Math 153 - Test 1

September 20, 2012

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) A scientist believes that the normal human body temperature is not 98.6°F as typically reported. He conducts a study to determine the "usual" human body temperature.

   (a) Is his study an experiment or an observational study? Briefly explain.

   OBSERVATIONAL STUDY --- HE IS NOT MODIFYING THE SUBJECTS, HE IS JUST MEASURING THEIR TEMPERATURES.

   (b) Are body temperatures discrete or continuous?

   CONTINUOUS

   (c) Identify the level of measurement (nominal, ordinal, interval, ratio) for the body temperatures.

   INTERVAL

   (d) He obtains a sample of 1000 temperatures by randomly selecting 50 people from each of 20 participating universities. Is his sample a simple random sample? Briefly explain.

   NO, NOT EVERY SAMPLE OF 1000 IS EQUALLY LIKELY. FOR EXAMPLE, A SAMPLE OF 1000 FROM A SINGLE UNIVERSITY IS IMPOSSIBLE.

   (e) Is the mean temperature of the 1000 subjects a parameter or statistic?

   STATISTIC

   (f) Suppose the scientist obtained his temperature measurements by mailing thermometers to people and asking them to measure and report their temperatures. State two possible problems with this approach.

   ① THOSE ASKED MAY NOT REPORT THEIR RESULTS.

   ② PARTICIPANTS MAY NOT ACCURATELY MEASURE/RECORD THEIR TEMPS.

2. (6 points) A sample of VZW stockholders is obtained as described. Identify the type of sampling (random, systematic, convenience, stratified, cluster).

(a) A complete list of stockholders is compiled and every 500th name is selected.

SYSTEMATIC

(b) Fifty different stockbrokers are randomly selected, and a survey is made of all their clients who own VZW shares.

CLUSTER

(c) At the annual stockholders' meeting, a survey is conducted of all who attend.

CONVENIENCE

(d) All of the stockholder zip codes are collected, and 8 stockholders are randomly selected from each zip code.

STRATIFIED

(e) Each stockholder is assigned a unique number. A computer randomly generates numbers to select the sample of stockholders.

RANDOM

(f) Friends who own VZW shares are selected.

CONVENIENCE

3. (4 points) A convenience sample of 10 PSC students is taken. Explain why this is not a simple random sample.

NOT EVERY SAMPLE OF 10 STUDENTS IS EQUALLY LIKELY. IN FACT, THE MORE "INCONVENIENT" A SAMPLE, THE LESS LIKELY IT IS.

4. (5 points) According to Chebyshev's Theorem, at least what percent of the data values from any set will lie within 2.5 standard deviations of the mean?

$$1 - \frac{1}{2.5^2} = 1 - \frac{1}{6.25} = \boxed{84\%}$$

5. (6 points) Fred received grades of B, A, A, B, C in classes with 5, 3, 4, 3, 1 credit hours, respectively. Compute Fred's GPA.

$$\frac{3(5) + 4(3) + 4(4) + 3(3) + 2(1)}{5 + 3 + 4 + 3 + 1} = \boxed{3.375}$$

6. (5 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*, HISTOGRAM, *ogive*, or *stem-and-leaf plot*. You may get partial credit if you offer brief explanations.

   (a) Biologists caught, weighed, and released 350 fish. They want to make a graph showing the numbers of fish in the different weight classes.

   Histogram

   (b) A couple wants to sketch a graph showing how they budget their monthly earnings. They'd like to show how their money is divided among 7 different categories.

   Pie chart

   (c) A teacher graded 25 tests, and they all had scores that were whole numbers between 17 and 55. She wants to display the entire set of scores.

   Stem-and-leaf plot

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

   Bar chart

   (e) Oscar 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

7. (14 points) The frequency distribution shown below gives the salaries (in thousands of dollars) of the employees at a small company.

Salary (thousands of $)	Frequency
12.5–27.9	5
28.0–43.4	13
43.5–58.9	27
59.0–74.4	8
74.5–89.9	3
90.0–105.4	2

TOTAL : 58

(a) What are the class boundaries associated with the first class listed above?

12.45 AND 27.95

(b) What is the class width?

15.5

(c) What is the relative frequency of the third class listed above?

$\frac{27}{58} \approx 46.6\%$

(d) If the frequency distribution was changed to a cumulative frequency distribution, what count would be associated with "Less than $89,950"?

