

ECE 341 - Homework #15

F-Test and ANOVA.

F-Test

1) The resistances of 1k 5% tolerance resistors were measured:

```
A: 989, 996, 993, 991, 993, 991, 997, 996, 995, 995, 991, 997, 1008, 995
B: 996, 995, 996, 995, 998, 996, 995, 990, 981, 988, 994, 999, 990, 992
C: 997, 992, 995, 994, 990, 990, 994, 992, 996, 992, 992, 994, 995, 988
```

Run and F-test to determine if there is a difference between the variance of

- Population A and B
- Population A and C

```
>> A = [989, 996, 993, 991, 993, 991, 997, 996, 995, 995, 991, 997, 1008, 995];
>> B = [996, 995, 996, 995, 998, 996, 995, 990, 981, 988, 994, 999, 990, 992];
>> C = [997, 992, 995, 994, 990, 990, 994, 992, 996, 992, 992, 994, 995, 988];
```

A vs. B:

```
>> F = var(A) / var(B)
F = 0.9376
```

This is less than one, so take the inverse:

```
>> 1/F
ans = 1.0666
```

From StatTrek, this corresponds to a probability of 0.545

There is a 54.5% chance that A and B have different variances

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)	13
Degrees of freedom (v_2)	13
f Statistic (f)	1.0666
Probability: $P(F \leq 1.0666)$	0.545
Probability: $P(F \geq 1.0666)$	0.455

Calculate

A vs. C:

```
>> F = var(A) / var(C)
F = 3.2601
```

From StatTrek, this corresponds to a probability of 0.979

There is a 97.9% chance that A and C have different variances

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)	13
Degrees of freedom (v_2)	13
f Statistic (f)	3.2601
Probability: $P(F \leq 3.2601)$	0.979
Probability: $P(F \geq 3.2601)$	0.021

Calculate

Global Temperatures (take 1)

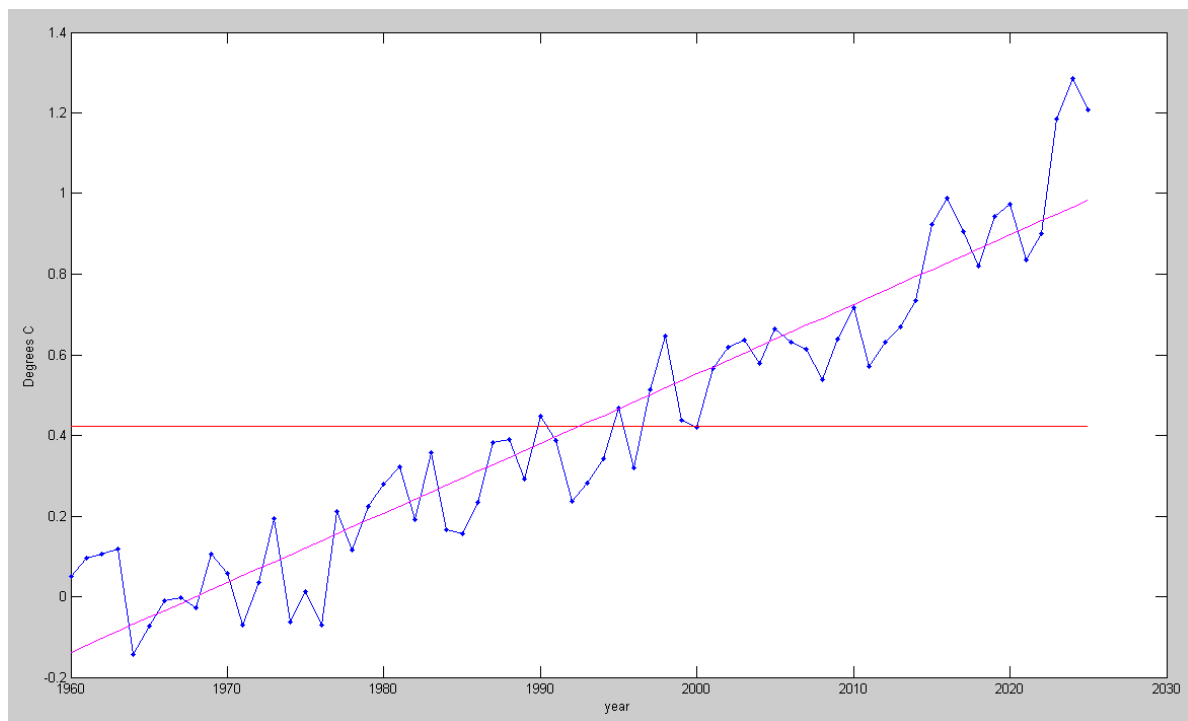
The global land and cocean temperature anomolies have been measured by NOAA since 1880

