IT3030 Biostatistics quiz #4 6/9/2020 ID:_____ Name:_____

1. (50%) The following table describes the relationship between two exam scores for 10 students in Biostatistics. Answer the following questions.

ID	Exam1	Exam2
1	46	68
2	59	57
3	75	68
4	84	91
5	82	76
6	79	84
7	57	53
8	75	84
9	71	80
10	61	70

(a) Compute the score difference defined by exam2 score minus exam1 score for each of the 10 students.

(b) Compute the mean value as well as the standard deviation of these 10 pairwise score differences.

(c) We want to see if students are making progress in exam scores. A null hypothesis is assumed that exam2 score is no different than exam1. Is this a 1-sided or 2-sided test? Why?

(d) Compute the t value for this test.

(e) Compute the p-value of this test (with α=0.05). Do you reject the null hypothesis or not? Does this mean the students are making progress or not?

Answer:

>> exam1 = [46 59 75 84 82 79 57 75 71 61];

>> exam2 = [68 57 68 91 76 84 53 84 80 70];

>> d=exam2-exam1 = 22 -2 -7 7 -6 5 -4 9 9 9

>> mean = mean(d) = 4.2000

>> std = std(d) = 9.0037

This is a 1-sided test since we only want to see the exam score gets higher.

>> t=mean/(std/sqrt(10)) = 1.4751

>> p_value=1-tcdf(t,10-1) = 0.0871

p-value is greater than 0.05, indicating that we do not reject the null hypothesis. This means we cannot conclude students are making progress.

2. (50%) Measurements of WBC (white- blood-cell count) for patients of 3 different diseases are shown below. Assume that WBC is normally distributed.

(a) Calculate the estimate of the within-groups variance S_W^2.

(b) Calculate the estimate of the between-group variance S_B^2.

(c) At the 0.05 level of significance, test the null hypothesis that the mean WBC of the 3 groups are identical. Compute and show the F-statistic, the two degree of freedoms used, and the p-value. What do you conclude?

(d) If you found that the population means are not identical, use the Bonferrino correction to determine where the differences lie (1 vs 2, 1 vs 3 and 2 vs 3). What is the significance level used in each individual test? Clearly compute your t-statistic and p-value for each of the pairwise t-tests. [Use the same within-groups variance computed from (a).]

Groups	Sample size n	Mean WBC	Standard deviation s
1: Common cold	269	11526	1341.3
2: Lung cancer	53	11445	1017.4
3: AIDS	9	12618	1225.2

Answer:

>> n1=269;n2=53;n3=9;x1=11526;x2=11445;x3=12618;s1=1341.3;s2=1017.4;s3=1225.2;

>> DF1=3-1 = **2**;DF2=n1+n2+n3-3 = **328**;

>> sw2=((n1-1)*s1^2+(n2-1)*s2^2+(n3-1)*s3^2)/DF2 = **1.6707e+006 = 1,670,700**

>> x=(n1*x1+n2*x2+n3*x3)/(n1+n2+n3) = 1.1543e+004

>> sb2=(n1*(x-x1)^2+n2*(x-x2)^2+n3*(x-x3)^2)/DF1 = **5.4937e+006 = 5,493,700**

>> F=sb2/sw2 = **3.2883**

>> 1-fcdf(F,DF1,DF2) = **0.0386**

This p-value is smaller than 0.05. We thus **reject the null hypothesis** that all three mean values are the same. In other words, the three are not identical.

The new level of significance used in the subsequent pairwise t-test will be 0.05/3 = **0.0167** according to Bonferroni.

Use the same sw2 = 166.5162 and DF2 = n1+n2+n3-3=328 for the following 3 pairwise t-tests.

1 vs 2:

>> t12=(x1-x2)/sqrt(sw2*(1/n1+1/n2)) = **0.4170**

>> 2*(1-tcdf(t12,DF2)) = **0.6770**

This is greater than 0.0167. We do not reject the null hypothesis that groups 1 and 2 have the same mean values. No significant difference between the two is concluded.

2 vs 3:

>> t23=(x2-x3)/sqrt(sw2*(1/n2+1/n3)) = **-2.5172**

>> 2*tcdf(t23,DF2) = **0.0123**

This is smaller than 0.0167. We believe there exists difference between groups 2 and 3.

1 vs 3:

>> t13=(x1-x3)/sqrt(sw2*(1/n1+1/n3)) = **-2.4932**

>> 2*tcdf(t13,DF2) = **0.0132**

This is smaller than 0.0167. We believe there exists difference between groups 1 and 3.