UNM Statistics Qualifying Exam Jan 2014 Name: ____________________

Due: 3 P.M., Tue Jan 21, 2014 UNM ID number

Qual Take Home (100 points) Complete both problems in this exam. Your report is to be typed, double-spaced, no smaller than ten-point font with one-inch margins, and should be identified by your UNM ID number (do not include your name). Each problem is to be no longer than three pages, and an additional five-page appendix is allowed for each problem but will be examined only at the discretion of the graders; the better constructed your appendix, the more likely it is to get examined.

Write your answers as they might appear in the methods, results, and conclusions sections of an academic paper (that is, do not include the common introduction and discussion sections). Insert tables and figures to support your points. Tables and figures should be well-labelled and cross-referenced from text, such as, "in Table 1 ...", or if in the appendix, "in Table A1 ...". Figures should include appropriate symbols suitable for black-and-white reproduction (that is, avoid use of color; consider symbols, line types, and distinct shades of gray). Computer output without explanation will not be reviewed. As necessary:

1. Plot and describe the data (that is, plot all the individual observations, in addition to summaries of data you might present with the results).
2. Clearly define population parameters and sample statistics.
3. Clearly specify hypotheses tested and explicitly state the associated model at least once (i.e., write model equation).
4. Define and assess method assumptions.
5. Write a coherent evidence-based conclusion that a layperson can understand.

You may **not** consult any other person when working on this exam or discuss your exam with anyone else regardless of whether or not the person is taking the exam. You may use your course notes as well as any available books or web resources for the exam. If including computer text tables where alignment is important, then please use a fixed-width font, such as `Courier`, for that text. Any points of clarification can be directed to Prof. Erik Erhardt, `erike@stat.unm.edu`.

Due: 3 P.M., Tue Jan 21, 2014, hand-delivered to Ana Parra Lombard in the main office of the Department of Mathematics and Statistics, MSC01 1115, 1 University of New Mexico, Albuquerque, New Mexico, 87131-0001. Please do not email your solutions.

(50pts) **1. Low birth weight, 1986**

The goal of this analysis is to understand the factors associated with birth weight (BWT). Data were collected on 189 women including several variables which were thought to be of importance.

Analyze the data. Find a good predictive model for birth weight. This should include variable selection and examination of residuals. Report the expected quality of predictions from this model, as well as any concerns you may have for prediction. Discuss the results of the analysis in a way that is clear to someone that does not have a lot of statistical training. In addition, provide a prediction for the birth weight of a new baby with the following characteristics: AGE = 30, LWT = 120, RACE = 1, SMOKE = 0, PTL = 1, HT = 0, UI = 0, FTV = 1.

Data description:
`http://www.umass.edu/statdata/statdata/data/lowbwt.txt`
Data:
`http://www.umass.edu/statdata/statdata/data/lowbwt.dat`
Alternatively:
`www.stat.unm.edu/~erike/exams/UNM_Stat_Exam_Qual_takehome_201401_pr2-DATA_lowbwt.txt`
`www.stat.unm.edu/~erike/exams/UNM_Stat_Exam_Qual_takehome_201401_pr2-DATA_lowbwt.dat`

```r
# One way of reading the data in R:
lbw <- read.table("http://www.umass.edu/statdata/statdata/data/lowbwt.dat",
    skip = 6)
colnames(lbw) <- read.table("http://www.umass.edu/statdata/statdata/data/lowbwt.dat",
    skip = 5, nrows = 1, stringsAsFactors = FALSE)
```

First 10 observations:

	AGE	LWT	RACE	SMOKE	PTL	HT	UI	FTV	BWT
1	19	182	Black	No	0	No	Yes	0	2523
2	33	155	Other	No	0	No	No	3	2551
3	20	105	White	Yes	0	No	No	1	2557
4	21	108	White	Yes	0	No	Yes	2	2594
5	18	107	White	Yes	0	No	Yes	0	2600
6	21	124	Other	No	0	No	No	0	2622
7	22	118	White	No	0	No	No	1	2637
8	17	103	Other	No	0	No	No	1	2637
9	29	123	White	Yes	0	No	No	1	2663
10	26	113	White	Yes	0	No	No	0	2665

