ECE 376 - Homework #7

Chi-Squared Test, Student t-Test.   Due Monday, March 7th

Chi-Squred Test

The following code implements a fair die and a loaded die (with the comment removed).

```
while(1) {
   while(!RB0);
   while(RB0) {
      d4 = (d4 + 1) % 4;
      d101 = (d101 + 1) % 101;
      }
    d4 = d4 + 1;
// Loaded Die
// if(d101 < 10) d4 = 4;
    LCD_Move(1,8); LCD_Out(d4, 1, 0);
    SCI_Out(d4, 1, 0);
    SCI_CRLF();
    }
```

1) Collect data for the fair 4-sided die. From your data, what is the probaility that the die is fair?

die roll	1	2	3	4
results	10	6	8	16

Set up a chi-squared table

	p	np	N	chi-squared
1	1/4	10	10	0
2	1/4	10	6	1.6
3	1/4	10	8	0.4
4	1/4	10	16	3.6
			Total	**5.6**

from StatTrek, a chi-squared score of 5.6 with 3 degrees of freedom corresponds to a probability of 0.83

Based upon this data, there is an 83% chance that the fair die is not fair

2) Remove the comment and collect data for the loaded die. From your data, what is the probaility that the die is fair?

die roll	1	2	3	4
results	8	11	6	15

Set up a chi-squared table

	p	np	N	chi-squared
1	1/4	10	8	0.4
2	1/4	10	11	0.1
3	1/4	10	6	1.6
4	1/4	10	15	2.5
			Total	**4.6**

From StatTrek, a chi-squared score of 4.6 with 3 degrees of freedom corresponds to a probability of 0.75

From this data, there is a 75% chance that the loaded die is not fair

Note: If you increased the number of rolls to something like 400, you could see the loaded die is loaded...

3) How loaded does the die have to be for you to be able to reliably detect that something is amiss?

Assume

- "detect" means "95% certainty"
- 40 die rolls
- x too many 4's
- Split shortage of 1's, 2's, and 3's evenly

95% certainty with 3 degrees of freedom translates to a chi-squared score of 7.82

	p	np	N	chi-squared
1	1/4	10	10 - x/3	x^2 / 90
2	1/4	10	10 - x/3	x^2 / 90
3	1/4	10	10 - x/3	x^2 / 90
4	1/4	10	10 + x	x^2 / 10
			Total	0.1333 x^2

To be 95% certain this is a loaded die

$$0.1333x^2 = 7.82$$

$$x = 7.65$$

$$p = \left(\frac{7.65}{40}\right) = 19.13\%$$

You can load the die with a 19% change of rolling a 4 without getting detected

- **with 40 die rolls**
- **with a probability of 95%**

Am I Psychic?

4) Determine whether or not you're psychic:

- Guess which number you're going to roll with the fair 4-sided die.
  - or take a deck of playing cards. Predct the suit of each card then record whether you were right or wrong.
- Roll the dice a bunch of times (>10)
- Record how many times you are correct

Use a chi-squared test to determine whether or not you're guessing (correct 25% of the time)

I rolled a 6-sided die 100 times and predicted the result before each roll.

Prediction (in order)

2 6 1 6 1 6 2 2 **5** 2 2 5 **2** 5 6 **4** 3 4 6 5 1 1 5 2 5 5 1 6 1 5 5 3 3 1 4 1 6 3 **1**
3 3 1 4 3 4 4 1 1 5 **3** 5 3 3 4 2 **3** 3 1 **4** **1** 1 **1** 6 4 6 5 **1** 1 6 5 4 3 6 1 4 **5** 3 5
4 4 6 **4** 4 4 **4** 2 4 1 5 4 3 **4** 1 4 3 **4** 1 6 2 5

Die Roll (in order)

