

Math 123, Fall 16, Midterm 1

Name: Solution

Instructor: Saba Gerami

- Directions:**
- Show all your work.
 - You only receive half of the points if you do not explain your reasoning.
 - You can use a non-graphing calculator.
 - You may not use cell phone, or notes.

1. Determine whether the data set is a population or a sample. Explain your reasoning.

a) A survey of 2311 U.S. adults about their health care provider

2 sample bc it is not all U.S. adults

b) The number of televisions in each U.S. household

2 Population bc it is all U.S. Households

2. Determine whether the numerical value is a parameter or statistics. Explain your reasoning.

a) Fifty percent of a sample of 1025 U.S. adults say that the best years for the United States are behind us.

2 Statistics bc comes from a sample

b) In 2012, Major league Baseball teams spent a total of \$2,940,657,545 on player' salaries.

2 Parameter bc comes from a population

3. Determine whether the data are qualitative or quantitative. Explain your reasoning.

a) A list of student ID numbers

2 qualitative. These #s have no meaning.

b) The final scores on a video game

2 quantitative bc the numbers make sense.

4. Determine the level of measurement of the data set (ordinal, nominal). Explain your reasoning.

a) Types of shows televised by a Netflix (Comedy, Drama, ...)

2

nominal → Cannot be ordered in a meaningful way.

b) The list of your top 5 best friends

2

Ordinal → can be ordered.

5. Identify the sampling technique used (random sample, simple random sampling, convenience sampling, cluster sampling, systematic sampling, stratified sampling).

a) Twenty-five students are randomly selected from each grade at a high school.

1

stratified sampling

b) A journalist waits at Delta airlines waiting line and asks them if they are happy with Delta.

1

Convenience sampling

c) Randomly surveying 10 students out of 876 where every sample of 10 people has the same chance of being selected.

1

simple random sampling

d) Randomly surveying 10 students out of 876 where every student has the same chance of being selected.

1

random sample

6. What is (are) the difference(s) between a histogram and a pareto (bar) chart?

2

histogram → bars are touching, for classes of data

Pareto → bars not touching, for exact data

Values

2

$min = 0$
 $max = 39$
 $range = 39 - 0 = 39$

$class\ width = \text{roundup}\left(\frac{39}{5}\right) \Rightarrow 8$
of classes

7. Use the data set, which represents the times (in minutes) spent reading a book by a high school student.

7 39 13 9 25 8 22 0 2 18 2 30 7
 35 12 15 8 6 5 29 0 11 39 16 15 1

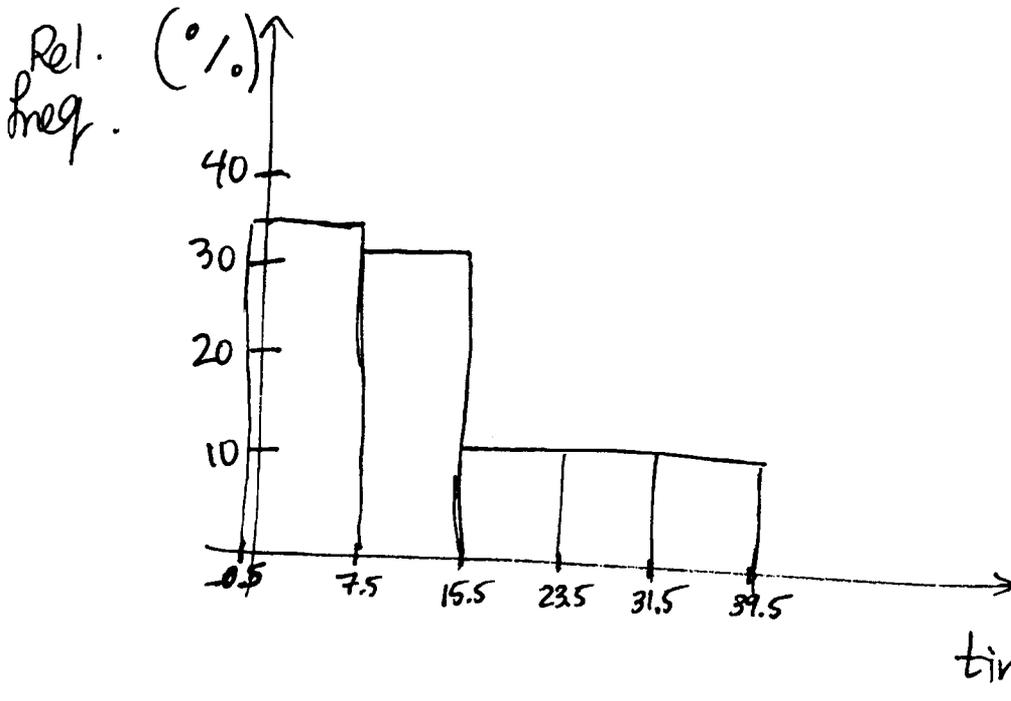
a) Construct a frequency distribution for the data set using five classes. Include class limits, midpoints, frequencies, relative frequencies, and cumulative frequencies.

