Examination Number:_________________________

Sign the Honor Pledge Below_____________
____

Last Name______________First__________
PID #________________________________
Write Your Section Number here:_________

University of North Carolina
Economics 400: Economic Statistics
Practice Final Examination

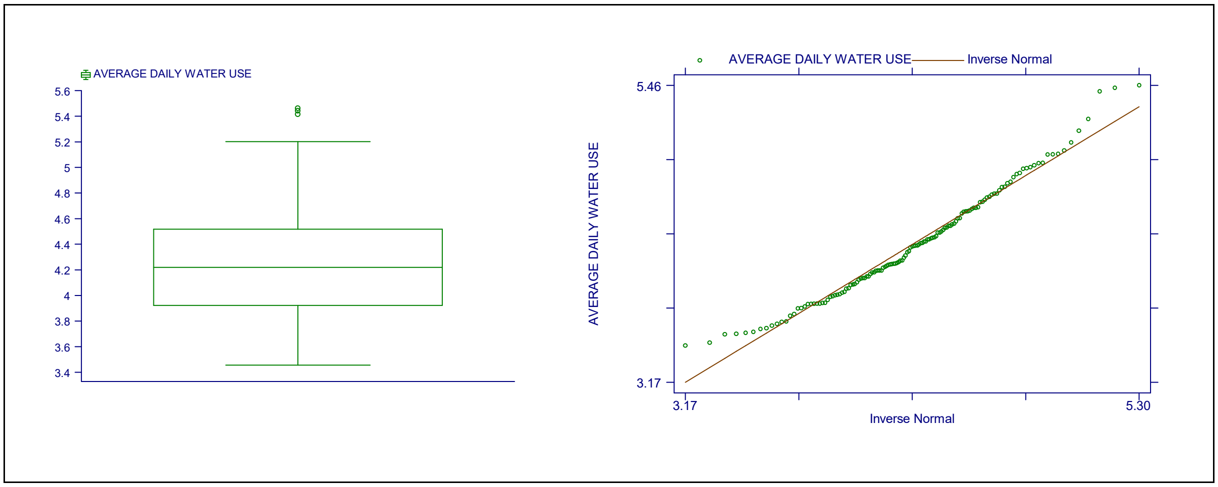
Prof. B. Turchi | Spring 2019

General Instructions: Answer all eleven (11) questions on this examination, writing your answers on the exam paper itself. Use the back of the pages for any work, if necessary. Sign the Honor Pledge above. **Be sure to note that tables and formulas are on the last page of the exam. The Equations on the actual final will be identical to those on the *second* *midterm*. .**

1. (8 points) Prove that the sample proportion, $\hat{p}$, computed from a random sample of size n is an unbiased estimator of the population proportion, p.

2. (8 points) Based on its experience with Hurricane Fran five years ago, Duke Power knows that, seven days after the onset of a power outage such as we are now experiencing, only 5% of homes will still be without power. Assuming that the random variable "number of days without power" follows the exponential distribution, what is the *mean number of days* that a home will be without power?

3. (12 points)The following graphs describe average daily water use (measured monthly) in Concord, Mass. between January, 1970 and May, 1981.



(a) Write the command that will produce the Stata 15 version of the graph on your left:

(b) Write the command that will produce the Stata 15 version of the graph on your right:

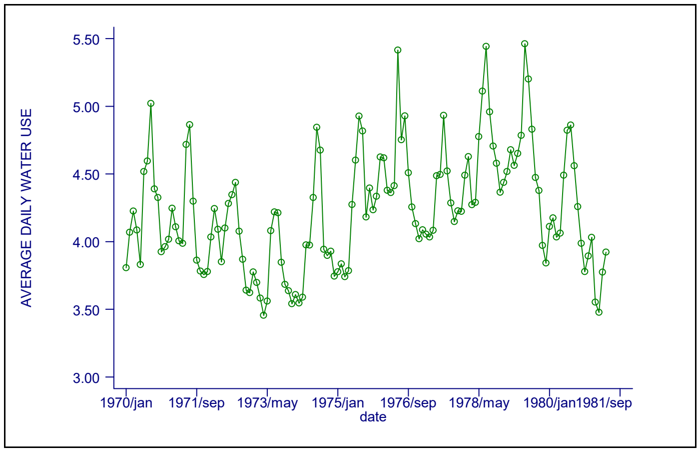
(c) What is the median value of average daily water use?

