Please put your name on the back of your answer book.
Do NOT put it on the front. Thanks.

- The exam is closed book, closed notes. Use only the formula sheet and tables I provide today. You may use a calculator.
- Write your answers in your blue book. Ask if you need a second (or third) blue book.
- You have 2 hours (120 minutes) to complete the exam.
  Stop working when the end of the exam is announced.
- Tables, a formula sheet, and scrap paper are provided.
- Points are indicated for each question. There are 120 total points.
- Important reminders:
  - budget your time. Some parts of each question should be easy; others may be hard. Make sure you do all parts you can.
  - notice that some parts do not require any computations.
  - show your work neatly so you can receive partial credit.
- Good luck!

1. 25 pts. A 1-way ANOVA was used to compare 5 treatments (labelled A, B, C, D, and E). Each treatment was randomly applied to 4 e.u's, so the total sample size, N, is 20. The investigator is especially interested in one a-priori contrast: the average of treatments A and B - treatment E. They used the 0.975 t quantile to calculate a 95% confidence interval for that contrast. The Mean Squared Error is used to compute the s.e. of the contrast. Each item in the following list identifies a change in the experimental design or a change in the data. For each item, please tell me whether it will INCREASE, DECREASE, or result in NO CHANGE in the **width** of the confidence interval.

   (a) Increase the number of replicates (e.g. from 4 to 8 e.u.s per treatment)

   (b) Find four blocks of five similar e.u.s, randomly assign treatments within blocks, and use an RCBD model to estimate the s.e. and calculate the confidence interval.

   (c) Increase the sample variance (MSE)

   (d) Increase the sample average for treatment E

   (e) Increase the coverage of the confidence interval (e.g. from 95% to 99%)

   (f) Increase the number of treatments (e.g. from 5 to 10). The number of replicates per e.u. is not changed. The contrast is still the average of treatments A and B - treatment E.

   (g) Use a Scheffe multiple comparisons adjustment instead of a t quantile.

   (h) Use a Bonferroni adjustment for 4 simultaneous confidence intervals inside of a t quantile.

2. 10 pts. Three research groups studied the prevalence of a specific disease on pig farms in Iowa. Each collected data from a simple random sample of n_i pig farms in Iowa, but each group used different methods of sampling on a farm. As a result, the standard deviation is expected to differ between studies. Each group also chose a different sample size, n_i. Here's what they found:

Study	Sample Size, n_i	Sample Average, $\overline{Y}_i$	Sample s.d., s_i
A	9	47%	15%
B	25	52%	20%
C	36	48%	36%

(a) Which study provides the most precise estimate of the mean prevalence in Iowa? Explain briefly.

(b) Estimate the standard error of the mean difference between Study A and Study B.

3. 20 pts. The following problem is based on an engineering study comparing two methods (A or B) for one step in making a component. The response is the length of time required for that step. Thirty components were used in the study: 20 were randomly assigned to method A; the remaining 10 used method B. Here are summary data for each group.

Method	n_i	Time to produce Average	s.d.
A	20	16.43	2.58
B	10	9.78	0.89

In studies like these, the ratio between the two times is the most reasonable way to summarize the treatment effect. The investigators expect that method A will take longer. They want to know if the ratio of means (Method A / Method B) is larger than 1.5.

(a) Is this a survey, an observational study or an experimental study? Briefly explain.

(b) The investigators used a randomization test to test the null hypothesis that the ratio of the two means equals 1.5. The test statistic is the ratio of the two sample averages (i.e. Method A average / Method B average). Here is the randomization distribution; this is the distribution of the test statistic when H0: ratio = 1.5 is true. The frequencies are written above each bar. There are a total of 1000 values in the randomization distribution. What is the p-value for a one-sided randomization test of ratio = 1.5? Note: the alternative hypothesis is H_A: ratio $>$ 1.5.

(c) The standard error of the estimated ratio is 0.0844. Calculate the t-statistic to test the hypothesis that the ratio = 1.5.

(d) Is a t-distribution a reasonable approximation to the randomization distribution, at least in the upper tail? Explain.

4. 65 pts. The U.S. Government has very strict rules for food safety. These require sterilization of most processed foods to reduce or eliminate unwanted bacteria. One common sterilization method is to heat food to a specified temperature and hold it at that temperature for a specified length of time. This kills any bacteria that are present in the food. Unfortunately, heating food also affects its quality. The following problem is based on a study of canned peas. This study compared five different combinations of time and temperature. Three (A, B, and D in the table below) meet the requirements for sterilization and two (C and E), labeled 'super-sterile', exceed the requirements. This study examined the quality of the peas after sterilization and storage for 6 months.

   Thirty (30) batches of canned peas were randomly assigned to one of the five temperature and time combinations. Treatments are labeled by the temperature / time, so 55/30 was heated to 55 degrees C for 30 minutes. Batches were then stored in one of six warehouses for six months. These warehouses were randomly chosen from a larger set of warehouses owned by the food processor. The warehouses had different storage conditions, which were expected to affect the quality of the canned peas, so the investigators made sure that one batch of each of the treatments was stored in each warehouse. The response is an average quality score measured for each batch of canned peas. This score ranges from 0 (very crisp = high quality) to 10 (very mushy = low quality). The data and treatment means are shown below:

			Warehouse						
Type	Treatment	Temp/Time	1	2	3	4	5	6	Mean
Sterile	A	50/60	3.2	3.8	1.2	0.6	0.6	1.5	1.82
Sterile	B	55/30	1.1	2.4	2.8	1.6	1.8	1.3	1.83
Super-sterile	C	55/45	4.9	4.5	2.2	3.0	2.0	1.1	2.95
Sterile	D	60/15	2.3	1.6	2.2	1.9	1.3	3.6	2.15
Super-sterile	E	60/20	4.6	6.5	5.6	3.8	4.1	4.3	4.82

   Additional SAS output that may be useful is included at the end of the exam.