2) Use the data from 1960-2025, use an F-test to determine if temperatures are rising in a linear fashion:

- $H_0: T = a$
- $H_1: T = a + by$

In Matlab, do a least-squares curve fit

```
>> year = Data(:,1);  
>> T = Data(:,2);  
  
>> B0 = [year.^0];  
>> A0 = inv(B0'*B0)*B0'*T  
  
a    0.4230  
  
>> B1 = [year, year.^0];  
>> A1 = inv(B1'*B1)*B1'*T  
  
b    0.0173  
a    -33.9591  
  
>> n = length(T)  
  
n =    66  
  
>> plot(year,T,'b.-',year,B0*A0,'r',year,B1*A1,'m')  
>> xlabel('year');  
>> ylabel('Degrees C')
```



Is the $b \cdot \text{year}$ term tracking data or noise? Use an F-test to find out

Compute the F statistic

```
>> F = (n-2) / (n-1) * var(T - B0*A0) / var(T - B1*A1)
F = 8.9584
```

From StatTrek, this corresponds to a probability > 0.999995

I am more than 99.999% certain that b*year term is tracking data rather than just noise (the y term is significant)

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)	64
Degrees of freedom (v_2)	65
f Statistic (f)	8.9584
Probability: $P(F \leq 8.9584)$	1.0
Probability: $P(F \geq 8.9584)$	0.0

Calculate

3) Use the data from 1960-2025, use an F-test to determine if temperatures are rising at an increasing rate:

- $H_0: T = a + by$
- $H_1: T = a + by + cy^2$

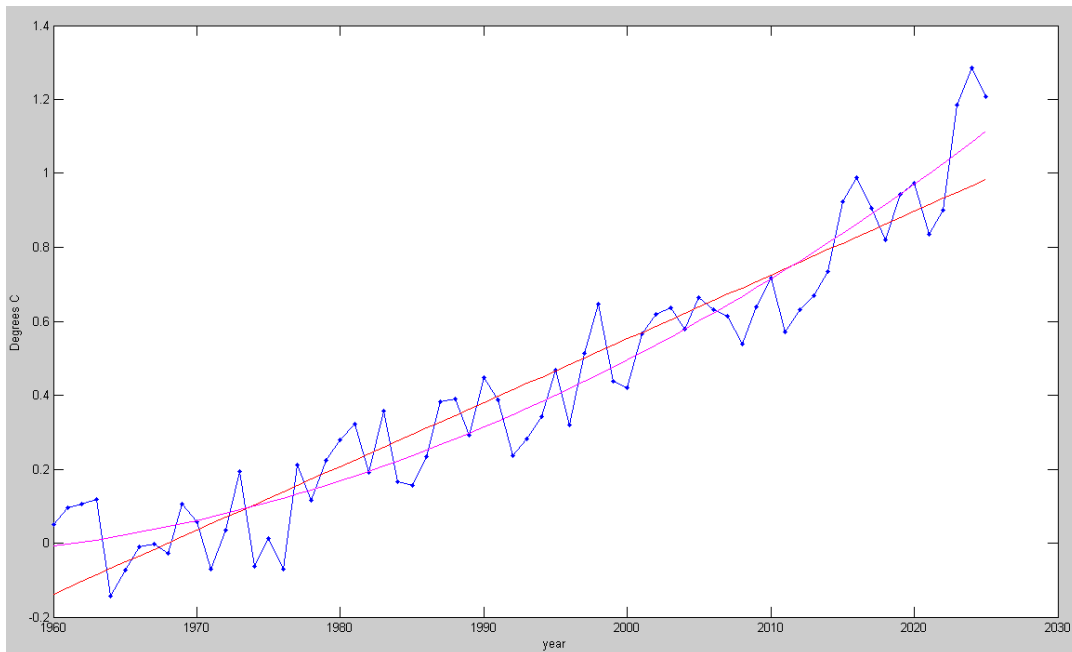
```
>> B0 = [year.^0, year];  
>> A0 = inv(B0'*B0)*B0'*T
```

```
a -33.9591  
b 0.0173
```

```
>> B1 = [year.^0, year, year.^2];  
>> A1 = inv(B1'*B1)*B1'*T
```

```
a 707.2150  
b -0.7268  
c 0.0002
```

```
>> plot(year, T, 'b.-', year, B0*A0, 'r', year, B1*A1, 'm')  
>> xlabel('year');  
>> ylabel('Degrees C')  
>>
```



Is the parabolic term tracking data or noise? Use an F test to find out

```
>> n = length(T)
```

```
n = 66
```

```
>> F = (n-2)/(n-1) * var(T - B0*A0) / var(T - B1*A1)
```

```
F = 1.3582
```

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)

Degrees of freedom (v_2)

f Statistic (f)

Probability: $P(F \leq 1.3582)$

Probability: $P(F \geq 1.3582)$

Calculate

I am 88.9% certain that the year² term is tracking data rather than just noise

- kind of a scary result...