(d) Does the boxplot suggest that the distribution of data is symmetrical or skewed? Be detailed and justify your answer.

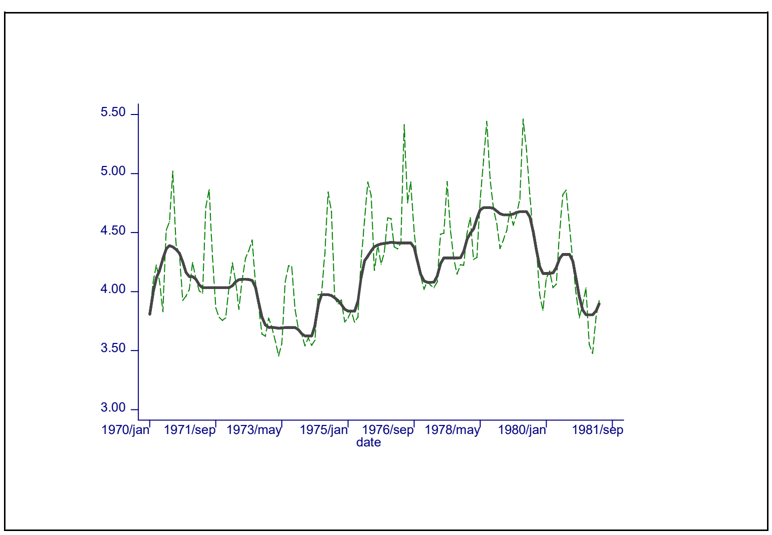
(e) What does the right-hand graph suggest about the shape of the distribution of these data?

(f) Do the two graphs agree or disagree about the shape of the data? Why/Why not?

4. (9 points) The graph below shows the time series of average daily water use in Concord, Mass. Based on your understanding of the components of time series, which of the following components does this series exhibit? seasonal, trend, irregular, cyclical. Explain your answer.



5. (10 points) (a) Looking at the graph of average daily water use in Concord, Mass. below, what does the solid line represent and what kind of Stata command might have been used to produce it? (b) What components of the time series does the solid line represent?



6. (6 points) The following Stata output shows the regression model:

$AvgWaterUse = \alpha + \beta \cdot AvgRainFall + \varepsilon$. Answer the following questions using these results as a basis for your answers.

```
. regress  h2ouse rain

      Source |       SS       df       MS              Number of obs =     137
-------------+------------------------------           F(  1,   135) =    1.57
       Model |  .296179057     1  .296179057           Prob > F      =  0.2117
    Residual |  25.3931185   135  .188097174           R-squared     =  0.0115
-------------+------------------------------           Adj R-squared =  0.0042
       Total |  25.6892976   136  .188891894           Root MSE      =   .4337

------------------------------------------------------------------------------
      h2ouse |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        rain |  -.0319033   .0254243    -1.25   0.212    -.0821848    .0183782
       _cons |   4.331621   .0854821    50.67   0.000     4.162564    4.500679
------------------------------------------------------------------------------
```

(a) Draw a causal diagram and explain what this regression is attempting to discover.

(b) What proportion of the total variation in water use does rainfall explain?

(c) Does rainfall have a significant impact on the level of water use? Why/Why Not?

7. (6 points) Here's another regression relating average daily water use to average daily temperature.

```
      Source |       SS       df       MS              Number of obs =     137
-------------+------------------------------           F(  1,   135) =   52.64
       Model |  7.20645737     1  7.20645737           Prob > F      =  0.0000
    Residual |  18.4828402   135  .136909927           R-squared     =  0.2805
-------------+------------------------------           Adj R-squared =  0.2752
       Total |  25.6892976   136  .188891894           Root MSE      =  .37001

------------------------------------------------------------------------------
      h2ouse |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        temp |   .0125993   .0017366     7.26   0.000     .0091648    .0160338
       _cons |   3.675907   .0832886    44.13   0.000     3.511187    3.840626
------------------------------------------------------------------------------
```

(a) Draw a causal diagram and write the equation that represents this equation.

(b) What proportion of the total variation in average daily water use does this regression represent?

(c) Does daily temperature have a significant impact on water use? Why?/Why not?

