

STATISTICS MASTERS/Ph.D.-QUALIFYING EXAM: TAKE HOME

January, 2022

General directions

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 “CODE WORD” in the header on each page; *do not include your name or UNM ID number*. Each problem is to be no longer than four pages, and an additional four-page appendix is allowed for each problem but will be examined only at the discretion of the graders; the better constructed your appendix with cross-references from the text, the more likely it is to get examined. In your data analysis, raw uninterpreted computer output will be graded as the dress it is.

Write your answers completely, but concisely. 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 . . .” and each should have a caption that describes it and briefly tells the reader why it is of value. Figures should include appropriate symbols suitable for black-and-white reproduction (that is, avoid use of color if possible; consider symbols, line types, and distinct shades of gray to distinguish categories or values).

Organize your sections to justify the validity of what you uncovered and the methods you used to uncover it. We want a summary of what you think is important, not a diary of how you spent your time. Remember that even the best data analysis is worthless if your reader does not understand it, so you are being graded on presentation as well as statistical content.

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, such as the mean and confidence intervals).
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 the model equation).
4. Define and assess method assumptions.
5. Write a coherent evidence-based conclusion that a layperson can understand.

You may use your course notes as well as any available books or web resources on general statistical methods for the exam. 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 nor are you allowed to use the internet to find analyses of these data.

Any points of clarification can be directed to Prof. Guoyi Zhang, gzhang12@math.unm.edu.

Email solutions by **3 PM, Fri Jan 15, 2022** to Ana Parra Lombard, aparra@math.unm.edu, Department of Mathematics and Statistics, University of New Mexico. Please do not turn in a physical copy of your solutions.

Problem 1, Wisconsin goat milk production

A Wisconsin family farm contacted a statistician to conduct an experiment to assess the effect of diet on milk production for a new breed of goats they had acquired. Four goat (“goat”) were selected at random from the flock and marked by a numeric ear tag. Four diets (“diet”) were chosen to reflect typical goat diets provided at farms in the area. Each goat is given a different diet during each of four consecutive lactation periods (“period”), and we will assume that the period between diets is sufficient so previous diets do not affect later milk production (“amount”).

Analyze the data provided by this experiment and make a recommendation to the farm. Please address the following questions as part of your write-up

1. What statistical design is being used, and why? Could a better design have been used, and why or why not?
2. Is there blocking? If so, what is/are the block(s)?
3. What is/are the nuisance factor(s) to be “averaged out” in the design?
4. What is/are the treatment(s)?
5. What is/are the outcome(s)/response(s)?
6. Plot the data (not only summaries of the data) in a way that helps you understand what the effects are.
7. Write out the statistical model (in notation) and fit the model parameters.
8. How many degrees-of-freedom are allocated to each source of variation?
9. 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.)
10. State and conduct statistical tests for the parameters, and interpret the test results.
11. Are the effects of the treatment(s) and block(s) independent? Provide evidence for your claim. Hint: review the covariance matrix structure for the coefficients.
12. Looking at the feed bags, the farmers noticed that the constituents of the diets are similar for A and B. The same is true for diets C and D. Write down and test the associated contrast.
13. Discuss anything else of interest and address the original goal of the experiment.

goat	period	diet	amount	goat	period	diet	amount
1	1	A	6.5	3	1	D	6.7
1	2	B	6.3	3	2	C	6.5
1	3	D	6.8	3	3	A	7.5
1	4	C	6.2	3	4	B	6.8
2	1	C	6.2	4	1	B	6.6
2	2	A	6.3	4	2	D	6.0
2	3	B	7.1	4	3	C	6.9
2	4	D	6.0	4	4	A	6.7

Download the data from

https://math.unm.edu/sites/default/files/files/qual-exams/stat/unm_exam_202201_stat_qual-takehome_dat1.csv.

Problem 2, body fat

The data were collected on 43 females. The purpose of the study was to determine the relationship between body fat and the other variables of the study. The response variable was bodyfat and other twelve predictor variables were considered in analysis as follows:

Variable	Description
fat	bodyfat (pound)
age	age (year)
weight	weight (pound)
height	height (inch)
chest	chest measurement (cm)
abd	abdominal measurement (cm)
hip	hip measurement (cm)
thigh	thigh measurement (cm)
knee	knee measurement (cm)
ankle	ankle measurement (cm)
biceps	biceps measurement (cm)
forearm	forearm measurement (cm)
wrist	wrist measurement (cm)

Using the data, build an appropriate regression model to predict fat (“bodyfat”) using a subset of the rest of the variables, making sure that you carefully assess all assumptions. Write a succinct, coherent, and complete summary of your analysis.

Download the data from

https://math.unm.edu/sites/default/files/files/qual-exams/stat/unm_exam_202201_stat_qual-takehome_dat2.csv.