IT3030 Quiz#2 2018.06.12 ID:_____ Name:_____

[Write down the MATLAB commands you used.]

1. (40%) Given a chi-square probability distribution of **DF=4**, answer the following questions.

(a) What is the probability density for χ^2 = 3, 5 and 8, respectively?

(b) What is the χ^2 value yielding the right-hand-side tail size equals 0.05?

(c) What is the size for the right-hand-side tail for χ^2 = 8?

(d) Determine the area under curve between χ^2 = 3 to 8.

>> chi2pdf(3,4), chi2pdf(5,4), chi2pdf(8,4)

ans = 0.1673, 0.1026, 0.0366

>> chi2inv(0.95,4) ans = 9.4877

>> 1-chi2cdf(8,4) ans = 0.0916

>> chi2cdf(8,4)-chi2cdf(3,4) ans = 0.4662

IT3030 Quiz#2 2018.06.12 ID:_____ Name:_____

2. (60%) Suppose a city has its residents classified as four neighborhoods: A, B, C, and D. A random sample of 650 residents has their occupation recorded as "white collar", "blue collar", or "no collar". The null hypothesis is that each person's neighborhood of residence is independent of the person's occupational classification. The data are tabulated as:

	A	B	C	D	total
White collar	90 e=_____	60 e=_____	104 e=_____	95 e=_____	349
Blue collar	30 e=_____	50 e=_____	51 e=_____	20 e=_____	151
No collar	30 e=_____	40 e=_____	45 e=_____	35 e=_____	150
Total	150	150	200	150	650

(a) (30%) Fill in the expected number of people (e=_____) in each of the 12 cells. Keep 1D (小數點後一位).

(b) (10%) Compute the χ^2 value between the original and expected tables.

(c) (10%) What is the degree of freedom for the χ^2 probability density function to use?

(d) (10%) Determine the p-value for this test. What is your conclusion?

```
>> O1=[90 60 104 95];O2=[30 50 51 20];O3=[30 40 45 35];
>> E1=[150 150 200 150]*349/650   =   80.5385   80.5385   107.3846   80.5385
>> E2=[150 150 200 150]*151/650   =   34.8462   34.8462    46.4615   34.8462
>> E3=[150 150 200 150]*150/650   =   34.6154   34.6154    46.1538   34.6154

>> chi2 = sum((O1 - E1).^2./E1) + sum((O2 - E2).^2./E2)+sum((O3 - E3).^2./E3)

chi2 =    24.5712

DF = (4-1)*(3-1) = 6

>> 1-chi2cdf(chi2,6) = 4.0984e-04
```

This is a p-value smaller than 0.05, so we'd reject the null hypothesis that the data is independent. Indeed one person's neighborhood of residence is dependent of the person's occupational classification.