(50 pts) **2. Pulse Oximetry**

Pulse oximetry is a method used for noninvasive measurement of blood oxygen. A device attached to a patient's finger emits light in two different known wavelengths (red and near-infrared) on one side of the finger, and the sensor on the other side of the finger records the amount of emitted light that passes through the finger tissues. The amount of oxygen in a patient's blood is determined by comparing the relative amount of absorbed light of the two different wavelengths. In a clinical environment, pulse oximetry is used in blood-oxygen monitors that continuously record measurements of a patient's blood-oxygen level and signal when the measurements fall below a safe level.

One problem that occurs with these monitors is false alarms. Some researchers suspect that movements of a patient's finger, such as might result from shivering, increase the incidence of false alarms.

In one research project to evaluate the causes of false alarms in pulse oximetry, biomedical engineers are testing whether shivering affects pulse oximetry measurements. Pulse oximeters are often used in two different environments in a hospital: inpatient rooms, where shivering is most often associated with a fever, and in a recovery room immediately after surgery. In the kind of shivering associated with a fever, which is called febrile shivering, is qualitatively different than the postoperative shivering experienced by surgery patients in the recovery room. In order to simulate these two types of shivering, the experimenters use a device for shaking a test subject's hand, called a shaker table. By altering the setup of the shaker table, the experimenters can simulate either febrile or postoperative shivering.

In the type of experiment we are considering, the standard pulse oximetry device is attached to the finger of a single test subject. The subject's hand is then attached to the shaker table, which vibrates the finger to simulate either postoperative or febrile shivering. In addition, the shaker table vibrations are set to one of three different intensities, which we will call mild, moderate, and severe. The response that is measured is the power of the electrical signal over a fixed range of frequencies. In this type of experiment, there are two factors: type of shivering, at two levels, and intensity of shivering, at three levels. There are six treatments, consisting of all combinations of type and intensity of shivering. The experimenters are interested in the effects of type of shivering and severity of shivering on the response.

In one particular experiment, the researchers used a single subject. From this subject, they obtained four measurements at each of the six treatments. The subject was allowed to rest for five minutes between runs. To more closely mimic a hospital environment, the apparatus was removed from the subjects finger after each run. The order in which the treatments were applied was randomized.

Data: `www.stat.unm.edu/~erike/exams/UNM_Stat_Exam_Qual_takehome_201401_pr1-DATA_pulse.dat`

	Intensity	Shivering	run1	run2	run3	run4
1	mild	febrile	3.78	5.27	7.28	3.19
2	mild	postoperative	12.41	8.58	8.39	7.22
3	moderate	febrile	5.19	10.29	7.37	7.64
4	moderate	postoperative	6.99	7.54	7.33	15.67
5	severe	febrile	9.95	6.54	11.40	12.49
6	severe	postoperative	2.64	3.10	2.16	1.98

Analyze the data provided by this experiment. In addition to analyses and comments arising from your own curiousity, please address the following as part of your write-up:

(a) What statistical design is being used, and why? Could a better design have been used, and why or why not?

(b) Is there blocking? If so, what is/are the block(s)?

(c) What is/are the nuisance factor(s) to be "averaged out" in the design?

(d) What is/are the treatment(s)?

(e) What is/are the outcome(s)/response(s)?

(f) Plot the data (not only summaries of the data) in a way that helps you understand what the effects are.

(g) Write out the statistical model (in notation) and fit the model parameters.

(h) How many degrees-of-freedom are allocated to each source of variation?

(i) State and assess model assumptions. (If assumptions are not met, try to address that. If you can not address unsatisfied model assumptions, mention this and continue as though the model assumptions are met.)

(j) State and conduct statistical tests for the parameters, and interpret the test results.

(k) Perform pairwise comparisons based on your final model and summarize which pairs of treatment combinations are different.

(l) Discuss anything else of interest, and address the original goal of the experiment.