8. (15 points) In 1997 the US Department of Agriculture reported that the average American consumed 28.0 lbs. of cheese. Last year, a randomly selected group of 35 Americans consumed the following amounts of cheese:

40	23	27	32	36	34	28
27	26	30	22	41	20	36
30	26	39	18	33	22	27
38	27	20	31	21	25	30
31	31	30	16	38	30	23

At the 5% significance level do the data provide sufficient evidence to conclude that last year's mean cheese consumption for Americans is different from the 1997 mean? Show your work and explain your procedure and the assumption(s) you make to find your answer.

9. (15 points) The UNC Career Services Center conducts surveys on the starting salaries of college graduates from UNC. The following table gives the starting annual salaries obtained from independent random samples of 35 liberal arts and 32 accounting graduates:

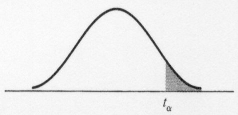
Liberal arts					Accounting			
33.0	29.8	34.3	33.6	34.0	35.9	31.4	34.3	31.9
31.7	36.8	31.3	30.7	31.0	34.4	36.3	37.8	33.6
32.1	34.1	32.6	32.0	31.7	35.0	33.0	32.7	34.0
33.8	33.4	30.1	32.5		37.2	36.9	35.8	36.3
32.9	32.2	32.1	30.2		36.5	31.4	33.4	35.5
31.3	35.7	33.0	32.2		35.9	36.4	36.8	35.2
33.9	32.1	33.8	33.5		33.2	33.1	32.0	35.7
34.4	29.3	29.3	33.5		32.5	33.5	36.4	37.6

At the 5% significance level, can you conclude that the mean starting salaries of liberal arts and accounting graduates differ? State your assumptions and show your work.

10. (5 points) In a least squares regression, what relationship do the mean of the dependent variable and the mean of the independent variable have with the estimated regression line?

11. (6 points) What is/are the sufficient statistics of (a) the normal distribution, (b) the exponential distribution and (c) the uniform distribution?

Equation Sheet follows: Remember, the actual equation sheets will be identical to Midterm 2



d.f.	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	d.f.
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
inf.	1.282	1.645	1.960	2.326	2.576	inf.

Source: From "Table of Percentage Points of the *t*-Distribution." *Biometrika*, Vol. 32 (1941), p. 300. Reproduced by permission of the Biometrika Trustees.

$$t=\frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{s_p\sqrt{\left(1/n_1\right)+\left(1/n_2\right)}}$$

$$df=n_1+n_2-2.$$

$$s_p=\sqrt{\frac{\left(n_1-1\right)s_1^2+\left(n_2-1\right)s_2^2}{n_1+n_2-2}}$$

$$\mathrm{t}=\frac{\left(\bar{x}_1-\bar{x}_2\right)-\left(\mu_1-\mu_2\right)}{\sqrt{\left(s_1^2/n_1\right)+\left(s_2^2/n_2\right)}}$$

$$\Delta=\frac{\left[\left(s_1^2/n_1\right)+\left(s_2^2/n_2\right)\right]^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1-1}+\frac{\left(s_2^2/n_2\right)^2}{n_2-1}}\text{ rounded down.}$$

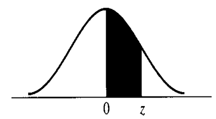
$$\hat{p}_p=\frac{x_1+x_2}{n_1+n_2}$$

$$z=\frac{\hat{p}-p_0}{\sqrt{\frac{p_0\left(1-p_0\right)}{n}}}$$

$$z=\frac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}_p\left(1-\hat{p}_p\right)}\bullet\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$\Pr\left(-t_{\alpha/2}<\frac{\bar{x}-\mu}{s/\sqrt{n}}<t_{\alpha/2}\right)=1-\alpha$$

Normal Curve Areas



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Source: This table is abridged from Table 1 of *Statistical Tables and Formulas*, by A. Hald (New York: John Wiley & Sons, Inc., 1952). Reproduced by permission of A. Hald and the publishers, John Wiley & Sons, Inc.

