ECE 341 - Homework #10

Testing with Normal Distributions. Summer 2024

Testing with Normal Distributions

Assume the monthly temperatures in Fargo, ND are normal distributions with the following mean and standard deviation:

Monthly High (Degrees F: Fargo, ND)												
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
mean	38.5363	41.0038	56.0625	78.1	87.8625	92.0138	94.625	94.6262	89.575	79.5	59.425	41.7875
st dev	6.4057	7.1528	10.7118	7.7909	4.5472	4.5281	4.0043	4.5967	5.6294	6.7842	7.4728	6.5327

Monthly Low (Degrees F: Fargo, ND)												
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
mean	-23.8725	-20.6238	-8.1475	15.1775	27.3413	40.425	46.4875	43.3387	30.6763	19.15	-1.0875	-17.025
st dev	8.2179	7.8559	10.0237	7.0423	4.3864	4.1576	4.0938	4.1522	4.8861	5.5212	9.0417	9.1069

1) How cold will this January get

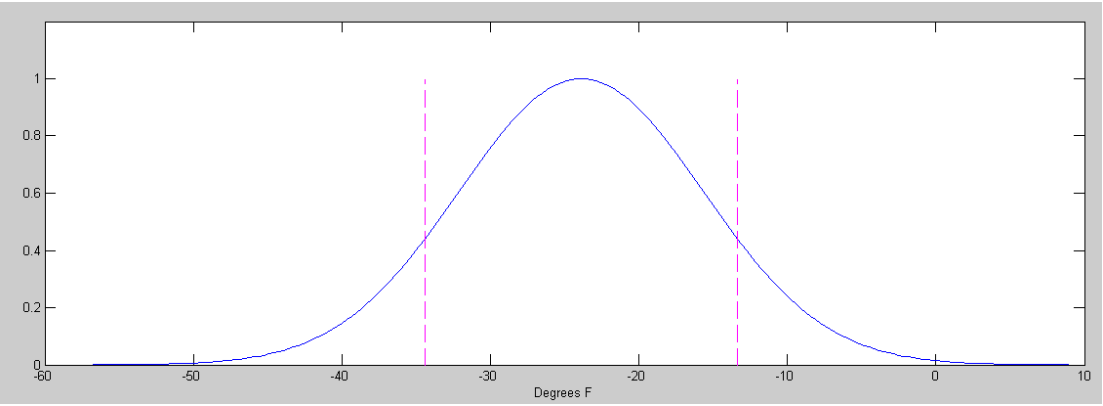
With a confidence level of 80%?

z-score = 1.28155 for 10% tails

$$T = \mu \pm 1.28155\sigma$$

$$T = -23.8725 \pm 1.28155 \cdot 8.2179$$

$$-34.40F < low < -13.34F \qquad p = 80\%$$

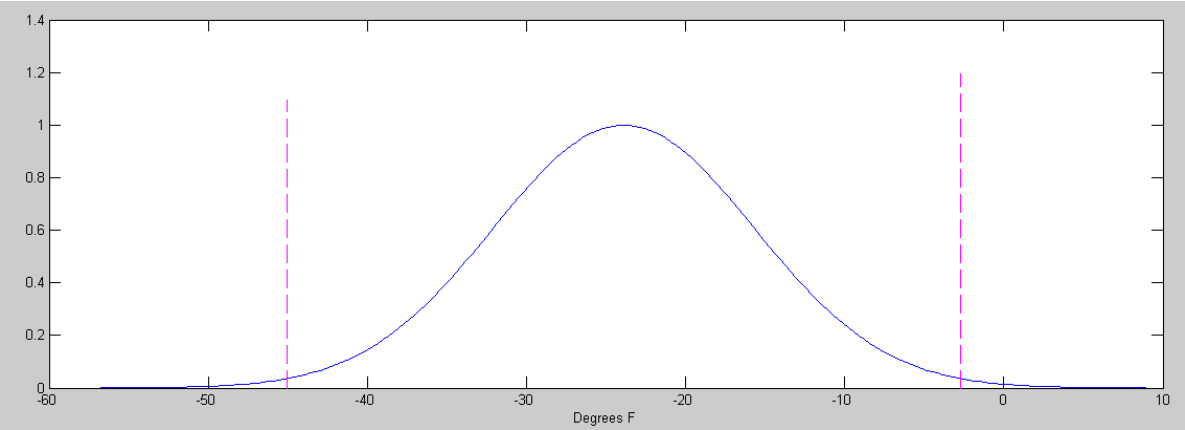


With a confidence level of 99%?