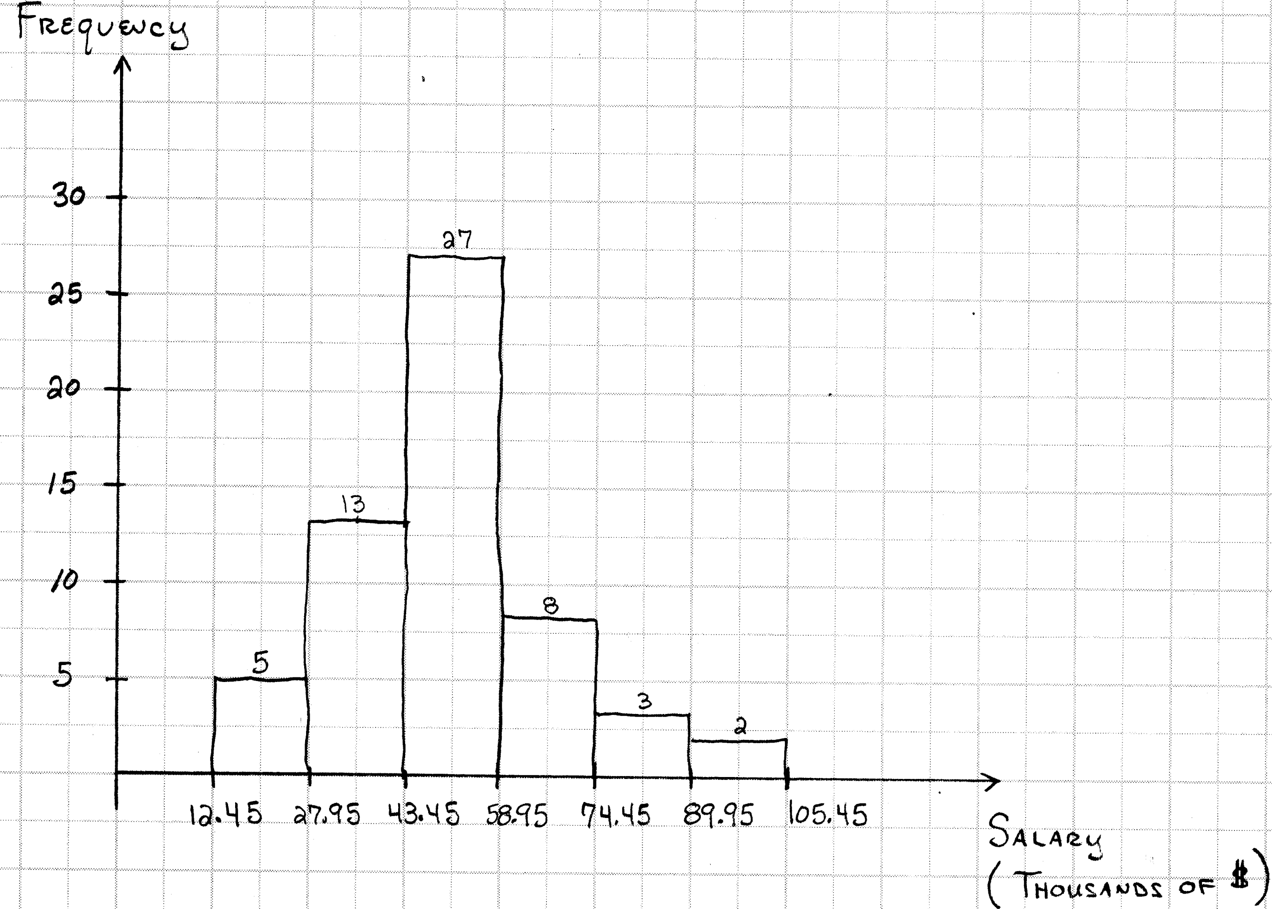
$5 + 13 + 27 + 8 + 3 = \boxed{56}$

(e) If class midpoints were used to compute a weighted mean, do you think the mean would be a good measure of center. (Hint: Do not compute the mean, but think about the distribution of data.)

YES, THE DATA LOOK ROUGHLY NORMALLY DISTRIBUTED.

I EXPECT MEAN ≈ MEDIAN ≈ MODE.

(f) Using class boundaries along the horizontal axis, construct the corresponding histogram. *Use the attached graph paper.*


Frequency
30
25
20
15
10
5
27
13
8
5
3
2
12.45
27.95
43.45
58.95
74.45
89.95
105.45
Salary
(Thousands of $)

8. (8 points) Mike scored 43 on a biology test with mean 38.1 and standard deviation 6.2. Julia scored 180 on a chemistry test with mean 157.9 and standard deviation 27.3.

   (a) Compute the corresponding z scores. Who scored better and why?

$$Z_{43} = \frac{43 - 38.1}{6.2} = \frac{4.9}{6.2} \approx 0.79$$

$$Z_{180} = \frac{180 - 157.9}{27.3} = \frac{22.1}{27.3} \approx 0.81$$

JULIA SCORED SLIGHTLY BETTER. HER SCORE WAS 0.81 STANDARD DEVIATIONS ABOVE THE MEAN, WHEREAS MIKE'S WAS 0.79.

   (b) Compute the coefficients of variation for the tests. Which tests (biology or chemistry) had greater variation?

$$\underbrace{\text{BIO: } \frac{6.2}{38.1} \approx 16.27\% \qquad \text{CHEM: } \frac{27.3}{157.9} \approx 17.29\%}_{\text{CHEM TESTS HAD SLIGHTY MORE VARIATION.}}$$

9. (8 points) If the data in a set are approximately normally distributed, then about 95% of the data values lie within 2 standard deviations of the mean. Suppose a normally distributed data set has mean 6.78 and standard deviation 1.95.

