

## Exam 1 – Descriptive Statistics

## Study Guide &amp; Review

## Chapters 1,2,3



Dr. Jorge Basilio

[gbasilio@pasadena.edu](mailto:gbasilio@pasadena.edu)

## General Exam Info

Exams are a way for you to show me what you have learned (and please show all your steps so I can see this!) and to give you a sense of accomplishment! They are meant to be challenging and not just homework problems with the numbers changed. I really want to prepare you for university level math classes—so some exams may be longer or more challenging than others. Remember that I do grade fairly and my goal is to push you to succeed and excel in this class. I often give hints in class as to exam problems (another great reason to come to class!), and I will post study guides along with the best way to review for each exam.

- Four exams are given during the semester—check our schedule for the exact dates.
- **Attendance required for all exams** and there are “**No Make-up Exams**” for any reason. However, I may replace the lowest exam score, regardless of the reason, with your final exam score provided the final exam score is higher than your lowest exam score and all assignments are turned-in on time.
- Your valid **PCC student ID** or a valid **government ID** is REQUIRED for all exams.
- During the exams—you will be required to leave your backpack and all non test items at the front of the room, including cell phones and smart watches. Only your pencil/eraser and calculator will be allowed during the exam, and there will be a calculator check. Should you need to leave during the exam please ask for permission first before leaving and leave your cell phone with me. Not doing these things could result in a 0 on your exam.
- Once the exam is returned, any problem you would like me to revisit must be brought to my attention by the next class session.
- **Always keep your exams!**

## Tentative Test Dates

- Exam 1

Monday, January 13

## Exam 1 Specific Info

- The exam is scheduled for the end of the class period, from 3:00 pm – 4:20 pm (80 minutes).
- Almost all questions have multiple parts
- This test will be closed book, no notes.
- You will need a calculator (only Ti84/+/-CE allowed)—you won’t be able to use your phone.
- You need to know what all the various terms in **bold** mean, but you don’t need to memorize definitions. I’m not going to ask you to “Define ...”. Instead I ask you questions that use those terms.
- **Material covered: Chapters 1, 2, 3, up to 3.3**
- Types of Questions to expect:
  - True-False
  - Circle the right answer
  - Fill-in the blank
  - Multiple-choice
  - Short response

## Chapter 1

### §1.1

- “Statistical Critical Thinking”
- Process of Statistics: prepare → analyze → conclude
- **Voluntary response sample**
- **Statistical Significance**
- Analyze common pitfalls:
  - **misleading conclusion**
  - **sample data reported not measures**
  - **loaded questions**
  - **order of the the questions**
  - **non-response**
  - **percentages**

### §1.2

- **population vs sample**
- **parameter vs statistics**
- **qualitative vs quantitative**
- **continuous vs discrete**
- Levels of Measurements
  - **nominal**
  - **ordinal**
  - **interval**
  - **ratio**

### §1.3

- collecting sample data
- experiment
- observational study
- experimental design
- **simple random sample**
- **random sample**
- Sampling:
  - **systematic**
  - **convenience**
  - **stratified**
  - **cluster**
- Experimental:
  - **completely randomized**
  - **randomized block**
  - **matched-pair**
- Observationall:
  - **retroactive**
  - **cross-sectional**
  - **prospective**

## Chapter 1

For the following I will ask you multiple choice/true-false/circle the right answer types of questions (you may need to provide examples):

- You should be able to distinguish between qualitative vs quantitative data.
- You should be able to distinguish between discrete vs continuous data.
- You should know about the different sampling methods (convenience, stratified, cluster, systematic)
- You should be able to identify whether data is nominal, ordinal, interval, or ratio (levels of measurement).
- Given a scenario, be able to identify the population, sample, parameter, statistic, variable, data and sampling method used.