z-score = 2.57583 for 0.5% tails

$$T = \mu \pm 2.57583\sigma$$

$$-45.04F < low < -2.70F \qquad p = 99\%$$



With a confidence level of 100%?

z-score = infinity

$$-\infty < low < +\infty \qquad p = 100\%$$

not really useful

2) What is the probability that it will break -35F this coming January?

$$z = \left(\frac{-35F - \mu}{\sigma}\right)$$

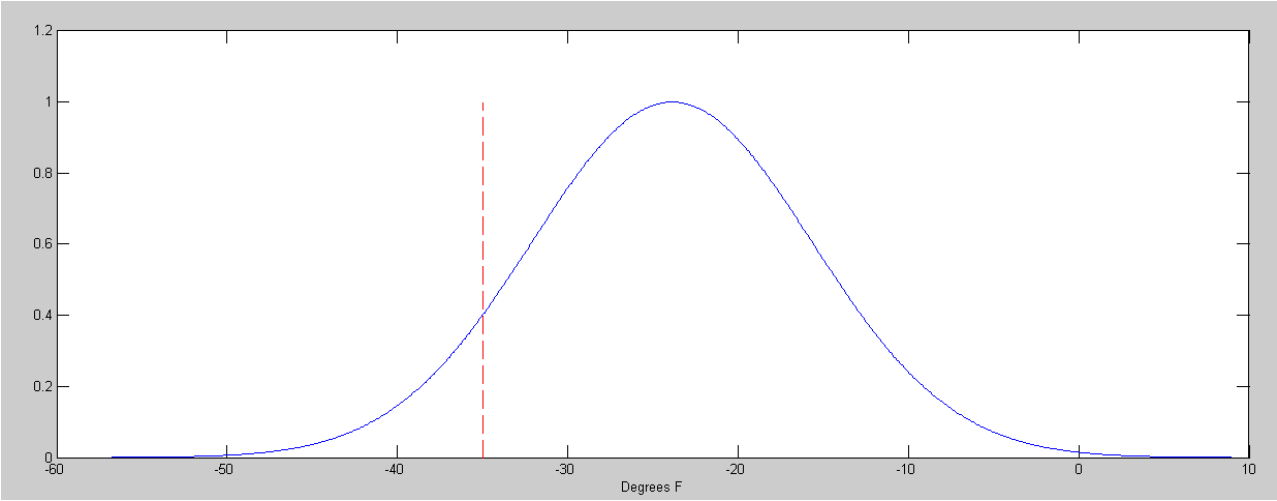
$$z = \left(\frac{-35F - (-23.8725)}{8.2179}\right)$$

$$z = 1.3541$$

From StatTrek

p = 0.08785

There is an 8.78% chance it will get colder than -35F this coming January



- Enter a value in three of the four textboxes.
- Leave the fourth textbox blank.
- Click the **Calculate** button to compute a value for the fourth textbox.

Standard score: z	-1.3541
Probability: P(Z≤-1.3541)	0.08785
Mean	0
Standard deviation	1

Calculate

3) What is the probability that there will be a killing frost in June (temperature gets below 28F)?

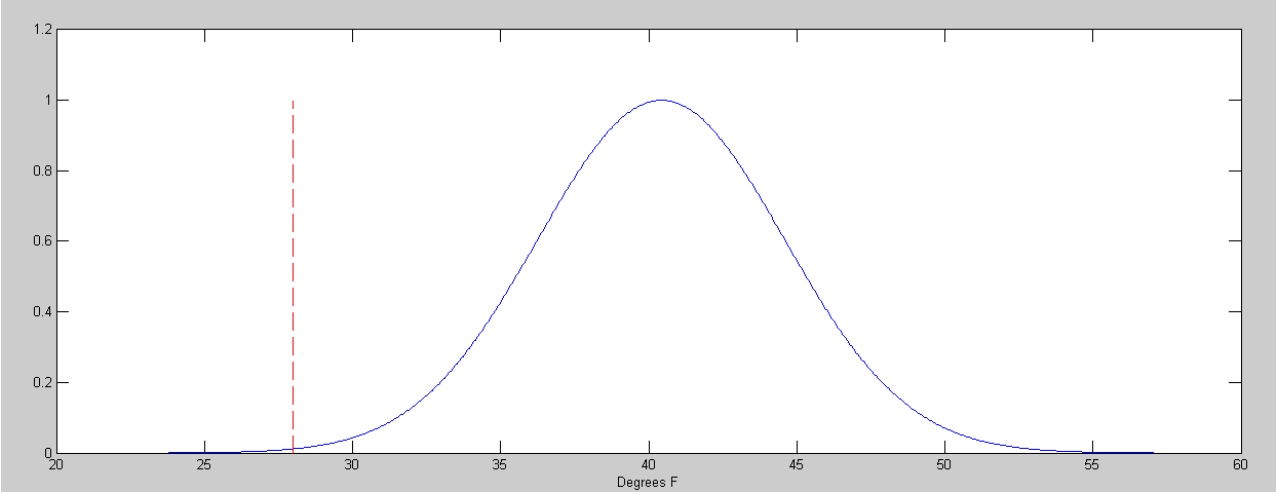
$$z = \left(\frac{28-\mu}{\sigma}\right)$$

