Unit IV

Probability and Risk

At the conclusion of each class, create a cheat sheet here to summarize the material.

1 Independent Events

2 Dependent Events and Conditional Probability

3 Permutations and Combinations

4 Expected value

Project: COVID-19 Testing in a Maternity Ward

Scenario: It is fall 2020. You are an epidemiologist at a large urban hospital in Omaha, Nebraska. The incidence of COVID-19 is rising locally, and your hospital is concerned about the effect of active infections on patient health and as well as becoming a hot spot for the spread of COVID. The labor and delivery (L&D) ward has asked for your assistance in designing COVID protocols. Currently, they test only symptomatic pregnant people for SARS-CoV-2, and they transfer positive patients to the COVID ward for labor, delivery, and recovery. They are considering implementing full testing of all patients at check-in.

Your tasks: Analyze the test under your local conditions. The hospital is using Abbott's RealTime SARS-CoV-2 test, a reverse-transcriptase polymerase chain reaction (RT-PCR) test. Abbott has not released the details, but a number of independent sources estimated sensitivity of the average PCR test to be up to 80% and specificity up to 99%.

- Construct a table to determine the positive predictive value (PPV) and negative predictive value (NPV) of the test under local conditions. (Note: County health officials estimate that there are approximately 323 actives SARS-CoV-2 infections per 100,000 residents.)
- Analyze the risk of missing a COVID infection in an L&D patient. (Note: Do a little research here; what risk does COVID pose to the patient and others?)
- Analyze the risk of incorrectly diagnosing an L&D patient with COVID. (Note: Do a little research here; what are COVID treatments in L&D like? What impact do they have on the patient and others?)
- Make a recommendation about a testing regime for this ward: Should all patients be tested? What actions should be taken if a patient is positive (or negative)?

Product: Construct a report (1-2 pages, double-spaced, with at least one table) for the hospital recommending a testing policy for SARS-CoV-2 in the L&D ward.

Submission Guidelines:

- Submit online under the appropriate assignment. Use good grammar and writing.
- It should include the usual sections - a good title/author list, an introduction outlining the scenario, an analysis of the test in light of local conditions, your recommendations.
- The report should be appropriate to your audience.
- Use at least three external sources and cite them.

Day 1 - Independent Events

Probability by Counting

Most of us have a pretty innate understanding of basic probability theory. We'll start with a few familiar types of examples that you can solve using common sense and some careful counting.

1. You roll a standard 6-sided die. What is the probability you get a 4? any even number?

2. You look at the clock. What is the probability that the minutes read less than 15?

3. A friend has 2 children named Sam and Alex. You'd like to know if they are boys or girls. (Note: assume that about half of kids are boys and half are girls; this is not actually correct, but it's pretty close.)

   (a) List out all the possible boy/girl combinations for these two children. What are the chances both are girls?

   (b) The friend tells you that Sam is actually a girl. What are the chances now that both are girls?

   (c) You misheard - the friend didn't say Sam is a girl, just that one child is a girl. Now what are the chances both are girls?

Can you do any of these calculations by formula rather than by counting? These are called **independent events**, two (or more) things that happen where the first thing doesn't affect the second thing

You roll two standard 6-sided dice (see all possible rolls below - this is the **sample space**).

1,1	1,2	1,3	1,4	1,5	1,6
2,1	2,2	2,3	2,4	2,5	2,6
3,1	3,2	3,3	3,4	3,5	3,6
4,1	4,2	4,3	4,4	4,5	4,6
5,1	5,2	5,3	5,4	5,5	5,6
6,1	6,2	6,3	6,4	6,5	6,6

4. What is the probability that they add to 12? add to 4? add to 4 or 12?

5. What is the probability that they add to 1?

6. What is the probability that the second is bigger than the first?

7. What is the probability both are 6s? What is the probability both are the same?

Probability by Calculation for Independent Events

8. Can you do any of the above by calculation?

9. Which of these are these independent events?

10. If two events are independent, how do you find the probability that both happen?

Complementary Events

You flip a fair coin.