## Chapter 2

### §2.1

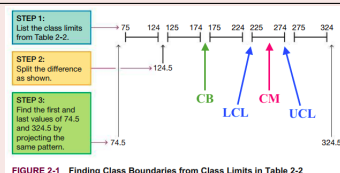
#### • Frequency Distributions

- LCL lower-class limit
- UCL upper-class limit
- CB class boundary  $CB = (UCL + LCL)/2$
- CM class midpoint

$$CM = (LCL + UCL)/2$$

- CW class width

$$CW = \text{difference b/w } LCLs \text{ or } UCLs$$

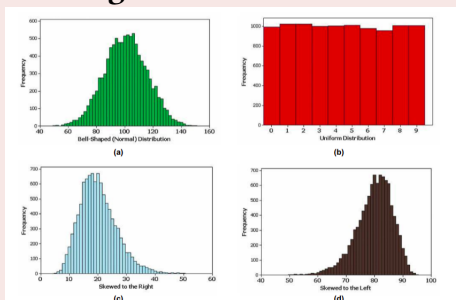


- Relative Frequency Distributions
- Percentage for Class
- Cumulative Frequency Distributions

### §2.2

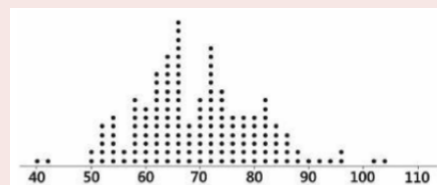
#### • Histograms

- Shape of a histogram/frequency distribution
  - Bell-shaped/Normal/Symmetric
  - Bimodal
  - Uniform
  - Skewed Left
  - Skewed Right



### §2.3

#### • dot plots



- time-series
- bar graphs (bar graphs vs histograms)
- Pareto chart
- Pie Chart
- Stemplot (stem-leaf plot)



- Graphs that deceive:
  - non-zero vertical axis
  - pictographs

## Chapter 2

For the following I will ask you multiple choice/true-false/circle the right answer types of questions (you may need to provide examples):

- Be able to fill out a frequency/relative frequency/cumulative relative frequency table and answer questions like those from the worksheets
- Be able to understand and read off information from histograms/dotplots/stemplots/bar graphs/pareto graphs/pie charts/time-series
- Understand graphs that deceive

## Chapter 3

### §3.1

#### • Measures of center

##### – Mean:

\* sample:  $\bar{x} = \frac{\sum x}{n}$

\* population:  $\mu = \frac{\sum x}{N}$

##### – Median: *Med*

\* when  $n$  is odd

\* when  $n$  is even

##### – Mode (bimodal, trimodal, multimodal)

##### – Midrange: $MR = \frac{\min + \max}{2}$

#### • Notation: $\sum, n, N, x, \bar{x}, \mu$

#### • Stat Law of Rounding

#### • Mean for a frequency distribution:

– sample:  $\bar{x} = \frac{\sum(f \cdot x)}{\sum f}$

– use CMs for the  $x$ s

#### • Weighted Means and GPAs

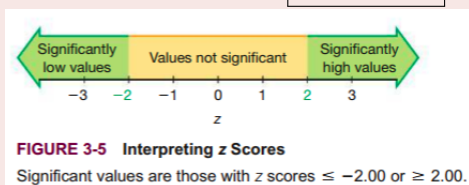
–  $\bar{x}_w = \frac{\sum(w \cdot x)}{\sum w}$

### §3.3

#### • z-scores

– z-score from sample:  $z = \frac{x - \bar{x}}{s}$

– z-score from population:  $z = \frac{x - \mu}{\sigma}$



#### • Percentiles

##### – $k^{th}$ Percentile:

$$P_k = \frac{\# \text{ scores} < \text{given score}}{\text{total \# scores}}$$

##### – Finding the score $L$ given a percentile $k$ :

$$L = \frac{k}{100} \cdot n$$

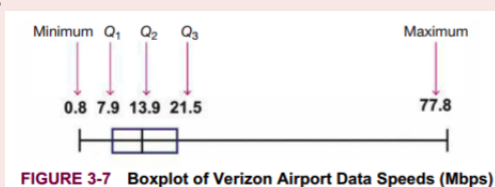
\* if  $L$  is a decimal, round up

\* if  $L$  is whole, then average the  $k^{th}$  score and the next higher score