$$z = \left(\frac{28-40.425}{4.1576}\right) = 2.9885$$

From StatTrek

p = 0.0014

There is a 0.14% chance of a killing frost in June (714 : 1 odds against)



- Enter a value in three of the four textboxes.
- Leave the fourth textbox blank.
- Click the **Calculate** button to compute a value for the fourth textbox.

Standard score: z	-1.3541
Probability: P(Z≤-1.3541)	0.08785
Mean	0
Standard deviation	1

Calculate

Testing with Two Populations

4) What is the probability that February will be colder than January?

Create a new variable, W

A = January

B = February

W = A - B

The mean and standard deviation are then

$$\mu_w = \mu_a - \mu_b$$

$$\mu_w = (-23.8275) - (-20.6328)$$

$$\mu_w = -3.2397$$

$$\sigma_w^2 = \sigma_a^2 + \sigma_b^2$$

$$\sigma_w^2 = (8.2179)^2 + (7.8559)^2$$

$$\sigma_w^2 = 129.249$$

$$\sigma_w = 11.3688$$

The z-score for W>0 is

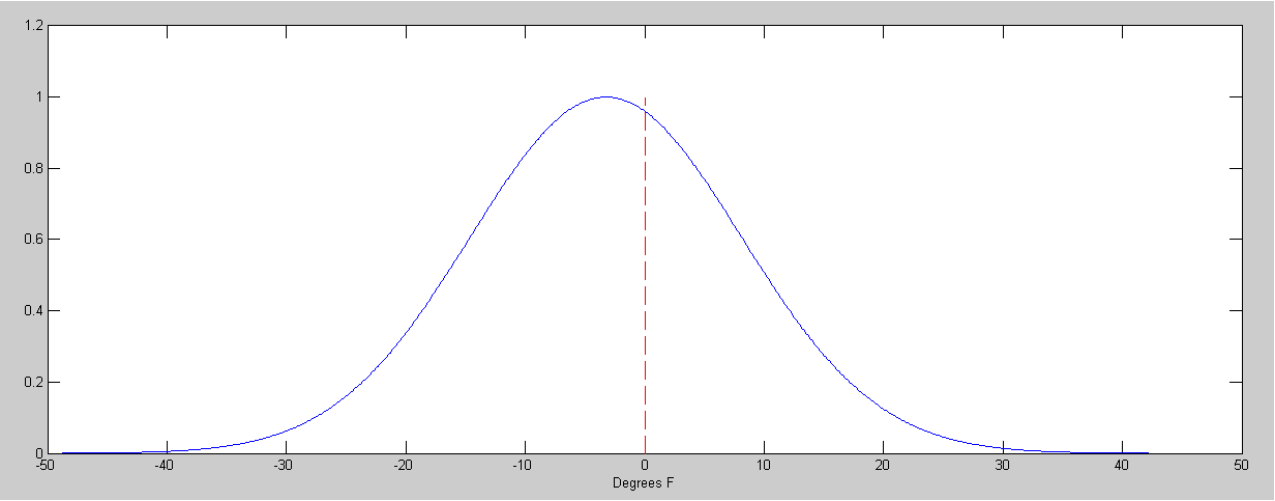
$$z = \left(\frac{\mu - 0}{\sigma}\right) = \left(\frac{-3.2397}{11.3688}\right)$$

$$z = -0.2850$$

This corresponds to a probability of 0.38782

There is a 38.782% chance that February will be colder than January

There is a 38.782% chance that January will be warmer than February (same thing)



W = January - February

The low for 20 months are as follows:

```
{ 5, 10, 13, 15, 17, 17, 20, 22, 23, 24, 25, 25, 28, 28, 30, 32, 33, 33, 35, 37 }
```

5) Which months are September and which ones are October? What threshold do you use for separating the data?

