

Name _____ Date _____

Test #1 – Chapters 1, 2 and 3

Chestnut Hill - Patrylak UEMA 111 Probability and Statistics Fall II, 2013

DESIGNING AN EXPERIMENT

Explain how you would design an experiment to evaluate whether praying for a hospitalized heart attack patient leads to a speedier recovery. What are the factors/treatments? Show how you would incorporate control, randomization, replication, comparison, and blinding. Discuss whether blocking is possible and/or necessary. Are there any issues with respect to realism and/or ethics that may need to be addressed?

Outline the W's of the experiment:

1. Control (Placebo)
2. Randomization – Specify and explain your method for picking subjects for the experiment and for assigning them to various treatment groups.
3. Replication
4. Blinding (or Double-Blinding)
5. Generalization (or Realism)
6. Blocking
7. Ethics
8. Comparison

Summary:

FUNDING STUDENTS' EDUCATION

A few years ago, the Department of Education published information about education in the United States. It reported the average amount (dollars per student) spent by public schools in each state during the prior school year. The following table divides the states according to location - whether they lie east or west of the Mississippi River.

Write a report comparing the amounts eastern and western states spend to educate their children. Remember! A complete report includes:

- a clear and concise discussion of the W's – who, what where, when, why and how - of the study under consideration
- well-labeled, correctly-constructed and properly-scaled graphs –frequency charts, relative frequency histograms (by location). Box plots will be invaluable!
- numeric summaries of the statistics/parameters. How and why you chose your particular measures. Include a simple, salient discussion of the shape, center, spread, gaps, clusters, outliers of the locations' distribution(s) above.
- a well-written comparison of the two distributions of expenditures per student stating your conclusion to the question, "Is the average amount spent per student by the states *independent of location (relative to the Mississippi River)"? You must offer numerical as well as graphical evidence to support your arguments.

EASTERN STATES	\$ SPENT PER STUDENT	WESTERN STATES	\$ SPENT PER STUDENT
Alabama	5,166	Alaska	9,074
Connecticut	9,221	Arizona	5,122
Delaware	7,963	Arkansas	4,999
D.C.	9,225	California	5,795
Florida	6,183	Colorado	6,099
Georgia	5,947	Hawaii	6,409
Illinois	6,858	Idaho	5,012
Indiana	6,786	Iowa	6,295
Kentucky	6,125	Kansas	6,406
Maine	7,238	Louisiana	5,645
Maryland	7,812	Minnesota	6,795
Massachusetts	8,299	Missouri	6,096
Michigan	7,717	Montana	6,448
Mississippi	4,575	Nebraska	6,584
New Hampshire	6,487	Nevada	5,758
New Jersey	10,233	New Mexico	4,984
New York	9,970	North Dakota	5,353
North Carolina	5,667	Oklahoma	5,398
Ohio	6,808	Oregon	7,348
Pennsylvania	7,777	South Dakota	5,281
Rhode Island	8,627	Texas	5,910
South Carolina	5,643	Utah	4,256
Tennessee	5,274	Washington	6,534
Vermont	7,500	Wyoming	6,718
Virginia	5,938		
West Virginia	6,779		
Wisconsin	7,680		

Name _____ Date _____

In addition to the above, answer these questions (and explain your answer as completely as necessary) regarding the above study.

- a. What is the population being observed?
- b. Was the above study a census or a sample of the population?
- c. What are the three variables measured for each state in this study?
- d. Are the numeric summaries that you presented statistics or parameters?
- e. Classify each of the variables as qualitative or quantitative?
- f. Is the quantitative variable discrete or continuous?
- g. What level of measurement does the quantitative variable reach?
- h. Was the above study an experiment or an observational study?
- i. If this was a study, was it a cross-sectional, retrospective or prospective study?
- j. We have learned that there are at least 12 errors (flaws) – some deliberate, some unintentional - from which studies such as this one can suffer. Name 5 such pitfalls to a good study, and discuss why you think they were or were not encountered in the course of gathering the above data.
- k. Instead of having used histograms above, dotplots or stemplot could have been used. Name some advantages and disadvantages of each type of graph.
- l. Pareto charts and pie charts may not be appropriate in this case study. Why?
- m. Enter the following measures:

Measure	Eastern States	Western States	All States Combined
Mean			
Median			
Mode			
Midrange			
R.M.S.			
Q1			
Q3			
Range			
S.D.			
Var.			
C.V.			

- n. When all of the data for 51 locations are combined, what percentage of the observations falls within one S.D. of the mean?
- o. What about - within two S.D. of the mean?
- p. What do the answers to the last two questions suggest about the “normality” of the distribution of Dollars for Students?

* Note: So far, the term “independent variables” is a term of art to you. Simply stated, do you feel, from all that you know about the situation that the amount of funding per student that the states accord their public schools is different in the east vs. the west. Later on the course, we will define independence and dependence of variables more

h

rigorously.