

Introduction (Rules and Suggestions)

The biostatistics lab activity is an important part of the biostatistics course. Complete all lab work as specified in this manual. Answer all questions each week. Your lab work will be graded for completeness every 5 or so weeks. Most labs can be completed within the allotted lab period. Occasional labs may require completion at home.

Bring this workbook to exams (including the final). The instructor or teaching assistant will check your workbook at that time. Honest but incorrect answers will NOT be marked down. However, incomplete or sloppy work will result in a deduction of up to 10 points on the exam.

In addition to completing the work in your lab manual, you are to keep notes and formulas in a separate *Procedure Notebook*. You will be allowed to use this *Procedure Notebook* on Part B of exams, but *not* on Part A. Compilation and use of the *Procedure Notebook* is documented on the course website.

The *premise* of the lab is this: You take a simple random sample from a population and, over the course of the semester, analyze the data in your sample. Data from the population ($N = 600$) are listed in Appendix 1 of this manual and are available in SPSS format as `populati.sav` on the course website. During the semester you will *not* be working with `populati.sav`. Instead, you will be working with data from your random sample in order to make inferences about the population.

POPULATI.SAV contains data on the following variables:

#	Variable	Description
1	ID	Identification number
2	AGE	Age in years ($\mu = 29.505$, $\sigma = 13.58$, min = 1, max = 65)
3	SEX	F = female (26.5%), M = male (66.7%), . = missing (6.8%)
4	HIV	Y = HIV+ (76.8%), N = HIV- (23.2%), . = missing (0.0%)
5	KAPOSISA	Kaposi sarcoma status: Y (52.8%), N (47.2%), . (0.0%)
6	REPORTDA	Report data reported as mm/dd/yy (min = 01/02/89, max = 02/05/90)
7	OPPORTUN	Opportunistic infection: Y (60.2%), N (35.3%), . (4.5%)
8	SBP1	Systolic blood pressure first reading. ($\mu = 120.13$, $\sigma = 18.53$)
9	SBP2	Systolic blood pressure second reading ($\mu = 119.95$, $\sigma = 19.07$)

During the first lecture or lab session you must sign up for a College of Applied Sciences and Arts (CASA) computer account. As you sign up for the account, write down your ID and password in a safe place. *You* are responsible for keeping track of your computer account and making certain it is functioning. If you experience difficulties with your computer account, contact the CASA technical staff via `tech@casa.sjsu.edu`.

The course assumes you already know how to use Windows® computers. If you do not know how to

Introduction

manage computer files on Windows computers, please take HPrf101 or some other basic computing course before enrolling in HS167.

After you complete each lab, print the homework exercises for the week and complete only those exercises assigned in class. Homework exercises are separate from the lab, and do *not* go into the workbook or *Procedure Notebook*. Assigned exercises will be graded for completeness, accuracy, and reporting style each week.

Lab 1: Measurement and Sampling

Lab 1: Measurement and Sampling

Purpose: To select a simple random sample from `populati.sav` and enter these data into an SPSS file.

1. **Random Numbers:** You want to select a simple random sample from the population listed in the appendix of this manual. The population consists of 600 individuals many of whom are HIV positive. The first step in selecting your sample is to generate 10 random numbers between 1 and 600. To generate 10 random numbers, start your Web browser and go to the <http://www.random.org/>. Then, in the middle column (labeled “How?”), click “randomized sequences.” In the “Smallest value” field type “1”. In the “Largest value” field type “600”. Click “Generate Sequence.” Write the *first ten* numbers in this sequence in the space below:

First random number: _____

Second random number: _____

Third random number: _____

Fourth random number: _____

Fifth random number: _____

Sixth random number: _____

Seventh random number: _____

Eighth random number: _____

Ninth random number: _____

Tenth random number: _____

Comment: The numbers in this random sequence do *not* repeat themselves. Therefore, the sample will be made *without* replacement.

2. **Data:** Identify the 10 people in the population with ID numbers that match your random numbers above. This comprises your unique random sample from the population. List these data here:

ID	AGE	SEX	HIV	KAPOSISA	REPORTDA	OPPORTUN	SBP1	SBP2

Lab 1: Measurement and Sampling

3. **Data Entry:** You are now going to enter your data into an SPSS .sav file. The .sav file format is native to SPSS and can be used only if you have SPSS on your computer. (All computers in the lab have SPSS.)
 - a. Go to the Windows Start Bar and select SPSS. (The college computers have the SPSS icon on the Start Bar. Your home installation may have the SPSS icon installed elsewhere.)
 - b. After starting SPSS, select “Type data into file.” Click on the Variable View tab at the bottom of your screen.
 - c. Type each variable’s name in the column labeled NAME. Create variables for ID, AGE, SEX, HIV, KAPOSISA, REPORTDA, OPPORTUN, SBP1, and SBP2.
 - d. In the column labeled TYPE, use “numeric” for scale variables (ID, AGE, SBP1, SBP2), “String” for nominal variables (SEX, HIV, KAPOSISA, OPPORTUN), and “Date” for the date variable (REPORTDA).
 - e. In the column labeled MEASURE, identify variables as scale, ordinal, or nominal, as appropriate.
 - f. You may leave the remaining columns blank.
 - g. Click the Data Table tab at the bottom of the screen and then carefully enter your data into the table.
 - h. After you’ve entered your data into the table, save these as LnameF10.sav (e.g., GerstmanB10.sav). Store your data in your **home directory** and on a **floppy disk***. If you save your data to the default directory, it will be gone the next time you sign onto the computer!
4. **Homework:** Go to the StatPrimer website and print the homework exercises for Chapter 1. Complete only those exercises assigned in class. These are done on separate sheets of paper and handed in for grading at the beginning of the next lecture. Do *not* put HW exercises in your *Procedure Notebook*.

See the syllabus for policies regarding homework assignments.

* If your personal CASA lab account is not yet functional store data on floppy only.