11. What is the probability that it's a head? What is the probability that it's a tail?

These are called **complementary events**, two things that are opposites.

12. You the coin three more times. What is the probability they are all heads? What is the probability that they are not all heads?

13. You flip the coin ten times total. What is the probability that there is at least one tail?

14. How is the probability that something happens related to the probability that it doesn't happen?

Cases

You own 6 blue shirts, 3 red shirts, and 2 white shirts. You also have 8 pairs of white socks and 4 pairs of blue socks.

15. You pull out a shirt in the dark. What is the probability that the shirt is blue? Red? Blue or red?

These are called **cases**, where there are several possible different situations that could happen (but not at once).

16. You pull out a pair of socks. What is the probability that both the shirt and the socks are white? that exactly one of them is white? that at least one of them is white?

17. How do you find the probability of something if you can break it down into separate cases?

Answers: 1. 1/6 = 17%, 1/2 = 50%. 2. 25% 3a. BB/BG/GB/GG; 25%. 3b. 50%. 3c. 33%. 4 2.8%, 8.3%, 11%. 5. 0%. 6. 42%. 7. 2.8%, 17%. 11. 50%, 50%. 12. 6%, 94%. 13. 99.9%. 15. 55%, 27%, 82%. 16. 12%, 61%, 73%.

Day 2 - Dependent Events and Conditional Probability

Dependent events

Last time, we talked about events which are independent, meaning the outcome of one of them doesn't affect the other. This time, we'll deal with **dependent events**, where the outcome of one does affect the other.

1. You roll a die twice.

   (a) What is the probability it will be even both times?

   (b) Are the two dice rolls independent or dependent? That is, does whether the first die is even or odd affect the chances that the second die is even or odd?

2. You draw two cards from a standard deck of playing cards.

   (a) What is the probability that they are both aces?

   (b) Are the two cards you draw independent events or dependent events? That is, does the identity of the first card affect your chances of drawing the second card?

3. You draw a poker hand of 5 cards.
   (a) What is the chance they are all hearts?
   (b) What is the chance you drew a flush? (Note: a flush is 5 cards of the same suit.)
   (c) What is the chance that you have no card higher than a 10?

Conditional Probability

The table below shows the number of survey subjects who have received and not received a speeding ticket in the last year, along with the color of their car.

	Ticket	No ticket	Total
Red car	15	135	150
Other color car	45	470	515
Total	60	605	665

4. Find the probability that a randomly chosen person has a speeding ticket given that they have a red car.

5. Find the probability that a randomly chosen person has a red car given that they have a speeding ticket.

Medical Testing

In 2018, there were approximately 327.2 million people in the US, and 1.1 million were estimated to be HIV positive. One of the better modern tests used to screen for HIV has a is 100% accurate for people with HIV and 99.9% accurate for people without HIV. [1]

6. Without running the numbers, guess: if you test positive, what is the chance you actually are positive?

7. Assume every person in the country is tested. Fill in the table.

	Positive Test	Negative Test	Total
HIV Positive			
HIV Negative			
Total			

8. Of the people with HIV:

   (a) How many will test correctly, both as a number and as a proportion? This is called the **sensitivity** of the test.

   (b) How many will test incorrectly? These are called **false negatives**.

9. Of the people without HIV:

   (a) How many will test correctly? This is called the **specificity** of the test.

   (b) How many will test incorrectly? These are called **false positives**.

[1] 4th generation ELISA. These numbers change constantly as tests improve. See, ex, T. Alexander (2016). Human Immunodeficiency Virus Diagnostic Testing: 30 Years of Evolution. *Clinical and Vaccine Immunology 23(4)*, 249-253.

10. How should a given person react to this test?

    (a) If someone tests positive, what is the probability that the test is correct? (This is the **positive predictive value** or **PPV** of the test.)

