Math 153 - Test 1

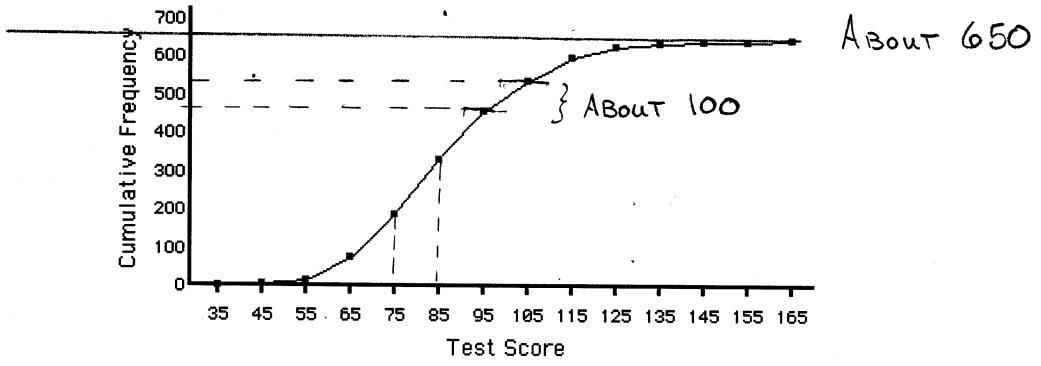
February 11, 2016

Name key

Score _____

Show all work to receive full credit. Supply explanations where necessary. You may use your calculator for all statistical computations.

1. (12 points) The following graph summarizes the scores on a widely administered test.



(a) What is the name of this type of graph?

Ogive

(b) About how many test scores are in the sample?

About 650

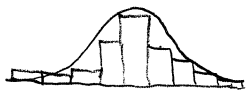
(c) Which interval along the horizontal axis contains the most test scores? Explain how you know.

The segment with steepest slope is from 75 to 85

(d) Assuming that the numbers along the horizontal axis are class midpoints, estimate the number of test scores between 100 and 110.

About 550 - 450 = 100

(e) Are the test scores approximately normally distributed? Explain. (Hint: Think about the corresponding histogram.)



Because the ogive is cumulative, it shows increase. The frequencies of test scores starts small, begins to increase rapidly with a peak between 75 & 85, then decreases. Yes, it looks roughly normal. (Perhaps a bit skewed right.)

2. (12 points) A data set contains decimal numbers between 1.7 and 40.1. Each number in the data set has exactly one digit to the right of the decimal point. Katrina constructs a frequency distribution whose lowest class limit is 1.4 and whose class width is 8.1.

(a) How many classes are there in Katrina's frequency distribution?

5 CLASSES (SEE LIMITS IN (b) AND (c))

(b) List all the lower class limits (starting with 1.4).

1.4, 9.5, 17.6, 25.7, 33.8

(c) List all the upper class limits.

9.4, 17.5, 25.6, 33.7, 41.8

(d) Determine the boundaries (on both sides) of the first class.

1.35 AND 9.45

(e) Make up some frequencies for each of the classes you found above so that your frequency distribution shows data that are skewed left.

TAIL TO THE LEFT.

NUMBERS	FREQUENCIES
1.4 - 9.4	1
9.5 - 17.5	3
17.6 - 25.6	4
25.7 - 33.7	10
33.8 - 41.8	16

3. (3 points) In the years following the end of World War II, it was found that there was a strong correlation between the number of human births and the world's stork population. Can we conclude storks cause babies? Explain your reasoning.

NO, CORRELATION DOES NOT IMPLY CAUSATION.

4. (7 points) Organize the following data into a stem-and-leaf plot. Then determine the mean, median, and mode.

