ECE 376 - Homework #6

Statistics and Data Collection

Data Collection

1) Use your PIC board to measure the temperature of the room

- Measure the temperature every 2 seconds (change to 100ms)
- Collect 50 data points (change to 72 data points)
- Plot or list your data

From your data determine the mean and standard deviation.

Main Routine:

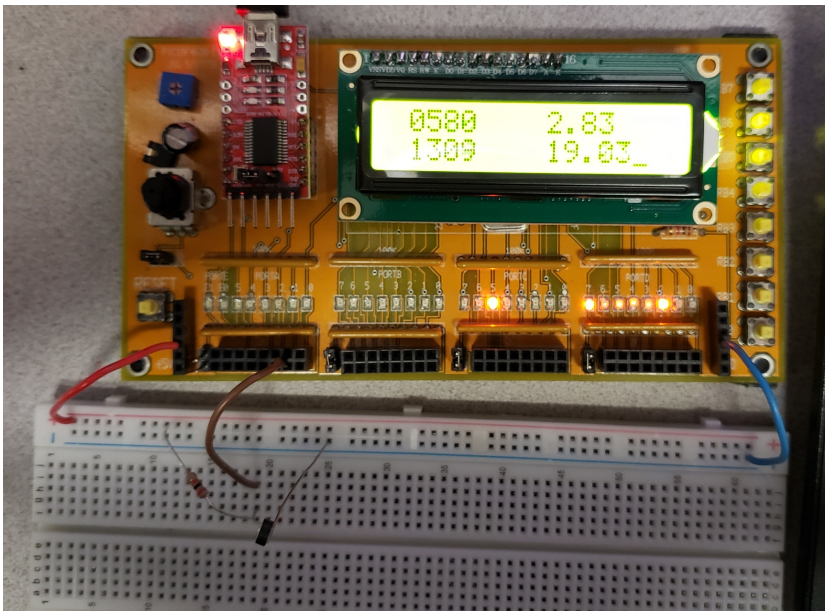
```
while(1) {

   A2D = A2D_Read(1);
   VOLT = 0.488 * A2D;
   OHM = 1000.0 * (A2D / (1023.0 - A2D) );
   CELSIUS = 3930. / ( log( A2D / (1023. - A2D) ) + 13.1879 ) - 273;

   LCD_Move(0,0);  LCD_Out(A2D, 4, 0);
   LCD_Move(0,8);  LCD_Out(VOLT, 3, 2);
   LCD_Move(1,0);  LCD_Out(OHM, 4, 0);
   LCD_Move(1,8);  LCD_Out(CELSIUS*100, 4, 2);

   SCI_Out(VOLT, 5, 2);
   SCI_Out(OHM, 4, 0);
   SCI_Out(CELSIUS*100, 4, 2);
   SCI_CRLF();

   Wait_ms(100);

   }
```

Compiler Results

```
Memory Summary:
    Program space        used  1E6Eh (  7790) of 10000h bytes   ( 11.9%)
    Data space           used    44h (    68) of   F80h bytes   (  1.7%)
    EEPROM space         used     0h (     0) of   400h bytes   (  0.0%)
    ID Location space    used     0h (     0) of     8h nibbles (  0.0%)
    Configuration bits   used     0h (     0) of     7h words   (  0.0%)
```

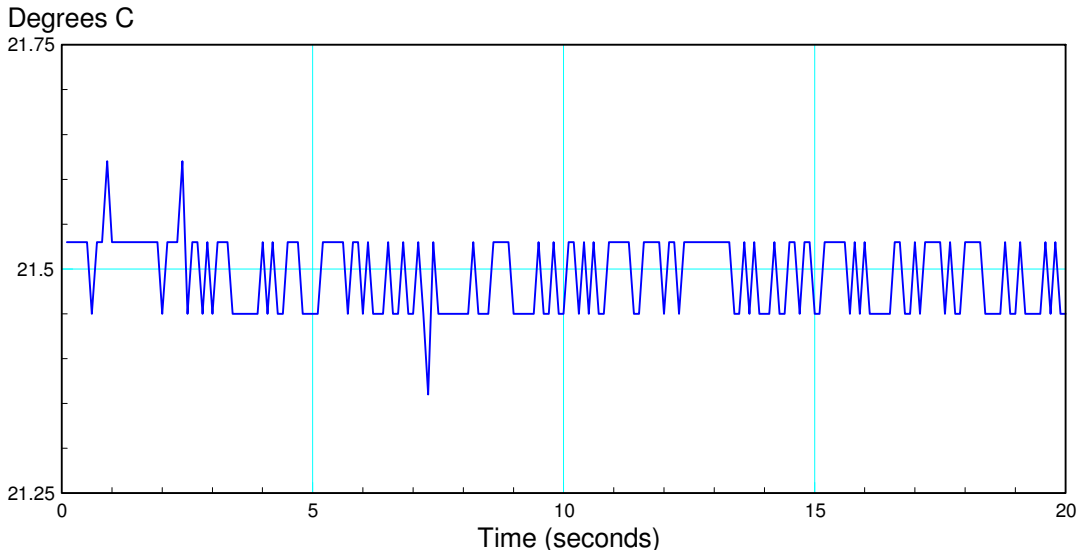
Data

```
Volts  Ohms  Degrees C
002.69 1171 21.45
002.69 1171 21.45
002.69 1171 21.45
002.69 1171 21.45
002.69 1171 21.45
002.69 1171 21.45
002.69 1171 21.45
```