2 4 6 2 3 2 4 5 5 1 4 1 2 6 4 4 4 2 5 2 4 3 6 6 4 4 4 2 5 1 4 4 6 2 1 4 2 1 1
6 6 2 4 6 1 3 6 6 3 3 4 6 2 5 5 3 4 2 4 1 5 1 4 5 3 6 1 6 5 6 2 6 2 3 6 5 4 1
5 5 1 4 3 3 4 4 3 6 2 2 4 4 4 6 2 4 4 5 4 2

Guess (bin)	p theoretical probability	np expected frequency	N actual frequency	$\chi^2 = \left(\frac{(np-N)^2}{np}\right)$
Correct	1/6	16.67	16	0.03
Incorrect	5/6	83.33	84	0.01
			Total:	0.03

From StatTrek, a chi-squared score of 0.03 with 1 degree of freedom corresponds to a probability of 0.14

Based upon these results, there is a 14% chance that I'm not just guessing

t-Test

5) Use your data from problem #7 for homework set #6 (data collection). Determing the 90% confidence interval for your data.

Data from homework #6

C = { 46.972uF, 46.904uF, 46.735uF }

```
>> C = [46.972,46.904,46.735]

C =   46.9720   46.9040   46.7350

>> x = mean(C)

x =   46.8703

>> s = std(C)

s =    0.1220
```

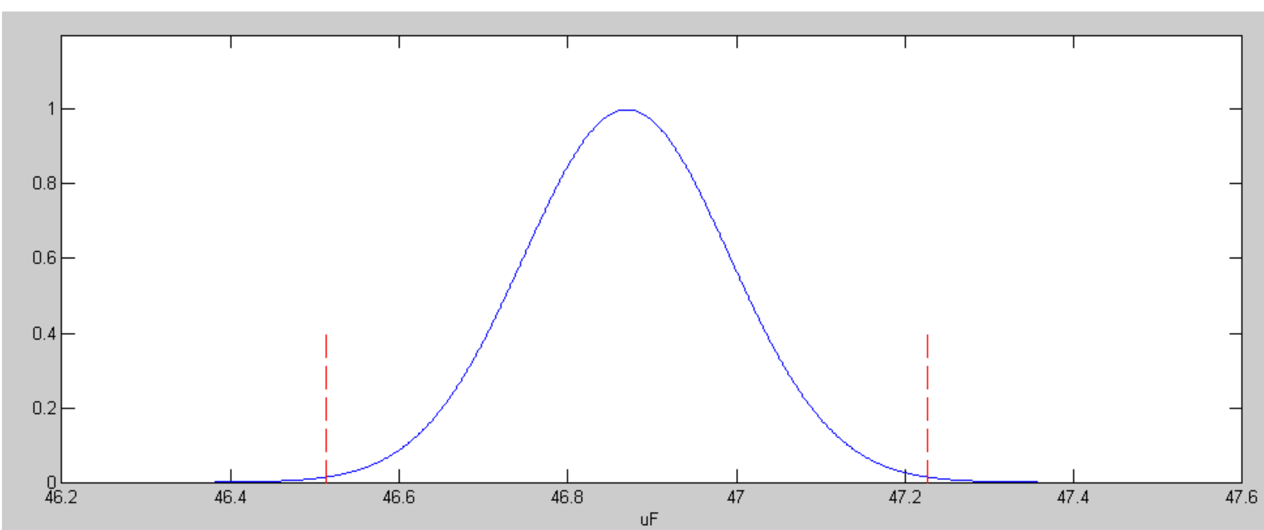
From StatTrek, the t-score that corresponds to 5% tails (90% confidence interval) with 2 degrees of freedom is

$t = 2.920$

The 90% confidence interval for C is thus

$\bar{x} - 2.920s < C < \bar{x} + 2.920s$ $\quad$ $p = 0.9$

46.514uF < C < 47.2267uF $\quad$ $p = 0.9$



pdf for the actual value of the 47uF capacitor

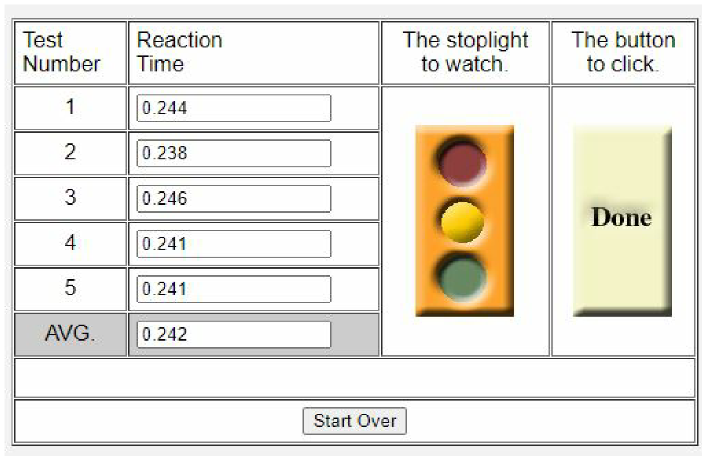
Reflex Times

The following web site allows you to record your reflex times:

https://faculty.washington.edu/chudler/java/redgreen.html

6) Record your reflex times using the above link. From your data, determine

- The 90% confidence interval for the next time you play this game

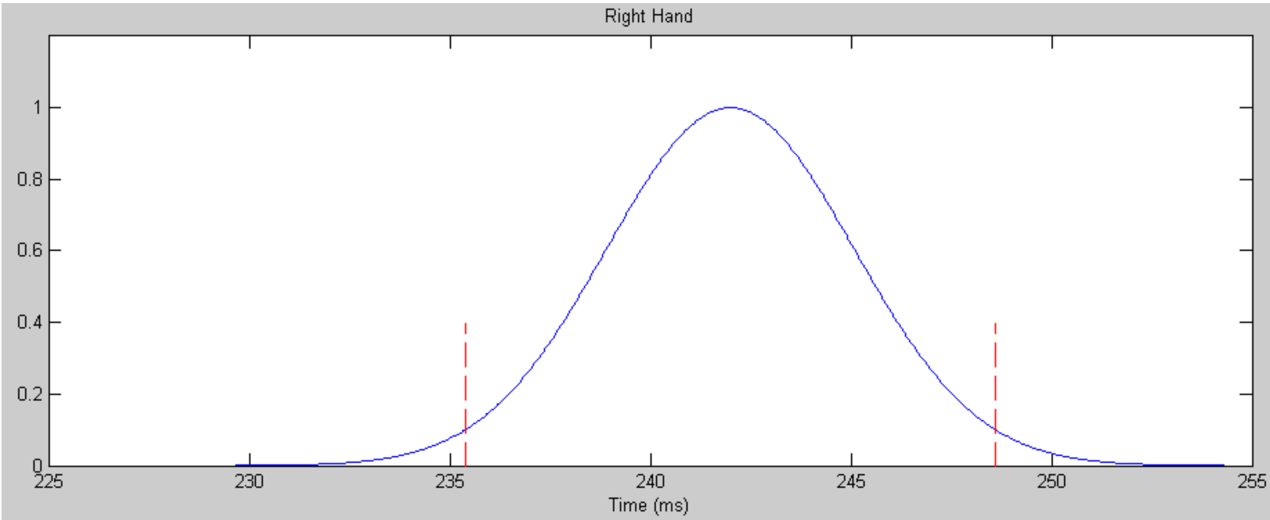


Test Number	Reaction Time	The stoplight to watch.	The button to click.
1	0.244		Done
2	0.238		
3	0.246		
4	0.241		
5	0.241		
AVG.	0.242		


```
>> Right = [0.244,0.238,0.246,0.241,0.241]';
>> xr = mean(Right)

xr =    0.2420

>> sr = std(Right)

sr =    0.0031

>> xr - 2.132*s

ans =    0.2354

>> xr + 2.132*s

ans =    0.2486
```