#### • Quartiles: $Q_1 = P_{25}, Q_2 = P_{50} = M, Q_3 = P_{75}$

#### • 5 Number Summary: $\min, Q_1, Q_2 = M, Q_3, \max$

#### • Box Plots



### §3.2

#### • Measures of Variation

– **Range:** range = max - min

– **Deviation** deviation =  $x - \bar{x}$

– **Standard Deviation:**

\* **sample standard deviation:**

$$s = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}$$

shortcut:  $s = \sqrt{\frac{n(\sum x^2) - (\sum x)^2}{n(n - 1)}}$

\* **population standard deviation:**

$$\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}$$

– **Variation**

\* **sample variation:**  $s^2$

\* **population variation:**  $\sigma^2$

#### • Importance of Standard Deviation

#### • Range Rule of Thumb:

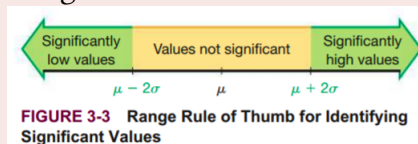
– **Significantly Low:** value  $< \mu - 2\sigma$

– **Significantly High:** value  $> \mu + 2\sigma$

– **Not Significant:**  $\mu - 2\sigma < \text{value} < \mu + 2\sigma$

– Approximation for standard deviation:

$$s \approx \text{range}/4$$



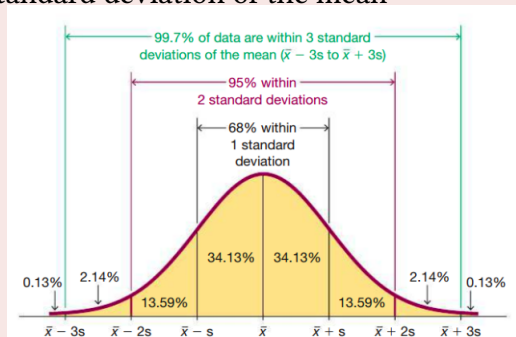
#### • Empirical Rule (69-95-99.7 Rule):

– For bell-shaped/normal distributions

– Approximately 68% of the values are within 1 standard deviation of the mean

– Approximately 95% of the values are within 2 standard deviation of the mean

– Approximately 99.7% of the values are within 3 standard deviation of the mean



## Chapter 3

For the following I will ask you multiple choice/true-false/circle the right answer types of questions (you may need to provide examples):

- You should know the statistical notation for population mean, sample mean, population standard deviation, sample standard deviation, population variance, sample variance, population size, and sample size.
- Be able to calculate mean, median, mode, sample standard deviation, and variance using your calculator
- Given a frequency distribution, be able to calculate the mean.
- Given a set of data, be able to create a box plot, write the 5 number summary (min, Q1, Q2 (median), Q3, and max).
- Be able to construct a box plot.
- Be able to interpret percentiles.
- Given two scenarios, be able to determine which situation is better using z-scores
- Be able to identify statistical significance using z-scores.

## Practice Test

**Note: solutions to all review exercises are available at the back of the book!**

- **Chapter 1**
  - Chapter Quiz Quiz, p. 35: 1-10 all
  - Review Exercises, p. 35-37: 2, 5, 7, 8, 9
  - Cumulative Review Exercises, p. 37: 1, 2, 3, 4, 8, 9, 10
- **Chapter 2**
  - Chapter Quiz Quiz, p. 75-76: 1-6, 8-10
  - Review Exercises, p. 76-77: 1, 2, 3, 4, 7, 8
  - Cumulative Review Exercises, p. 77-78: 1, 2, 3, 4, 5, 6
- **Chapter 3**
  - Chapter Quiz Quiz, p. 127: 1, 2, 3, 4, 5, 6, 7, 8, 10
  - Review Exercises, p. 127-128: 1, 2, 3, 4, 5, 6, 7
  - Cumulative Review Exercises, p. 128: