ECE 341 - Homework #15

F-Test and ANOVA. Due Friday, June 10th

Test of a 3+ Populations

1) The average temperature in Fargo for three different months is:

		mean	std	n
A	June	65.8032	3.0791	80
B	July	70.9427	2.5143	80
C	Aug	69.0227	2.6740	80

Determine if the means are the same using an ANOVA test.

Determine the global mean

$$\overline{G} = \left(\frac{1}{N}\right)\left(n_a\overline{A} + n_b\overline{B} + n_c\overline{C}\right)$$

Determine MSSb and MSSw

$$MSS_b = \left(\frac{1}{k-1}\right)\left(n_a\left(\overline{A}-\overline{G}\right)^2 + n_b\left(\overline{B}-\overline{G}\right)^2 + n_c\left(\overline{C}-\overline{G}\right)^2\right)$$

$$MSS_w = \left(\frac{1}{N-k}\right)\left((n_a-1)s_a^2 + (n_b-1)s_b^2 + (n_c-1)s_c^2\right)$$

Matlab Code:

```
Xa = 65.8032;
Sa = 3.0791;
Xb = 70.9427;
Sb = 2.5143;
Xc = 69.0227;
Sc = 2.6740;
Na = 80;
Nb = 80;
Nc = 80;
k = 3;
N = Na + Nb + Nc
G = (Na*Xa + Nb*Xb + Nc*Xc) / N
MSSb = (Na*(Xa-G)^2 + Nb*(Xb-G)^2 + Nc*(Xc-G)^2) / (k-1)
MSSw = ((Na-1)*Sa^2 + (Nb-1)*Sb^2 + (Nc-1)*Sc^2) / (N-k)
F = MSSb / MSSw
```

Result:

```
N =      240
G =       68.5895
MSSb =  539.5472
MSSw =    7.6509
F =      70.5203
```

You can also get the same answer with an ANOVA table

A	B	C	A	B	C
			3.0791 std(A)	2.5143 std(B)	2.6740 std(C)
Na = 80	Nb = 80	Nc = 80	748.98 sum of squares	499.41 sum of squares	564.87 sum of squres
N = 240			1813.3 sum of squares		
65.8032 mean(A)	70.9427 mean(B)	69.0027 mean(C)	MSSw = 7.6509		
68.5895 G = global mean					
621.09 Na $(A - G)^2$	422.91 Nb $(B - G)^2$	15.01 Nc $(C - G)^2$			
1079.1 sum of squres					
MSSb = 539.54					

```
F = MSSb / MSSw
F =      70.5203
```

Now use an F table with

- numerator = 2 degrees of freedom (k-1)
- denominator = 237 degrees of freedom (N-k)

This corresponds to a probability of 1 (> 99.995%)

I am more than 99.995% that the three data sets have a different mean

You'd have to do 1 on 1 t-tests to determine which one (or more) is the outlier.

2) The global average for three decades are:

		mean	std	n
A	1880-1899	-0.1766	0.121	240
B	1960-1969	0.0233	0.1161	240
C	2010-2019	0.7944	0.1685	240

Determine if the means are the same using an ANOVA test.

Matlab Code:

```
Xa = -0.1766;
Sa =  0.121;
Xb =  0.0233;
Sb =  0.1161;
Xc =  0.7944;
Sc =  0.1685;
Na =  240;
Nb =  240;
Nc =  240;
k = 3;
N = Na + Nb + Nc
G = (Na*Xa + Nb*Xb + Nc*Xc) / N
MSSb = (Na*(Xa-G)^2 + Nb*(Xb-G)^2 + Nc*(Xc-G)^2) / (k-1)
MSSw = ((Na-1)*Sa^2 + (Nb-1)*Sb^2 + (Nc-1)*Sc^2) / (N-k)
F = MSSb / MSSw
```

Result:

```
N =     720
G =       0.2137
MSSb =   63.0958
MSSw =    0.0188
F =    3349.5e
```

Now use an F table with

- numerator = 2 degrees of freedom (k-1)
- denominator = 717 degrees of freedom (N-k)

This corresponds to a probability of 1 (> 99.995%)

(note: An F-score of 5 corresponds to a probability of 99.3%, so this is *way* off the chart)

3) The scores for three players playing *Hungry Hungry Hippo* are:

```
A:  73    63    79    59    60
B:  52    31    75    64    53    74
C:  53    69    68    74    74    62    70
```

Determine if the means are the same using an ANOVA test.

Matlab Code:

```
A = [73    63    79    59    60];
B = [52    31    75    64    53    74];
C = [53    69    68    74    74    62    70];
Xa = mean(A);
Sa =  std(A);
Xb =  mean(B);
Sb =  std(B);
Xc =  mean(C);
Sc =  std(C);
Na =  length(A);
Nb =  length(B);
Nc =  length(C);
k = 3;
N = Na + Nb + Nc
G = (Na*Xa + Nb*Xb + Nc*Xc) / N
MSSb = (Na*(Xa-G)^2 + Nb*(Xb-G)^2 + Nc*(Xc-G)^2) / (k-1)
MSSw = ((Na-1)*Sa^2 + (Nb-1)*Sb^2 + (Nc-1)*Sc^2) / (N-k)
F = MSSb / MSSw
```

Result

```
N =      18
G =      64.0556
MSSb =  156.2270
MSSw =  134.1660
F =       1.1644

>>
```

Now use an F table with

- numerator = 2 degrees of freedom (k-1)
- denominator = 15 degrees of freedom (N-k)

This corresponds to a probability of 0.66

There is a 66% chance that these populations have different means