Not asked for but interesting to plot the resulting pdf for my reflex time with my right hand

```
>> s1 = [-4:0.01:4]';
>> p = exp(-s1.^2 / 2);
>> plot((s1*sr + xr)*1000, p)
>> xlabel('Time (ms)');
>> title('Right Hand');
```



My next trial will be in the range of (235.4ms, 248.6ms) with a probability of 0.9

The probability that your reflex time will be less than 200ms

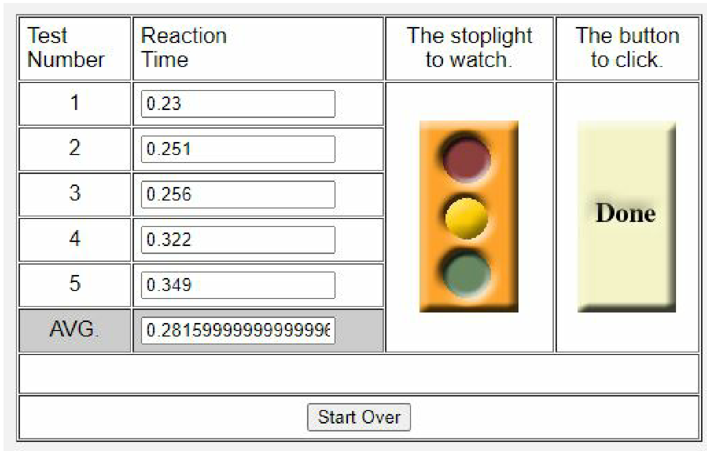
```
>> t = (xr - 0.2)/s

t =   13.6266
```

From StatTrek, a t-score of 13.6266 corresponds to a probability of 0.0001

There is a 0.01% chance my next reaction time will be less than 200ms

7) Collect a second data set from the above link with a diferent condition (pick one)

- Measure with my left hand



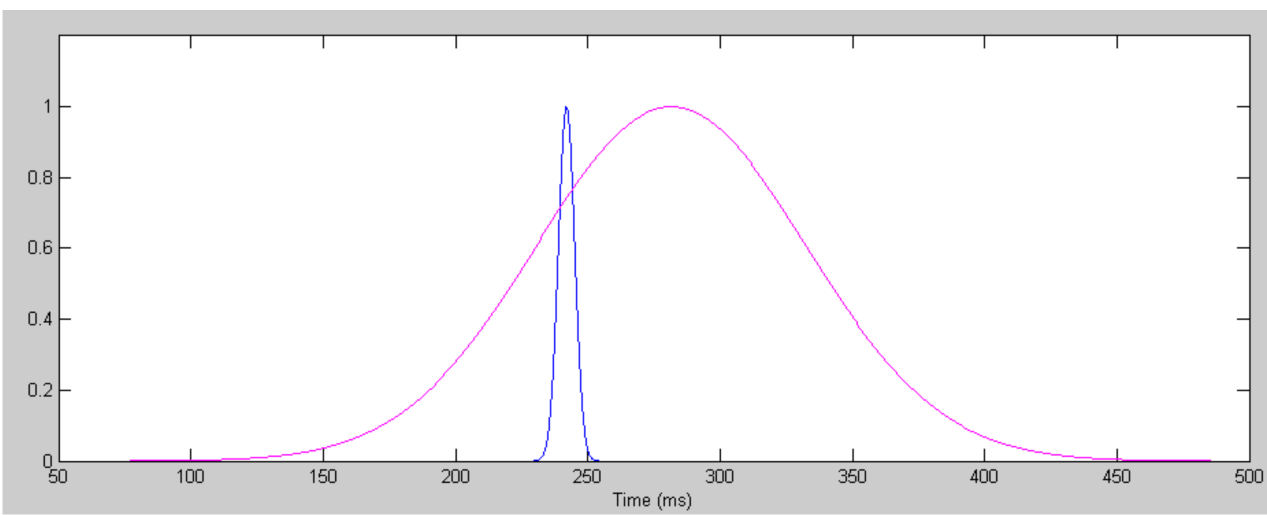
Test Number	Reaction Time	The stoplight to watch.	The button to click.
1	0.23		Done
2	0.251		
3	0.256		
4	0.322		
5	0.349		
AVG.	0.2815999999999999		

