Please place your name on Each page now **NAME: ______________________________**

IE 303 Statistics I **Winter, 2002**
Mark S. Daskin **Dept. of IE/MS** **Northwestern University**

Statistics I

Quiz 2

Problem 1: (35 percent)

Suppose you are in the business of producing milk. At the beginning of each hour you select bottles of milk to be tested for purity. Two different testing procedures (A and B) are available. Both procedures yield an unbiased estimate of the number of parts per million (PPM) of a particular contaminant (e.g., lead) that might be in the milk. However, the two procedures differ in their inherent variability and in the amount of time that it takes to do the test.

In particular, assume that the variance of the number of parts per million of the contaminant using method A is 100 (PPM^2) while the variance for method B is 300 (PPM^2). Method A takes 30 minutes per sample and method B takes only 15 minutes per sample.

Thus, every hour we can test 6 bottles of milk: 2 using method A and 4 using method B. Let the observations be $X_1^A, X_2^A, X_1^B, X_2^B, X_3^B$, and X_4^B where, for example, X_3^B denotes the third observation made using method B.

Consider the following estimator of the actual number of parts per million in the milk in any hour:

$$\hat{X} = w_1 X_1^A + w_1 X_2^A + w_2 X_1^B + w_2 X_2^B + w_2 X_3^B + w_2 X_4^B$$

In other words, the two observations taken using method A are each weighted by w_1 while each of the 4 taken using method B are weighted by w_2. Clearly for $\hat{X}$ to be an unbiased estimator of the population level of contaminants, we require $2w_1 + 4w_2 = 1$.

Please place your name on Each page now **NAME:** ____________________________________

a) Find values of w_1 and w_2 so that the variance of $\hat{X}$ is minimized.

Hint: recall that the variance of $\hat{X}$ will be given by

$$Var(\hat{X}) = w_1^2 Var(X_1^A) + w_1^2 Var(X_2^A) + w_2^2 Var(X_1^B) + w_2^2 Var(X_2^B) + w_2^2 Var(X_3^B) + w_2^2 Var(X_4^B)$$

b) Using the values of w_1 and w_2 that you found in part (a), what is the variance of $\hat{X}$?

c) What would be the variance of the unweighted average,

$\overline{X} = \frac{1}{6}\left(X_1^A + X_2^A + X_1^B + X_2^B + X_3^B + X_4^B\right)$?

d) What is the percentage reduction in the variance of the estimator that results from using $\hat{X}$ instead of the unweighted average $\overline{X}$? (i.e. what is the value of the answer to (c) minus (a) all divided by (a) as a percentage?)

Please place your name on Each page now **NAME:** ____________________________________

Problem 2: (35 Percent)

Suppose you are interested in determining the mean household income in a city. To do so, you decide to sample 100 households randomly. Letting X_j be the income of the j^{th} household **measured in $1000**, you obtain the following information.

$$\sum_{j=1}^{100} X_j = 4850 \qquad\qquad \sum_{j=1}^{100} X_j^2 = 241{,}561$$

a) Find the (1) sample average, the (2) sample variance and (3) sample standard deviation of household incomes measured in $1000.

b) Assuming that the household incomes are from a Normal distribution, find a two-sided 95% confidence interval for the true population mean household income. (Again, report your answer in terms of $1000).

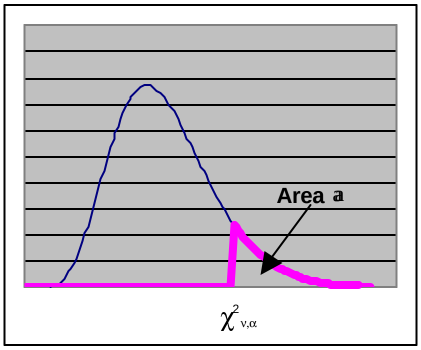
c) Consider the following hypothesis:

$$H_0: \quad m \geq 50$$
$$H_1: \quad m < 50$$

For what values of the sample average would you reject the null hypothesis at $a = 0.05$?
Do you reject the null hypothesis at this level of significance?

d) Suppose the true mean is $47.65 thousand dollars. What is the (approximate) probability of a type II error for the test outlined in part (c) above?

e) Find a 90% confidence interval for the population variance. You might find the following table useful.



