IT3030 Biostatistics **Final exam 2019.06.18** ID:_____

Name:_____

【Show all your computational/MATLAB steps.】

1. (25%) Below is the systolic blood pressure for 10 subjects, measured before and after taking a drug X.

N	Before taking X (x_1)	After taking X (x_2)
1	134	125
2	130	128
3	118	105
4	123	101
5	126	111
6	108	118
7	110	107
8	117	119
9	121	115
10	114	101

We'd like to perform a paired t-test at the level α=0.05 to the random variable $d=x_1-x_2$ to know whether the blood pressure drops after taking the drug. (a) State your null hypothesis. (b) Is this a 1-sided or 2-sided test? Why? (c) What is the value of t? (d) What is the p-value for this test? (e) Your conclusion? Why?

Answer:

(a) H_o: $d=x_1-x_2 <= 0$ that the blood pressure rises or does not change after taking the drug. The reason to use less or equal zero is due to that we asked whether the blood pressure "drops" or not. In other words, we only concern if it decreases significantly or not.

(b) **1-sided test**, since we are asking whether blood pressure drops after taking the drug. Only d > 0 significantly would reject the null hypothesis.

(c)

>> x1 = 134 130 118 123 126 108 110 117 121 114

>> x2 = 125 128 105 101 111 118 107 119 115 101

>> d=x1-x2= 9 2 13 22 15 -10 3 -2 6 13

>> mu=mean(d)= 7.1000

>> std=std(d)= 9.2910

>> t=(mu-0)/(std/sqrt(10))= **2.4166**

(d) Computing p-value (2-sided):

>> p = 1 - tcdf(t,10 – 1)= **0.0194**. Note that we did not multiply it by 2, because we concern only one tail (1-sided test).

(e) Since p-value is **smaller than** α=0.05, we would **reject** the null hypothesis. That is, taking the drug Z **would decrease** the blood pressure significantly.

2. (25%) Given 4 medical centers each in northern, central, southern and eastern Taiwan. Each center recruited a number of patients with their head counts, mean value and standard deviation of blood pressure shown in the table. (a) Compute the grand mean of blood pressure for all patients from these 4 centers. (b) Determine the within-group variance S_w^2 and between-group variance S_B^2. (c) Determine the F value so that we may perform an F-test. (d) What is the null hypothesis H_0, and the two degrees of freedom DF1 and DF2 used for the F-test? (e) Determine the p-value for the F-test given α=0.05, and your conclusion to reject or not to reject the null hypothesis.

Region	North	Central	South	East
Head count	N_1=12	N_2=23	N_3=31	N_4=12
Mean value	X_1=111.1	X_2=120.4	X_3=109.8	X_4=106.4
Standard deviation	S_1=5.5	S_2=10.4	S_3=7.5	S_4=7.7

Answer:

```
>> n1=12;n2=23;n3=31;n4=12;
>> x1=111.1;x2=120.4;x3=109.8;x4=106.4;
>> s1=5.5;s2=10.4;s3=7.5;s4=7.7;
>> X=(n1*x1+n2*x2+n3*x3+n4*x4)/(n1+n2+n3+n4)
X =    112.6026
>> >> sw2=((n1-1)*s1^2+(n2-1)*s2^2+(n3-1)*s3^2+(n4-1)*s4^2)/(n1+n2+n3+n4-4)
sw2 =      68.2697
>> >> sb2=(n1*(X-x1)^2+n2*(X-x2)^2+n3*(X-x3)^2+n4*(X-x4)^2)/(4-1)
sb2 =    710.2132
>> F=sb2/sw2
F =      10.4030
>>
```

The null hypothesis is that the blood pressures are all the same among these 4 groups.

```
>> df1=4-1=          3
>> df2=(n1+n2+n3+n4)-4=        74
>> p=1-fcdf(F,df1,df2)=      8.6219e-06
>>
```

This is smaller than 0.05, so we will **reject H_0**. That is, there exists significant blood pressure difference in at least one pair of centers among the four.

3. (25%) The side effects of a new drug are being tested against a placebo (安慰劑). A simple random sample of 565 patients yields the results shown below. At a significance level of α = 0.05, is there enough evidence to conclude that the treatment is independent of the side effect of nausea (噁心想吐)? (a) Shown the expected table. (b) Determine the χ^2 value. (c) Determine the p-value and your answer whether the treatment is independent of the side effect of nausea.

Result	Drug	Placebo	Total
Nausea	36	13	49
No nausea	254	262	516
Total	290	275	565

Ans:

>> O1=36;O2=13;O3=254;O4=262;

>> E1=290*49/565 = 25.1504

>> E2=49-E1 = 23.8496

>> E3=290-E1 = 264.8496

>> E4=275-E2 = 251.1504

The expected contingency table is:

25.1504	23.8496	49
264.8496	251.1504	516
290	275	565

>> O=[O1 O2 O3 O4]; E=[E1 E2 E3 E4];

>> chi2=sum((O-E).^2./E) = **10.5291**

>> 1-chi2cdf(chi2,1) = **0.0012**

This p-value is smaller than 0.05, we will reject the null hypothesis that there is no difference between using drug and using placebo. That is, the drug **causes strong side effect** of nausea.

4. (25%) The following table shows statistics for 6 people.

Measurements Taken on 6 normal people between the ages of 14 and 24 years		
Serum Cholesterol Y	Weight X_1	Systolic Blood Pressure X_2
162.2	51.0	108
158.0	52.9	111
157.0	56.0	115
155.0	56.5	116
156.0	58.0	117
154.1	60.1	120

(a) Assuming that $Y = mX_2 + c$. Determine the values for m and c.

(b) Assuming that $Y = b_0 + b_1X_1 + b_2X_2$. Determine the values for b_0 , b_1 and b_2.

Answer:

```
>> x1' =     51.0000    52.9000    56.0000    56.5000    58.0000    60.1000
>> x2' =     108    111    115    116    117    120
>> y' =    162.2000   158.0000   157.0000   155.0000   156.0000   154.1000
>> polyfit(x2,y,1)
ans =     -0.6283   228.9952
>>
```

Here, **m = -0.6283** and **c = -228.9952**

```
>> A=[ones(size(x2)) x1 x2];
>> b=regress(y,A)
b =
   319.9618
     2.8434
    -2.8072
>>
```

Here, b_0 **= 319.9618**, b_1 **= 2.8434**, and b_2 **= -2.8072**.