~~47~~ ~~32~~ ~~48~~ ~~51~~ ~~10~~ ~~27~~ ~~50~~ ~~21~~ ~~24~~ ~~12~~
~~21~~ ~~32~~ ~~48~~ ~~12~~ ~~28~~ ~~32~~ ~~36~~ ~~37~~ ~~38~~ ~~49~~

Stem	Leaf
1	0 2 2
2	1 1 4 7 8
3	2 2 2 6 7 8
4	7 8 8 9
5	0 1

3 | 6 means 36

$\text{Mean} = \frac{\sum x}{20} = \frac{655}{20} = 32.75$

$\text{Median} = \frac{10^{TH} + 11^{TH}}{2} = \frac{32+32}{2} = 32$

Mode = Most Frequent = 32

5. (3 points) PSC students were asked by email to participate in a survey regarding student involvement. Several hundred students participated. Do you think the sampling method is flawed? Explain.

Yes flawed. It is a voluntary response survey. Such surveys are typically biased.

6. (6 points) Determine the level of measurement. Choose from *nominal*, *ordinal*, *interval*, or *ratio*.

(a) Rankings of National Football League quarterbacks

Ordinal

(b) Amount of time (in minutes) spent studying

Ratio

(c) Illinois automobile license plate numbers

Nominal

(d) Years in which AT&T chief executive officers retired

Interval

7. (6 points) A professor separated her students' lab reports into two piles—those of the passing students and those of the failing students. Twenty students passed, and their average score was 82.5. Eight students failed, and their average score was 62.2. What was the average score of all the students?

$$\frac{20(82.5) + 8(62.2)}{28} \approx 76.7$$

8. (5 points) Determine whether the given value is a statistic or a parameter.

   (a) There are 58 national parks in the National Park Service.

   ALL OF THEM

   PARAMETER

   (b) Among the students who volunteered to take the CCSSE survey, 47% were full-time students.

   SAMPLE

   STATISTIC

   (c) The mean atomic weight of all elements in the periodic table is 134.355 atomic mass units.

   PARAMETER

9. (3 points) Determine whether each value comes from a discrete collection or a continuous collection of data.

   (a) Sam owns 210 video games.

   DISCRETE

   (b) One of the dogs at the animal shelter weighs 210 lbs.

   CONTINUOUS

10. (12 points) The frequency distribution shown below gives the salaries (in thousands of dollars) of the employees at a small company.

CLASS WIDTH = 19.4

Salary (thousands of $)	Frequency
12.8–32.1	13
32.2–51.5	8
51.6–70.9	4
71.0–90.3	1
90.4–109.7	1

$13+8+4+1+1 = 27$

(a) Describe the distribution of salaries.

SKEWED RIGHT - TAIL TO RIGHT

(b) Based on your answer from above, which do you expect to be greater, the mean or the median? Why?

I EXPECT MEAN > MEDIAN

A SMALL NUMBER OF LARGE SALARIES PULLS MEAN UP.

NOTICE BELOW THAT THIS DIDN'T ACTUALLY HAPPEN.

(c) What are the class midpoints?

$\frac{12.8+32.1}{2} = 22.45,\ 41.85,\ 61.25,\ 80.65,\ 100.05$

(d) Use class midpoints to estimate the mean salary.

$$\bar{X} \approx \frac{13(22.45)+8(41.85)+4(61.25)+80.65+100.05}{27}$$

$\approx \boxed{38.98}$ (THOUSANDS OF $)

(e) Use class midpoints to estimate the median.

MEDIAN = 14TH SALARY (MIDPT)

= $\boxed{41.85}$ (THOUSANDS OF $)

11. (10 points) For each of the following situations, tell which type of graph would best display the data. Choose from *dot plot*, *bar graph*, *time-series graph*, *scatterplot*, *pie chart*, *ogive*, *histogram*, or *stem-and-leaf plot*. You may get partial credit if you offer brief explanations.

(a) A CEO wants to sketch a graph showing how the company budget is split up among 10 different categories.