   (a) What are the minimum and maximum "usual" values (i.e. the values that are located two standard deviations from the mean)?

$6.78 - 2(1.95) = 2.88 \leftarrow$ MIN "USUAL" VALUE

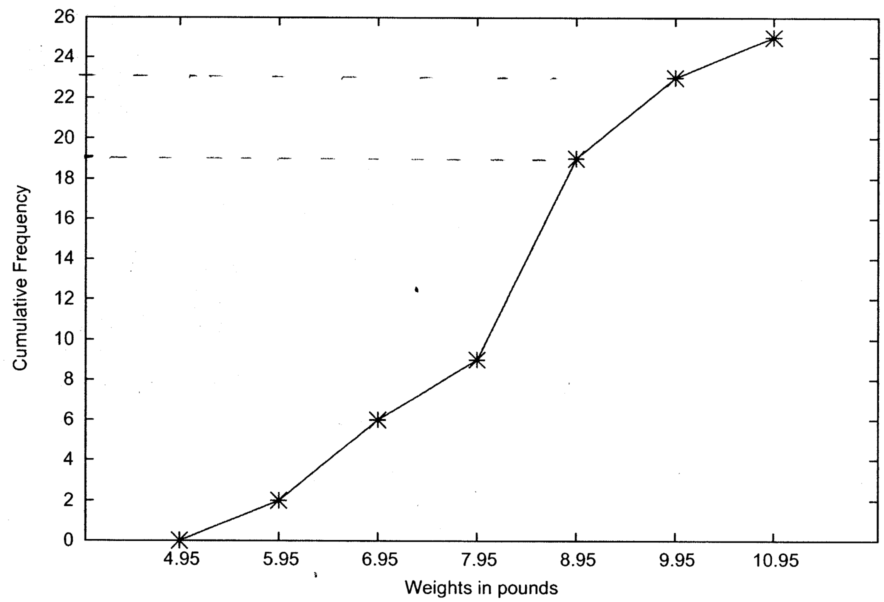
$6.78 + 2(1.95) = 10.68 \leftarrow$ MAX "USUAL" VALUE

   (b) If the data set contains 1250 numbers, about how many are between the values you computed above?

95% OF 1250 IS 1187.5

$\boxed{\text{ABOUT } 1188}$

10. (10 points) The following ogive shows the distribution of birth weights of the full-term babies born last month at a local hospital.



(a) How many babies are in the sample described by the ogive?

LOOKS LIKE 25

(b) How many babies had birth weights between 8.95 lbs and 9.95 lbs?

23 - 19 = 4

(c) In which range of birth weights were there the most babies?

7.95 lbs - 8.95 lbs

(d) In which range of birth weights were there the fewest babies?

4.95 lbs - 5.95 lbs OR 9.95 lbs - 10.95 lbs

BOTH HAVE 2 BABIES

(e) Are birth weights continuous or discrete? Are numbers of babies continuous or discrete?

WEIGHTS: CONTINUOUS BABIES: DISCRETE

11. (14 points) In the following stem-and-leaf plot, 3 | 2 means 32.

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				

(a) Are the data values shown above approximately normally distributed? Briefly explain.

IT LOOKS LIKE IT. THE OUTLINE OF THE STEM-AND-LEAF PLOT ROUGHLY HAS THE SHAPE OF A SMOOTH BELL-CURVE.

(b) Compute the mean, median, and mode.

$$\text{MEAN} = \frac{10+12+12+\cdots+50+51}{20} = \frac{655}{20} = 32.75$$

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

$$\text{MODE} = 32$$

(c) Compute the range.

$$51-10=41$$

(d) Use the range to estimate the standard deviation.

$$S \approx \frac{\text{RANGE}}{4} = \frac{41}{4} = 10.25$$

(e) Briefly explain how the values of the mean, median, and mode support your conclusion in part (a).

SINCE MEAN $\approx$ MEDIAN = MODE, THE DISTRIBUTION IS PEAKED IN THE MIDDLE AND NOT SKEWED.

12. (8 points) The Insurance Institute for Highway Safety conducted tests with crashes of new cars traveling at 6 mi/h. The costs of damage from a simple random sample are shown below. Compute the sample mean and standard deviation. Based on your result, is damage of \$10,000 unusual? Briefly explain.

\$7448 \$4911 \$9051 \$6374 \$4277

FROM CALCULATOR...

$\bar{X} = \$6412.^{20}$

$S = \$1926.^{80}$

$\bar{X} + 2S = \$10,265.^{80}$ = MAXIMUM "USUAL" VALUE

SINCE \$10,000 IS LESS THAN $\bar{X} + 2S$, \$10,000 IS NOT UNUSUAL.

13. (5 points extra credit) Some teens and adults were asked how much cash they were carrying. 15 teens were carrying a mean amount of \$9.75, and 25 adults were carrying a mean amount of \$28.15. What was the mean amount carried by all 40 people?

$$\frac{15(9.75) + 25(28.15)}{40} = \frac{850}{40} = \boxed{\$21.^{25}}$$