   (a) 5 pts. What is the experimental unit in this study? Briefly explain.
   What is the observational unit? Briefly explain.

   (b) 10 pts. Complete the entries in the ANOVA table :

Source	df	SS	MS	F
Warehouses	??	11.68		
Treatments	??	??	??	??
Error	??	??	0.968	
Total	??	69.29		

   (c) 10 pts. The F statistic for Treatments in the previous part tests a specific null hypothesis (H0):

      i. identify that null hypothesis
      ii. approximate the p-value for that test
      iii. write an appropriate one sentence conclusion.

(d) 10 pts. The analysis represented by the ANOVA table in part 4b makes some standard assumptions. Here are a residual plot, a labeled plot of the responses, and a normal quantile plot. List these assumptions, then evaluate them using these plots and other available information. Formal tests are not required.

(e) 5 pts. Which of the following is the best approach to answer the question 'which pairs are treatments are significantly different?'. Briefly explain your choice.
T-tests (LSD), Protected LSD, Tukey-Kramer HSD, Bonferroni, Scheffe, Something else.

(f) 10 pts. The investigators are specifically interested in two questions:
How large is the mean difference between the two super-sterile treatments (C and E)?
How large is the difference between the mean quality of the three sterile treatments (A, B, and D) and the mean quality of the two super-sterile treatments (C and E).

i. Write out the coefficients to estimate each of these contrasts.
ii. Are these two contrasts orthogonal? Explain why or why not.

(g) 5 pts. Construct a 95% confidence interval for the difference between the two 'super-sterile' treatments, (C and E)

(h) 5 pts. The investigators are also interested in whether there are any differences in quality among the three 'sterile' treatments, (A, B, and D). That is, they want to test H0: $\mu_A = \mu_B = \mu_D$. If this is possible **without computing Sums-of-Squares from the raw data**, please test H0: no differences between treatments A, B, and D. If this is not possible, say so.
I do not want you to compute SS from the raw data. That may take a lot of time.

(i) 5 pts. The investigators are intrigued by the apparent slight decrease in quality in treatment D (60/15), compared to treatments A and B. They want to conduct another study using these three treatments. Again, they will use a block design and want you to suggest a sample size (number of warehouses). They are concered primarily about the contrast between treatment D and the average of treatments A and B. You agree to use a standard deviation of 1. They want an $\alpha = 5\%$ test to have 80% power to detect a difference of 0.35. How many warehouses should they use?

SAS code and part of the SAS output for problem 4 (quality of canned peas)

```
data food;
  infile 'foodqual.txt';
  input warehouse trt quality;

proc glm;
  class warehouse trt;
  model quality = warehouse trt;

  lsmeans trt /stderr;

  estimate 'ave of A,B,D - ave of C,E' trt 2 2 -3 2 -3 /divisor = 6;
  contrast 'ave of A,B,D - ave of C,E' trt 2 2 -3 2 -3;

  estimate 'ave of A,B,C - ave of D,E' trt 2 2 2 -3 -3 /divisor = 6;
  contrast 'ave of A,B,C - ave of D,E' trt 2 2 2 -3 -3;

  estimate 'difference C - E' trt 0 0 1 0 -1;
  estimate 'difference D - E' trt 0 0 0 1 -1;
  contrast 'difference C - E' trt 0 0 1 0 -1;
  contrast 'difference D - E' trt 0 0 0 1 -1;

run;
```

```
-------------------------------------------------------------------------
                          Class Level Information

          Class          Levels    Values

          warehouse           6    1 2 3 4 5 6

          trt                 5    50/60 55/30 55/45 60/15 60/20

                  Number of Observations Read          30
                  Number of Observations Used          30
```

Continuation of the SAS output for problem 4 (quality of canned peas)

Least Squares Means

trt	quality LSMEAN	Standard Error	Pr > |t|
50/60	1.81666667	0.40163555	0.0002
55/30	1.83333333	0.40163555	0.0002
55/45	2.95000000	0.40163555	<.0001
60/15	2.15000000	0.40163555	<.0001
60/20	4.81666667	0.40163555	<.0001

Dependent Variable: quality

Contrast	DF	Contrast SS	Mean Sq.	F Value	Pr > F
ave of A,B,D - ave of C,E	1	27.378	27.378	28.29	<.0001
ave of A,B,C - ave of D,E	1	11.858	11.858	12.25	0.0023
difference C - E	1	10.453	10.453	10.80	0.0037
difference D - E	1	21.333	21.333	22.04	0.0001

Parameter	Estimate	Standard Error	t Value	Pr > |t|
ave of A,B,D - ave of C,E	-1.95000000	0.36664141	-5.32	<.0001
ave of A,B,C - ave of D,E	-1.28333333	0.36664141	-3.50	0.0023
difference C - E	-1.86666667	0.56799844	-3.29	0.0037
difference D - E	-2.66666667	0.56799844	-4.69	0.0001