does not have to be %

Class	Midpoint	Frequency $n=26$	Relative Frequency	Cumulative Frequency
0-7	3.5	9	$\frac{9}{26} \times 100 = 34.6$	9
8-15	11.5	8	$\frac{8}{26} \times 100 = 30.8$	17
16-23	19.5	3	$\frac{3}{26} \times 100 = 11.5$	20
24-31	27.5	3	11.5	23
32-39	35.5	3	11.5	26 = n ✓

2 pts 2 pts $\sum f = 26$ 2 pts 1 pts 1 pts

b) Construct a relative frequency histogram using part (a). Do not forget to label the axes.



2 pts

8. Use the weighted mean formula to find the GPA for someone who got A in two 3-unit classes, a B in a 5 unit class and C in three 4-unit classes. (A=4.0, B=3.0, C=2.0)

4 pts

Grade	pts(x)	Weights(units) w	x*w
A	4.0	6 (bc two 3-unit)	24
B	3.0	5	15
C	2.0	12 (three 4-unit)	24
D	1.0		
		$\Sigma w = 23$	$\Sigma xw = 63$

$$\bar{X} = \frac{\Sigma xw}{\Sigma w} = \frac{63}{23} = \boxed{2.74}$$

9. Use the data given in the stemplot.

KEY: 4 | 2 = 420

a) Find the median, mode, and mean.

400, 400, 450, 490, 490, 490, 580, 600, 650, 670
 680, 700, 700, 890, 890, 900, 920, 930

Q_1 (circled)
 Q_2 (660)
 Q_3 (890)

```

4 | 0 0 5 9 9 9
5 | 8
6 | 0 (5 7) 8
7 | 0 0
8 | 9 9
9 | 0 2 3
  
```

3 pts

bc $\frac{650+670}{2}$

Median: 660

Mode: 490

Mean: $\frac{\Sigma x}{n} = \frac{11830}{18} = 657.22$

b) Find Q_1 and Q_2 = median = 660

2 pts

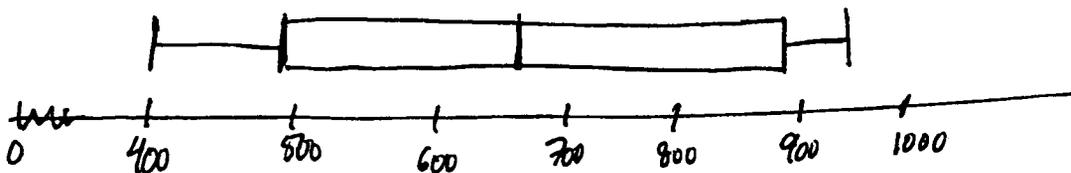
490

$Q_3 = 890$

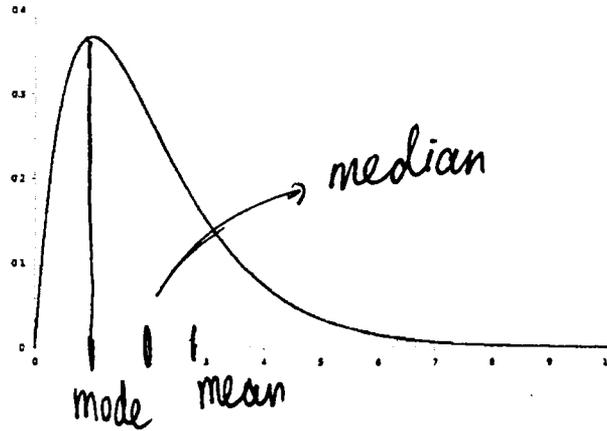
c) Draw a box-and-whisker plot.

min=400, $Q_1=490$, $Q_2=660$, $Q_3=890$, max=930

1 pt



10.



a) Determine whether the approximate shape of the distributions is symmetric, skewed left, skewed right, or none of these.

