ECE 341 - Test #3

Markov Chains and Data Analysis

Open-Book, Open Notes. Calculators, Matlab, Tarot cards, Internet allowed. Just not other people.

Please sign if possible (i.e. you did not get help from someone else).

No aid given, received, or observed: _____

Due Monday, June 15th, 8am

Please make the subject "ECE 341 Test3" if submitting homework electronically to Jacob_Glower@yahoo.com (or on blackboard)

1) Markov Chains: Two people, A and B, are playing a game.

- A has a 60% chance of winning A gains +1 point on a win
- There is a 15% chance of a tie A loses one point on a tie
- A has a 25% chance of losing A loses two points on a loss

If A reaches +3 points, A wins the match

If A reaces -3 points, B wins the match

1a) What is the state transition matrix (going from k games to k+1 games)

$$\begin{bmatrix} p3(k+1) \\ p2(k+1) \\ p1(k+1) \\ e(k+1) \\ m1(k+1) \\ m2(k+1) \\ m3(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0.6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0 & 0 & 0 \\ 0 & 0.15 & 0 & 0.6 & 0 & 0 & 0 \\ 0 & 0.25 & 0.15 & 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0.25 & 0.15 & 0 & 0.6 & 0 \\ 0 & 0 & 0 & 0.25 & 0.15 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.4 & 1 \end{bmatrix} \begin{bmatrix} p3(k) \\ p2(k) \\ p1(k) \\ e(k) \\ m1(k) \\ m2(k) \\ m3(k) \end{bmatrix}$$

1b) What is the chance that A will win the match assuming they start out at even (A has zero points)?

```
A = [1,0,0,0,0,0,0;0.6,0,0.15,0.25,0,0,0];
A = [A;0,0.6,0,0.15,0.25,0,0];
A = [A;0,0,0.6,0,0.15,0.25,0];
A = [A;0,0,0,0.6,0,0.15,0.25];
A = [A;0,0,0,0,0.6,0,0.4];
A = [A;0,0,0,0,0,0,1];
A = A'
```

```
    1.0000    0.6000         0         0         0         0         0
         0         0    0.6000         0         0         0         0
         0    0.1500         0    0.6000         0         0         0
         0    0.2500    0.1500         0    0.6000         0         0
         0         0    0.2500    0.1500         0    0.6000         0
         0         0         0    0.2500    0.1500         0         0
         0         0         0         0    0.2500    0.4000    1.0000
```

```
A^100

    1.0000    0.8157    0.6402    0.4788    0.3157    0.1894         0
         0    0.0000    0.0000    0.0000    0.0000    0.0000         0
         0    0.0000    0.0000    0.0000    0.0000    0.0000         0
         0    0.0000    0.0000    0.0000    0.0000    0.0000         0
         0    0.0000    0.0000    0.0000    0.0000    0.0000         0
         0    0.0000    0.0000    0.0000    0.0000    0.0000         0
         0    0.1843    0.3598    0.5212    0.6843    0.8106    1.0000
```

A has a 47.88% chance of winning the match

1c) What is the z-transform for the probability of A winning after k games?

```
X0 = [0;0;0;1;0;0;0];
C = [1,0,0,0,0,0,0];
G = ss(A,X0,C,0,1);
zpk(G)


                      0.216 (z+0.3) (z-0.3)
-------------------------------------------------------------------
(z-1) (z-0.7933) (z-0.2547) (z+0.2306) (z^2 + 0.8174z + 0.3477)

Sampling time (seconds): 1
```

Times z to get the z-transform

```
                      0.216 (z+0.3) (z-0.3) z
-------------------------------------------------------------------
(z-1) (z-0.7933) (z-0.2547) (z+0.2306) (z^2 + 0.8174z + 0.3477)
```

1d) From the z-transform, determine the probability of A winning

$$\left(\frac{0.216(z+0.3)(z-0.3)z}{(z-0.7933)(z-0.2547)(z+0.2306)(z^2+0.8174z+0.3477)}\right)_{z=1} = 0.4788$$