$$n=\frac{z_{\alpha/2}^2\sigma^2}{D^2}.\qquad n=\frac{z_{\alpha/2}^2}{4D^2}$$

Binomial Coefficients

n	$\binom{n}{0}$	$\binom{n}{1}$	$\binom{n}{2}$	$\binom{n}{3}$	$\binom{n}{4}$	$\binom{n}{5}$	$\binom{n}{6}$	$\binom{n}{7}$	$\binom{n}{8}$	$\binom{n}{9}$	$\binom{n}{10}$
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
0	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1
11	1	11	55	165	330	462	462	330	165	55	11
12	1	12	66	220	405	792	924	792	495	220	66
13	1	13	78	286	715	1287	1716	1716	1287	715	286
14	1	14	91	364	1001	2002	3003	3432	3003	2002	1001
15	1	15	105	455	1365	3003	5005	6435	6435	5005	3003
16	1	16	120	560	1820	4368	8008	11440	12870	11440	8008
17	1	17	136	680	2380	6188	12376	19448	24310	24310	19448
18	1	18	153	816	3060	8568	18564	31824	43758	48620	43758
19	1	19	171	969	3876	11628	27132	50388	75582	92378	92378
20	1	20	190	1140	4845	15304	38760	77520	125970	167960	184756

If necessary, use the identity $\binom{n}{k} = \binom{n}{n-k}$.

$$f(\tilde{x}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$$

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$$

$$p(x) = \frac{C_x^r C_{n-x}^{N-r}}{C_n^N} = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

Mean: $\mu = n\left(\frac{r}{N}\right)$

Variance: $\sigma^2 = n\left(\frac{r}{N}\right)\left(\frac{N-r}{N}\right)\left(\frac{N-n}{N-1}\right)$

Standard deviation: $\sigma = \sqrt{\sigma^2}$

$$f(x) = \begin{cases} \frac{1}{(b-a)}, & a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

$$\mu = \frac{1}{2}(b+a) \text{ and } \sigma = \frac{(b-a)}{\sqrt{12}}$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \lambda > 0, x \ge 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\mu = \frac{1}{\lambda} \text{ and } \sigma = \frac{1}{\lambda}$$

$$P(x \ge a) = e^{-\lambda a}, a \ge 0 \text{ and } \lambda > 0$$

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{\text{all } k} P(B|A_k)P(A_k)}$$

$$F_{(n_A-1;n_B-1)} = \frac{s_A^2/\sigma_A^2}{s_B^2/\sigma_B^2}$$

$$\chi^2_{(n-1)} = \frac{n-1}{\sigma_0^2}s^2$$

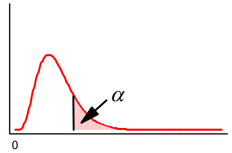
$$P(x) = \begin{cases} \frac{e^{-\lambda t}(\lambda t)^x}{x!}, & \text{for } x = 0,1,2,,,\infty, \quad \lambda > 0, \\ 0, & \text{otherwise.} \end{cases}$$

λ = the mean number of events in a given segment of time ($t = 1$)

t = the length of a particular subsegment ($t \le 1$)

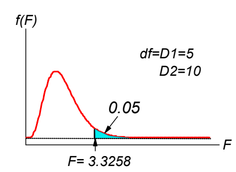
$E[x] = \mu_x = \lambda t$ = the expected number of events in one subsegment length t

Critical values of χ^2_α



df	$\chi^2_{0.995}$	$\chi^2_{0.99}$	$\chi^2_{0.975}$	$\chi^2_{0.95}$	$\chi^2_{0.90}$	$\chi^2_{0.10}$	$\chi^2_{0.05}$	$\chi^2_{0.025}$	$\chi^2_{0.001}$	$\chi^2_{0.005}$	df
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879	1
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597	2
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838	3
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860	4
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750	5
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548	6
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278	7
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955	8
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589	9
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188	10
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757	11
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300	12
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819	13
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319	14
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801	15
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267	16
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718	17
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156	18
19	6.844	7.633	8.907	10.117	11.651	27.204	30.143	32.852	36.191	38.582	19
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997	20
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401	21
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.290	42.796	22
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181	23
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559	24
25	10.520	11.524	13.120	14.611	16.473	34.382	37.653	40.647	44.314	46.928	25
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290	26
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645	27
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.994	28
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336	29
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672	30
40	20.707	22.164	24.433	26.509	29.051	51.805	55.759	59.342	63.691	66.767	40
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490	50
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.381	91.955	60
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.424	104.213	70
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.628	112.328	116.320	80
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.135	124.115	128.296	90
100	67.328	70.065	74.222	77.930	82.358	118.499	124.343	129.563	135.811	140.177	100