    (b) If someone tests negative, what is the probability that the test is correct? (This is the **negative predictive value** or **NPV**.)

    (c) Interpret the PPV (in other words, write a coherent English sentence about what it means).

11. A number of people have proposed mandatory testing for HIV and other illnesses; for example, all blood banks test for HIV, and most will notify donors if they test positive. Different organizations use different levels of screening involving more tests or fewer tests and tests with varying accuracy. What are the pros and cons of mandatory testing? Would your answer be different for deciding whether to include a donation in a blood bank vs. notifying a person that they had tested HIV-positive?

Note: The definitions in your book are a little vague. The definitions here are correct.

Answers: 1a. 25%. 2a. 0.45%. 3a. 0.0495%. 3b. 0.198%. 3c. 14.5%. 4. 10%. 4. 25% 8a. 1.1 million (100%). 8b. 0 (0%). 9a 325.77 million. 9b. 0.33 (33%). 10a. 77%. 10b. 100%.

Day 3 - Permutations and Combinations

Permutations vs. Combinations

We often run into situations where we have some items, and we want to take a few of them and do something with them.

1. Sally, Bob, and Gilbert are running for class office.

   (a) One needs to be elected President and one Treasurer. How many election outcomes are there? List them.

   (b) What is the probability that Sally will be elected (as either President or Treasurer)?

   (c) Two will be selected to go to the student Senate. How many election outcomes are there? List them.

   (d) What's the probability that Sally will be elected to the Senate?

   (e) Did you get the same answers for the President/Treasurer and Senate? If yes, why are they equivalent? If no, what's the difference?

If we take some things and line them up in order, then we say we **pick** them, and we now have an **arrangement** or a **permutation**. If we take some things and throw them together without paying attention to their order, we say we **choose** them, and we now have a **selection** or **combination**.

2. There are 25 students in your class.

    (a) You need to pick two, one to be class President and one to be Treasurer. How many ways can you do this?

    (b) You need to choose two to go to the student Senate. How many ways can you do this?

3. There are 25 students in your class.

    (a) How many ways can you pick a President, a Treasurer, and a Secretary?

    (b) How many ways can you choose three to go to the student Senate? (Hint: how many ways can you order the three students you choose?)

Formulas

Let's summarize what we've done so far.

4. What is a permutation (or arrangement)?
   (a) Explain in your own words.

   (b) Give a formula for the number of permutations of k items picked out of n items. This is commonly written nPk or P(n,k).

   (c) How do you calculate a permutation on your calculator?

5. How many ways can you order 4 people? 5? n?

6. What is a combination (or selection)?
   (a) Explain in your own words.

   (b) Give a formula for the number of combinations of k items selected out of n items. This is commonly written nCk or C(n,k).

   (c) How do you calculate a combination on your calculator?

A shortened form of notation is the **factorial**. Give versions of the formulas for combinations and permutations using factorials.

Repetition

We have already seen one final type of situation, where we may be able to reuse an item.

7. There are 3 students in a class. You need to choose a President, Treasurer, and Secretary. After a recent change to the bylaws, one student may hold more than one office.

   (a) How many ways can you choose the President, Treasurer, and Secretary?

   (b) What is the probability that the three offices will be held by different people?

8. A different class has 25 students in it. You need to choose a President, Treasurer, and Secretary, and a student may hold multiple offices.

   (a) How many ways can you do this?

   (b) What is the probability that the three offices will be held by different people?

We call this a **permutation with repetition**. There is also a **combination with repetition**, but we will not cover the formula in this class.

Formulas, continued

9. What is a permutation with repetition?

   (a) Explain in your own words.

   (b) Give a formula for the number of permutations (with repetition) of k items picked out of n items.

   (c) How do you calculate it on a calculator?

10. CHALLENGE: There is a track and field meet where 25 athletes compete in 5 events.

    (a) You want to record a list of first place winners for the events. How many possible lists are there?

    (b) Your friend also keeps a tally for the student paper but only cares who got gold (and how many), not which event it was. How many possible lists are there?