Start Over

```
>> Left = [0.23,0.251,0.256,0.322,0.349]';
>> xl = mean(Left)

xl =     0.2816

>> sl = std(Left);
>> xw = xr - xl

xw =   -0.0396
```

Just for fun, plot the two pdf's together

```
>> plot((s1*sr + xr)*1000, p, 'b', (s1*sl + xl)*1000, p, 'm')
>> xlabel('Time (ms)');
>> ylim([0,1.2])
```



Right Hand (blue) & Left Hand (magenta)

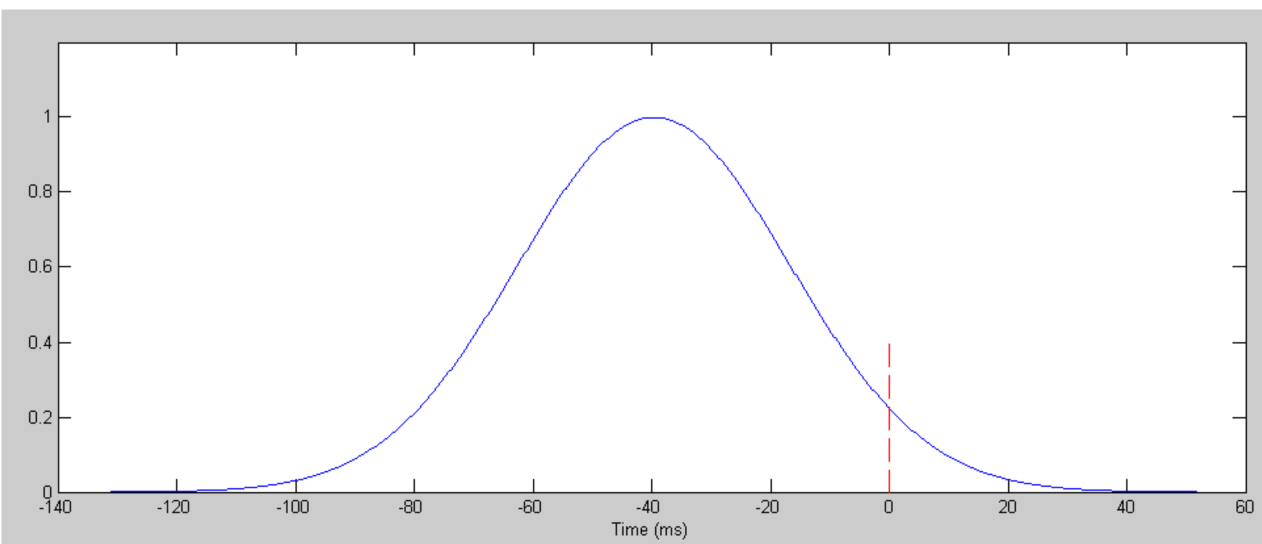
8) Do a comparison of means test (t-test with W = A - B) to determine the probability that population A has a mean that is less than population B.

- *That coffee reduces your reaction time*
- *That your dominant hand has a faster reaction time, etc.*

```
>> xw = xr - xl

xw =   -0.0396

>> sw = sqrt(sr^2/5 + sl^2/5)

sw =    0.0229

>> tw = xw / sw

tw =   -1.7310

>> den = (sl^2/5)^2/4 + (sr^2/5)^2/4;
>> df = num/den

df =    4.0291
```

From StatTrek, a t-score of -1.731 with 4 degrees of freedom corresponds to a probability of 7.92%

There is 7.92% chance that my right hand has a larger (worse) reaction time than my left hand



pdf of w: (Reaction Time of Right Hand) - (Reaction Time of Left Hand)