STAT 50 – Elementary Statistics — Spring 2019

Exam 1 – Descriptive Statistics

Study Guide & Review

Dr. Jorge Basilio

Pasadena City College

gbasilio@pasadena.edu

Chapters 1,2,3

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.

- Attendance required for all exams–I do NOT drop the lowest exam score.
- **"Make-up Exams"** are given only in extreme cases and at instructor's discretion; a student is allowed at most one make-up exam. (Documentation must be provided for the instructor to even consider a make up exam. This means you would need a doctor's note, etc.) A "Make-Up Exam" means you will be allowed to replace the missing score with the percentage you earn on the final exam. Please contact your instructor as soon as possible should there be a problem.
- **Your student 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 graded and returned, any problem you would like me to revisit must be brought to my attention by the next class session.
- Always keep your exams!

Exam 1 Date & Time		
• Exam 1	Tuesday, March 12	1:15 pm – 2:35 pm

ALSO, after the exam:

- As the test is only 80 minutes long, we will have a 10 minutes break, then continue with new material.
- I take attendance at the end of class on test days.

Exam 1 Specific Info

- Almost all questions have multiple parts
- You will need a calculator (only Ti83/84 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
- I will ask you multiple choice/true-false/circle the right answer/fill-in the blanks/short response types of questions (you may need to provide examples).

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 $CW =$ difference b/w $LCLs$ or $UCLs$

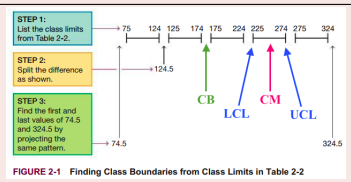
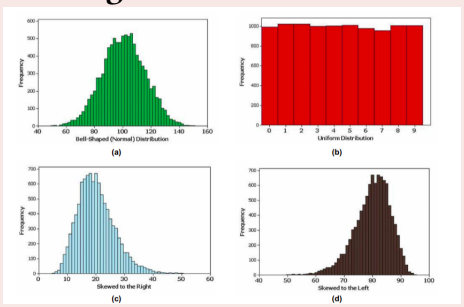


FIGURE 2-1 Finding Class Boundaries from Class Limits in Table 2-2

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

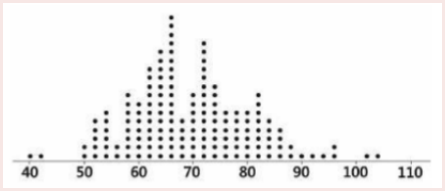
§2.2

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

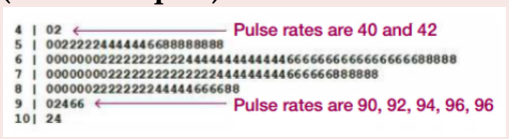


§2.3

- **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: $\boxed{\bar{x} = \frac{\sum x}{n}}$
    - population: $\boxed{\mu = \frac{\sum x}{N}}$
  - **Median:** M
    - when n is odd
    - when n is even
  - **Mode** (bimodal, trimodal, multimodal)
  - **Midrange:** $MR = \frac{\min + \max}{2}$
- **Notation**: $\sum$, n, N, x, $\bar{x}$, μ
- **Stat Law of Rounding**
- Mean for a frequency distribution:
  - sample: $\boxed{\bar{x} = \frac{\sum(f \cdot x)}{\sum f}}$
  - use CMs for the xs
- **Weighted Means** and GPAs
  - $\boxed{\bar{x}_w = \frac{\sum(w \cdot x)}{\sum w}}$

§3.3

- **z-scores**
  - z-score from sample: $\boxed{z = \frac{x - \bar{x}}{s}}$
  - z-score from population: $\boxed{z = \frac{x - \mu}{\sigma}}$

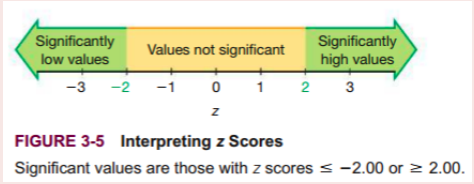


FIGURE 3-5 Interpreting z Scores
Significant values are those with z scores ≤ -2.00 or ≥ 2.00.

- **Percentiles**
  - k^{th} **Percentile:**
    $$\boxed{P_k = \frac{\text{\# scores < given score}}{\text{total \# scores}}}$$
  - **Finding the score L given a percentile k:**
    $$\boxed{L = \frac{k}{100} \cdot n}$$
    - 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**

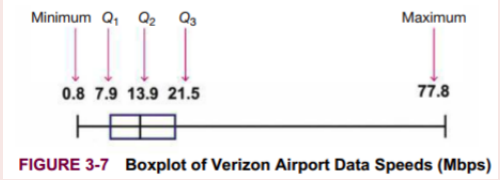


FIGURE 3-7 Boxplot of Verizon Airport Data Speeds (Mbps)

§3.2

- **Measures of Variation**
  - **Range:** range = max - min
  - **Deviation** deviation = $x - \bar{x}$
  - **Standard Deviation:**
    - **sample standard deviation:**
      $$\boxed{s = \sqrt{\frac{\sum(x - \bar{x})^2}{n}}}$$
      shortcut: $\boxed{s = \sqrt{\frac{n(\sum x^2) - (\sum x)^2}{n(n-1)}}}$
    - **population standard deviation:**
      $$\boxed{\sigma = \sqrt{\frac{\sum(x - \mu)^2}{N}}}$$
  - **Variation**
    - **sample variation:** s^2
    - **population variation:** σ^2
- Important of Standard Deviation
- **Range Rule of Thumb:**
  - **Significantly Low:** value $< \mu - 2\sigma$
  - **Significantly High:** value $> \mu + 2\sigma$
  - **Not Significant:** $\mu - 2\sigma <$ value $< \mu + 2\sigma$
  - Approximation for standard deviation: $s \approx$ range/4

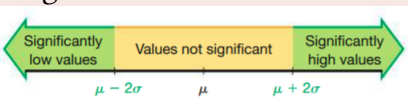


FIGURE 3-3 Range Rule of Thumb for Identifying Significant Values

- **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

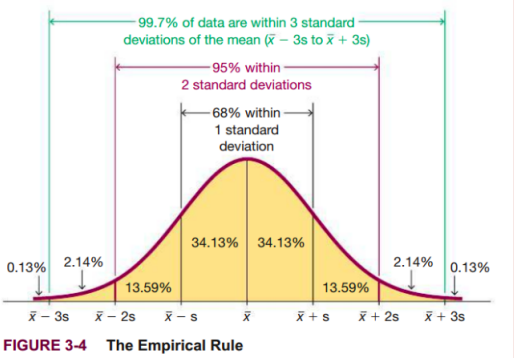


FIGURE 3-4 The Empirical Rule

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.