2) t-Test (One data set). A Monte-Carlo simulation was run 10 times. Each simulation dealt 10,000 hands for 5-card draw and counted the number of times you got a 3-of-a-kind.

hands = { 788 752 755 800 748 787 777 758 828 796 }

2a) If I run this experiment an 11th time, what number will I get with a confidence level of 90%? (5% tails)

- t = 1.8333

```
DATA = [788   752   755   800   748   787   777   758   828   796 ]

x = mean(DATA)

x =  778.9000

s = std(DATA)

s =   25.7917

[x - 1.833*s,  x + 1.833*s]


  731.6238  826.1762
```

2b) What is the 90% confidence interval for the actual probability of getting 3-of-a-kind with 5-card draw based upon this data?

```
x = mean(DATA)

x =  778.9000

s = std(DATA) / sqrt(10)

s =    8.1560

[x - 1.833*s,  x + 1.833*s]

ans =  763.9500  793.8500
```

3) t-Test (Two data sets): The average temperatures in March and November from 1942 - 2019 in Fargo, ND are:

(web site: http://www.bisonacademy.com/ECE111/Code/Fargo_Weather_Monthly_Avg.txt)

March: { 33.8 17.6 20.2 35.1 34.7 24.7 17.7 22.1 20.7 15.7 20.4 28.0 25.5 16.8 19.5 27.5
30.9 30.9 15.5 34.5 22.7 29.0 20.8 13.7 29.7 26.8 34.0 15.2 18.6 27.6 23.9 36.0 22.9 18.4
22.9 32.0 23.5 20.4 20.7 33.5 22.9 29.8 23.4 32.8 31.6 31.4 29.5 20.0 31.4 30.3 32.6 25.6
30.5 28.3 17.3 20.1 26.5 31.1 35.2 22.9 20.0 24.7 30.3 28.0 27.4 31.6 22.8 24.0 35.4 20.5
41.6 17.3 22.3 33.5 38.3 29.5 26.2 19.8 }

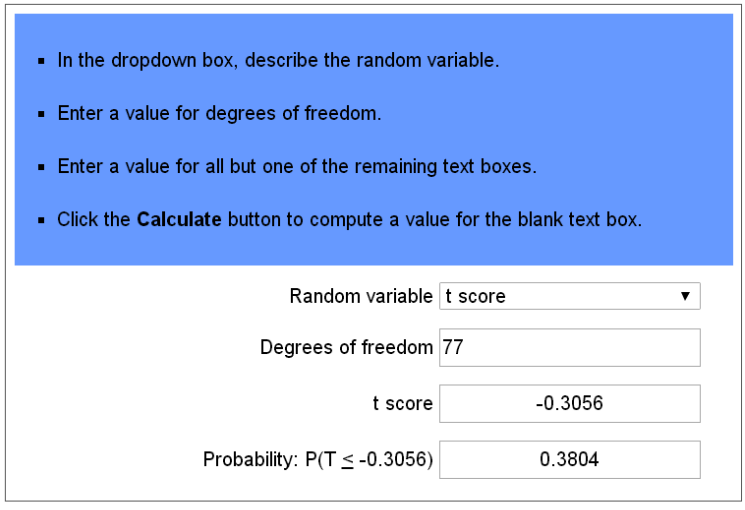
November: { 27.5 29.6 32.9 26.0 27.1 22.2 29.1 34.1 22.3 21.0 30.8 35.1 34.5 18.8 30.5 28.9
29.7 20.5 29.9 30.1 35.0 34.7 29.2 26.0 23.1 29.0 31.0 30.5 27.8 29.6 28.5 25.1 29.2 31.1
23.2 25.6 22.7 24.5 33.0 35.4 24.1 31.2 29.7 15.4 23.1 33.4 27.5 24.0 32.1 22.0 27.3 26.6
34.0 21.2 17.7 23.2 29.3 37.1 25.9 39.7 27.9 24.9 34.2 31.2 31.9 31.1 31.7 38.4 29.2 33.8
30.3 28.2 22.8 36.2 41.8 28.3 22.4 22.0 }