PORTIONS OF WHOLE → PIE CHART

(b) A teacher graded 20 tests, and they all had scores that were whole numbers between 15 and 60. He wants to display the entire set of scores.

STEM-AND-LEAF

TOO FEW SCORES WITH TOO BIG RANGE FOR DOT PLOT

(c) Biologists caught, weighed, and released 500 fish. They want to make a graph showing the numbers of fish in the different weight classes.

HISTOGRAM

(d) Among other things, Pike's dairy sells ice cream, milk, butter, yogurt, and cheese. A manager would like to show a graph displaying last year's total sales of these products.

BAR GRAPH

(e) Craig randomly selected 100 women. For each woman, he recorded her age and the number of minutes each day that she read. He formed ordered pairs and plotted the data.

SCATTERPLOT

12. (10 points) What type of sampling is described in each situation. Choose from *random*, *systematic*, *convenience*, *stratified*, or *cluster*. You may get partial credit if you offer brief explanations.

(a) Fifteen midwestern universities are selected at random to take part in a survey. All students at those universities participate in the survey.

ALL FROM SOME → CLUSTER

(b) Every fourth student is selected to participate in a survey.

SYSTEMATIC

(c) The first eight people to walk through the door are given a free gift card.

CONVENIENCE

(d) One hundred business cards are placed into a box. The cards are mixed up and five are selected.

EACH IS EQUALLY LIKELY

RANDOM

(e) Students are divided into groups according to age, and twenty people are selected at random from each group.

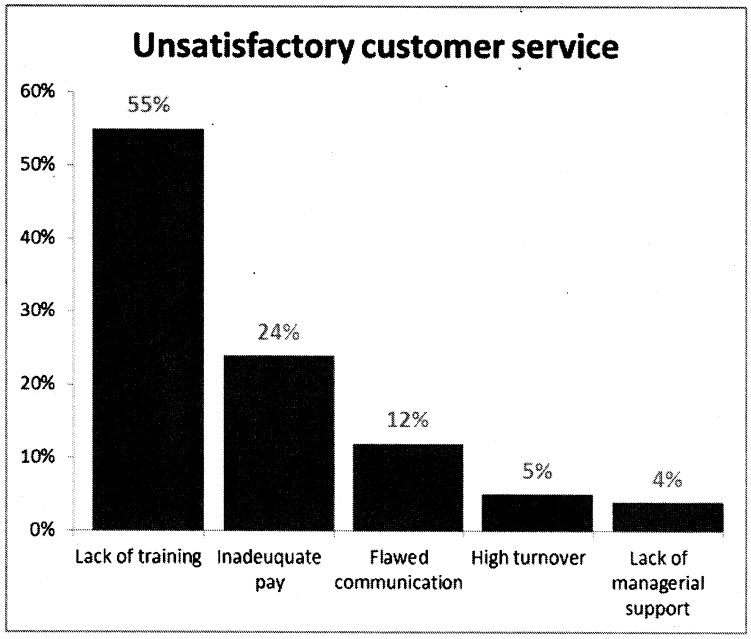
SOME FROM ALL → STRATIFIED

13. (4 points) What is the difference between an experiment and an observational study.

IN AN OBSERVATIONAL STUDY, SUBJECTS ARE MERELY OBSERVED AND OBSERVATIONS RECORDED. IN AN EXPERIMENT, SUBJECTS ARE GIVEN A TREATMENT, AND THE EFFECTS OF THE TREATMENT ARE OBSERVED.

14. (3 points) What is the name of this type of graphical display? Be specific.

PARETO CHART



15. (4 points) A convenience sample of 5 students is selected. Explain why this is not a simple random sample.

A SAMPLE OF 5 WAS PICKED BY CONVENIENCE. SINCE NOT ANY GROUP OF 5 IS AS CONVENIENT AS ANOTHER, NOT EVERY SAMPLE OF 5 IS AS LIKELY AS EVERY OTHER.