In Matlab

```
>> Data = [  < paste > ];
>> T = Data(:,3);
>> x = mean(T)

x =   21.4925

>> s = std(T)

s =    0.0425

>> n = length(T)

n =   228
```



2) Use a chi-squared test to determine if your data has a uniform distribution over the range of [a, b] where 'a' is the minimum of your data and "b' is the maximum of your data.

```
>> sum(T == 21.36)
ans =     1

>> sum(T == 21.45)
ans =   107

>> sum(T == 21.53)
ans =   118

>> max(T)
ans =   21.6200

>> sum(T == 21.62)
ans =     2
```

Computing the chi-squared score assuming a uniform distribution

Degrees C	p()	expected	actual	chi-squared
21.36	1/4	57	1	55.02
21.45	1/4	57	107	43.86
21.53	1/4	57	118	65.28
21.62	1/4	57	2	53.07
			Total	164.16

From StatTrek, a chi-squared value of 164 with 3 degrees of freedom corresponds to a probability of 1.000

- **I am more than 99.9995% certain this is not a uniform distributionb**

Chi-Squred Test

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

```
while(1) {
   while(!RB0);
   while(RB0) {
      d7 = (d7 + 1) % 7;
      d100 = (d100 + 1) % 100;
      }
    d7 = d7 + 1;
// Loaded Die
// if(d100 < 10) d7 = 7;
    LCD_Move(1,8); LCD_Out(d7, 1, 0);
    SCI_Out(d7, 1, 0);
    SCI_CRLF();
    }
```

3) Collect data for the fair 7-sided die. From your data, what is the probability that the die is fair?

Raw Data: 185 die rolls

1	2	3	4	5	6	7
21	24	14	29	30	32	35

Setting up a table to compute the Chi-Sqquared score

Die Roll	p()	expected	actual	chi-squared
1	1/7	26.43	21	1.12
2	1/7	26.43	24	0.22
3	1/7	26.43	14	5.85
4	1/7	26.43	29	0.25
5	1/7	26.43	30	0.48
6	1/7	26.43	32	1.17
7	1/7	26.43	35	2.78
			Total	11.87

From StatTrek,

- A chi-squared score of 11.87
- With 6 degrees of freedom
- Has a probability of 0.94

From this data, there is a 94% chance that the die is not fair

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

Raw Data: 218 die rolls

1	2	3	4	5	6	7
26	38	27	25	32	20	50

Setting up a chi-squared table

Die Roll	p()	expected	actual	chi-squared
1	1/7	31.34	26	0.91
2	1/7	31.34	38	1.42
3	1/7	31.34	27	0.6
4	1/7	31.34	25	1.28
5	1/7	31.34	32	0.01
6	1/7	31.34	20	4.1
7	1/7	31.34	50	11.11
			Total	**19.44**

From StatTrek, a chi-squared score of 19.44 corresponds to a probability of 0.997

I am 99.7% certain that this is a loaded die

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

Assume

- 210 die rolls
- 95% certain that the die is loaded
- X too many 7's
- Each other number is X/6 too few

From Stat-Trek, 95% certainty means the chi-squared score should be 3.8410

Die Roll	p()	expected	actual	chi-squared
1 - 6	6/7	180	180 - X	$\left(\frac{X^2}{180}\right)$
7	1/7	30	30 + X	$\left(\frac{X^2}{30}\right)$
			Total	0.03889 X^2

$$0.03889X^2 = 3.8410$$

$$X = 9.938$$

I can load the die so that I get 10 more 7's than I should and get away with it 95% of the time. 10/210 corresponds to 4.76%

$$loading = \left(\frac{\text{10 extra 7s}}{\text{210 rolls}}\right) = 4.76\%$$

Note: The loading depends upon the sample size

- Smaller sample sizes can have more loading
- Larger sample sizes can have less loading
- As the sample size goes to infinity, the loading goes to zero for detecting at 95% probability

Am I Psychic?

6) Take a coin and flip it 10+ times.

- Each time you flip the coin, predict the result.
- Count the number of times you're right and wrong.

Guess	T	H	H	T	H	H	T	H	T	H
Actual	H	H	T	T	T	H	H	H	T	H

7) Use a chi-squared test to determine if you're just guessing (p = 0.5)

Result	p()	expected	actual	chi-squared
Correct	1/2	5	6	0.2
Incorrect	1/2	5	4	0.2
			Total	0.4

From StatTrek, the corresponds to a probability of 0.47

There is a 47% chance I'm not just guessing

- Sadly, there's no evidence I'm psychic