Answers: 1a. 6. 1b. 67%. 1c. 3. 1d. 67%. 2a. 600. 2b. 300. 3a. 13,800. 3b. 2300. 5. 24, 120, n! 7a. 27. 7b. 22%. 8a. 15,625. 8b. 88%. 10a. 15,625. 10b. 2925.

Day 4 - Expected value

We can also calculate the expected value, or the amount we expect to gain/lose on average from something.

Games

1. A street gambler flips a coin. A bystander can pay \$1 to bet on the outcome: if the coin lands on heads, they get \$2 back; if tails, they get nothing (so they lose the \$1).

   (a) What should they expect to win/lose overall?

   (b) Someone plays the game 9 times, and the coin lands on tails 7 times. What is their average win/loss?

   (c) This person continues playing for awhile and estimates the coin lands on tails approximately 70% of the time. Estimate their wins/losses.

   (d) After being confronted, the street gambler admits that the coin is not fair: it lands on tails 75% of the time, on average. Now determine the average wins/losses a player can expect.

This is called the **expected value** of the game.

1. In roulette, a wheel has 38 spaces: 18 red numbered spaces, 18 black numbered spaces, and 2 green numbered spaces. In one way to play, a player picks one of the spaces, pays \$1, and the wheel is spun. If it lands on the chosen space, they get \$36 back; otherwise, they get nothing.

   (a) Without calculating anything, does this game sound fair? Will wins and losses average out?

   (b) What is the probability the player will win?

   (c) What is the probability the player will lose?

   (d) What is the expected value of a roulette bet?

   (e) The dealer would like to change the cost to play to make roulette fair. How much should you pay each play?

   (f) The dealer decides instead to change the payout to make roulette fair. How much should a win pay?

Other Costs

2. You buy a \$5 raffle ticket for charity. If 2000 tickets are sold and the 1 lucky winner gets \$4000, what is the expected value of this ticket?

3. A 40-year-old man in the US has a 0.242% risk of dying during the next year. If he buys a \$275 insurance policy that pays a \$100,000 death benefit, what is the expected value of the policy?

4. About 75% of the 140 million housing units in the US have working smoke detectors, and about 7,510 people are injured or die each year in a fire in a unit without a working smoke detector.

   (a) If having a working smoke detector in a house reduces the risk of injury or death from a fire by 55%, how many deaths would be prevented each year by equipping every housing unit with a working detector?

   (b) If a smoke detector costs \$10 on average, what is the **cost per life saved** if all units are equipped with detectors?

Risk Analysis

5. A plane holding 13 passengers flies 75 miles and has a crash landing, killing 2 passengers. How many passenger-miles has the plane flown? How many fatalities per passenger mile is that?

6. The fatality rate on US airlines in the 21st century has averaged 1 fatality per 3.4 billion passenger-miles. A proposed change to aviation regulations will decrease accidents but increase prices. We can analyze the impact this will have on safety.

   (a) The fatality rate for driving in the 21st century has average 34 times that for flying. What is the fatality rate per passenger-mile for driving?

   (b) A proposed change to air traffic safety would decrease the fatality rate for flying by a factor of 10. What is the new fatality rate for flying?

   (c) This proposed change would also increase cost, resulting in 10% of the people who would fly driving instead. What would the new fatality rate for this population be? Would it be higher or lower than the original rate?

Answers: 1a. \$0. 1b. -\$0.55 1c. -\$0.40. 1d. -\$0.50. 1b. 1/38. 1c. 37/38. 1d. -\$0.053. 1e. \$1.89. 1f. \$4.22. 2. -\$3. 3. -\$33. 4a. 4,130.5. 4b. 5. 975 pass-mi; 0.00205 fatal/pass-mi. 6a. 1 fatal per 100 mill pass-mi. 6b. 1 fatal per 34 bill pass-mi. 6c. 1 fatal per 2.7 bill pass-mi; higher.