3a) What is the probability that March 2021 will be warmer than November 2021?

```
Xw = mean(March) - mean(Nov)

Xw =   -2.5013

Sw = sqrt( var(March) + var(Nov) )

Sw =    8.1856

t = Xw / Sw

t =   -0.3056
```

From StatTrek, this t-score with 77 degrees of freedom (78 data points) corresponds to a probability of 0.3084

There is a 38.04% chance that March 2021 will be warmer than November 2021

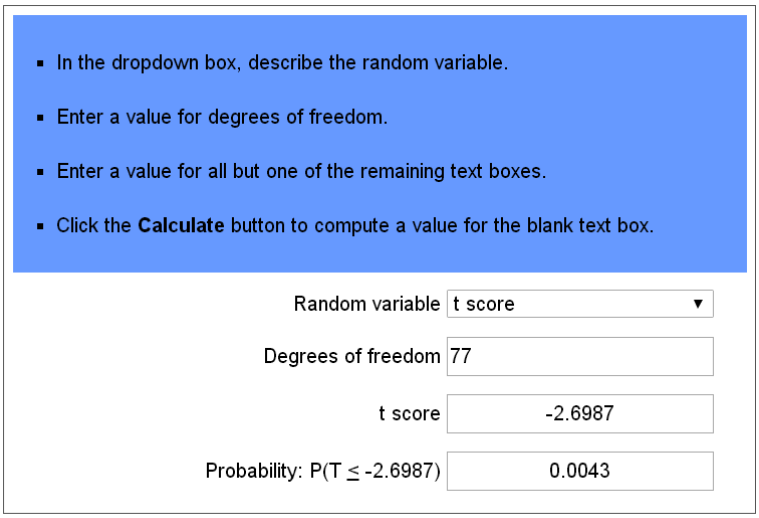


3b) What is the probability that March is warmer than November?

```
Xw = mean(March) - mean(Nov)

Xw =   -2.5013

Sw = sqrt( var(March)/78 + var(Nov)/78 )

Sw =    0.9268

t = Xw / Sw

t =   -2.6987
```

From StatTrek, a t-score of -2.6987 corresponds to a probability of 0.0043

There is a 0.43% chance that March is warmer than November



4) Chi-Squared Test: The following Matlab code generated 100 random values for X:

```
X = zeros(1,100);
for i=1:100
    X(i) = sum(-5*log(1 - rand(3,1)));
    end
```

It is conjectured that X has an exponential distribution with a mean of 15

$$f(x) = \left(\frac{1}{15}\right) e^{-t/15}\, u(t)$$

4a) Generate 100 values for X and give the sorted results (Matlab command *sort(X)*)

```
X = sort(X)
```

```
3.1138 4.4660 4.7383 4.7624 4.7786 4.8436 4.8866 4.9346 5.0978
5.7065 6.2805 6.4941 6.8421 6.8523 7.0184 7.0591 7.2585 7.3949
7.5951 7.7765 7.9319 7.9553 8.3896 8.4584 8.5460 8.5630 8.8428
9.1138 9.1169 9.1393 9.5615 9.6115 9.9383 10.1052 10.2525 10.2809
10.4788 10.6490 10.8053 11.0624 11.3872 11.7600 11.8243 11.8715 12.1728
12.3226 12.3487 12.4203 12.9772 12.9952 13.2287 13.4827 13.4863 13.5290
13.5295 14.0771 14.2606 14.4555 14.4559 14.5053 15.3895 15.3941 15.4975
15.6418 15.7049 15.7082 15.7389 16.2524 17.0743 17.1574 17.1707 17.9278
18.4223 18.9780 19.1629 19.1834 19.3254 19.7832 19.7932 20.4452 20.8665
21.7754 21.9181 22.1857 22.4459 24.3366 24.7308 24.7654 24.8263 24.9744
25.5571 26.5399 26.6183 27.1296 27.2432 27.4440 29.2756 30.0066 31.5335
66.2146
```

4b) Determine if X does or does not have this exponential pdf using a Chi-squared test.

