

The University of Utah, Spring 2002
Mathematics 3070–5, Solutions to Midterm 1

1. During the course of his discovery of the element argon, Lord Rayleigh (1842–1919) prepared volumes of pure nitrogen by 2 different chemical methods: (1) From an air sample, from which he removed all oxygen; and (2) by chemical decomposition of nitrogen from nitrous oxide or ammonium nitrite. The following table shows the masses, measured in grams and rounded to 3 digits, of nitrogen gas obtained by the second method:

Mass (in grams)
2.301
2.299
2.299
2.302
2.299
2.299

- (a) Write the order statistics of Lord Rayleigh's measurements.

Solution: From smallest to largest they are: 2.299, 2.299, 2.299, 2.299, 2.301, 2.302.

- (b) Construct an ordered stem and leaf plot for the measurements. Explain your choice of stem and leaves.

Solution: Here is a stemleaf plot: The decimal point is 3 digit(s) to the left of the vertical stem:

```

2299 | 0000
2300 |
2301 | 0
2302 | 0

```

- (c) Is 2.302 an outlier? Explain your reasoning carefully.

Solution: This is a close call. Therefore, in light of the fact that we know very little about this data, as given, any reasonable answer is correct here. Here is the official, very careful, reasoning; a less careful answer might be sufficient though: First let us recall your textbook's definition. An outlier as a number that is more than $Q_3 + 1.5 \times \text{IQR}$. Here, $Q_1 = 2.299$ and $Q_3 = 2.301$; therefore, $\text{IQR} = 2.301 - 2.299 = 0.002$ and $Q_3 + (1.5 \times \text{IQR}) = 2.304$. Therefore, 2.302 is [very very] barely not an outlier, according to this definition.

2. A random-number generator produced the following data set; these are integers drawn at random from 0 to 99:

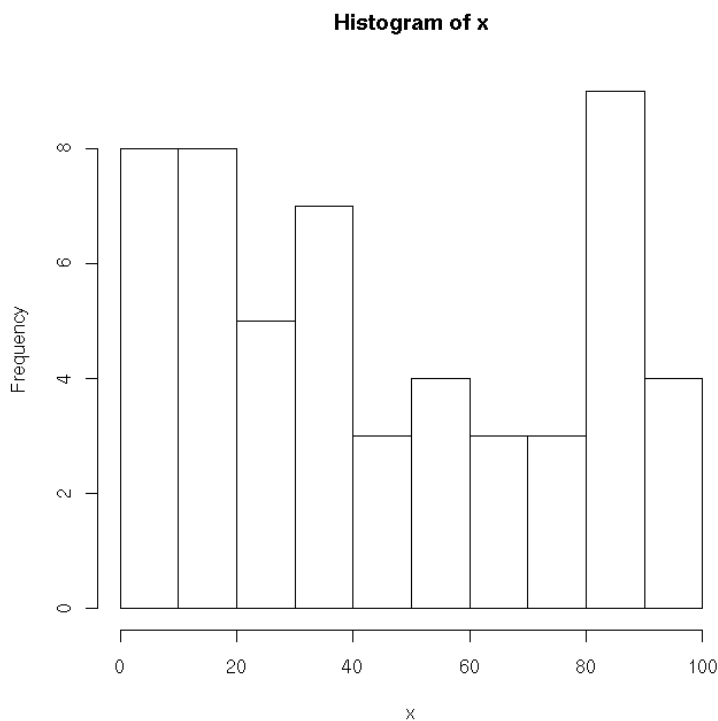
```

6 95 89 32 7 18 8 93 58 45 12 48 24 37 24 16 83 18
35 79 55 80 0 4 85 94 16 49 81 5 85 4 25 16 60 90
63 39 87 65 32 16 28 72 99 23 81 35 87 61 12 51 9 32

```

Carefully plot a histogram using class markers 10, 20, 30, ..., 90.

Solution: Here is one possibility, using frequencies:



3. *True or False:*

(a) *Is it always true that $P(A \cup B) = P(A) + P(B)$?*

Solution: No. We know that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. Therefore, the statement is false every time $P(A \cap B) \neq 0$. For instance, there are 10 people in a room; 2 of them are employed men; 3 are unemployed men; 3 are employed women; and 2 are unemployed women. You select a person at random. Let $A := \{\text{man selected}\}$ and $B := \{\text{employed selected}\}$. Then, $P(A) = 0.5$, $P(B) = 0.5$, and $P(A \cap B) = P\{\text{employed man}\} = 0.2$.

(b) *Is it always true that $P(A \cap B) = P(A) \cdot P(B)$?*

Solution: No. For instance, suppose we draw two cards, without replacement, from a standard deck of 52 cards. Let $A := \{\text{first card is a King}\}$ and $B := \{\text{second card is a King}\}$. Then, as we have seen in the lectures,

$$P(A \cap B) = \frac{4}{52} \times \frac{3}{51} \neq P(A) \times P(B) = \frac{4}{52} \times \frac{4}{52}.$$

(c) *Is it always true that if $A \subset B$ then $P(A) \leq P(B)$?*

Solution: Yes. Draw a Venn diagram.

(d) Is it always true that if $P(A) \leq P(B)$ then $A \subset B$?

Solution: No. For instance: We roll a fair die at random. Let $A := \{\text{odd number of dots rolled}\}$ and $B := \{\text{the no. of rolled dots is at least 3}\}$. Then, $P(A) = 1/2$ and $P(B) = 4/6 = 2/3$. So $P(A) < P(B)$, but A is manifestly not a subset of B .

(e) *Experimental data suggests that the probability that it rains tomorrow is 0.5, the probability that it is cold tomorrow is 0.3, and the probability that it neither rains nor is cold tomorrow is 0.3. Is it true that the probability that it rains and is cold tomorrow is 0.1?*

Solution: Yes. Draw a Venn diagram.

4. Consider the following synthetic data:

8 19 93 18 21 22 17 11 24 40 93 26 47 58 62 71 93 50 12 97 40.

(a) *What is the mean? Explain your computation carefully.*

Solution: The mean is

$$\frac{8 + 19 + 93 + 18 + 21 + 22 + 17 + 11 + 24 + 40 + 93 + 26 + 47 + 58 + 62 + 71 + 93 + 50 + 12 + 97 + 40}{21} \approx 43.90476.$$

(b) *What is the median? Explain your computation carefully.*

Solution: There are an odd [21] number of data points. So the 11th is the median [10 on each side]. But first, we sort:

8 11 12 17 18 19 21 22 24 26 40 40 47 50 58 62 71 93 93 93 97.

The 11th number is 40; this is the median.

(c) *What is the IQR? Explain your computation carefully.*

Solution: First, we need Q_1 and Q_3 . Since the data size is odd [21], any number between 18 and 19 can be viewed as Q_1 , and any number between 62 and 71 can be Q_3 . To be concrete, I average: $Q_1 = (18 + 19)/2 = 18.5$ and $Q_3 = (62 + 71)/2 = 66.5$. Therefore, $\text{IQR} = 66.5 - 18.5 = 48$.

(d) *Is "8" an outlier?*

Solution: It would be if it were $\leq Q_1 - (1.5 \times \text{IQR}) = 18.5 - (1.5 \times 48) < 0$. It most certainly is not; therefore, 8 is not an outlier.

(e) *It turns out that the standard deviation of this data set is 30.42516. Explain briefly what this means.*

Solution: That would mean that the typical deviation of the data points from the mean of approx. 44 is roughly 30 points.