LAST NAME (Please Print): _____

FIRST NAME (Please Print): _____

HONOR PLEDGE (Please Sign): _____

Statistics 111

Midterm 4

- This is a closed book exam.
- You may use your calculator and a single page of notes.
- The room is crowded. Please be careful to look only at your own exam. Try to sit one seat apart; the proctors may ask you to randomize your seating a bit.
- **Report all numerical answers to at least two correct decimal places** or (when appropriate) write them as a fraction.
- All question parts count for 1 point.

1. Consider the following design and data.

run	A	B	C	D	E	F	obs
1	-	-	-	-	-	-	3
2	+	-	-	-	+	-	17
3	-	+	-	-	+	+	-2
4	+	+	-	-	-	+	-4
5	-	-	+	-	+	+	6
6	+	-	+	-	-	+	20
7	-	+	+	-	-	-	11
8	+	+	+	-	+	-	13
9	-	-	-	+	-	+	2
10	+	-	-	+	+	+	-8
11	-	+	-	+	+	-	6
12	+	+	-	+	-	-	4
13	-	-	+	+	+	-	10
14	+	-	+	+	-	-	-8
15	-	+	+	+	-	+	-3
16	+	+	+	+	+	+	4

List the defining relations. (1 point)

What interactions are confounded with A?

_____ What is your numerical estimate of the AB interaction?

_____ Give the symbolic name for this design (include the resolution).

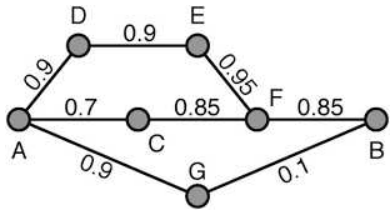
In this context, how do we interpret things when a main effect is confounded with a two-way interaction?

If none of these factors affects the response, then under the usual assumptions, what is the distribution of the effect estimates?

2. _____ Let $f(x) = (x/3)\exp(-x^2/6)$ for $x > 0$, and zero else. What is the value of the hazard function when $x = 2$?

_____ Does this have increasing failure rate?

3. _____ Consider the network shown below. If all edges work independently, with probabilities as shown, what is the chance that you can link from A to F?



4. When does multicollinearity occur and why is it bad? (2 pts)

5. Explain what ten-fold cross-validation does and how to do it. (5 pts)

6. You want to test whether there are differences in IQ between gender and major. You do not have resources to study all possible majors, so, on a whim, you focus upon Statistics, Pyschology, and Economics. You observe two people for each combination of treatment levels. Complete the following table. Only the blanks in the parentheses count.

Source	df	SS	MS	F
Gender	_____	10	_____	(_____)
Major	_____	5	_____	(_____)
Interaction	_____	_____	6	(_____)
Error	_____	_____	_____	
Total	_____	33		

_____ What is your critical value for the 0.05 level test on Major?

_____ Do you conclude that Major is significant?

7. _____ The value of a car at time t is $14,000/(1+t^2)$ (t is measured in years). Each year, the chance of being totalled is 0.1. At $t = 0$, you purchase a year of insurance for \$100. What is the last year for which you should purchase insurance?

8. You study whether IQs differ by gender with an RCBD. To control for the effect of major, you consider three majors: Statistics, Psychology, and Economics. You take one observation for each gender/major combination. Suppose the sum of squares due to Gender is 20, the sum of squares due to Major is 6, and the total sum of squares is 30. Complete the table. Only blanks in parentheses count.

Source	df	SS	MS	F
Gender	_____	_____	_____	(_____)
Major	_____	_____	_____	(_____)
Error	_____	_____	_____	
Total	_____	_____		

_____ What is your critical value for testing the Gender effect?

_____ If Major were significant at the 0.05 level, what would be your critical value for the upper tail least significant difference test?

_____ Was it wise to control for Major?

_____ If you had not, what would be your test statistic?

_____ If you had not, what would be your critical value?

9. List all, and only, the true statements. _____

A. Running line smoothers are better than fixed bin-width smoothers.
B. Simmelian ties enforce normative social behavior.
C. In greenlighting the *Challenger* launch, the managers interpolated.
D. Cox proportional hazards models describe for the lifespans of wild songbirds.
E. Bag-of-words models lose semantic information.
F. Commonplace risks are overestimated.