Use the midpoint

September low = 30.6763F

October low = 19.15F

Set the threshold to the midpoint (somewhat arbitrary)

Threshold = 24.91315F

above = September

below = October

6) With your threshold, what is the probability of

- A false positive ? (the temperature was assigned to September but actually came from October)
- A false negative? (the temperature was assigned to October but actually came from September)

False Positive:

Month is October but it was assigned to September (low was more than 24.91F)

The z-score is

$$z = \left(\frac{24.91315 - \mu}{\sigma}\right)$$

$$z = \left(\frac{24.91315 - 19.15}{5.5212}\right) = 1.0438$$

From StatTrek, the tail is 0.14829

There is a 14.829% chance of a false positive

False Negative

Month = September and it was assigned to October (low was less than 24.91F)

The z-score is

$$z = \left(\frac{24.91315 - \mu}{\sigma}\right)$$

$$z = \left(\frac{24.91315 - 30.6763}{4.8861}\right) = -1.1795$$

From StatTrek, the tail has an area of 0.1191

There is an 11.795% chance of a false negative

Note: You could use a differen threshold and make these two probabilities match if you wanted to...

Regression Analysis

The average temperature in March in Fargo, ND is available at

`http://www.bisonacademy.com/ECE111/Code/Fargo_Weather_Monthly_Avg.txt`

7) Find the least-squares curve fit for this data as

$$T = ay + b$$

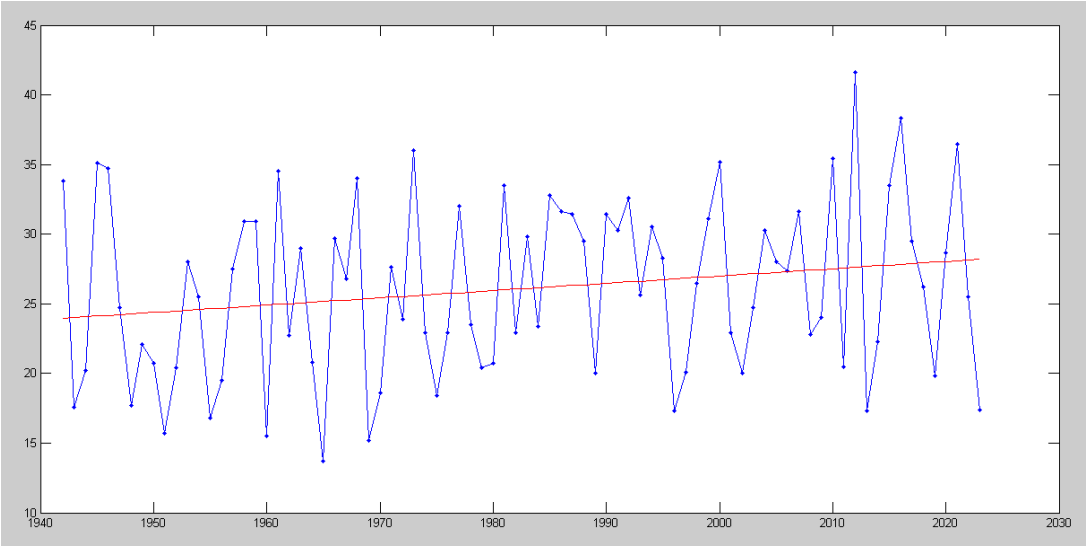
where T is the temperature in degrees F and y is the year.

```
>> DATA = [ <paste> ];
>> year = DATA(:,1);
>> March = DATA(:,4);
>> B = [year, year.^0];
>> A = inv(B'*B)*B'*March

a    0.0523
b  -77.6216

>> plot(year,March,'b.-',year,B*A,'r')
>>
```

From this curve fit, how much has March in Fargo warmed up since 1942?

There are 81 years worth of data. The temperature went up 4.2370F over these 81 years

```
>> a = A(1);
>> dT = a * 81

dT =    4.2370
```



8) Determine the correlation coefficient between

February vs. March

```
>> Feb = DATA(:,3);
>> Mar = DATA(:,4);
>> Jul = DATA(:,8);
>> Cov = mean(Feb .* Mar) - mean(Feb)*mean(Mar)

Cov =   14.4248

>> Correlation = Cov / (std(Feb) * std(Mar))

Correlation =    0.3370
```

There is a 33.70% correlation between February and March

- ***If February is warm, it is 33% more likely that March will also be warm***

March vs. July

```
>> Cov = mean(Mar .* Jul) - mean(Mar)*mean(Jul)

Cov =    1.0230

>> Correlation = Cov / (std(Mar) * std(Jul))

Correlation =    0.0639
```

There is a 6.39% correlation between March and July

- ***If March is warm, it is 6% more likely that July will also be warm (not much correlation)***