

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

January, 2021

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, 2021** 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

Measuring hardness

Prior to engraving in metal, an assessment of the metal's hardness is made in order to choose the appropriate engraving bit. In a discreet location on the material to be engraved, a hardness testing machine presses a testing tip into the metal with a fixed force and the depth of the resulting depression is measured which corresponds to the metal's hardness. Lately, inconsistent engraving results has raised questions about differences in testing tips and the following experiment was designed.

Four pairs of testing tips (8 total) from different vendors are compared to see whether they produce significantly different readings; for each pair, one tip is called "Rep 1" and the other is called "Rep 2" for each tip type. One vendor provided four metal plates of unknown hardnesses for testing purposes, and each tip is tested on each metal plate. The order of testing each combination of tip and metal plate is randomized and the testing machine is completely reset between each test. The results are provided in the table.

Assess whether the four types of testing tips measure hardness equally. Describe the statistical experimental design. Fit the simplest linear model that adequately explains the data in this experiment. Write the model using the same parametrization that appears in the computer output, explain the terms in the model, and list all the assumptions. Examine the main effects, interaction terms, and do multiple comparisons of the interested treatments or combination of treatments. Make sure that you carefully assess all the assumptions. Write a succinct, coherent, and complete summary of your analysis.

Table 1: Data for Problem 1

Tip	Rep	Metal 1	Metal 2	Metal 3	Metal 4
Tip 1	Rep 1	9.25	9.77	8.74	9.55
Tip 1	Rep 2	9.42	9.98	8.26	9.06
Tip 2	Rep 1	9.00	9.30	8.59	9.10
Tip 2	Rep 2	8.98	8.82	8.24	9.04
Tip 3	Rep 1	9.88	10.57	9.32	10.01
Tip 3	Rep 2	10.01	10.49	9.07	9.62
Tip 4	Rep 1	9.96	10.31	9.38	10.06
Tip 4	Rep 2	9.70	10.34	9.67	10.38

Download the data from

https://math.unm.edu/sites/default/files/files/qual-exams/stat/unm_exam_202101_stat_qual-takehome_dat1.csv.

Problem 2

Housing value

This data set contains a randomly selected subset of 405 observations from a housing data with 12 variables. The first variable ID is the random number used to generate the data, and the other variables are described as follows:

1. V1: per capita crime rate by town
2. V2: proportion of residential land zoned for lots over 25,000 sq.ft.
3. V3: proportion of non-retail business acres per town
4. V4: nitric oxides concentration (parts per 10 million)
5. V5: average number of rooms per dwelling
6. V6: proportion of owner-occupied units built prior to 1940
7. V7: index of accessibility to radial highways
8. V8: full-value property-tax rate per USD 10,000
9. V9: pupil-teacher ratio by town
10. V10: percentage of lower status of the population
11. RES: median value of owner-occupied homes in USD 1000's, target (or response) variable

Using the data, build an appropriate regression model to predict RES using a subset of the rest 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_202101_stat_qual-takehome_dat2.csv.