Statistics 341 Name _____
Winter 2019
J. Morita

Midterm Exam Rewrite

Instructions:

For many of the problem parts for which you did not earn full credit, you may do the following "rewrite" & "earn back" some of the points you lost on the original midterm exam. Note that "rewriting" involves more than merely writing up the correct solution to each problem. For each problem, you may earn either ½ of what you originally lost, or else zero.

The last problem in this document is for each student to work, regardless of your midterm exam scores. It is labeled "Problem SS".

Note that your "rewrites" must be entirely complete, correct & clear. Grading standards will be very high.

Problem 1.

(a) No rewrite

(b) If you lost a full point for part (b), then clearly define bias (in the statistical sense). Give a statistical definition (theoretical) as well as a clear verbal description, including why we might be interested in this property of an estimator. Describe what the bias is assessing here in this problem. In doing so, you will need to identify the parameter of interest, the data that are being used to estimate the parameter, the relevant statistic(s), the relevant probability distribution and what it is the probability distribution of, and the estimator. And, of course, solve part (b), clearly indicating your final answer.

(c) If you lost at least one point for part (c), then refer to (b) above and do the same tasks, but for "efficiency" AND "relative efficiency" instead of for "bias. Then work part (c) on the original exam, showing & justifying all steps. Also, solve part (c), clearly indicating your final answer. Don't forget the interpretation of your result.

(d) If you lost at least one point for part (d), then rework this part, justifying each step, and clearly indicating your final answer. Then describe what the Cramer-Rao bound tells us (in general, not just for this problem). Write your description for someone who has studied introductory statistics, but not theoretical statistics, as in STAT 341. Why do we care about the C-R Bound?

(e) If you lost at least one point for part (e.) then rework the problem part. Give a clear description of why this result DOES make **intuitive** sense. Also, address the question: "Which, if either, of these estimators is Minimum Variance Unbiased?

Problem 2.

If you missed more than 1 point on Problem 2, then rework the problem entirely. Also, give a clear description to someone who has studied introductory statistics, but not theoretical statistics, of what relative efficiency measures. Along the way, you will need to describe MSE, its relation to Variance, and their connections to sample size in this problem.

Problem 3.

(i) Clearly & completely define MSE, variance & bias. Give a statistical definition (theoretical) as well as a clear verbal description, including why we might be interested in this property of an estimator. Your verbal description should be understandable to someone who is intelligent but has not studies statistics. That means you must also describe all relevant terminology, including parameter, statistic, sampling distribution, estimation error, mean, variance, ... Be sure to include a description of why/how the name "mean squared error"

describe what MSE measures (The mean of what? What is squared? What is the error that is referred to?).

(ii) Write a complete, clear proof as requested in the original exam problem.

Problem 4.

(a) If you lost one or more points on Problem 4(a) on our Midterm Exam, then work the following.

Consider $X_1, X_2, \ldots, X_n$ ~iid Geometric, with pdf

$$f_X(x;\theta) = \theta (1-\theta)^x \quad \text{for } x = 0,1,2,\ldots, \quad 0 < \theta < 1$$
$$= 0 \quad \text{elsewhere.}$$

Recall that we have two "usual" ways to define Geometric random variables, so be careful here.

(i) Compare & contrast the pdf here with the one for our exam Problem 4.

(ii) Find a sufficient statistic for θ using this parameterization of the pdf.
Is the distribution of your sufficient statistic, or a simple function of your sufficient statistic, one of our "familiar" distributions? If so, give the name of the distribution, and identify the parameter values. If not, then state "None of these". In either case, you must justify your answer.

(iii) Give an intuitive non-mathematical explanation for why your answers to (a) & (b) do or do not make sense in the context of this problem.

(b) If you lost one or more points on Problem 4(b)&(c) combined on our Midterm Exam, then use the distribution above to solve our Midterm Exam problem parts 4(b) & (c).

Problem 5.

(a) If you missed any points on part (a), then clearly define what each of parts (i) through (v) is, and for each part, give an example that clearly matches your definition.

(b) If you missed a full point on part (b), then clearly describe, compare and contrast the types of processes that result in Binomial and Hypergeometric random variables. Be sure that you clearly describe the differences & similarities between those processes. Support your descriptions with a realistic, illustrative example.

(c) If you missed 1 or more points on part (c), then do the following. Describe the idea behind Method of Moments estimation. To do so, you will need to define & describe moments. Describe what they measure. Compare & contrast population & sample moments.
Then derive the Method of Moments estimator for the exam problem, clearly matching it with your answer to part (c).

(d) Clearly describe why your estimator in (c) does "make intuitive sense" in this situation.

Problem "SS".

(a) Derive the MLE for the situation in Problem 5. You may ask Prof. Morita for hints, after you have thought about this for a few days.

(b) Comment on whether and how your answer does or does not match intuition. Your comments should be clear and written in complete sentences.