Split the data into N bins. Space the bins every 5 seconds (somewhat arbitrary)

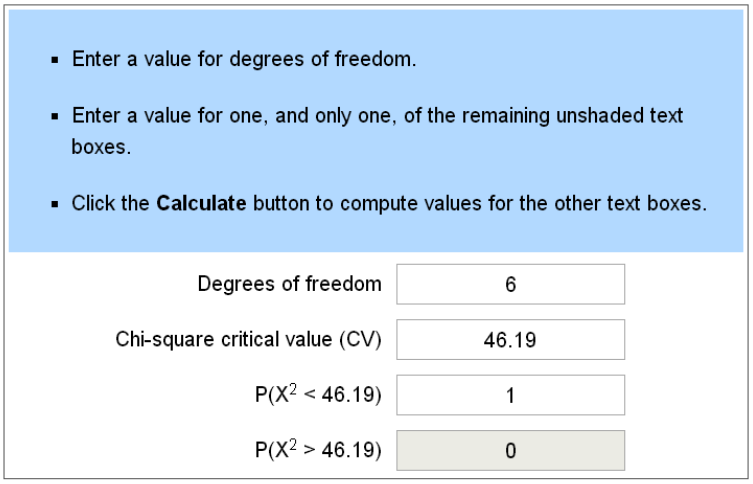
0-5	5-10	10-15	15-20	20 - 25	25 - 30	30 +
3.1138	5.0978	10.1052	15.3895	20.4452	25.5571	31.5335
4.4660	5.7065	10.2525	15.3941	20.8665	26.5399	66.2146
4.7383	6.2805	10.2809	15.4975	21.7754	26.6183	
4.7624	6.4941	10.4788	15.6418	21.9181	27.1296	
4.7786	6.8421	10.6490	15.7049	22.1857	27.2432	
4.8436	6.8523	10.8053	15.7082	22.4459	27.4440	
4.8866	7.0184	11.0624	15.7389	24.3366	29.2756	
4.9346	7.0591	11.3872	16.2524	24.7308	30.0066	
	7.2585	11.7600	17.0743	24.7654		
	7.3949	11.8243	17.1574	24.8263		
	7.5951	11.8715	17.1707	24.9744		
	7.7765	12.1728	17.9278			
	7.9319	12.3226	18.4223			
	7.9553	12.3487	18.9780			
	8.3896	12.4203	19.1629			
	8.4584	12.9772	19.1834			
	8.5460	12.9952	19.3254			
	8.5630	13.2287	19.7832			
	8.8428	13.4827	19.7932			
	9.1138	13.4863				
	9.1169	13.5290				
	9.1393	13.5295				
	9.5615	14.0771				
	9.6115	14.2606				
	9.9383	14.4555				
		14.4559				
		14.5053				

Compute the chi-squred score:

bin	p	np	N	
0 - 5	0.2835	28.35	8	14.6075
5 - 10	0.2031	20.31	25	1.0830
10 - 15	0.1455	14.55	27	10.6531
15 - 20	0.1043	10.43	19	7.0417
20 - 25	0.0747	7.47	11	1.6681
25 - 30	0.0535	5.35	8	1.3126
30+	0.1353	13.53	2	9.8256
			Total	46.1917

From StatTrek. a chi-squred value of 46.19 corresponds to a probability of 1.000 (> 0.99995)

I am more than 99.9995% certain that this data is not from an exponential distribution



5) F-Test (Three data sets): The average temerature in Fargo since 1942 in the months of June, July, and August are: (http://www.bisonacademy.com/ECE111/Code/Fargo_Weather_Monthly_Avg.txt)