F-Distribution Table: Upper 5% Probability (or 5% Area) under F-distribution Curve



			F-Table for alpha = 0.05																
/	df1=1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	inf
df2=1	161.4476	199.5	215.7073	224.5832	230.1619	233.986	236.7684	238.8827	240.5433	241.8817	243.906	245.9499	248.0131	249.0518	250.0951	251.1432	252.1957	253.2529	254.3144
2	18.5128	19	19.1643	19.2468	19.2964	19.3295	19.3532	19.371	19.3848	19.3959	19.4125	19.4291	19.4458	19.4541	19.4624	19.4707	19.4791	19.4874	19.4957
3	10.128	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123	8.7855	8.7446	8.7029	8.6602	8.6385	8.6166	8.5944	8.572	8.5494	8.5264
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.041	5.9988	5.9644	5.9117	5.8578	5.8025	5.7744	5.7459	5.717	5.6877	5.6581	5.6281
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725	4.7351	4.6777	4.6188	4.5581	4.5272	4.4957	4.4638	4.4314	4.3985	4.365
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.099	4.06	3.9999	3.9381	3.8742	3.8415	3.8082	3.7743	3.7398	3.7047	3.6689
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.866	3.787	3.7257	3.6767	3.6365	3.5747	3.5107	3.4445	3.4105	3.3758	3.3404	3.3043	3.2674	3.2298
8	5.3177	4.459	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881	3.3472	3.2839	3.2184	3.1503	3.1152	3.0794	3.0428	3.0053	2.9669	2.9276
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789	3.1373	3.0729	3.0061	2.9365	2.9005	2.8637	2.8259	2.7872	2.7475	2.7067
10	4.9646	4.1028	3.7083	3.478	3.3258	3.2172	3.1355	3.0717	3.0204	2.9782	2.913	2.845	2.774	2.7372	2.6996	2.6609	2.6211	2.5801	2.5379
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.948	2.8962	2.8536	2.7876	2.7186	2.6464	2.609	2.5705	2.5309	2.4901	2.448	2.4045
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964	2.7534	2.6866	2.6169	2.5436	2.5055	2.4663	2.4259	2.3842	2.341	2.2962
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144	2.671	2.6037	2.5331	2.4589	2.4202	2.3803	2.3392	2.2966	2.2524	2.2064
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458	2.6022	2.5342	2.463	2.3879	2.3487	2.3082	2.2664	2.2229	2.1778	2.1307
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876	2.5437	2.4753	2.4034	2.3275	2.2878	2.2468	2.2043	2.1601	2.1141	2.0658
16	4.494	3.6337	3.2389	3.0069	2.8524	2.7413	2.6572	2.5911	2.5377	2.4935	2.4247	2.3522	2.2756	2.2354	2.1938	2.1507	2.1058	2.0589	2.0096
17	4.4513	3.5915	3.1968	2.9647	2.81	2.6987	2.6143	2.548	2.4943	2.4499	2.3807	2.3077	2.2304	2.1898	2.1477	2.104	2.0584	2.0107	1.9604
18	4.4139	3.5546	3.1599	2.9277	2.7729	2.6613	2.5767	2.5102	2.4563	2.4117	2.3421	2.2686	2.1906	2.1497	2.1071	2.0629	2.0166	1.9681	1.9168
19	4.3807	3.5219	3.1274	2.8951	2.7401	2.6283	2.5435	2.4768	2.4227	2.3779	2.308	2.2341	2.1555	2.1141	2.0712	2.0264	1.9795	1.9302	1.878
20	4.3512	3.4928	3.0984	2.8661	2.7109	2.599	2.514	2.4471	2.3928	2.3479	2.2776	2.2033	2.1242	2.0825	2.0391	1.9938	1.9464	1.8963	1.8432
21	4.3248	3.4668	3.0725	2.8401	2.6848	2.5727	2.4876	2.4205	2.366	2.321	2.2504	2.1757	2.096	2.054	2.0102	1.9645	1.9165	1.8657	1.8117
22	4.3009	3.4434	3.0491	2.8167	2.6613	2.5491	2.4638	2.3965	2.3419	2.2967	2.2258	2.1508	2.0707	2.0283	1.9842	1.938	1.8894	1.838	1.7831
23	4.2793	3.4221	3.028	2.7955	2.64	2.5277	2.4422	2.3748	2.3201	2.2747	2.2036	2.1282	2.0476	2.005	1.9605	1.9139	1.8648	1.8128	1.757
24	4.2597	3.4028	3.0088	2.7763	2.6207	2.5082	2.4226	2.3551	2.3002	2.2547	2.1834	2.1077	2.0267	1.9838	1.939	1.892	1.8424	1.7896	1.733
25	4.2417	3.3852	2.9912	2.7587	2.603	2.4904	2.4047	2.3371	2.2821	2.2365	2.1649	2.0889	2.0075	1.9643	1.9192	1.8718	1.8217	1.7684	1.711
26	4.2252	3.369	2.9752	2.7426	2.5868	2.4741	2.3883	2.3205	2.2655	2.2197	2.1479	2.0716	1.9898	1.9464	1.901	1.8533	1.8027	1.7488	1.6906
27	4.21	3.3541	2.9604	2.7278	2.5719	2.4591	2.3732	2.3053	2.2501	2.2043	2.1323	2.0558	1.9736	1.9299	1.8842	1.8361	1.7851	1.7306	1.6717
28	4.196	3.3404	2.9467	2.7141	2.5581	2.4453	2.3593	2.2913	2.236	2.19	2.1179	2.0411	1.9586	1.9147	1.8687	1.8203	1.7689	1.7138	1.6541
29	4.183	3.3277	2.934	2.7014	2.5454	2.4324	2.3463	2.2783	2.2229	2.1768	2.1045	2.0275	1.9446	1.9005	1.8543	1.8055	1.7537	1.6981	1.6376
30	4.1709	3.3158	2.9223	2.6896	2.5336	2.4205	2.3343	2.2662	2.2107	2.1646	2.0921	2.0148	1.9317	1.8874	1.8409	1.7918	1.7396	1.6835	1.6223
40	4.0847	3.2317	2.8387	2.606	2.4495	2.3359	2.249	2.1802	2.124	2.0772	2.0035	1.9245	1.8389	1.7929	1.7444	1.6928	1.6373	1.5766	1.5089
60	4.0012	3.1504	2.7581	2.5252	2.3683	2.2541	2.1665	2.097	2.0401	1.9926	1.9174	1.8364	1.748	1.7001	1.6491	1.5943	1.5343	1.4673	1.3893
120	3.9201	3.0718	2.6802	2.4472	2.2899	2.175	2.0868	2.0164	1.9588	1.9105	1.8337	1.7505	1.6587	1.6084	1.5543	1.4952	1.429	1.3519	1.2539
inf	3.8415	2.9957	2.6049	2.3719	2.2141	2.0986	2.0096	1.9384	1.8799	1.8307	1.7522	1.6664	1.5705	1.5173	1.4591	1.394	1.318	1.2214	1

Answers to Final Exam

1.

$$E\left[\hat{p}\right] = u_{\hat{p}} = E\left[\frac{x}{n}\right] = \frac{1}{n}E\left[x\right] = \frac{1}{n}\times(np) = p$$

Where x is the number of successes in n trials and is distributed binomally. *(np)* is the mean of the binomial distribution.

2. This problem requires the use of the exponential distribution, in particular the formula showing *right tail probabilities:*

$$P\left(x \geq a\right) = e^{-\lambda a} \Rightarrow P\left(x \geq 7\right) = 0.05 = e^{-\lambda 7}$$

$$\ln\left(0.05\right) = -\lambda \times 7 \Rightarrow -\left(\frac{\ln\left(0.05\right)}{7}\right) = \lambda = 0.428$$

$$\mu = \frac{1}{\lambda} = \frac{1}{0.428} = 2.337 \text{ days.}$$

We use the formula for right-tail probabilities to solve for λ, the reciprocal of which is equal to the mean of the distribution.

3. (a) graph box *varname*, ylabel(3.4 3.6 to 5.6) (Stata 13)

(b) qnorm *varname*

(c) 4.2 units

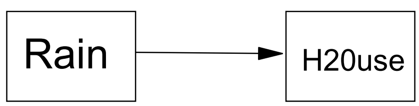
(d) The boxplot suggests that the data, while reasonably symmetrical in the middle is slightly skewed right, given the longer whisker at the upper end and the 3 outliers.

(e) The qnorm plot suggests also that while the data track the normal distribution well in the center, that the tails are thicker than they would be with the normal. The skew appears to be to the right.

(f) The two graphs basically agree because both show a symmetrical center but a longer upper tail .

4. These data show seasonal, irregular and cyclical components. There does not appear to be much of a trend component, but the student might say something like "trend-cycle" which would be OK.
5. (a) The solid line is clearly the result of some sort of smoothing procedure: perhaps a *moving average,* or the *smooth* command.

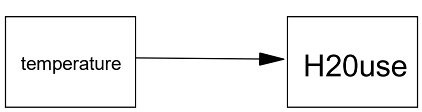
   (b) The solid line represents the cyclical component ... there does not appear to be much of a trend, so if the student says cyclical *and* trend, then take off one point.
6. (a) The causal diagram would look like:



   This regression is attempting to determine if there is a causal relationship between average rainfall and water use by households.

   (b) The $R^2 = 0.0115$, which says that only about 1.1% of the variation in water use is accounted for by the regression.

   (c) No, rainfall appears not to have a significant effect on water use. The coefficient on "rain" is small (-0.0319033) and it has a small t-statistic, which allows us to reject the null hypothesis at only the 21.2% level.
7. (a) The causal diagram would look like:



   And the equation representing this relationship would be:

$$H2OUse = \alpha + \beta \bullet temp + \varepsilon$$

   (b) This regression explains 28.05% of the total variation in water use ($R^2 = 0.2805$)

(c) Yes, temperature has a significant positive effect on water use. The estimated beta coefficient is 0.0125993 with a t-statistic of 7.26, which allows us to reject the null hypothesis at less than the 0.000 level.

8. The student needs to compute the sample mean, and the sample standard deviation:

$$\bar{x} = \frac{\sum_{i=1}^{35} weight_i}{n = 35}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} \left(weight_i - \bar{x}\right)^2}{n-1}}$$

Then he/she needs to compute the t-statistic:

$$t = \frac{\bar{x} - 28.0}{\frac{6.479833}{\sqrt{35}}} = \frac{28.8 - 28.0}{\frac{6.479833}{\sqrt{35}}} = 0.7304$$

Then, using the t-table (or the normal table since the df > 30) we see that one cannot reject the null hypothesis that Americans are consuming the same amount of cheese as before.

Remember, the student should have assumed that the underlying population was normally distributed in order to perform this t-test.

9. Here, the student needs to calculate means and standard errors for both populations (or the pooled standard error if assuming equal population variances):

$$\bar{x}_{la} = 32.51143$$
$$\bar{x}_{ac} = 34.7375$$
$$s_p = 0.4417233 \text{ and } df = 65$$
$$s_{la} = 0.2895889$$
$$s_{ac} = 0.3363331 \text{ and } df = 62.6214 \text{ rounded down}$$

Then, using the formulas for pooled variance two-sample t-statistic or unequal variance t-statistic (see the formulas for both versions on the formula sheet at the back of the test) the student needs to calculate the t-statistic. The student can use either the pooled or the unequal variance version; however, he must do so consistently and correctly.

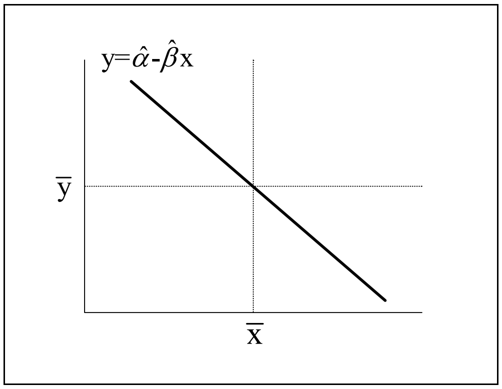
$t = -5.0395$ (for pooled test)

$t = -5.0156$ (for unequal variance version)

In either case, we can easily reject the null hypothesis. Note also, that the student could use the normal distribution, since the $df > 30$.

In either case, the student should assume that both underlying populations are normally distributed, which they appear to be.

10. The regression line passes through the means of both the x- and y-variables.



11. a) Normal Distribution: mean and standard deviation

    b) Exponential Distribution: λ.

    c) Uniform Distribution: *a & b. (i.e., the two endpoints of the distribution.*