2pt

skewed right

b) Indicate in the x-axis, where the mean, mode, and median are approximately located.

3pt not necessary

mode → the peak
 mean → towards the tail
 median → bw mean & mode

11. You scored 65 points on a midterm for which the mean was 70 and the standard deviation was 5. Your friend scored 70 on another midterm for which the mean was 80 and the standard deviation was 2. Who did worse as compared to the rest of class? Explain your reasoning. You may use z-scores.

You: 65, mean 70, std. dev. 5 → so you are 1 std. dev. below the mean.

4pt

friend: 70, mean 80, std dev. 2 → friend is 5 std. dev. below mean.

You did better!

OR

$$\text{You: } z_{\text{score}} = \frac{65-70}{5} = -1$$

$$\text{friend: } z_{\text{score}} = \frac{70-80}{2} = -5$$

you did better bc your z-score is higher.

8

12. Calculate the population mean, population variance, population standard deviation and population coefficient of variation for the following data. You must show your work by making a table as shown in class.

5 6 14 11 7 9

$$N=6 \rightarrow \mu = \frac{\sum x}{N} = \frac{52}{6} = 8.67$$

8
pts

x	$x - \mu$	$(x - \mu)^2$
5	-3.67	13.44
6	-2.67	7.11
14	5.33	28.44
11	2.33	5.44
7	-1.67	2.78
9	0.33	0.11
$\sum x = 52$		$\sum = 57.32$

$$\text{Var.} = \frac{\sum (x - \mu)^2}{N} = \frac{57.32}{6} = 9.55$$

$$\text{std dev.} = \sqrt{9.55} = 3.09$$

Population mean: 8.67

Population standard deviation: 3.09

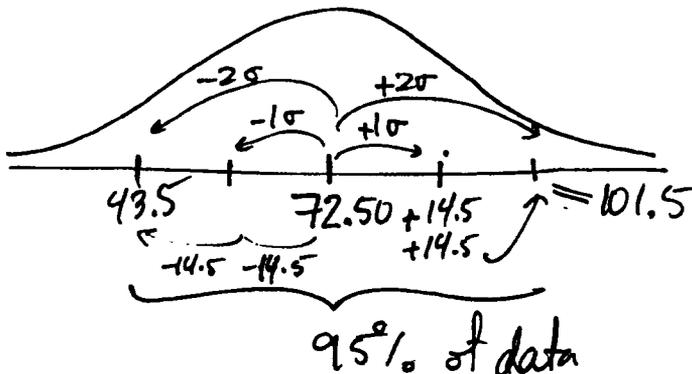
Population variance: 9.55

Population coefficient of variation: 35.64%

$$\frac{\sigma}{\mu} \times 100 = \frac{3.09}{8.67} \times 100$$

13. The mean for satellite television for a sample of households was \$72.50 per month, with a standard deviation of \$14.50 per month. What two values do 95% of the data lie? Assume the data set has a bell-shaped distribution.

4



bw \$43.5
& \$101.5

14. The frequency distribution shows the numbers of a golf membership per household for a sample of households in San Luis Obispo. Calculate the sample mean, sample variance, sample standard deviation and sample coefficient of variation for the following data. You must show your work by making a table as shown in class.

$x \cdot f$	Number of X membership	Frequency f	$x - \bar{x}$	$(x - \bar{x})^2$	$(x - \bar{x})^2 \cdot f$
0	0	13	-1.55	2.40	31.2
9	1	9	-0.55	0.30	2.70
38	2	19	0.45	0.20	3.80
24	3	8	1.45	2.10	16.80
8	4	2	2.45	6.00	12.00
$\sum x \cdot f = 79$		$\sum f = 51 = n$			$\sum (x - \bar{x})^2 \cdot f = 66.5$

$$\text{Variance} = \frac{\sum (x - \bar{x})^2 \cdot f}{n - 1} = \frac{66.5}{50} = 1.33 \rightarrow S = \sqrt{1.33} = 1.15$$

$$\text{where } \bar{x} = \frac{\sum x \cdot f}{n} = \frac{79}{51} = 1.55$$

$$\text{coef. of variance} = \frac{S}{\bar{x}} \times 100 = \frac{1.15}{1.55} \times 100 = 74.19\%$$

sample mean: 1.55

sample variance: 1.33

sample standard deviation: 1.15

sample coefficient of variation: 74.19%

10pts

7
/ 10