June:

62.6 63.7 65.5 60.3 64.7 61.4 63.8 65.6 65.5 60.7 67.3 64.9 65.3 65.4 70.7 62.7 59.7 68.1 63.3 68.6 66.2 68.8 67.2 63.9 66.1 62.6 64.1 57.3 67.8 67.4 66.9 64.7 64.4 65.2 68.5 66.5 64.7 65.4 65.7 62.8 59.1 66.0 65.7 60.0 67.5 69.1 73.8 64.1 66.9 70.1 61.9 63.1 68.2 71.4 67.0 68.9 63.4 66.3 62.7 65.9 69.0 65.4 62.5 68.2 68.5 69.8 63.6 63.6 66.5 66.8 69.9 67.5 67.2 67.1 68.8 67.8 70.3 67.1

July:

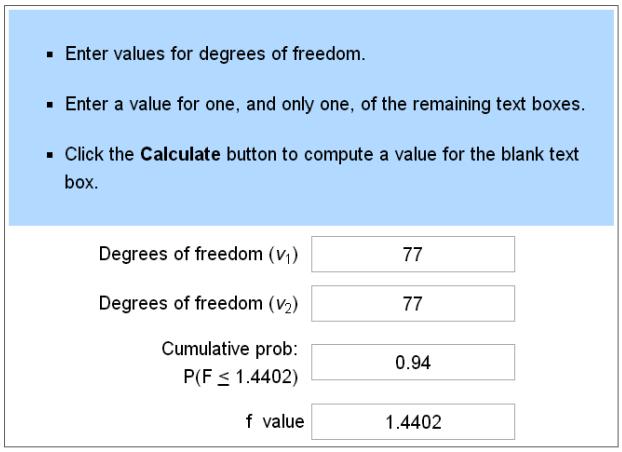
68.8 73.0 69.3 68.6 71.2 71.5 71.2 71.3 67.9 68.5 69.7 69.9 71.8 74.5 68.0 75.3 67.3 72.2 71.4 70.4 68.8 73.6 74.0 68.5 73.8 67.9 69.6 68.3 71.8 65.0 68.3 68.2 73.6 74.2 71.7 72.2 69.5 71.9 71.8 71.1 70.9 73.5 70.6 68.9 71.5 74.0 75.8 75.8 70.0 70.2 64.3 67.0 67.6 70.0 67.8 69.2 71.7 71.5 70.6 72.5 73.0 70.4 68.2 71.3 74.9 74.0 70.3 66.5 72.0 74.5 76.6 71.7 69.4 72.5 71.9 72.1 70.9 72.2

5a) What is the probability that the variance of June is different than the variance of July?

```
F = var(June) / var(July)

F = 1.4402
```

Frrom StatTrek, this corresponds to a probability of 0.94

It is 94% likely that June has a higher variance than July



5b) What is the probability that

- June 1942 - 1967
- June 1968 - 1993
- June 1994 - 2019

all have the same average temeprature using an ANOVA test?

```
A = June(1:26);
B = June(27:52);
C = June(53:78);

Na = length(A);
Nb = length(B);
Nc = length(C);
Xa = mean(A);
Xb = mean(B);
Xc = mean(C);
Na = length(A);
Nb = length(B);
Nc = length(C);
G = mean([A,B,C])

k = 3;
N = Na + Nb + Nc

MSSb = (Na*(Xa-G)^2 + Nb*(Xb-G)^2 + Nc*(Xc-G)^2) / (k-1)

MSSw = ((Na-1)*var(A) + (Nb-1)*var(B) + (Nc-1)*var(C)) / (N-k)

F = MSSb / MSSw

G =   65.7269
N =    78
MSSb =   36.2458
MSSw =    8.4158
F =    4.3069
```

From StatTrek, it is 98% certain that these three time periods do **not** have the same average temperature