CHI SQUARED DISTRIBUTION

	a									
d.f.	**0.995**	**0.990**	**0.975**	**0.950**	**0.900**	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**
95	63.250	65.898	69.925	73.520	77.818	113.038	118.752	123.858	129.973	134.247
96	64.063	66.730	70.783	74.401	78.725	114.131	119.871	125.000	131.141	135.433
97	64.878	67.562	71.642	75.282	79.633	115.223	120.990	126.141	132.309	136.619
98	65.693	68.396	72.501	76.164	80.541	116.315	122.108	127.282	133.476	137.803
99	66.510	69.230	73.361	77.046	81.449	117.407	123.225	128.422	134.641	138.987
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.170

Problem 3: (30 Percent)

You are interested in determining whether a new spray-on coating can help skaters skate faster in speed skating. To determine the answer, you select 10 world-class athletes. Each athlete will skate the same course on two successive days. On one of those days (chosen at random for each athlete), the blades will be spayed with the new material. On the other day, the blades will be spayed with a simple air spray (like a placebo) so that the athlete does not know on which day she is skating with the new coating. The data are shown in the table below.

Skater	Without coating	With coating	Difference
Alice	145.31	144.10	1.21
Barbara	147.34	145.37	1.97
Carla	146.24	145.37	0.87
Diane	141.66	141.42	0.24
Faith	141.91	139.10	2.81
Gail	139.01	139.84	-0.83
Heather	144.29	145.68	-1.39
Joan	144.42	141.62	2.80
Keren	148.77	144.12	4.65
Tamar	143.64	139.40	4.24
TOTAL	**1,442.59**	**1,426.02**	**16.57**
SUM OF SQUARED VALUES	**208,181.70**	**203,415.20**	**64.15**
Average	**144.26**	**142.60**	**1.66**
Standard deviation	**2.889**	**2.623**	**2.019**
Variance	**8.345**	**6.878**	**4.077**

Throughout this problem you can assume that the times are Normally distributed.

a) Treat the observations of times as if they were independent samples. Test the following hypothesis using a significance level of α=0.05. Assume that the underlying variances of the times are equal.

Be sure to state clearly whether or not you reject the null hypothesis

$$H_0: \quad m_{without\ coating} \le m_{with\ coating}$$
$$H_1: \quad m_{without\ coating} > m_{with\ coating}$$

b) Now perform the same test, but do **not** assume that the underlying variances are the same.

Be sure to state clearly whether or not you reject the null hypothesis

c) Now perform the same hypothesis test treating the data as if they came from a matched pairs design (as they do since each woman skates both with and without the coating).

Be sure to state clearly whether or not you reject the null hypothesis

d) Perform the following test using α=0.10.

$$H_0: \quad s^2_{without\,coating} = s^2_{with\,coating}$$

$$H_0: \quad s^2_{without\,coating} \neq s^2_{with\,coating}$$

Be sure to state clearly whether or not you reject the null hypothesis

Please place your name on Each page now

NAME: ______________________________

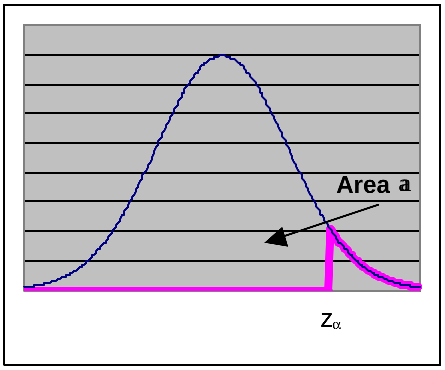


Table below gives Probability that a N(0,1) variable is less than or equal to z (or the area a above). NOTE WHICH AREA IS BEING GIVEN ABOVE!!

