Exam (4:20 – 5:45 pm)**

**Name (한국어): 이인석 KHU ID (학번): 2018110659**

**ALLOWED:**

* Lecture materials provided by the instructor.
* Other lecture materials prepared by you.

**NOT ALLOWED:**

* All other materials, including any online resources.
* Sharing of materials.
* Any kind of communications: no messaging or emails.

**NOTE:**

* Provide your answers on this Word file for 5 questions and submit it to the e-campus.
* Consider just three digits after the decimal point: only up to three decimal places (ex: 0.123).
* Use the chatting box in the private mode for questions – no verbal questions.
* You cannot leave the Zoom meeting by 5:45 pm.

1. **(20 pts)** An experiment to compare *k* = 3 factor levels has = 17 and = 32.30, = 20 and = 34.06, and = 18 and = 32.02. Also, = 59843.21.
2. Populate the analysis of variance table below, where you need to consider only grey cells to fill in.

Overall mean = (32.3\*17+34.06\*20+32.02\*18)/(17+20+18) = 32.8484

SSTr = (5.112+29.361+12.351)

SSR = 59843.21 – (17\*32.3^2 + 20\*34.06^2+18\*32.02^2)

MSTr = SSTr/k-1 = 46.825/2

MSE = SSE/nt-k = 450.561/52

F-statistic = MSTr/MSE = 23.412/8.665

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Degree of Freedom** | **Sum of Squares** | **Mean Squares** | ***F*-Statistic** | **P-Value** |
| **Treatments** | 3-1=2 | 46.825 | 23.412 | 2.702 | 0.0765 |
| **Errors** | 55-3=52 | 450.561 | 8.665 | (empty) | |
| **Total** | 2+52=54 | 497.385 | (empty) | | |

1. Show H0 for the F test above.

Null hypothesis : mu1=mu2=mu3

1. Show the test result of the F test with = 0.1.

If alpha = 0.1, reject H0. Because alpha > p-value

1. **(20 pts)** You are given the data in the form of the table below. Please answer the following questions.

|  |  |  |  |
| --- | --- | --- | --- |
| **A** | **B** | **C** | **D** |
| 154.4 | 184.3 | 161.2 | 185.8 |
| 145 | 142.1 | 110.8 | 135.4 |
| 152.5 | 149.6 | 168.6 | 154.1 |
| 138.5 | 137.2 | 124.9 | 123.2 |
| 161.7 | 160.4 | 166.1 | 169.9 |

1. Fill in the ANOVA table below with the randomized block design. Each column is for each treatment, and each row is for blocks.You only need to populate grey cells on the ANOVA table.

SSTr = 3.741 + 58.996 + 123.256 + 28.68

SSE = 9.7 + 1089.99 +…. + 22887.15

SSBI = 7.259 + 29.03 +… + 37.925

MSTr = 214.674/3

MSBi = 96913.2/4

MSE = 7320.906/12

F\_t = MStr/MSE

F\_b = MSBi/MSE

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Degree of Freedom** | **Sum of Squares** | **Mean Squares** | ***F*-statistic** | **P-Value** |
| **Treatments** | 4-1=3 | 214.674 | 71.558 | 0.117 | 0.948 |
| **Blocks** | 5-1=4 | 96913.2 | 24228.3 | 39.714 | 7.89701E-07 |
| **Error** | 3\*4=12 | 7320.906 | 610.075 | (empty) | |
| **Total** | 3+4+12=19 | 104448.777 | (empty) | | |

1. Show H0 for the F-test of Treatments.

Null hypothesis: alpha1=alpha2=alpha3=alpha4=0

1. Show the F-test result of Treatments with = 0.05.

If alpha = 0.05, accept H0. They are equal.

1. **(20 pts)** Answer the following questions with the data in the table below:

|  |  |
| --- | --- |
| **x** | **y** |
| 37 | 65 |
| 36.4 | 67.2 |
| 35.8 | 70.3 |
| 34.3 | 71.9 |
| 33.7 | 73.8 |
| 32.1 | 75.7 |
| 31.5 | 77.9 |

1. Build a regression model for explaining y as x and find and .

Beta\_1 = -2.119, Beta\_0 = 144.588

y = 144.588 - 2.119\*x

1. Calculate the 95% confidence interval for .

Sample error = 0.1944, alpha = 0.05, t\_0.025,5=2.571

Lower = -2.619

Upper = -1.620

1. Provide H0 for the two-sided t-test of (with *b1* = 0).

Null hypothesis: beta\_1 = 0

1. Conduct the two-sided t-test for (with = 0.05 and *b1* = 0) and provide the result with the p-value.

t-statistic = (beta\_1-0)/s.e = -2.119/0.1944 = -10.902

p-value = 2, so, accept H0.

1. **(20 pts)** Assume that you were given some data with building heights and stories. When you built a regression model with Heights (Y) and Stories (X), you have obtained the ANOVA table below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ANOVA** |  |  |  |  |  |
|  | ***df*** | ***SS*** | ***MS*** | ***F*** | ***P-value*** |
| **Regression** | 1 | 545452.2 | 545452.2 | 174.7843 | 4.96E-11 |
| **Residual** | 19 | 59293.6 | 3120.716 |  |  |
| **Total** | 20 | ? |  |  |  |

1. What is SS of Total in the table above?

= 545452.2 + 59293.6 = 604745.8

1. What is R2 value of the table above?

R-squared = SSR/SST = 0.902

1. Assume that you made another regression model with Heights (X) and Stories (Y). If you obtained 312.248 for SS of Residual in the new regression model, calculate SS of Regression in the new regression model.
2. **(20 pts)** The two tables below are the results of a multiple regression model with x1, x2, x3, x4, and y. Answer the following questions with the two tables below.

|  |
| --- |
| **Coefficients** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Term** | **Coef** | **SE Coef** | **95% CI** | **T-Value** | **P-Value** |
| Constant | -0.7560 | 0.7361 | (-2.2665, -0.1544) | -0.03 | 0.0135 |
| x1 | 0.15453 | 0.06334 | (0.02457, 0.28448) | 2.44 | 0.0215 |
| x2 | 0.21705 | 0.03164 | (0.15214, 0.28197) | 6.86 | <0.0001 |
| x3 | 0.010806 | 0.004622 | (0.001323, 0.020290) | 2.34 | 0.0270 |
| x4 | 0.09464 | 0.05455 | (-0.01729, 0.20657) | 1.73 | 0.0942 |

|  |  |
| --- | --- |
| |  | | --- | | **Analysis of Variance** | |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Source** | **DF** | **SS** | **MS** | **F-Value** | **P-Value** | | Regression | 4 | 47.9096 | 11.9774 | 18.17 | <0.0001 | | Error | 27 | 17.7953 | 0.6591 |  |  | | Total | 31 | 65.7049 |  |  |  | |

a) What is H0 of the F test?

Null hypothesis : Beta\_1 = beta\_2 = beta\_3 = beta\_4 = 0

b) What is the result of the F test with = 0.05?

c) If you are doing a model fitting with the result above, explain which variable you want to drop in the next model with = 0.05.