ANOVA

4) For the three populations of 1k resistors in problem #1, run an ANOVA test to see if the means of these resistors are different

Matlab Code

```
A = [989, 996, 993, 991, 993, 991, 997, 996, 995, 995, 991, 997, 1008, 995];
B = [996, 995, 996, 995, 998, 996, 995, 990, 981, 988, 994, 999, 990, 992];
C = [997, 992, 995, 994, 990, 990, 994, 992, 996, 992, 992, 994, 995, 988];

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 = 42
G = 993.6429
MSSb = 14.0000
MSSw = 16.4524
F = 0.8509
```

This is less than 1.000, so invert it. Find the probability using StatTrek

$$p = 0.696$$

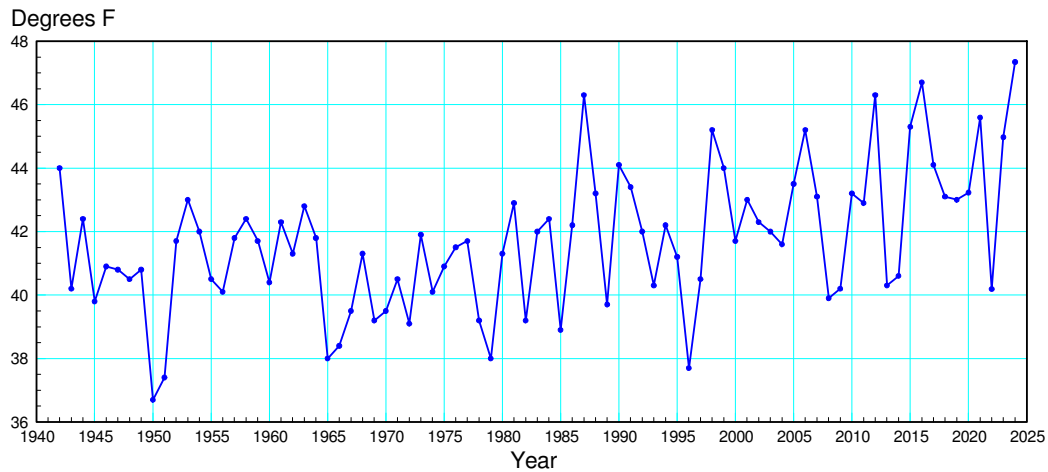
There is a 69.6% chance the populations have different means.

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)	40
Degrees of freedom (v_2)	41
f Statistic (f)	1.17522
Probability: $P(F \leq 1.17522)$	0.696
Probability: $P(F \geq 1.17522)$	0.304

Calculate

4) The average temperature in Fargo has been reported by Hector airport since 1942



Use ANOVA to determine if the mean temperature in Fargo is different over the following time spans:

- A: 1942-1970
- B: 1980-2000
- C: 2010-2025

Matlab Code: Bring in the data into Matlab

```
A = [44.0 40.2 42.4 39.8 40.9 40.8 40.5 40.8 36.7 37.4 41.7 43.0 42.0 40.5 40.1
41.8 42.4 41.7 40.4 42.3 41.3 42.8 41.8 38.0 38.4 39.5 41.3 39.2 39.5];

B = [41.3 42.9 39.2 42.0 42.4 38.9 42.2 46.3 43.2 39.7 44.1 43.4 42.0 40.3 42.2
41.2 37.7 40.5 45.2 44.0 41.7];

C = [43.2 42.9 46.3 40.3 40.6 45.3 46.7 44.1 43.1 43.0 43.23 45.59 40.19 44.97
47.34 43.52];

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
```

Results:

```
N = 66
G = 41.8476
MSSb = 47.7418
MSSw = 3.8628
F = 12.3594
```

Use StatTrek and an F test to see if the populations have different means

$$p > 0.99995$$

I'm 99.995% certain that these populations have different means

more than 99.9995% certain that Fargo's temperatures are changing

- Enter values for degrees of freedom (v_1 and v_2).
- Enter a value for one, and only one, of the other textboxes.
- Click **Calculate** to compute a value for the last textbox.

Degrees of freedom (v_1)	65
Degrees of freedom (v_2)	64
f Statistic (f)	12.3594
Probability: $P(F \leq 12.3594)$	1
Probability: $P(F \geq 12.3594)$	0

Calculate