Stat 50 - Elementary Statistics

Exam 2 – Probability, Part 1

Study Guide & Review

Chapters 4 & 5

Dr. Jorge Basilio

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!**

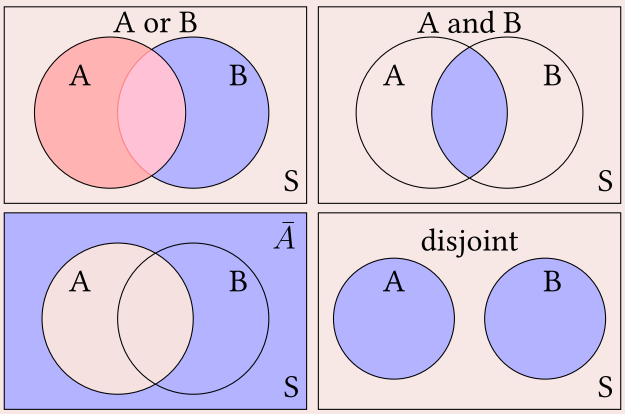
Exam 2 Specific Info

- The exam is scheduled for the end of the class period, the last 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: 4.1, 4.2, 4.3, 5.1, & 5.2**
- Old material: You still need to know Chapter 2 and Chapter 3, especially Chapter 3 material.
- $(M \to E)$ **"Math to English"** If you see this, it just means that I would like you to write your answer in a complete sentence and include all relevant **units**.
- Types of Questions to expect:
  - True-False
  - Circle the right answer
  - Fill-in the blank
  - Multiple-choice
  - Short response

Chapter 4

§4.1

- **Set Theory**
  - **set notation**: { things }.
  - **sample space** S.
    Ex: $S = \{HH, HT, TH, TT\}$,
    Ex: $S = \{1, 2, 3, 4, 5, 6\}$, etc
- **events**: $A, B, C, \ldots$
  - Ex: $E = \{2, 4, 6\}$= "rolling an even #"
- **simple events**
  - Ex: $A = \{5\}$= "rolling a 5"
- **Venn Diagrams**



- **Probability**
  - **Classical/Theoretical Probability**
    $$P(A) = \frac{\#\text{outcomes in } A}{\#\text{outcomes in } S}$$
  - **Relative Frequency Probability**
    $$P(A) = \frac{\#\text{occurences of } A}{\#\text{of trials } n}$$
  - **Subjective Probability**
  - Rounding rule for probability
  - Probabilities can be written as decimals, fractions, and percentages. But decimals are preferred.
  - $0 \leq P(A) \leq 1$
  - Impossible events: $P(A) = 0$
  - Certain events: $P(A) = 1$

§4.2

- **Independent** vs **Dependent** events
- **Probability Rules**
  - **Sum Rule**
    General:
    $$P(A \textbf{ or } B) = P(A) + P(B) - P(A \textbf{ and } B)$$
    Disjoint: $P(A \textbf{ or } B) = P(A) + P(B)$
  - Using "and" in one selection:
    $$P(A \textbf{ and } B) = \frac{\#A \text{ and } B}{\#S}$$
  - **Multiplication Rule**
    A, B indep. $\rightarrow$ $P(A \textbf{ and } B) = P(A) * P(B)$
    A, B dep. $\rightarrow$ $P(A \textbf{ and } B) = P(A) * P(B|A)$
  - **Completments Rule** $P(\bar{A}) = 1 - P(A)$
  - Probability of Disjoing sets: $P(A\textbf{and}B) = 0$ when A and B are disjoint
- Using Probability to determine significance:
  - $P(x \text{ or more}) \leq 0.05 \rightarrow x$ is significantly high
  - $P(x \text{ or less}) \leq 0.05 \rightarrow x$ is significantly low

§4.3

- **"At least one" Trick**
  $$P(\textbf{at least one}) = 1 - P(\textbf{none})$$
- **Conditional Probability**
  $$P(B|A) = \frac{P(A \textbf{ and } B)}{P(A)} = \frac{\frac{\#A \text{ and } B}{\#S}}{\frac{\#A}{\#S}} = \frac{\#A \text{ and } B}{\#A}$$

Chapter 4

For the following I may ask you multiple choice/true-false/circle the right answer types of questions but mostly you'll see free-response questions similar to the worksheet problems:

- Be able to distinguish between classical/theoretical, relative frequency, and subjective probabilities.
- Be able to identify events as impossible or certain.
- Be able to use set notation to express: sample space, events, and simple events
- Be able to compute probabilities from the definition and using the Sum rule, Multiply rule, and Complement rule
- Understand the difference between the **two formulas using "and":** Use $P(A\text{and}B) = \frac{\#A\text{and}B}{\#S}$ when making **one selection** (ex: probability of drawing one card that's a heart and Queen); Use $P(A \text{ and } B) = P(A) * P(B|A)$ when making **two selections** (ex: probability of selecting two blue balls from a bag that contains....)

- Understand when events are independent vs dependent and how this effects probabilities using conditional probabilities
- Study all of the examples from the worksheets carefully! Be be prepared to information given to you in table.

Chapter 5

§5.1

- **Random Variables** X
- Discrete vs Continuous Random Variables
- **Probability Distributions** for a RV X
  - as a list of all probabilities
  - as a table of values
  - as a histogram
  - **Requirements**
    1. $\sum P(x) = 1$
    2. $0 \leq P(x) \leq 1$
- Formulas for Probability Distributions

  mean: $\boxed{\mu = \sum [x \cdot P(x)]}$

  variance: $\boxed{\sigma^2 = \sum (x-\mu)^2 \cdot P(x)]}$

  variance: $\boxed{\sigma^2 = \sum [x^2 \cdot P(x)] - \mu^2}$

  standard deviation: $\boxed{\sigma = \sqrt{\sum [x^2 \cdot P(x)] - \mu^2}}$
- **Expectation** $E(X) = \sum [x \cdot P(x)]$
  - key point: the "long run behavior" or "many/large sample"
  - expectation is also called the mean. So the mean tells us the long run expectations
- **Range Rule of Thumb:**
  - **Significantly Low:** value $< \mu - 2\sigma$
  - **Significantly High:** value $> \mu + 2\sigma$
  - **Not Significant:** $\mu - 2\sigma <$ value $< \mu + 2\sigma$

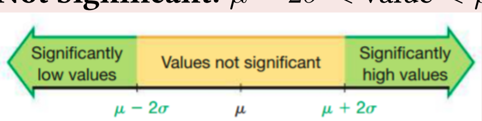


FIGURE 3-3 Range Rule of Thumb for Identifying Significant Values

§5.2

- **Binomial Probability Distributions**
  - **Requirements**
    1. fixed number of trials
    2. trials are independent
    3. **only two** outcomes ("success" or "failure")
    4. probability of success of each trial remains the same
  - **Notation**
    - n total number of trails
    - p be the probability of a single success
    - q be the probability of a single failure. (NOTE: $q = 1 - p$)
    - x be the number of successful trials of X. So notice that x can take values from 0 up to n, i.e. $x = 0, 1, 2, 3, \ldots, n$.
- **Binomial Probability Formula**
  - Let $P(X = x)$ denote the **probability of exactly x successful trails out n in a binomial probability distribution**, then

    $P(X = x) = {}_n C_x p^x q^{n-x}$

    $\boxed{P(X = x) = \text{binompdf}(n, p, x)}$
- Formulas for Binomial Distribution
  - $\boxed{\mu = np}$ $\boxed{\sigma^2 = npq}$ $\boxed{\sigma = \sqrt{npq}}$
  - Recall: the range rule of thumb
- Cumulative Probabilities

  $\boxed{01 \cdots x} \cdots n$ $P(X \leq x) = \text{binomcdf}(n, p, x)$

  $\boxed{01 \cdots} x \cdots n$ $P(X < x) = \text{binomcdf}(n, p, x - 1)$

  $01 \cdots \boxed{x \cdots n}$ $P(X \geq x) = 1 - \text{binomcdf}(n, p, x)$

  – $01 \cdots x \boxed{\cdots n}$ $P(X \leq x) = 1 - \text{binomcdf}(n, p, x)$

Chapter 5

For the following I may ask you multiple choice/true-false/circle the right answer types of questions but mostly you'll see free-response questions similar to the worksheet problems:

- Be able to distinguish between discrete and continuous random variables.
- Be able to determine whether or not information given represents a probability distribution.
- Be able to build a probability distribution for a given random variable. (Ex: a) X = number of Heads that come up out of 4 coin flips, b) X = number of girls that a family has if they plan to have 4 children)
- Be able to recognize if a probability distribution is binomial so that you can use binompdf or binomcdf. Also, know the four requirements needed for a binomial distribution.
- Be able to compute probabilities of binomial distributions for a single value of x and multiple values of x using the rules of probability from Ch 4 (including "at least one", "or", "and" etc).

- Be able to create a table for a binomial probability distribution using your calculator.
- Be able to compute expectations and know the significance of expectation.
- Be able to compute the mean, standard deviation, and variance given a probability distribution.
- Be able to recognize if a value is significantly low or high using the range rule of thumb.

Calculator Skills you must know

- STATS Edit lists, clear lists
- 1-Var Stats use to compute mean, standard deviation, 5 number summary
- 2nd+VARS Distributions → binompdf(n,p,x) and binomcdf(n,p,x)
- 2nd+ENTER Trick!

Practice Test

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

- **Chapter 4**
  - Chapter Quick Quiz, p. 178: 1, 2, 4-10 all
  - Review Exercises, p. 178-179: # 1-10 all, 12, 13
  - Cumulative Review Exercises, p. 180-181: # 1, 2, 3
- **Chapter 5**
  - Chapter Quick Quiz, p. 220: # 1-10 all
  - Review Exercises, p. 220-221: # 1-5 all, 7, 8, 9
  - Cumulative Review Exercises, p. 221-222: # 1

Formula Sheet for Exam 2

- $P(A \textbf{ or } B) = P(A) + P(B) - P(A \textbf{ and } B)$
- $P(A \textbf{ and } B) = P(A) * P(B)$
- $P(A \textbf{ and } B) = P(A) * P(B|A)$
- $P(B|A) = \dfrac{P(A \textbf{ and } B)}{P(A)}$
- $\mu = \sum[x \cdot P(x)]$
- $\sigma^2 = \sum(x-\mu)^2 \cdot P(x)]$
- $\sigma = \sqrt{\sum[x^2 \cdot P(x)] - \mu^2}$
- $E(X) = \sum[x \cdot P(x)]$

NOTE ON "AND"

There's two different formulas that we use when computing probabilities of events using the word "and." the distinction is easy to remember if you differentiate between them using the following:

- when making **one selection using two categories** use:

$$P(A\text{and}B) = \frac{\#A \text{ and } B}{\#S}$$

For example, you use this formula to compute the probability of drawing one card that is a Club and Queen.

- when making **two selections (repeating the same thing)** use:

$$P(A \text{ and } B) = P(A) * P(B|A)$$

For example, let's say you have a bag with 50 candies of three different colors (R, G, B). You use this formula to compute the probability of getting two red candies when you make two selections.

Note that to compute $P(B|A)$ you need to know whether the events A and B are independent or dependent since it effects the probability. In the candy example, it depends whether or not you are making selections with (independent) or without replacement (dependent).