Normal Distribution

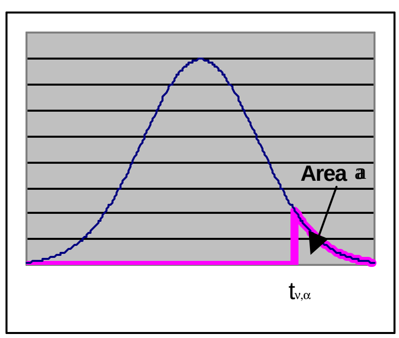
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.500	0.504	0.508	0.512	0.516	0.520	0.524	0.528	0.532	0.536
0.10	0.540	0.544	0.548	0.552	0.556	0.560	0.564	0.567	0.571	0.575
0.20	0.579	0.583	0.587	0.591	0.595	0.599	0.603	0.606	0.610	0.614
0.30	0.618	0.622	0.626	0.629	0.633	0.637	0.641	0.644	0.648	0.652
0.40	0.655	0.659	0.663	0.666	0.670	0.674	0.677	0.681	0.684	0.688
0.50	0.691	0.695	0.698	0.702	0.705	0.709	0.712	0.716	0.719	0.722
0.60	0.726	0.729	0.732	0.736	0.739	0.742	0.745	0.749	0.752	0.755
0.70	0.758	0.761	0.764	0.767	0.770	0.773	0.776	0.779	0.782	0.785
0.80	0.788	0.791	0.794	0.797	0.800	0.802	0.805	0.808	0.811	0.813
0.90	0.816	0.819	0.821	0.824	0.826	0.829	0.831	0.834	0.836	0.839
1.00	0.841	0.844	0.846	0.848	0.851	0.853	0.855	0.858	0.860	0.862
1.10	0.864	0.867	0.869	0.871	0.873	0.875	0.877	0.879	0.881	0.883
1.20	0.885	0.887	0.889	0.891	0.893	0.894	0.896	0.898	0.900	0.901
1.30	0.903	0.905	0.907	0.908	0.910	0.911	0.913	0.915	0.916	0.918
1.40	0.919	0.921	0.922	0.924	0.925	0.926	0.928	0.929	0.931	0.932
1.50	0.933	0.934	0.936	0.937	0.938	0.939	0.941	0.942	0.943	0.944
1.60	0.945	0.946	0.947	0.948	0.949	0.951	0.952	0.953	0.954	0.954
1.70	0.955	0.956	0.957	0.958	0.959	0.960	0.961	0.962	0.962	0.963
1.80	0.964	0.965	0.966	0.966	0.967	0.968	0.969	0.969	0.970	0.971
1.90	0.971	0.972	0.973	0.973	0.974	0.974	0.975	0.976	0.976	0.977
2.00	0.977	0.978	0.978	0.979	0.979	0.980	0.980	0.981	0.981	0.982
2.10	0.982	0.983	0.983	0.983	0.984	0.984	0.985	0.985	0.985	0.986
2.20	0.986	0.986	0.987	0.987	0.987	0.988	0.988	0.988	0.989	0.989
2.30	0.989	0.990	0.990	0.990	0.990	0.991	0.991	0.991	0.991	0.992
2.40	0.992	0.992	0.992	0.992	0.993	0.993	0.993	0.993	0.993	0.994
2.50	0.994	0.994	0.994	0.994	0.994	0.995	0.995	0.995	0.995	0.995
2.60	0.995	0.995	0.996	0.996	0.996	0.996	0.996	0.996	0.996	0.996
2.70	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997	0.997
2.80	0.997	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.998
2.90	0.998	0.998	0.998	0.998	0.998	0.998	0.998	0.999	0.999	0.999
3.00	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.10	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.20	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
3.30	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Please place your name on
Each page now **NAME: ________________________________**

Table below gives the t value such that the area to the right of that value is a when the t-distribution has the indicated number of degrees of freedom. See figure below.

t-distribution

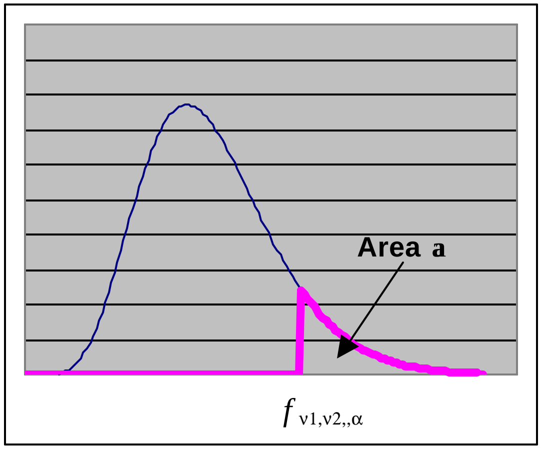
	α						
d.f.	**0.100**	**0.050**	**0.025**	**0.010**	**0.005**	**0.001**	**0.001**
1	3.078	6.314	12.706	31.821	63.656	318.289	636.578
2	1.886	2.920	4.303	6.965	9.925	22.328	31.600
3	1.638	2.353	3.182	4.541	5.841	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	5.894	6.869
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	1.311	1.699	2.045	2.462	2.756	3.396	3.660
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460
120	1.289	1.658	1.980	2.358	2.617	3.160	3.373
infinite	1.282	1.645	1.960	2.326	2.576	3.090	3.290



Please place your name on
Each page now

NAME: ___________________________________

These two tables give the critical F statistics for the indicated tail probabilities and the degrees of freedom in the numerator and denominator. See the figure below.



F distribution values for a right hand tail of 0.10

	d.f. numerator											
d.f denom	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**
1	39.864	49.500	53.593	55.833	57.240	58.204	58.906	59.439	59.857	60.195	60.705	61.220
2	8.526	9.000	9.162	9.243	9.293	9.326	9.349	9.367	9.381	9.392	9.408	9.425
3	5.538	5.462	5.391	5.343	5.309	5.285	5.266	5.252	5.240	5.230	5.216	5.200
4	4.545	4.325	4.191	4.107	4.051	4.010	3.979	3.955	3.936	3.920	3.896	3.870
5	4.060	3.780	3.619	3.520	3.453	3.405	3.368	3.339	3.316	3.297	3.268	3.238
6	3.776	3.463	3.289	3.181	3.108	3.055	3.014	2.983	2.958	2.937	2.905	2.871
7	3.589	3.257	3.074	2.961	2.883	2.827	2.785	2.752	2.725	2.703	2.668	2.632
8	3.458	3.113	2.924	2.806	2.726	2.668	2.624	2.589	2.561	2.538	2.502	2.464
9	3.360	3.006	2.813	2.693	2.611	2.551	2.505	2.469	2.440	2.416	2.379	2.340
10	3.285	2.924	2.728	2.605	2.522	2.461	2.414	2.377	2.347	2.323	2.284	2.244
12	3.177	2.807	2.606	2.480	2.394	2.331	2.283	2.245	2.214	2.188	2.147	2.105
15	3.073	2.695	2.490	2.361	2.273	2.208	2.158	2.119	2.086	2.059	2.017	1.972

Please place your name on
Each page now

NAME: ___________________________________

F distribution values for a right hand tail of 0.05												
	d.f. numerator											
d.f denom	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**
1	161.446	199.499	215.707	224.583	230.160	233.988	236.767	238.884	240.543	241.882	243.905	245.949
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396	19.412	19.429
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.785	8.745	8.703
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.912	5.858
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.678	4.619
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.000	3.938
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.575	3.511
8	5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388	3.347	3.284	3.218
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.073	3.006
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.913	2.845
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.687	2.617
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.475	2.403